Compendium: Mathematics & Physics

Nicolas Dewolf

June 20, 2019

Contents

Chapter 1

Introduction

This compendium has it roots in the need of a (then) compact summary of important theorems and formulas during physics and mathematics classes at university. When the interest in more (and more exotic) subjects grew, this collection lost its compactness and became the chaos it now is.

1.1 Conventions

Definitions, properties and formulas marked by a dagger symbol † are explained and/or derived in one of the appendices. This has been done such that the summary itself contains only core notions and theorems.

Definitions of words in the middle of a text will be indicated by the use of **bold font**. Terminology that has been defined in the past but that receives a new meaning/nuance will be indicated by *italic text*. Notions that have not been defined in this summary but that are relevant are also indicated by *italic text*.

Vectors in Euclidean space will be denoted by a bold font letter with an arrow above: $\vec{\boldsymbol{a}}$. Vectors in Minkowski space (4-vectors) and differential forms will be written without the arrow: $\mathbf{a}$. Matrices and tensors will always be represented by capital letters and dependent on the context we will use bold font or normal font. Objects from a general category will be denoted by a lower or upper case letter (depending on the context) while the categories themselves will be denoted by names in **bold font**.

Important references will be indicated at the beginning of each chapter.

Part I

Set Theory & Algebra

Chapter 2

Set Theory

2.1 Collections

Notation 2.1.1. Let X, Y be two sets. The set of maps $f : X \to Y$ is denoted by Y^X.

Definition 2.1.2 (Power set). Let S be a set. The power set is defined as the set of all subsets of S and is (often) denoted by $P(S)$ or 2^S. The existence of this set is stated by the *axiom of power set*.

Corollary 2.1.3. $S \subset P(S)$

Definition 2.1.4 (Collection). Let A be a set. A collection of elements in A is a subset of A.

Definition 2.1.5 (Family). Let A be a set and let I be another set, called the **index set**. A family of elements of A is a map $f : I \to A$. A family with index set I is often denoted by $(x_i)_{i\in I}$. In contrast to collections a family can 'contain' multiple copies of a single element.

Definition 2.1.6 (Helly family). A Helly family of order k is a pair (X, F) with $F \subset 2^X$ such that for every finite $G \subset F$:

$$\bigcap_{V\in G} V = \emptyset \implies \exists H \subseteq G : \left(\bigcap_{V\in H} V = \emptyset\right) \wedge \Big(|H| \leq k\Big) \tag{2.1}$$

A Helly family of order 2 is sometimes said to have the **Helly property**.

Definition 2.1.7 (Diagonal). Let S be a set. The diagonal of S is defined as follows:

$$\Delta_S = \{(a, a) \in S \times S : a \in S\} \tag{2.2}$$

Definition 2.1.8 (Partition). A partition of X is a family of disjoint subsets $(A_i)_{i\in I} \subset X$ such that $\bigcup_{i\in I} A_i = X$.

Definition 2.1.9 (Refinement). Let P be a partition of X. A refinement P' of P is a finite collection of subsets such that every $A \in P$ can be written as a disjoint union of elements in P'. Hence P' is also a partition.

Definition 2.1.10 (Cover). A cover of S is a collection of sets $\mathcal{F} \subseteq 2^S$ such that

$$\bigcup_{V\in\mathcal{F}} V = S \tag{2.3}$$

2.2 Set operations

Definition 2.2.1 (Symmetric difference).

$$A\Delta B = (A\backslash B) \cup (B\backslash A) \tag{2.4}$$

Definition 2.2.2 (Complement). Let Ω be the universal set . Let $E \subseteq \Omega$. The complement of E is defined as:

$$E^c = \Omega\backslash E \tag{2.5}$$

Formula 2.2.3 (de Morgan's laws).

$$\left(\bigcup_i A_i\right)^c = \bigcap_i A_i^c \tag{2.6}$$

$$\left(\bigcap_i A_i\right)^c = \bigcup_i A_i^c \tag{2.7}$$

2.3 Ordered sets

2.3.1 Posets

Definition 2.3.1 (Preordered set). A preordered set is a set equipped with a reflexive and transitive binary relation.

Definition 2.3.2 (Partially ordered set). A set P equipped with a binary relation $\leq$ is called a partially ordered set (**poset**) if the following 3 axioms are fulfilled for all elements $a, b, c \in P$:

1. Reflexivity: $a \leq a$
2. Antisymmetry: $a \leq b \wedge b \leq a \implies a = b$
3. Transitivity: $a \leq b \wedge b \leq c \implies a \leq c$

It is a preordered set for which the binary relation is also anti-symmetric.

Definition 2.3.3 (Totally ordered set). A poset P with the property that for all $a, b \in P$: $a \leq b$ or $b \leq a$ is called a (non-strict) totally ordered set. This property is called **totality**.

Definition 2.3.4 (Strict total order). A non-strict order $\leq$ has an associated strict order $<$ that satisfies $a < b \iff a \leq b \wedge a \neq b$.

Definition 2.3.5 (Maximal element). An element m of a poset P is maximal if for every $p \in P$, $m \leq p$ implies that $m = p$.

Definition 2.3.6 (Chain). A totally ordered subset of a poset is called a chain.

Theorem 2.3.7 (Zorn's lemma[1]). *Let $(P, \leq)$ be a poset. If every chain in P has an upper bound in P, then P has a maximal element.*

Definition 2.3.8 (Directed[2] set). A directed set is a set X equipped with a preorder $\leq$ and with the additional property that every 2-element subset has an upper bound, i.e. for every two elements $a, b \in X$ there exists an element $c \in X$ such that $a \leq c \wedge b \leq c$.

Definition 2.3.9 (Net). A net on a topological space X is a subset of X indexed by a directed set I.

[1] This theorem is equivalent to the *axiom of choice*.

[2] Sometimes called an *upward* directed set. Downward directed sets are analogously defined with a lower bound for every two elements. Directed sets are also sometimes called **filtered sets**.

2.3.2 Bounds

Definition 2.3.10 (Supremum). The supremum $\sup(X)$ of a poset X is the smallest upper bound of X.

Definition 2.3.11 (Infimum). The infimum $\inf(X)$ of a poset X is the greatest lower bound of X.

Definition 2.3.12 (Maximum). If $\sup(X) \in X$ the supremum is called the maximum of X. This is denoted by $\max(X)$.

Definition 2.3.13 (Minimum). If $\inf(X) \in X$ the supremum is called the minimum of X. This is denoted by $\min(X)$.

2.3.3 Lattices

Definition 2.3.14 (Semilattice). A poset $(P, \leq)$ for which every 2-element subset has a supremum (also called a **join**) in P is called a join-semillatice. Similarly, a poset $(P, \leq)$ for which every 2-element subset has an infimum (also called a **meet**) in P is called a meet-semilattice.

Notation 2.3.15. The join of $\{a, b\}$ is denoted by $a \wedge b$. The meet of $\{a, b\}$ is denoted by $a \vee b$.

Definition 2.3.16 (Lattice). A poset $(P, \leq)$ is called a lattice if it is both a join- and a meet-semilattice.

Definition 2.3.17 (Bounded lattice). A lattice $(P, \leq)$ is called bounded if it contains a greatest element (denoted by $\top$ or 1) and a smallest element (denoted by $\bot$ or 0) such that:

$$\bot \leq x \leq \top \tag{2.8}$$

for all $x \in P$.

Definition 2.3.18 (Heyting algebra). A bounded lattice H such that for every two elements $a, b \in H$ there exists a greatest element $x \in H$ for which

$$a \wedge x \leq b \tag{2.9}$$

This element is denoted by $a \to b$. The **pseudo-complement** $\neg a$ of an element $a \in H$ is then defined as $a \to \bot$.

Definition 2.3.19 (Frame). A poset $(P, \leq)$ which admits all joins[3] and all finite limits and for which the **infinite distributivity law** is satisfied:

$$y \wedge \left(\bigvee_{i \in I} x_i \right) = \bigvee_{i \in I} (y \wedge x_i) \tag{2.10}$$

2.3.4 Real numbers

Property 2.3.20 (First axiom). The set of real numbers is an ordered field $(\mathbb{R}, +, \cdot, <)$

Property 2.3.21 (Dedekind completeness[4]). Every non-empty subset of $\mathbb{R}$ that is bounded above has a supremum.

[3] When working with categories this has to be restricted to "all small joins", hence the index category should be a set.

[4] This form of the completeness axiom is also called the supremum property or the Dedekind completeness.

Property 2.3.22. $\mathbb{Q} \subset \mathbb{R}$

Remark. There is only one way to extend the field of rational numbers to the field of reals such that it satisfies the two previous axioms. This means that for every possible construction, their exists a bijection (isomorphism) between the two.

Definition 2.3.23 (Extended real line).

$$\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\} = [-\infty, \infty] \tag{2.11}$$

2.3.5 Filter

Definition 2.3.24 (Filter). Let X be a partially ordered set. A family $\mathcal{F} \subseteq 2^X$ is a filter on X if it satisfies following conditions:

1. $\emptyset \notin \mathcal{F}$
2. $\forall A, B \in \mathcal{F} : A \cap B \in \mathcal{F}$
3. If $A \in \mathcal{F}$ and $A \subseteq B$ then $B \in \mathcal{F}$

2.4 Algebra of sets

Definition 2.4.1 (Algebra of sets). A collection $\mathcal{F}$ of subsets of X is a called an algebra over X if it is closed under finite unions, finite intersections and complements. The pair $(X, \mathcal{F})$ is also called a **field of sets**.

2.4.1 σ-algebra

Definition 2.4.2 (σ-algebra). A collection of sets Σ is a σ-algebra over a set X if it satisfies the following 3 axioms:

1. $X \in \Sigma$
2. Closed under complements: $\forall E \in \Sigma : E^c \in \Sigma$
3. Closed under countable unions: $\forall \{E_i\}_{i=1}^n \subset \Sigma : \bigcup_{i=1}^n E_i \in \Sigma$

Remark 2.4.3. Axioms (2) and (3) together with de Morgan's laws[5] imply that a σ-algebra is also closed under countable intersections.

Corollary 2.4.4. Every algebra of sets is also a σ-algebra.

Property 2.4.5. The intersection of a family of σ-algebras is again a σ-algebra.

Definition 2.4.6. A σ-algebra $\mathcal{G}$ is said to be generated by a collection of sets $\mathcal{A}$ if

$$\mathcal{G} = \bigcap \{\mathcal{F} : \mathcal{F} \text{ is a } \sigma\text{-algebra that contains } \mathcal{A}\} \tag{2.12}$$

It is the smallest σ-algebra containing $\mathcal{A}$.

Notation 2.4.7. The σ-algebra generated by a collection of sets $\mathcal{A}$ is often denoted by $\mathcal{F}_\mathcal{A}$ or $\sigma(\mathcal{A})$.

[5] See equations 2.6 and 2.7.

Definition 2.4.8 (Borel set). Let $\mathcal{B}$ be the σ-algebra generated by all open[6] sets $O \subset X$. The elements $B \in \mathcal{B}$ are called Borel sets.

Definition 2.4.9 (Product σ-algebra). The smallest σ-algebra containing the products $A_1 \times A_2$ for all $A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2$ is called the product σ-algebra of $\mathcal{F}_1$ and $\mathcal{F}_2$.

Notation 2.4.10. The product σ-algebra of $\mathcal{F}_1$ and $\mathcal{F}_2$ is denoted by $\mathcal{F}_1 \times \mathcal{F}_2$.

Alternative Definition 2.4.11. The product σ-algebra $\mathcal{F}$ can also be equivalently defined in the following two ways:

1. $\mathcal{F}$ is generated by the collection
$$\mathcal{C} = \{A_1 \times \Omega_2 : A_1 \in \mathcal{F}_1\} \cup \{\Omega_1 \times A_2 : A_2 \in \mathcal{F}_2\}$$
2. $\mathcal{F}$ is the smallest σ-algebra such that the following projections are measurable (see 14.1.34):
$$\mathrm{Pr}_1 : \Omega \to \Omega_1 : (\omega_1, \omega_2) \mapsto \omega_1$$
$$\mathrm{Pr}_2 : \Omega \to \Omega_2 : (\omega_1, \omega_2) \mapsto \omega_2$$

Remark. Previous definitions can easily be generalized to higher dimensions.

2.4.2 Monotone class

Definition 2.4.12 (Monotone class). Let $\mathcal{A}$ be a collection of sets. $\mathcal{A}$ is called a monotone class if it has the following two properties:

- For every increasing sequence $A_1 \subset A_2 \subset ... :$
$$\bigcup_{i=1}^{+\infty} A_i \in \mathcal{A}$$
- For every decreasing sequence $A_1 \supset A_2 \supset ... :$
$$\bigcap_{i=1}^{+\infty} A_i \in \mathcal{A}$$

Theorem 2.4.13 (Monotone class theorem). *Let $\mathcal{A}$ be an algebra of sets 2.4.1. If $\mathcal{G}_\mathcal{A}$ is the smallest monotone class containing $\mathcal{A}$ then it coincides with the σ-algebra generated by $\mathcal{A}$.*

2.5 Functions

2.5.1 Domain

Definition 2.5.1 (Domain). Let $f : X \to Y$ be a function. The set X, containing the arguments of f, is called the domain of f.

Notation 2.5.2. The domain of f is denoted by $\mathrm{dom}(f)$.

Definition 2.5.3 (Support). Let $f : X \to \mathbb{R}$ be a function with an arbitrary domain X. The support of f is defined as the set of points where f is non-zero.

Notation 2.5.4. The support of f is denoted by $\mathrm{supp}(f)$

Remark. The support of a function is a subset of its domain.

Notation 2.5.5. Let X, Y be two sets. The set of functions $\{f : X \to Y\}$ is often denoted by X^Y.

[6] For $X = \mathbb{R}$ we find that open, closed and half-open (both types) intervals generate the same σ-algebra.

2.5.2 Codomain

Definition 2.5.6 (Codomain). Let $f : X \to Y$ be a function. The set Y, containing (at least) all the output values of f, is called the codomain of f.

Definition 2.5.7 (Image). Let $f : X \to Y$ be a function. The following subset of Y is called the image of f:

$$\{y \in Y \mid \exists x \in X : f(x) = y\} \tag{2.13}$$

It is denoted by $\text{im}(f)$.

Definition 2.5.8 (Level set). Let $f : X \to \mathbb{R}$ be a real-valued function and let $c \in \mathbb{R}$. The following set is called the level set of f:

$$L_c(f) = \{x \in X : f(x) = c\} \tag{2.14}$$

For $X = \mathbb{R}^2$ the level set is called a **level curve** and for $X = \mathbb{R}^3$ it is called the **level surface**.

2.6 Axiomatization

2.6.1 ZFC

The following set of axioms and axiom schemes gives a basis for axiomatic set theory whih fixes a number of issues in naieve set theory where one takes the notion of set for granted. This theory is called **Zermelo-Frenkel** set theory (ZF). When extended with the axiom of choice (see further) it is called ZFC, where the C stand for "choice".

Axiom 2.1 (Extensionality).

$$\forall x, y : \forall z \big[z \in x \iff z \in y\big] \implies x = y \tag{2.15}$$

Axiom 2.2 (Regularity).

$$\forall x : \exists z \big[z \in x\big] \implies \exists a \big[a \in x \wedge \neg \exists b (b \in a \wedge b \in x)\big] \tag{2.16}$$

This axiom says that for every non-empty set x one can find an element $a \in x$ such that x and a are disjoint. Among other things this axiom implies that no set can contain itself.

The following axiom is technically not an axiom but an axiom schema.

Axiom 2.3 (Specification). For every predicate φ one obtains an axiom:

$$\forall w_1, ..., w_n, A : \exists B : \forall x \big[x \in B \iff (x \in A \wedge \varphi(x, w_1, ..., w_n, A)\big] \tag{2.17}$$

This axiom (schema) says that for every set x one can build another set of elements in x which satisfy a given predicate. By the axiom of extensionality this subset $B \subseteq A$ is unique.

2.6.2 Material set theory

ZF(C) is an instance of material set theory. Every element of a set is a set itself and hence has internal structure.

Definition 2.6.1 (Pure set). A set U is pure if for every sequence $x_n \in x_{n-1} \in \cdots \in x_1 \in U$ all the elements x_i are also sets.

Definition 2.6.2 (Urelement[7]). An object which is not a set.

[7] Sometimes called an **atom**.

2.6.3 Von Neumann-Bernays-Gödel

2.6.4 Universes

Definition 2.6.3 (Grothendieck universe). A Grothendieck universe U is a set satisfying following axioms:

1. Transitivity: If $x \in U$ and $y \in x$ then $y \in U$
2. Power set: If $x \in U$ then $P(x) \in U$
3. If $x, y \in U$ then $\{x, y\} \in U$
4. Unions: If $I \in U$ and $\{x_i\}_{i \in I} \subset U$ then $\bigcup_{i \in I} x_i \in U$

2.6.5 Structural set theory

In contrast to material set theory the fundamental notions are sets and relations between them. An element of a set does not have any internal structure and so only becomes relevant if one specifies axtra structure (or relations) on the sets. This impies that elements of sets are not sets themselves as this is a meaningless statement since by default they lack internal structure. Even stronger, it is meaningless to compare two elements if one does not provide relations or extra structure on the sets.

2.6.6 ETCS

As this axiomatization

Remark. ETCS is the abbreviation of ''Elementary Theory of the Category of Sets''.

Axiom 2.4. ETCS The category of sets is a well-pointed (elementary) topos

Axiom 2.5. ETCS

Chapter 3

Algebra

3.1 Groups

Definition 3.1.1 (Semigroup). Let G be a set equipped with a binary operation $\star$. $(G, \star)$ is a semigroup if it satisfies following axioms:

1. G is closed under $\star$
2. $\star$ is asssociative

Definition 3.1.2 (Monoid). Let M be a set equipped with a binary operation $\star$. $(M, \star)$ is a monoid if it satisfies following axioms:

1. M is closed under $\star$
2. $\star$ is associative
3. M contains an identity element with respect to $\star$

Property 3.1.3 (Eckmann-Hilton argument). Let $(M, \circ), (M, \otimes)$ be two monoidal structures on a set M such that:

$$(a \circ b) \otimes (c \circ d) = (a \otimes c) \circ (b \otimes d) \tag{3.1}$$

for all $a, b, c, d \in M$. Then the two monoidal structures coincide and are in fact Abelian.

Definition 3.1.4 (Group). Let G be a set equipped with a binary operation $\star$. $(G, \star)$ is a group if it satisfies following axioms:

1. G is closed under $\star$
2. $\star$ is asssociative
3. G has an identity element with respect to $\star$
4. Every element in G has an inverse element with respect to $\star$

Definition 3.1.5 (Abelian group). Let $(G, \star)$ be a group. If $\star$ is commutative, then G is called an Abelian or **commutative** group.

Construction 3.1.6 (Grothendieck completion). Let the couple $(A, \boxplus)$ be an Abelian monoid. From this monoid one can construct an Abelian group $G(A)$, called the Grothendieck completion of A, as the quotient of $A \times A$ by the equivalence relation

$$(a_1, a_1') \sim (a_2, a_2') \iff \exists c \in A : a_1 + \boxplus a_2' \boxplus + c = a_1' \boxplus + a_2 \boxplus + c \tag{3.2}$$

The identity element is given by the equivalence class of $(0,0)$, which will be denoted by 0. By the definition of $G(A)$, this class contains all elements $\alpha \in \Delta_A$. From this last remark it follows that $[(a,b)] + [(b,a)] = 0$ which implies that the additive inverse of $[(a,b)]$ is given by $[(b,a)]$.

Example 3.1.7. The Grothendieck completion of the natural numbers $G(\mathbb{N})$ coincides with the additive group of integers $\mathbb{Z}$. The positive integers are then given by the equivalence classes $[(n,0)]$ and the negative integers are given by the classes $[(0,n)]$.

Universal property 3.1.8. Let $G(A)$ be the Grothendieck completion of A. For every monoid morphism $m : A \to B$ between an Abelian monoid and an Abelian group, there exists a group morphism $\varphi : G(A) \to B$.

Definition 3.1.9 (Totally ordered group). A totally ordered group (sometimes) is a group G together with a total order[1] such that:

- $a \leq b \implies ac \leq bc$
- $a \leq b \implies ca \leq cb$

for all $a, b, c \in G$. If only the first (resp. second) property holds then the group is said to be left- (resp. right-) ordered.

3.1.1 Cosets

Definition 3.1.10 (Coset). Let G be a group and H a subgroup of G. The left coset of H with respect to $g \in G$ is defined as the set

$$gH = \{gh : h \in H\} \tag{3.3}$$

The right coset is analogously defined as Hg. If for all $g \in G$ the left and right cosets coincide then the subgroup H is said to be a **normal subgroup**[2]. The sets of left and right cosets are denoted by G/H and $H\backslash G$ respectively.

Alternative Definition 3.1.11 (Normal subgroup). Let G be a group with subgroup H. Consider the conjugate classes gHg^{-1} for all $a \in G$. If all classes coincide with H itself, then H is called a normal subgroup.

Notation 3.1.12. Let N be a normal subgroup of G. This is often denoted by $N \lhd G$.

Definition 3.1.13 (Quotient group). Let G be a group and N a normal subgroup. The coset space[3] G/N can be turned into a group by equipping it with a product such that the product of aN and bN is $(aN)(bN)$. The fact that N is a normal subgroup can be used to rewrite this as $(aN)(bN) = (ab)N$.

Definition 3.1.14 (Center). The center of a group is defined as follows:

$$Z(G) = \{z \in G : \forall g \in G, zg = gz\} \tag{3.4}$$

This set is a normal subgroup of G.

[1] See definition 2.3.3.

[2] Also called a **normal divisor** or **invariant subgroup**.

[3] Left and right cosets coincide for normal subgroups.

3.1.2 Abelianization

Definition 3.1.15 (Commutator subgroup[4]). The commutator subgroup $[G, G]$ of G is defined as the group generated by the elements

$$[g, h] = g^{-1}h^{-1}gh$$

for all $g, h \in G$. This group is a normal subgroup of G.

Property 3.1.16. G is Abelian if and only if $[G, G]$ is trivial.

Definition 3.1.17 (Abelianization). The quotient group $G/[G, G]$ is an Abelian group, called the Abelianization of G.

Property 3.1.18. A quotient group G/H is Abelian if and only if $[G, G] \leq H$.

3.1.3 Order

Definition 3.1.19 (Order of a group). The number of elements in the group. It is denoted by $|G|$ or $\text{ord}(G)$.

Definition 3.1.20 (Order of an element). The order of an element $a \in G$ is the smallest integer n such that

$$a^n = e \tag{3.5}$$

where e is the identity element of G.

Definition 3.1.21 (Torsion group). A torsion group is a group for which all element have finite order. The torsion set $\text{Tor}(G)$ of a group G is the set of all elements $a \in G$ that have finite order. For Abelian groups, $\text{Tor}(G)$ is a subgroup.

Theorem 3.1.22 (Lagrange). *Let G be a finite group with subgroup H. Then $|H|$ is a divisor of $|G|$.*

Corollary 3.1.23. The order of any element $g \in G$ is a divisor of $|G|$.

3.1.4 Symmetric and alternating groups

Definition 3.1.24 (Symmetric group). The symmetric group S_n or Sym_n of the set $V = \{1, 2, ..., n\}$ is defined as the set of all permutations of V. The number n is called the **degree** of the symmetric group. The symmetric group $\text{Sym}(X)$ of a finite set X is analogously defined.

Theorem 3.1.25 (Cayley's theorem). *Every finite group is isomorphic to a subgroup of S_n where $n = |G|$.*

Definition 3.1.26 (Alternating group). The alternating group A_n is the subgroup of S_n containing all even permutations.

Definition 3.1.27 (Cycle). A k-cycle is a permutation of the form $(a_1\ a_2\ ...\ a_k)$ sending a_i to a_{i+1} (and a_k to a_1). A **cycle decomposition** of an arbitrary permutation is the decomposition into a product of disjoint cycles.

Formula 3.1.28. Let τ be a k-cycle. Then τ is k-cyclic (hence the name *cycle*):

$$\tau^k = \mathbb{1}_G \tag{3.6}$$

[4]Also called the **derived subgroup**.

Example 3.1.29. Consider the set $\{1, 2, 3, 4, 5, 6\}$. The permutation $\sigma : x \mapsto x + 2 \pmod 6$ can be written using the cycle decomposition $\sigma = (1\ 3\ 5)(2\ 4\ 6)$.

Definition 3.1.30 (Transposition). A permutation which exchanges two elements but lets the other ones unchanged.

3.1.5 Group presentations

Definition 3.1.31 (Relations). Let G be a group. If the product of a number of elements $g \in G$ is equal to the identity e then this product is called a relation on G.

Definition 3.1.32 (Complete set of relations). Let H be a group generated by a subgroup G. Let R be a set of relations on G. If H is uniquely (up to an isomorphism) determined by G and R then the set of relations is said to be complete.

Definition 3.1.33 (Presentation). Let H be a group generated by a subgroup G and a complete set of relations R on G. The pair (G, R) is called a presentation of H.

It is clear that every group can have many different presentations and that it is (very) difficult to tell if two groups are isomorphic by just looking at their presentations.

Notation 3.1.34. The presentation of a group G is often denoted by $\langle S|R\rangle$, where S is the set of generators and R the set of relations.

3.1.6 Direct product

Definition 3.1.35 (Direct product). Let G, H be two groups. The direct product $G \otimes H$ is defined as the set-theoretic Cartesian product $G \times H$ equipped with a binary operation $\cdot$ such that:

$$(g_1, h_1) \cdot (g_2, h_2) = (g_1 g_2, h_1 h_2) \tag{3.7}$$

where the operations on the right hand side are the group operations in G and H. The structure $G \otimes H = (G \times H, \cdot)$ forms a group.

Remark 3.1.36. This definition can be generalized to any number of groups, even infinity (where one has to replace the n-tuples by infinite Cartesian products).

Definition 3.1.37 (Weak direct product). Consider the direct product of a number (finite or infinite) of groups. The subgroup consisting of all elements for which all components, except finitely many, are the identity is called the weak direct product or, in the case of Abelian groups, the **direct sum**.

Notation 3.1.38. The direct sum is often denoted by $\oplus$, in accordance with the notation for vector spaces (and other algebraic structures).

Remark 3.1.39. For a finite number of groups, the direct product and direct sum coincide.

Definition 3.1.40 (Inner semidirect product). Let G be a group, H a subgroup of G and N a normal subgroup of G. G is said to be the inner semidirect product of H and N, denoted by $N \rtimes H$, if it satifies the following equivalent statements:

- $G = NH$ where $N \cap H = \{e\}$.
- For every $g \in G$ there exist unique $n \in N, h \in H$ such that $g = nh$.
- For every $g \in G$ there exist unique $h \in H, n \in N$ such that $g = hn$.

- There exists a group morphism $\rho : G \to H$ which satisfies $\rho|_H = e$ and $\ker(\rho) = N$.
- The composition of the natural embedding $i : H \to G$ and the projection $\pi : G \to G/N$ is an isomorphism between H and G/N.

G is also said to **split** over N.

Property 3.1.41. If both H and N are normal in the above definition, the inner semidirect product coincides with the direct product. For a finite number of groups $\{G_i\}$ we see that the direct product is generated by the elements of the groups G_i.

If the subgroups H and N have presentations $\langle S_H | R_H \rangle$ and $\langle S_N | R_N \rangle$ then the (inner) direct product is given by:

$$H * N = \langle S_H \cup S_N | R_H \cup R_N \cup R_C \rangle \tag{3.8}$$

where R_C is the set of relations that give the commutativity of H and N.

Definition 3.1.42 (Outer semidirect product). Let G, H be two groups and let $\varphi : H \to \mathrm{Aut}(G)$ be a group morphism. The outer semidirect product $G \rtimes_\varphi H$ is defined as the set-theoretic Cartesian product $G \times H$ equipped with a binary relation $\cdot$ such that:

$$(g_1, h_1) \cdot (g_2, h_2) = (g_1 \varphi(h_1)(g_2), h_1 h_2) \tag{3.9}$$

The structure $(G \rtimes_\varphi H, \cdot)$ forms a group.

By noting that the set $N = \{(g, e_H) | g \in G\}$ is a normal subgroup isomorphic to G and that the set $B = \{(e_G, h) | h \in H\}$ is a subgroup isomorphic to H, we can also construct the outer semidirect product $G \rtimes_\varphi H$ as the inner semidirect product $N \rtimes B$.

Remark 3.1.43. The direct product of groups is a special case of the outer semidirect product where the group morphism is given by the trivial map $\varphi : h \mapsto e_G$.

3.1.7 Free product

Definition 3.1.44 (Free product). Consider two groups G, H. The free group $G * H$ is deifned as the set consisting of all words composed of all elements in G and H together with the concatenation (and reduction[5]) as multiplication. Due to the reduction, every element in $G * H$ is of the form $g_1 h_1 g_2 h_2 ....$

Remark 3.1.45. For non-trivial groups the free product is always infinite.

Property 3.1.46. If the groups G and H have presentations $\langle S_G | R_G \rangle$ and $\langle S_H | R_H \rangle$ then the free product is given by:

$$G * H = \langle S_G \cup S_H | R_G \cup R_H \rangle \tag{3.10}$$

From 3.8 we see that the free product is a generalization of the direct product.

Definition 3.1.47 (Free product with amalgamation). Consider the groups F, G, H and two group morphisms $\phi : F \to G$ and $\psi : F \to H$. The free product with amalgamation $G *_F H$ is defined by adding the following set of relations to the presentation of the free product $G * H$:

$$\{\phi(f)\psi(f)^{-1} = e : f \in F\} \tag{3.11}$$

Alternatively, the free product with amalgamation can be constructed as

$$G *_F H = (G * H)/N_F \tag{3.12}$$

where N_F is the normal subgroup generated by elements of the form $\phi(f)\psi(f)^{-1}$.

[5]Two elements of the same group, written next to eachother, are replaced by their product.

3.1.8 Free groups

Definition 3.1.48 (Free Abelian group). An Abelian group G with generators $\{g_i\}_{i\in I}$ is said to be freely generated if every element $g \in G$ can be uniquely written as a formal linear combination of the generators:

$$G = \left\{ \sum_i a_i g_i \middle| a_i \in \mathbb{Z} \right\} \tag{3.13}$$

The set of generators $\{g_i\}_{i\in I}$ is then called a **basis**[6] of G. The number of elements in the basis is called the **rank** of G.

Property 3.1.49. Consider a free group G. Let $H \subset G$ be a subgroup. Then H is also free.

Theorem 3.1.50. *Let G be a finitely generated Abelian group of rank n. This group can be constructed in two different ways:*

$$G = F/H \tag{3.14}$$

where both F, H are free and finitely generated Abelian groups. The second decomposition is:

$$G = A \oplus T \qquad \text{where} \qquad T = Z_{h_1} \oplus \cdots \oplus Z_{h_m} \tag{3.15}$$

where A is a free and finitely generated group of rank $n - m$ and all Z_{h_i} are cyclic groups of order h_i. The group T is called the ***torsion subgroup***[7]*.*

Property 3.1.51. The rank $n - m$ and the numbers h_i from previous theorem are unique.

3.1.9 Group actions

Definition 3.1.52 (Group morphism). A group morphism $\Phi : G \to H$ is a map satisfying $\forall g, h \in G$

$$\Phi(gh) = \Phi(g)\Phi(h) \tag{3.16}$$

Definition 3.1.53 (Kernel). The kernel of a group morphism $\Phi : G \to H$ is defined as the set

$$K = \{g \in G : \Phi(g) = \mathbb{1}_H\} \tag{3.17}$$

Theorem 3.1.54 (First isomorphism theorem). *Let G, H be a groups and let $\varphi : G \to H$ be a group morphism. If φ is surjective than $G/\ker\varphi \cong H$.*

Definition 3.1.55 (Group action). Let G be a group. Let V be a set. A map $\rho : G \times V \to V$ is called an action of G on V if it satisfies the following conditions:

- Identity: $\rho(\mathbb{1}_G, v) = v$
- Compatibility: $\rho(gh, v) = \rho(g, \rho(h, v))$

For all $g, h \in G$ and $v \in V$. The set V is called a (left) **G-space**.

Remark 3.1.56. A group action can alternatively be defined as a group morphism from G to $\mathrm{Sym}(V)$. It assigns a permutation of V to every element $g \in G$.

Notation 3.1.57. The action $\rho(g, v)$ is often denoted by $g \cdot v$ or even gv.

[6] In analogy with the basis of a vector space.
[7] See also definition 3.1.21.

Definition 3.1.58 (Orbit). The orbit of an element $x \in X$ with respect to a group G is defined as the set:

$$G \cdot x = \{g \cdot x | g \in G\} \tag{3.18}$$

The relation $p \sim q \iff \exists g \in G : p = g \cdot q$ induces an equivalence relation for which the equivalence classes coincide with the orbits of G. The set of equivalence classes $X/\sim$ (sometimes denoted by X/G) is called the **orbit space**.

Definition 3.1.59 (Stabilizer). The stabilizer group or **isotropy group** of an element $x \in X$ with respect to a group G is defined as the set:

$$G_x = \{g \in G | g \cdot x = x\} \tag{3.19}$$

This is a subgroup of G.

Theorem 3.1.60 (Orbit-stabilizer theorem). *Let G be a group acting on a set X. Let G_x be the stabilizer of some $x \in X$. The following relation holds:*

$$|G \cdot x||G_x| = |G| \tag{3.20}$$

Definition 3.1.61 (Free action). A group action is free if $g \cdot x = x$ implies $g = e$ for every $x \in X$. Equivalently, a group action is free if the stabilizer group of all elements is trivial.

Definition 3.1.62 (Faithful action). A group action is faithful or **effective** if the morphism $G \to \mathrm{Sym}(X)$ is injective. Alternatively, a group action is faithful if for every two group elements $g, h \in G$ there exists an element $x \in X$ such that $g \cdot x \neq h \cdot x$.

Definition 3.1.63 (Transitive action). A group action is transitive if for every two elements $x, y \in X$ there exists a group element $g \in G$ such that $g \cdot x = y$. Equivalently we can say that there is only one orbit.

Definition 3.1.64 (Homogeneous space). If the group action of a group G on a G-space X is transitive, then X is said to be a homogeneous space.

Property 3.1.65 (†). Let X be a set and let G be a group such that the action of G on X is transitive. Then there exists a bijection $X \cong G/G_x$ where G_x is the stabilizer of any element $x \in X$.

Definition 3.1.66 (Principal homogenous space). If the group action of a group G on a homogeneous space X is also free, then X is said to be a principal homogeneous space or **G-torsor**.

An example of an $\mathbb{R}^n$-torsor is the n-dimensional affine space.

Definition 3.1.67 (G-module). Let G be a group. Let M be a commutative group. M equipped with a left group action $\varphi : G \times M \to M$ is a (left) G-module if φ satisfies the following equation (distributivity):

$$g \cdot (a + b) = g \cdot a + g \cdot b \tag{3.21}$$

where $a, b \in M$ and $g \in G$.

Definition 3.1.68 (G-module morphism). A G-module morphism is a map $f : V \to W$ satisfying

$$g \cdot f(v) = f(g \cdot v) \tag{3.22}$$

where the $\cdot$ symbol represents the group action in W and V respectively. It is sometimes called a **G-map**, a **G-equivariant map** or an **intertwining map**.

Definition 3.1.69 (Crossed module). A crossed module is a quadruple (G, H, t, α) where:

- G, H are two groups.
- t is a group morphism $H \to G$.
- α is a group morphism $G \to \mathrm{Aut}(H)$, i.e. G acts by automorphisms on H.

These structures satisfy two compatibility conditions:

- t is G-equivariant:
$$t(\alpha(g)h) = gt(h)g^{-1} \tag{3.23}$$
- **Peiffer identity**:
$$\alpha(t(h))h' = hh'h^{-1} \tag{3.24}$$

3.1.10 Group cohomology

Definition 3.1.70 (Group cohomology). Consider a group G together with a G-module A. First we define the chain group as
$$C^k(G; A) = \{\text{all functions}^8 \text{ from } G^k \text{ to } A\} \tag{3.25}$$
The coboundary map $d^k : C^k(G; A) \to C^{k+1}(G; A)$ is then defined as follows:
$$\begin{aligned} d^k f(g_1, ..., g_k, g_{k+1}) = g_1 \cdot f(g_2, ...,&g_k, g_{k+1}) + (-1)^{k+1} f(g_1, ..., g_k) \\ &+ \sum_{i=1}^{k} (-1)^{i+1} f(g_1, ..., g_i g_{i+1}, ..., g_k, g_{k+1}) \end{aligned} \tag{3.26}$$
The cohomology groups are then defined by
$$H^k(G; A) = \frac{\ker(d^k)}{\mathrm{im}(d^k)} \tag{3.27}$$

3.2 Rings

Definition 3.2.1 (Ring). Let R be a set equipped with two binary operations $+, \cdot$ (called addition and multiplication). $(R, +, \cdot)$ is a ring if it satisfies the following axioms:

1. $(R, +)$ is a commutative group.
2. $(R, \cdot)$ is a monoid.
3. Multiplication is distributive with respect to addition.

Definition 3.2.2 (Unit). An invertible element of ring $(R, +, \cdot)$. The set of units forms a group under multiplication.

Construction 3.2.3 (Localization). Let R be a commutative ring and let S be a multiplicative monoid in R. We first define an equivalence relation $\sim$ on $R \times S$ in the following way:
$$(r_1, s_1) \sim (r_2, s_2) \iff \exists t \in S : t(r_1 s_2 - r_2 s_1) = 0 \tag{3.28}$$
The set $R^* = (R \times S)/\sim$, called the localization of R with respect to S, can now be turned into a ring by defining an addition and a multiplication. By writing $(r, s) \in R^*$ as the formal fraction $\frac{r}{s}$ we obtain the familiar operations of fractions:

[8] This means all set-theoretic functions.

- $\frac{r_1}{s_1} + \frac{r_2}{s_2} = \frac{r_1 s_2 + r_2 s_1}{s_1 s_2}$
- $\frac{r_1}{s_1} \cdot \frac{r_2}{s_2} = \frac{r_1 r_2}{s_1 s_2}$

Remark 3.2.4. The localization of R with respect to the monoid S can be interpreted as the ring obtained by collapsing S into a single unit of R.

Notation 3.2.5. The localization of R with respect to S is often denoted by $S^{-1}R$.

Definition 3.2.6 (Integral domain). A commutative ring R in which the product of two nonzero elements is again nonzero.

Definition 3.2.7 (Valuation). Let K be a field and let Γ be a totally ordered Abelian group[9]. First we extend the group law of Γ to the union $\Gamma \cup \{\infty\}$ in the following way:

- $g + \infty = \infty + g = \infty$ for all $g \in \Gamma$
- $g \leq \infty$ for all $g \in \Gamma$

A valuation of K is a map $\nu : K \to \Gamma \cup \{\infty\}$ such that:

- $\nu(a) = \infty \iff a = 0$
- $\nu(ab) = \nu(a) + \nu(b)$
- $\min(\nu(a), \nu(b)) \leq \nu(a + b)$ where the equality holds if and only if $\nu(a) \neq \nu(b)$

3.2.1 Ideals

Definition 3.2.8 (Ideal). Let $(R, +, \cdot)$ be a ring with $(R, +)$ its additive group. A subset $I \subseteq R$ is called an ideal[10] of R if it satisfies the following conditions:

1. $(I, +)$ is a subgroup of $(R, +)$
2. $\forall n \in I, \forall r \in R : (n \cdot r), (r \cdot n) \in I$

Definition 3.2.9 (Unit ideal). Let $(R, +, \cdot)$ be a ring. R itself is called the unit ideal.

Definition 3.2.10 (Proper ideal). Let $(R, +, \cdot)$ be a ring. A subset $I \subset R$ is said to be a proper ideal if it is an ideal of R and if it is not equal to R.

Definition 3.2.11 (Prime ideal). Let $(R, +, \cdot)$ be a ring. A proper ideal I is a prime ideal if for any $a, b \in R$ the following relation holds:

$$ab \in I \implies \text{a} \in I \vee b \in I \tag{3.29}$$

Definition 3.2.12 (Maximal ideal). Let $(R, +, \cdot)$ be a ring. A proper ideal I is said to be maximal if there exists no other proper ideal T in R such that $I \subset T$.

Definition 3.2.13 (Minimal ideal). A proper ideal is said to be minimal if it contains no other nonzero ideal.

Construction 3.2.14 (Generating set of an ideal). Let R be a ring and let X be a subset of R. The two-sided ideal generated by X is defined as the intersection of all two-sided ideals containing X. An explicit construction is given by:

$$I = \left\{ \sum_{i=1}^{n} l_i x_i r_i \;\middle|\; \forall n \in \mathbb{N} : \forall l_i, r_i \in R \text{ and } x_i \in X \right\} \tag{3.30}$$

Left and right ideals are generated in a similar fashion.

[9] See definition 3.1.9.

[10] More generally: two-sided ideal

Definition 3.2.15 (Principal ideal). An ideal which is generated by a single element.

Definition 3.2.16 (Principal ideal domain). An integral domain[11] in which every ideal is principal.

Construction 3.2.17 (Extension). Let I be an ideal of a ring R and let $\iota : R \to S$ be a ring morphism. The extension of I with respect to ι is the ideal generated by the set $\iota(I)$.

Definition 3.2.18 (Local ring). A local ring is a ring for which a unique maximal left ideal exists.[12]

Property 3.2.19. The localization of a ring R with respect to a prime ideal P is a local ring, where the maximal ideal is the extension of P with respect to the ring morphism $\iota : R \to R^*$

Definition 3.2.20 (Residue field). Consider a local ring R and let I be its maximal ideal. The quotient ring R/I forms a field, called the residue field.

3.2.2 Modules

Definition 3.2.21 (R-Module). Let $(R, +, \cdot)$ be a ring. A set X is an R-module if it satisfies the same axioms as those of a vector space 19.2.1 but where the scalars are only elements of a ring instead of a field.

Definition 3.2.22 (Free module). A module is said to be free if it admits a basis.

Property 3.2.23. For a general R-module the existence of a basis is not guaranteed unless R is a division ring. See construction 19.2.10 to see how this basis can be constructed.

Corollary 3.2.24. As every field is in particular a division ring, the existence of a basis follows from the above property for R-modules.

Definition 3.2.25 (Projective module). A module P is said to be projective if:

$$P \oplus M = F \tag{3.31}$$

where M is a module and F is a free module.

Example 3.2.26 (Dual numbers). Let R be a ring. The R-algebra of dual numbers over R, often denoted by $R[\varepsilon]$, is defined as the free R-module with basis $1, \varepsilon$ where $\varepsilon^2 = 0$.

3.2.3 Semisimple modules

Definition 3.2.27 (Semisimple ring). A ring is said to be semisimple if it is semisimple as a module over itself.

Theorem 3.2.28. *A ring is semisimple if and only if it is Artinian and its Jacobson radical vanishes.*

[11] See definition 3.2.6.

[12] This also implies that there exists a unique maximal right ideal and these ideals coincide.

3.2.4 Graded rings

Definition 3.2.29 (Graded ring). Let R be a ring that can be written as the direct sum of Abelian groups A_k:

$$R = \bigoplus_{k\in\mathbb{N}} A_k \tag{3.32}$$

If R has the property that for every $i, j \in \mathbb{N} : A_i \star A_j \subseteq A_{i+j}$, where $\star$ is the ring multiplication, then R is said to be a graded ring. The elements of the space A_k are said to be **homogeneous of degree** k.

Formula 3.2.30 (Graded commutativity). Let $m = \deg v$ and let $n = \deg w$. If

$$vw = (-1)^{mn} wv \tag{3.33}$$

for all elements v, w of the graded ring then it is said to be a graded-commutative ring.

3.3 Other algebraic structures

3.3.1 Direct systems

Definition 3.3.1 (Direct system). Let $(I, \leq)$ be a directed set[13]. Let $\{A_i\}_{i\in I}$ be a family of algebraic objects (groups, rings, ...) and let $\{f_{ij} : A_i \to A_j\}_{i,j\in I}$ be a set of morphisms with the following properties:

- For every $i \in I$: $f_{ii} = e_i$, where e_i is the identity in A_i.
- For every $i \leq j \leq k \in I$: $f_{ik} = f_{jk} \circ f_{ij}$.

The pair (A_i, f_{ij}) is called a direct system over I.

Definition 3.3.2 (Direct limit[14]). Consider a direct system (A_i, f_{ij}) over a (directed) set I. The direct limit A of this direct system is defined as follows:

$$\varinjlim A_i = \bigsqcup_{i\in I} A_i \Big/ \sim \tag{3.34}$$

where the equivalence relation is given by $x \in A_i \sim y \in A_j \iff \exists k \in I : f_{ik}(x) = f_{jk}(y)$. Informally put: two elements are equivalent if they eventually become the same.

The algebraic operations on A are defined such that the inclusion maps $\phi_i : A_i \to A$ are morphisms.

3.3.2 Inverse systems

Definition 3.3.3 (Inverse system). Let $(I, \leq)$ be a directed set[15]. Let $\{A_i\}_{i\in I}$ be a family of algebraic objects (groups, rings, ...) and let $\{f_{ij} : A_j \to A_i\}_{i,j\in I}$ be a set of morphisms with the following properties:

- For every $i \in I$: $f_{ii} = e_i$, where e_i is the identity in A_i.

[13] See definition 2.3.8.
[14] Also called a **inductive limit**.
[15] See definition 2.3.8.

- For every $i \leq j \leq k \in I$: $f_{ik} = f_{ij} \circ f_{jk}$.

The pair (A_i, f_{ij}) is called an inverse system over I.

Definition 3.3.4 (Inverse limit[16]). Consider an inverse system (A_i, f_{ij}) over a (directed) set I. The inverse limit A of this inverse system is defined as follows:

$$\varprojlim A_k = \left\{ \vec{a} \in \prod_{i \in I} A_i \,\middle|\, a_i = f_{ij}(a_j), \forall i \leq j \in I \right\} \tag{3.35}$$

For all $i \in I$ there exists a **natural projection** $\pi_i : \varprojlim A_k \to A_i$.

Universal property 3.3.5. Let A be the inverse limit of a given inverse system. For any other inverse limit A' over the same inverse system, there exists a unique isomorphism $\phi : A \to A'$.

Remark 3.3.6. The direct and inverse limit are each other's (categorical) dual. The former is a colimit while the latter is a limit in category theory.

3.3.3 Exact sequences

Definition 3.3.7 (Exact sequence). Consider a sequence (finite or infinite) of algebraic structures and their corresponding homomorphisms:

$$A_0 \xrightarrow{\Phi_1} A_1 \xrightarrow{\Phi_2} \cdots \xrightarrow{\Phi_n} A_n \tag{3.36}$$

The sequence is exact if for every $k \in \mathbb{N} : \mathrm{im}(\Phi_k) = \ker(\Phi_{k+1})$. This implies that $\Phi_{k+1} \circ \Phi_k = 0$ for all $h \in \mathbb{N}$. It follows that exact sequences are a special type of chain complex as defined in definition 6.1.1.

Definition 3.3.8 (Short exact sequence). A short exact sequence is an exact sequence of the form:

$$0 \to A_0 \xrightarrow{\Phi_1} A_1 \xrightarrow{\Phi_2} A_3 \to 0 \tag{3.37}$$

A long exact sequence is an infinite exact sequence.

Property 3.3.9. Looking at some small examples we can derive some important constraints for certain exact sequences and especially for short exact sequences. Consider the sequence

$$0 \to A \xrightarrow{\Phi} B$$

This sequence can only be exact if Φ is an injective homomorphism (**monomorphism**). This follows from the fact that the only element in the image of the map $0 \to A$ is 0 because the map is a homomorphism. The kernel of Φ is thus trivial which implies that Φ is injective.

Analogously, the sequence

$$A \xrightarrow{\Psi} B \to 0$$

is exact if Ψ is a surjective homomorphism (**epimorphism**). This follows from the fact that the kernel of the map $B \to 0$ and thus the image of Ψ is all of B which implies that Ψ is surjective.

It follows that the sequence

$$0 \to A \xrightarrow{\Sigma} B \to 0$$

is exact if Σ is a **bimorphism** (which is often an isomorphism).

[16] Also called a **projective limit**.

3.4 Integers

3.4.1 Partition

Definition 3.4.1 (Composition). Let $n \in \mathbb{N}$. A k-composition of n is a k-tuple $(t_1, ..., t_k)$ such that $\sum_{i=1}^{k} t_k = n$.

Definition 3.4.2 (Partition). Let $n \in \mathbb{N}$. A partition of n is an ordered composition of n.

Definition 3.4.3 (Young diagram[17]). A Young diagram is a visual representation of the partition of an integer n. It is a left justified system of boxes, where every row corresponds to a part of the partition.

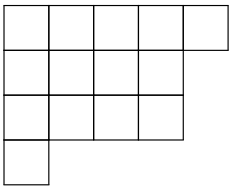


Figure 3.1: A Young diagram representing the partition $(5, 4, 4, 1)$ of 14.

Definition 3.4.4 (Conjugate partition). Let λ be a partition of n with Young diagram $\mathcal{D}$. The conjugate partition λ' is obtained by reflecting $\mathcal{D}$ across its main diagonal.

Example 3.4.5. Conjugating diagram 3.1 gives us the partition $(4, 3, 3, 3, 1)$ which is represented by:

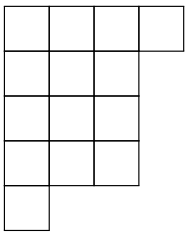


Figure 3.2: A Young diagram representing the partition $(4, 3, 3, 3, 1)$ of 14.

3.4.2 Superpartition

Definition 3.4.6 (Superpartition). Let $n \in \mathbb{N}$. A superpartition in the m-fermion sector is a sequence of integers of the following form:

$$\Lambda = (\Lambda_1, ..., \Lambda_m; \Lambda_{m+1}, ..., \Lambda_N) \tag{3.38}$$

where the first m numbers are strictly ordered, i.e. $\Lambda_i > \Lambda_{i+1}$ for all $i < m$, and the last $N - m$ numbers form a normal partition.

Both sequences, separated by a semicolon, form in fact distinct partitions themself. The first one represents the antisymmetric fermionic sector (this explains the strict order) and the second one represents the symmetric bosonic sector. This amounts to the following notation:

$$\Lambda = (\lambda^a; \lambda^s)$$

The degree of the superpartition is given by $n \equiv |\Lambda| = \sum_{i=1}^{N}$.

[17] Sometimes called a **Ferrers diagram**.

Notation 3.4.7. A superpartition of degree n in the m-fermion sector is said to be a superpartition of $(n|m)$. To every superpartition Λ we can also associate a unique partition Λ^* by removing the semicolon and reordering the numbers such that they form a partition of n. The superpartition Λ can then be represented by the Young diagram belonging to Λ^* where the rows belonging to the fermionic sector are ended by a circle.

Part II

Category Theory

Chapter 4

Category theory

4.1 Categories and subcategories

Definition 4.1.1 (Subcategory). Let $\mathbf{C}$ be a category. A subcategory $\mathbf{S}$ of $\mathbf{C}$ consists of a subcollection of objects ob_S and a subcollection of morphisms hom_S that satisfy following conditions:

- For every object in $\mathrm{ob}(\mathbf{S})$ the identity morphism is an element of hom_S.
- For every morphism in hom_S both the source and target are elements of $\mathrm{ob}(\mathbf{S})$.
- For every pair of morphisms in hom_S the composition is also an element of hom_S.

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

$$\mathbf{S}(X, Y) = \mathbf{C}(X, Y) \tag{4.1}$$

Definition 4.1.2 (Small category). A category $\mathbf{C}$ is said to be small if both $\mathrm{ob}(\mathbf{C})$ and $\mathrm{hom}(\mathbf{C})$ 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 to a small category is said to be **essentially small**.

Definition 4.1.3 (Opposite category). Let $\mathbf{C}$ be a category. The opposite category $\mathbf{C}^{op}$ is defined by reversing all arrows in $\mathbf{C}$.

Property 4.1.4. From the definition of the opposite category it easily follows that

$$(\mathbf{C}^{op})^{op} = \mathbf{C} \tag{4.2}$$

i.e. *op* is an involution.

Definition 4.1.5 (Skeletal category). A category is said to be skeletal if isomorphic objects are identical, i.e. every isomorphism is an identity morphism. The **skeleton** of a category is an equivalent skeletal category (often taken to be a subcategory).

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

Definition 4.1.6 (Enriched category). Let $(\mathbf{M}, \otimes, \mathbf{1})$ be a monoidal category[1]. An enriched category over $\mathbf{M}$, also called an $\mathbf{M}$-category, consists of following elements:

[1] See definition 4.9.1 further below.

- A collection of objects $\mathrm{ob}(\mathbf{C})$.
- For every pair of object $A, B \in \mathrm{ob}(\mathbf{C})$ there is an object $\mathbf{C}(A, B) \in \mathrm{ob}(\mathbf{M})$ for which the following morphisms exist:
  - $\mathrm{id}_A : \mathbf{1} \to \mathbf{C}(A, A)$
  - $\circ_{ABC} : \mathbf{C}(B, C) \otimes \mathbf{C}(A, B) \to \mathbf{C}(A, C)$ replacing the usual composition

The associativity and identity morphisms from ordinary categories are given by commutative diagrams of the id and $\circ$ morphisms together with the associators and unitors in $\mathbf{M}$.

Definition 4.1.7 (Category with weak equivalences). A category $\mathbf{C}$ with a subcategory $\mathbf{W}$ such that:

- $\mathbf{W}$ contains all isomorphisms in $\mathbf{C}$.
- Any two composable morphisms $f, g \in \hom_W$ satisfy the 2-out-of-3 property: If any two of $\{f, g, f \circ g\}$ are in $\mathbf{W}$ then so is the third.

4.2 Functors

Definition 4.2.1 (Covariant functor). Let $\mathbf{A}, \mathbf{B}$ be categories. A (covariant) functor F is a map $\mathbf{A} \to \mathbf{B}$ satisfying following conditions:

- F maps every object $X \in \mathrm{ob}(\mathbf{A})$ to an object $FX \in \mathrm{ob}(\mathbf{B})$.
- F maps every morphism $\phi \in \mathbf{A}(X, Y)$ to a morphism $F(\phi) \in \mathbf{B}(FX, FY)$.
- $F(\mathbb{1}_X) = \mathbb{1}_{FX}$
- $F(\phi \circ \psi) = F(\phi) \circ F(\psi)$

Definition 4.2.2 (Contravariant functor). Let A, B be categories. A contravariant functor F is a map $A \to B$ satisfying following conditions:

- F maps every object $X \in \mathrm{ob}(\mathbf{A})$ to an object $FX \in \mathrm{ob}(\mathbf{B})$.
- F maps every morphism $\phi \in \mathbf{A}(X, Y)$ to a morphism $F(\phi) \in \mathbf{B}(FY, FX)$.
- $F(\mathbb{1}_X) = \mathbb{1}_{FX}$
- $F(\phi \circ \psi) = F(\psi) \circ F(\phi)$

Remark 4.2.3. A contravariant functor can also be defined as a covariant functor from the opposite category. Therefore from now on we will drop the word *covariant* when talking about functors. Furthermore, a contravariant functor $G : \mathbf{C}^{op} \to \mathbf{Set}$ is often called a **presheaf**.

Definition 4.2.4 (Dagger category[2]). A category equipped with a contravariant endofunctor, i.e. a functor $\dagger : \mathbf{C}^{op} \to \mathbf{C}$, such that:

- $\forall C \in \mathrm{ob}(\mathbf{C}) : \mathbb{1}_C^\dagger = \mathbb{1}_C$
- $\dagger \circ \dagger = \mathbb{1}_C$

The second property says that $\dagger$ is an **involutive** functor.

Remark 4.2.5. The concept of a dagger structure allows the usual definition of **unitary** and **self-adjoint** morphisms.

[2] Also called a **†-category**.

Property 4.2.6. The unitary morphisms in a dagger category form a groupoid[3].

Definition 4.2.7 (Faithful functor). A functor $F : \mathbf{C} \to \mathbf{D}$ is said to be faithful if the map

$$\mathbf{C}(X, Y) \to \mathbf{D}(FX, FY)$$

is injective for all objects $X, Y \in \text{ob}(\mathbf{C})$.

Definition 4.2.8 (Full functor). A functor $F : \mathbf{C} \to \mathbf{D}$ is said to be full if the map

$$\mathbf{C}(X, Y) \to \mathbf{D}(FX, FY)$$

is surjective for all objects $X, Y \in \text{ob}(\mathbf{C})$.

Example 4.2.9 (hom-functor). Let $\mathbf{C}$ be a locally small category. Every object $X \in \text{ob}(\mathbf{C})$ induces a functor $h^X : \mathbf{C} \to \mathbf{Set}$ defined as follows:

- h^X maps every object $Y \in \text{ob}(\mathbf{C})$ to the set $\mathbf{C}(X, Y)$.
- For all $Y, Z \in \text{ob}(\mathbf{C})$, h^X maps every morphism $f \in \mathbf{C}(Y, Z)$ to the morphism $f \circ - : \mathbf{C}(X, Y) \to \mathbf{C}(X, Z) : g \mapsto f \circ g$.

Remark 4.2.10. The contravariant hom-functor h_X is defined by replacing $\mathbf{C}(X, -)$ by $\mathbf{C}(-, X)$.

Definition 4.2.11 (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 are triples (A, B, γ) where $A \in \text{ob}(\mathbf{A}), B \in \text{ob}(\mathbf{B})$ and $\gamma : FA \to GB \in \text{hom}(\mathbf{C})$.
- Morphisms $(A, B, \gamma) \to (K, L, \sigma)$ are pairs (f, g) where $f : A \to K \in \text{hom}(\mathbf{A})$ and $g : B \to L \in \text{hom}(\mathbf{B})$ such that $\sigma \circ F(f) = G(g) \circ \gamma$.
- Composition of morphisms is defined componentwise.

Definition 4.2.12 (Slice category). Let $\mathbf{C}$ be a category and let $c \in \text{ob}(\mathbf{C})$. The slice category $\mathbf{C}/c$ of $\mathbf{C}$ over c is defined as follows:

- The objects are morphisms in $\mathbf{C}$ with codomain c.
- The morphisms $f \to g$ are morphisms h in $\mathbf{C}$ such that $g \circ h = f$.

4.3 Natural transformations

Definition 4.3.1 (Natural transformation). Let F, G be two functors between the categories $\mathbf{C}$ and $\mathbf{D}$. A natural transformation ψ from F to G consists of a collection of morphisms satisfying two conditions:

- For every object $X \in \text{ob}(\mathbf{C})$ there exists a morphism $\psi_X : FX \to GX$ in $\text{hom}(\mathbf{D})$. This morphism is called the **component** of ψ at X.
- For every morphism $f \in \mathbf{C}(X, Y)$ we have $\psi_Y \circ F(f) = G(f) \circ \psi_X$.

It is often said that ψ_X **is natural in** X.

Notation 4.3.2. A natural transformation ψ from a functor F to a functor G is denoted by $\psi : F \Rightarrow G$.[4]

[3] See definition 4.12.2.

[4] This is in analogy with the notation for general 2-morphisms. See section 4.11 for more information.

Example 4.3.3 (Representations). When considering representations as a functor $\rho : \mathrm{Grp} \to \mathrm{FinVect}$ we see that the intertwiners[5] arise as natural transformations.

Definition 4.3.4 (Functor category). Let C be a small category and let D be a category. The functors $F : \mathbf{C} \to \mathbf{D}$ form the objects of a category with the natural transformations as morphisms. This category is denoted by $[\mathbf{C}, \mathbf{D}]$ or $\mathbf{D}^{\mathbf{C}}$ (analogous to 2.1.1).

Definition 4.3.5 (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, ψ), where ψ is the natural isomorphism, is called a **representation** of F.

Remark 4.3.6. A similar definition holds for presheafs $G : \mathbf{C}^{op} \to \mathbf{Set}$. In this case $\hom(X, -)$ has to be replaced by $\hom(-, X)$.

Theorem 4.3.7 (Yoneda's lemma). *Let $\boldsymbol{C}$ be a locally small category and let $F : \mathbf{C} \to \mathbf{Set}$ be a functor. For every object $X \in ob(\boldsymbol{C})$ there exists a natural isomorphism[6] between the set of natural transformations $\mathrm{Nat}(h^X, F)$ and FX.*

Remark. The image of a natural transformation $\psi \in \mathrm{Nat}(h^X, F)$ is given by $\psi_X(\mathbb{1}_X)$.

Corollary 4.3.8 (Yoneda embedding). When F is another hom-functor h^Y we obtain the following result:

$$\mathrm{Nat}(h^X, h^Y) \cong \mathbf{C}(Y, X) \tag{4.3}$$

where one should pay attention to the right hand side where Y appears in the first argument.

Let $\hom(f, -)$ 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 $\hom_C(f, -)$ can also be interpreted as a covariant functor $G : \mathbf{C}^{op} \to \mathbf{Set}^{\mathbf{C}}$. This way we see that Yoneda's lemma gives us a fully faithful functor (i.e. an embedding) h^- from the opposite category $\mathbf{C}^{op}$ to the functor category $\mathbf{Set}^{\mathbf{C}}$.

As usual all of this can be done for contravariant functors. This gives us an embedding $h_- : \mathbf{C} \hookrightarrow \mathbf{Set}^{\mathbf{C}^{op}}$, called the Yoneda embedding.

4.4 Adjunctions

Definition 4.4.1 (Hom-set adjunction). Let $F : \mathbf{C} \to \mathbf{D}$ and $G : \mathbf{D} \to \mathbf{C}$ be two functors. These functors form a hom-set adjunction (often just called an adjunction) if the following isomorphism is natural in a, b:

$$\hom_D(Fa, b) \cong \hom_C(a, Gb) \tag{4.4}$$

The functor F (resp. G) is called the left (resp. right) adjoint[7].

Notation 4.4.2. An adjunction(F, G) between categories $\mathbf{C}, \mathbf{D}$ is often denoted by:

$$\mathbf{C} \underset{G}{\overset{F}{\underset{\longrightarrow}{\overset{\longleftarrow}{\bot}}}} \mathbf{D}$$

[5] See definition 3.1.68.

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

[7] Sometimes the word **adjunct** is used (French versus Latin).

Definition 4.4.3 (Unit-counit adjunction). Let $F : \mathbf{C} \to \mathbf{D}$ and $G : \mathbf{D} \to \mathbf{C}$ be two functors. These functors form a unit-counit adjunction if there exist natural transformations

$$\varepsilon : F \circ G \Rightarrow 1_D \tag{4.5}$$
$$\eta : 1_C \Rightarrow G \circ F \tag{4.6}$$

such that the following compositions are identity morphisms:

$$F \xrightarrow{F\eta} FGF \xrightarrow{\varepsilon F} F \tag{4.7}$$
$$G \xrightarrow{\eta G} GFG \xrightarrow{G\varepsilon} G \tag{4.8}$$

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

Property 4.4.4. Every hom-set adjuction induces a unit-counit adjunction. Let $\Phi_{a,b}$ be the natural isomorphism associated to the hom-set adjunction $F \dashv G$. The unit ε_d is obtained as the adjunct $\Phi^{-1}_{Gd,d}(1_{Gd})$ of the identity morphism on $Gd \in \mathrm{ob}(\mathbf{C})$ and the counit η_c is analogously defined as the adjunct $\Phi_{c,Fc}(1_{Fc})$ of the identity morphism at $Fc \in \mathrm{ob}(\mathbf{D})$.

Similarly every unit-counit adjunction induces a hom-set adjunction. Consider a morphism $f : Fc \to d$. The (right) adjunct $\tilde{f}$ is defined as the composition

$$Gf \circ \eta_c : c \to (G \circ F)c \to Gd$$

Similarly, consider a morphism $\tilde{g} : c \to Gd$. The (left) adjunct g is defined as the composition

$$\varepsilon_d \circ F\tilde{g} : Fc \to (F \circ G)d \to d$$

Now it should be obvious that the above definition of a unit-counit adjunction can be generalized to general 2-categories:

Definition 4.4.5 (Adjunction). Let $\mathbf{C}$ be a 2-category. An adjunction in $\mathbf{C}$ is a pair of 1-morphisms $F : a \to b$ and $G : b \to a$ together with 2-morphisms $\varepsilon : F \circ G \Rightarrow 1_b$ and $\eta : 1_a \Rightarrow G \circ F$ that satisfy the zig-zag identities.

Remark 4.4.6 (Duals and adjunctions). If one looks at the defining relations of duals in a rigid monoidal category one should see 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.

4.5 Initial and terminal objects

Definition 4.5.1 (Initial object). An object O in a category $\mathbf{C}$ is called initial if for every other object P there exists a unique morphism $\iota_P : O \to P$.

Definition 4.5.2 (Terminal object). An object O in a category $\mathbf{C}$ is called terminal if for every other object P there exists a unique morphism $\tau_P : P \to O$. This object is sometimes denoted by $\mathbf{1}$.

Property 4.5.3. If an initial (resp. terminal) object exists, then it is unique (possibly up to isomorphism).

Definition 4.5.4 (Well-pointed category). A category is said to be well-pointed if the terminal object is a generator[8].

Definition 4.5.5 (Zero object). An object which is both initial and terminal. The zero object is often denoted by **0**.

Property 4.5.6 (Zero morphism). From the definition of the zero object it follows that for any two objects A, B there exists a unique morphism $0_{AB} : A \to 0 \to B$.

Definition 4.5.7 (Pointed category). A category is said to be pointed if it contains a zero object.

Definition 4.5.8 (Global element). Let $\mathbf{C}$ be a category with terminal object $\mathbf{1}$. A global element of an object $X \in \mathrm{ob}(\mathbf{C})$ is a morphism $\mathbf{1} \to X$.

Remark 4.5.9. In the category **Set** the elements of a set S are in one-to-one correspondence with the global elements of S and one has the important property (axiom) that two functions $f, g : S \to S'$ coincide if their evaluation at every element $s \in S$ is equal or equivalently if the precompositions with global elements coincide.

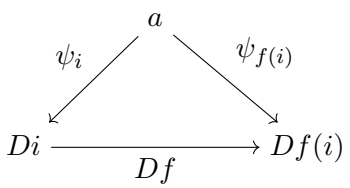
However this way of checking equality can fail in other categories. Consider for example **Grp**. In this category the terminal object is $0 = \{e\}$. The only morphism from this group to any other group G is the one mapping e to the unit in G (0 is also an initial object in **Grp**). It is obvious that precomposition with this morphism tells us nothing about the equality of other morphisms. To recover the technique used in **Set** one needs to generalize the notion of "element":

Definition 4.5.10 (Generalized element). Let $\mathbf{C}$ be category and consider an object X in $\mathbf{C}$. For any object $Y \in \mathrm{ob}(\mathbf{C})$ one calls a morphism $Y \to X$ a generalized element of X. The morphisms $Y \to X$ are also called **Y-elements** in X or elements of **shape** Y in X.

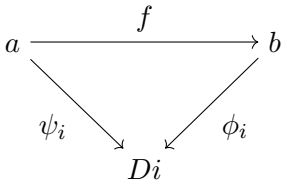
4.6 Diagrams and limits

Definition 4.6.1 (Diagram). A diagram in $\mathbf{C}$ with index category $\mathbf{I}$ is a (covariant) functor $D : \mathbf{I} \to \mathbf{C}$.

Definition 4.6.2 (Cone). Let $D : \mathbf{I} \to \mathbf{C}$ be a diagram. A cone from $a \in \mathrm{ob}(\mathbf{C})$ to D consists of a family of morphisms $\psi_i : a \to Di, \forall i \in \mathrm{ob}(\mathbf{I})$ such that $\psi_j = Df \circ \psi_i$ for all morphisms $f : i \to j \in \hom(\mathbf{I})$, as depicted in figure 4.1a.



(a) Component of cone over D.



(b) Morphism of cones.

Alternative Definition 4.6.3. This definition can be reformulated by defining an additional functor[9] $\Delta_a : \mathbf{I} \to \mathbf{C}$ which maps every element $i \in \mathrm{ob}(\mathbf{I})$ to a and every morphism $g \in \hom(\mathbf{I})$ to 1_a. The morphisms ψ_c can then be seen as the components of a natural transformation $\psi : \Delta_a \Rightarrow D$. Hence a cone (a, ψ) is an element of $[\mathbf{I}, \mathbf{C}](\Delta_a, D)$.

[8] See definition 4.7.12.

[9] The notation Δ_a tells us that $\Delta : C \to [\mathbf{I}, \mathbf{C}]$ is the **diagonal functor**, i.e. $\Delta(c)$ is the constant functor from $\mathbf{I}$ to $\mathbf{C}$ with target object c.

Definition 4.6.4 (Morphism of cones). Let $D : \mathbf{I} \to \mathbf{C}$ be a diagram and let $(a, \psi), (b, \phi)$ be cones to D. A morphism between these cones is a morphism of the apexes $f : a \to b$ such that the diagrams of the form 4.1b commute for all $i \in \mathrm{ob}(\mathbf{I})$. The cones to 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 4.6.5 (Limit). Consider a diagram $D : \mathbf{I} \to \mathbf{C}$. The limit $\lim D$ of this diagram, if it exists, is the terminal object of the category $\mathbf{Cone}(D)$.

This definition gives us following universal property:

Universal property 4.6.6. Let $D : \mathbf{I} \to \mathbf{C}$ be a diagram and let $\lim D$ be its limit. For every cone $(a, \psi) \in \mathbf{Cone}(D)$ there exists a unique morphism $f : c \to \lim D$.

Example 4.6.7 (Terminal object). A terminal object $\mathbf{1}$ is a limit over the empty diagram.

Definition 4.6.8 (Equalizer). Consider following diagram:

$$X \underset{g}{\overset{f}{\rightrightarrows}} Y$$

The limit of this diagram is called the equalizer of f and g. Explicitly the equalizer is the universal object E together with a morphism $e : E \to X$ such that $f \circ e = g \circ e$.

Definition 4.6.9 (Finitely complete category). A category is said to be finitely complete if it has a terminal object and if all equalizers and all finite products exist.

Alternative Definition 4.6.10 (Finitely complete category). A category is said to be finitely complete if all finite limits exist.

Definition 4.6.11 (Span). A span in a category C is a diagram of the form 4.2a.

Let Λ be the category with three objects $\{-1, 0, 1\}$ and two morphisms $i : 0 \to -1$ and $j : 0 \to 1$. By the above definition of a diagram a span in C is equivalent to a functor $S : \mathbf{\Lambda} \to \mathbf{C}$.

Definition 4.6.12 (Pullback[10]). The pullback of two morphisms $f : A \to C$ and $g : B \to C$ is defined as the limit of cospan 4.2b.

(a) Span (category theory). (b) Cospan.

Figure 4.2

Notation 4.6.13 (Pullback). The pullback of two morphisms $f : A \to C$ and $g : B \to C$ is often denoted by $A \times_C B$.

Property 4.6.14. If a terminal object $\mathbf{1}$ exists then the pullback $A \times_{\mathbf{1}} B$ is equal to the (Cartesian) product $A \times B$.

Definition 4.6.15 (Pushout). The dual notion of a pullback, i.e. the colimit of a span.

[10] Also called a **fibre product** or **Cartesian square**.

4.7 Morphisms

Definition 4.7.1 (Monomorphism). Let $\mathbf{C}$ be a category. A morphism $\mu \in \mathbf{C}(A, B)$ is called a mono[11] if for every object $X \in \text{ob}(\mathbf{C})$ and every two morphisms $\alpha_1, \alpha_2 \in \mathbf{C}(X, A)$ such that $\mu \circ \alpha_1 = \mu \circ \alpha_2$ we can conclude that $\alpha_1 = \alpha_2$.

Definition 4.7.2 (Epimorphism). Let $\mathbf{C}$ be a category. A morphism $\varepsilon \in \mathbf{C}(A, B)$ is called an epimorphism[12] if for every object $X \in \text{ob}(\mathbf{C})$ and every two morphisms $\alpha_1, \alpha_2 \in \mathbf{C}(B, X)$ such that $\alpha_1 \circ \varepsilon = \alpha_2 \circ \varepsilon$ we can conclude that $\alpha_1 = \alpha_2$.

Definition 4.7.3 (Balanced category). A category is said to be balanced if every monic epi is an isomorphism.

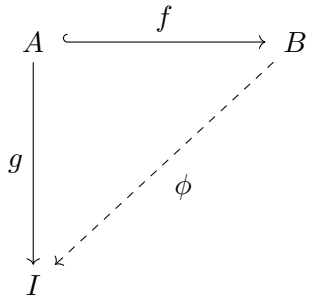
Property 4.7.4 (Global elements). Every global element[13] is monic.

Property 4.7.5 (Equalizing morphism). Consider an equalizer (E, e). The equalizing morphism e is monic. Similarly a coequalizing morphism is epic.

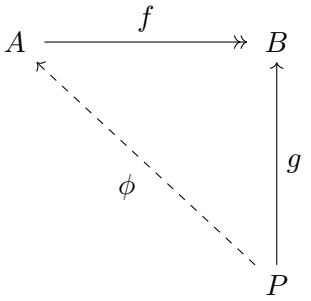
Definition 4.7.6 (Regular monomorphism). A mono is said to be regular if it arises as an equalizer of two parallel morphisms.

Property 4.7.7 (Regular bimorphism). A monic regular epimorphism is an isomorphism. Analogously so is an epic regular monomorphism is an isomorphism.

Definition 4.7.8 (Injective object). Let $\mathbf{C}$ be an category. An object $I \in \text{ob}(\mathbf{C})$ is said to be injective if for every $A, B \in \text{ob}(\mathbf{C})$, mono $f : A \to B$ and morphism $g : A \to I$ there exists a morphism $\phi : B \to I$ such that $\phi \circ f = g$.



(a) Injective object I.



(b) Projective object P.

Dually one can construct:

Definition 4.7.9 (Projective object). Let $\mathbf{C}$ be an category. An object $P \in \text{ob}(\mathbf{C})$ is said to be projective if for every $A, B \in \text{ob}(\mathbf{C})$, epi $f : A \to B$ and morphism $g : P \to A$ there exists a morphism $\phi : P \to B$ such that $f \circ \phi = g$.

Definition 4.7.10 (Subobject). Let $\mathbf{C}$ be a category and let $A \in \text{ob}(\mathbf{C})$ be any object. A subobject B of A is a mono $B \hookrightarrow A$.

In fact one should work up to isomorphism and hence the complete definition goes as follows. A subobject B of A in the category $\mathbf{C}$ is an isomorphism class of monos $i : B \hookrightarrow A$ in the slice category $\mathbf{C}/A$.

[11] Sometimes just a **mono** or a **monic** morphism.
[12] Sometimes just an **epi** or an **epic** morphism.
[13] See definition 4.5.8.

Definition 4.7.11 (Well-powered category). A category $\mathbf{C}$ is said to be well-powered if for every object $A \in \text{ob}(\mathbf{C})$ the class of subobjects $\mathbf{Sub}(A)$ is small.

Definition 4.7.12 (Generator[14]). Let $\mathbf{C}$ be a category. A collection of objects $\{O_i \in \text{ob}(\mathbf{C})\}$ is called a collection of generators for $\mathbf{C}$ if, given any two objects $A, B \in \text{ob}(\mathbf{C})$ and any two morphisms $f, g : A \to B$, we have that $f \neq g \implies \exists j \in I, h_j \in \mathbf{C}(O_i, A) : f \circ h_j \neq g \circ h_j$.

Definition 4.7.13 (Decategorification). Let $\mathbf{C}$ be a (essentially) small category. The set of isomorphism classes of $\mathbf{C}$ is called the decategorification of $\mathbf{C}$. This is given by a functor

$$K : \mathbf{Cat} \to \mathbf{Set}$$

Definition 4.7.14 (Lift). A lift of a morphism $f : A \to B$ along an epi $e : X \to B$ is a morphisms $\tilde{f} : A \to X$ satisfying $f = e \circ \tilde{f}$.

Definition 4.7.15 (Lifting property). A morphism $f : A \to B$ has the left lifting property with respect to a morphism $g : C \to D$ if for every commutative diagram there exists a morphism

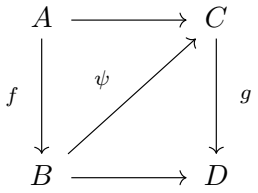


Figure 4.4: Left lifting property.

$\psi : B \to C$ such that the triangles commute. The morphism g is also said to have the right lifting property with respect to f. If the morphism ψ is unique then f and g are said to be **orthogonal**.

Definition 4.7.16 (Weak factorization system). Consider a category $\mathbf{C}$. A pair (L, R) of classes of morphisms in $\mathbf{C}$ is called a weak factorization system (WFS) if it satisfies the folllowing 3 properties:

- Every morphism in $\mathbf{C}$ factors as a composition $g \circ f$ where $f \in L$ and $g \in R$.
- L consists of the morphisms in $\mathbf{C}$ that have the left lifting property with respect to morphisms in R.
- R consists of the morphisms in $\mathbf{C}$ that have the right lifting property with respect to morphisms in L.

Definition 4.7.17 (Localization). Consider a category $\mathbf{C}$ with a collection of morphism $M \subset \text{mor}(\mathbf{C})$. The localization of $\mathbf{C}$ with respect to M is constructed by adding for each morphism $f \in M$ a formal inverse to $\text{mor}(\mathbf{C})$.

4.8 Abelian categories

Definition 4.8.1 (Pre-additive category). A (locally small) category enriched over $\mathbf{Ab}$, i.e. every hom-set is an Abelian group and composition is bilinear.

Property 4.8.2. Let $\mathbf{A}$ be a pre-additive category and let $X \in \text{ob}(\mathbf{A})$. The following statements are equivalent:

[14] Also called a **separator**.

- X is an initial object.
- X is a final object.
- $\mathbb{1}_X = 0$

Any initial/terminal object is hence a zero object[15].

Property 4.8.3. In a pre-additive category the finite products $\prod_{i\in I} X_i$ are isomorphic to the finite coproducts $\bigsqcup_{i\in I} X_i$ (which are called **direct sums** in this context). If a product $X \times Y$ exists then so does the coproduct $X \sqcup Y$ and if the coproduct exists then so does the product. In general if a product and coproduct exist and are equal then one speaks of a **biproduct**.

Definition 4.8.4 (Additive category). A pre-additive category in which all finite biproducts exist.

If one speaks of additive categories, one generally assumes that the associated functors are of a specific type:

Definition 4.8.5 (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 all biproducts, i.e. if it satisfies the following conditions:

- It preserves zero objects: $F0_A \cong 0_{A'}$.
- There exists a natural isomorphism $F(X \oplus Y) \cong FX \oplus FY$.

One can generalize this notion 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.

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

Definition 4.8.6 (Kernel). Let $f : X \to Y$ be a morphism. A[16] kernel is a morphism $k : Z \to X$ such that:

- $f \circ k = 0$
- Universal property: for every other morphism $k' : Z' \to X$ such that $f \circ k' = 0$ there exists a unique morphism $h : Z' \to Z$ such that $k \circ h = k'$.

Hence a kernel of f is an equalizer of f and 0.

Notation 4.8.7 (Kernel). If the kernel of $f : X \to Y$ exists then it is denoted by $\ker(f) \to X$.

Definition 4.8.8 (Cokernel). Let $f : X \to Y$ be a morphism. A cokernel is a morphism $p : Y \to Z$ such that:

- $p \circ f = 0$
- Universal property: for every other morphism $p' : Y \to Z'$ such that $p' \circ f = 0$ there exists a unique morphism $h : Z \to Z'$ such that $h \circ p = p'$.

Hence a cokernel of f is a coequalizer of f and 0.

Notation 4.8.9 (Cokernel). If the cokernel of $f : X \to Y$ exists then it is denoted by $Y \to \mathrm{coker}(f)$.

[15] See definition 4.5.5.

[16] Note the word *a*. The kernel of a morphism is determined up to an isomoprhism.

Remark 4.8.10. The name and notation of the kernel[17] (in the categorical sense) is explained by remarking that by Yoneda's lemma the morphism

$$\ker(f) \to X$$

represents the functor

$$F : Z \mapsto \ker\Big(\mathbf{C}(Z, X) \to \mathbf{C}(Z, Y)\Big)$$

Definition 4.8.11 (Pre-Abelian category). An additive category is pre-Abelian if every morphism has a kernel and cokernel.

Definition 4.8.12 (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 $\operatorname{coker}(\ker(f)) \cong \ker(\operatorname{coker}(f))$.

Property 4.8.13. 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 4.8.14 (Simple object). Let $\mathbf{C}$ be an Abelian category. An object $A \in \operatorname{ob}(\mathbf{C})$ is said to be simple if the only subobjects of A are $\mathbf{0}$ and A itself. An object is said to be semisimple if it is a direct sum of simple obejcts.

Definition 4.8.15 (Semisimple category). A category is said to be semisimple if every object is semisimple (where generally the direct sums are taken over finite index sets).

Definition 4.8.16 (Exact functor). Let $F : \mathbf{A} \to \mathbf{A}'$ be an additive functor between additive categories. We use the following definitions:

- 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 4.8.17. 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

  $$0 \longrightarrow A \longrightarrow B \longrightarrow C$$

  to an exact sequence of the form

  $$0 \longrightarrow FA \longrightarrow FB \longrightarrow FC$$

- If F is right exact it maps an exact sequence of the form

  $$A \longrightarrow B \longrightarrow C \longrightarrow 0$$

  to an exact sequence of the form

  $$FA \longrightarrow FB \longrightarrow FC \longrightarrow 0$$

- If F is exact it maps short exact sequences to short exact sequences.

[17] Similarly for the cokernel.

4.9 Monoidal categories

Definition 4.9.1 (Monoidal category[18]). A category $\mathbf{C}$ equipped with a bifunctor

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

called the **tensor product** or **monoidal product**, together with a distinct object $\mathbf{1}$, called the **unit object**, and 3 natural isomorphisms, called the **coherence maps**:

- **Associator**: $\alpha_{A,B,C} : (A \otimes B) \otimes C \cong A \otimes (B \otimes C)$
- **Left unitor**: $\lambda_A : \mathbf{1} \otimes A \cong A$
- **Right unitor**: $\rho_A : A \otimes \mathbf{1} \cong A$

such that the **triangle** and **pentagon** diagrams commute. (See figures 4.5 and 4.6.)

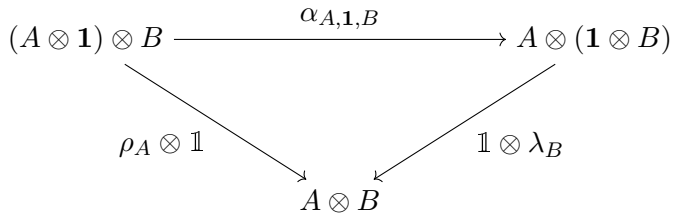


Figure 4.5: Triangle diagram.

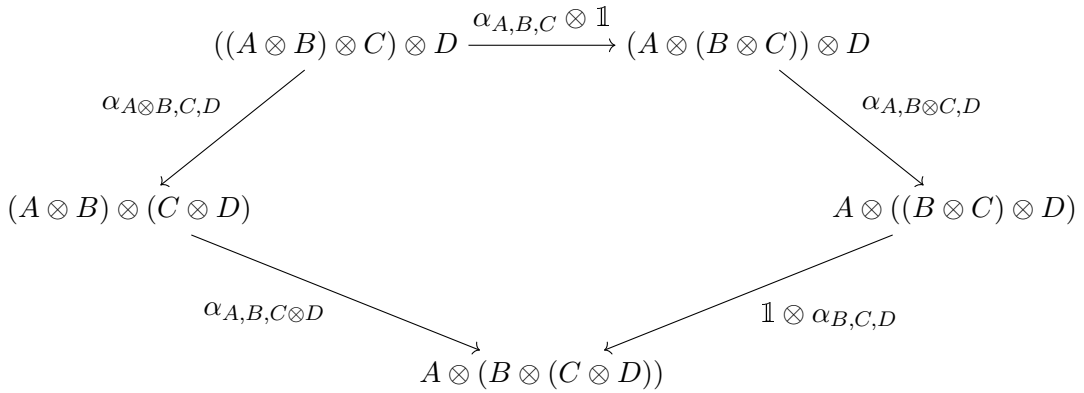


Figure 4.6: Pentagon diagram.

Definition 4.9.2 (Strict monoidal category). A monoidal category is called strict if the associator α and the unitors λ, ρ are identity morphisms.

Definition 4.9.3 (Scalar). In a monoidal category the scalars are defined as the endomorphisms $\mathbf{1} \to \mathbf{1}$.

Property 4.9.4. The set of scalars forms a commutative monoid.

Property 4.9.5. Every scalar $s : \mathbf{1} \to \mathbf{1}$ induces a natural transformation

$$s_A : A \cong \mathbf{1} \otimes A \xrightarrow{s \otimes \mathbb{1}} \mathbf{1} \otimes A \cong A$$

satisfying the "usual" rules of scalar multiplication in linear algebra:

[18] Sometimes called a **tensor** category.

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

where $s \diamond f$ denotes $f \circ s_A = s_B \circ f$.

Definition 4.9.6 (Weak inverse). Let $(\mathbf{C}, \otimes, \mathbf{1})$ be a monoidal category. Consider an object $X \in \text{ob}(\mathbf{C})$. An object $Y \in \text{ob}(\mathbf{C})$ is called a weak inverse of X if it satisfies $X \otimes Y \cong \mathbf{1}$.

Remark 4.9.7. 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. $Y \otimes X \cong \mathbf{1}$.

Theorem 4.9.8 (MacLane's coherence theorem). *Let $\boldsymbol{C}, \boldsymbol{D}$ be monoidal categories. Any two natural transformations $\eta, \varepsilon : F \Rightarrow G$ constructed solely from the associator α and the unitors λ, ρ coincide.*

Definition 4.9.9 (Closed monoidal category). A monoidal category $(\mathbf{C}, \otimes, \mathbf{1})$ is said to be closed monoidal if for every object $B \in \text{ob}(\mathbf{C})$ there exists a a right adjoint[19] to the tensor product functor $- \otimes B : \mathbf{C} \to \mathbf{C}$, i.e.:

$$\forall A, C \in \text{ob}(\mathbf{C}) : \exists B \Rightarrow C \in \text{ob}(\mathbf{C}) : \mathbf{C}(A \otimes B, C) \cong \mathbf{C}(A, B \Rightarrow C) \tag{4.9}$$

where the isomorphism is natural in $A, C \in \text{ob}(\mathbf{C})$.

4.9.1 Monoidal functors

Definition 4.9.10 (Monoidal functor). Let $(\mathbf{C}, \otimes, \mathbf{1}_C), (\mathbf{D}, \circledast, \mathbf{1}_D)$ be two monoidal categories. A functor $F : \mathbf{C} \to \mathbf{D}$ is said to be monoidal if there exists:

- A natural isomorphism $\psi_{A,B} : FA \circledast FB \Rightarrow F(A \otimes B)$ such that the diagram in figure 4.7 commutes.

$$\begin{array}{ccc}
(FA \circledast FB) \circledast FC & \xrightarrow{\alpha_{\mathbf{D}}} & FA \circledast (FB \circledast FC) \\
\downarrow{\scriptstyle \psi_{A,B} \circledast \mathbb{1}} & & \downarrow{\scriptstyle \mathbb{1} \circledast \psi_{B,C}} \\
F(A \otimes B) \circledast FC & & FA \circledast F(B \otimes C) \\
\downarrow{\scriptstyle \psi_{A \otimes B, C}} & & \downarrow{\scriptstyle \psi_{A, B \otimes C}} \\
F\Big((A \otimes B) \otimes C\Big) & \xrightarrow[F(\alpha_{\mathbf{C}})]{} & F\Big(A \otimes (B \otimes C)\Big)
\end{array}$$

Figure 4.7

- An isomorphism $\phi : \mathbf{1}_D \to F\mathbf{1}_C$ which makes the two diagrams in figure 4.8 commute.

Remark 4.9.11. The morphisms $\psi_{A,B}$ and ϕ are also called **coherence maps** or **structure morphisms**.

Property 4.9.12. For every monoidal functor F there exists a canonical isomorphism $\phi : \mathbf{1}_D \to F\mathbf{1}_C$ defined by the commutative diagram in figure 4.9.

[19] See definition 4.10.1 of an *internal hom* for more information.

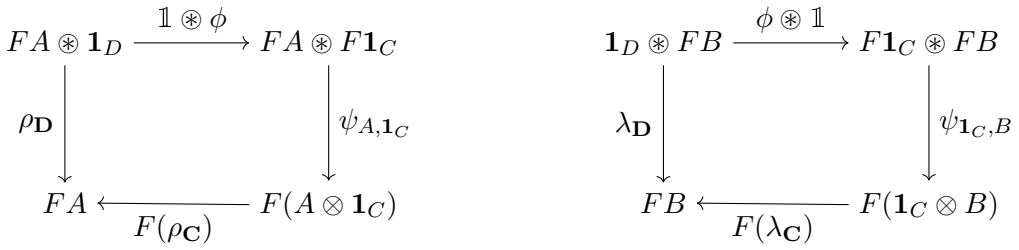


Figure 4.8: Unitality diagrams.

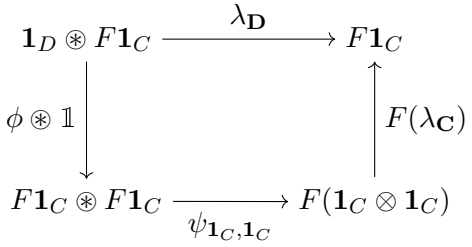


Figure 4.9

Definition 4.9.13 (Lax monoidal functor). A lax monoidal functor is defined as a monoidal functor for which the coherence maps are merely morphisms and not isomorphisms.

Definition 4.9.14 (Monoidal natural transformation). A natural transformation η between (lax) monoidal functors (F, ψ, ϕ_F) and (G, σ_G, ϕ_G) is said to be (lax) monoidal if it makes the diagrams in figure 4.10 commute.

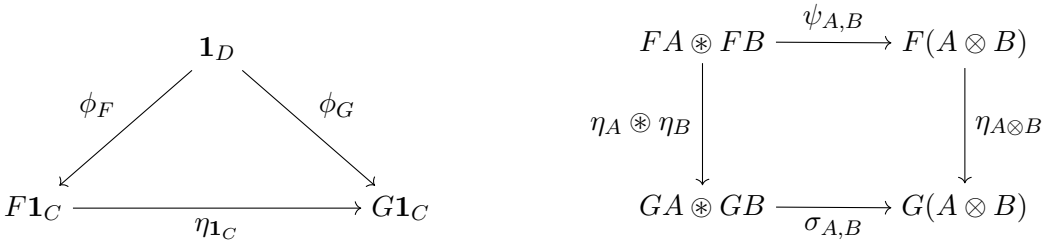


Figure 4.10: Monoidal natural transformation.

Definition 4.9.15 (Monoidal equivalence). Two monoidal categories $\mathbf{C}, \mathbf{D}$ are monoidally equivalent if there exist monoidal functors $F : \mathbf{C} \to \mathbf{D}$ and $G : \mathbf{D} \to \mathbf{C}$ such that there exist monoidal natural isomorphisms $\eta : \mathbb{1}_C \Rightarrow G \circ F$ and $\varepsilon : F \circ G \Rightarrow \mathbb{1}_D$.

Theorem 4.9.16 (MacLane's strictness theorem). *Every monoidal category is monoidally equivalent to a strict monoidal category.*

4.9.2 Braided categories

Definition 4.9.17 (Braided monoidal category). Let $(\mathbf{C}, \otimes, \mathbf{1})$ be a monoidal category. $\mathbf{C}$ is called a braided monoidal category if it comes equipped with a natural isomorphism

$$\sigma_{A,B} : A \otimes B \cong B \otimes A$$

such that the two **hexagon** diagrams in figures 4.11 and 4.12 commute. The isomorphism σ is called the **braiding** (morphism).

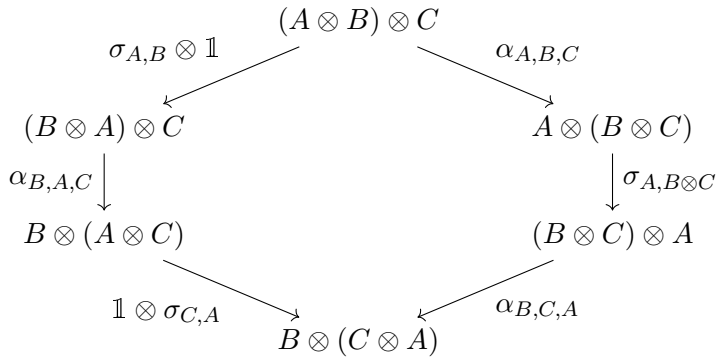


Figure 4.11: Hexagon diagram 1.

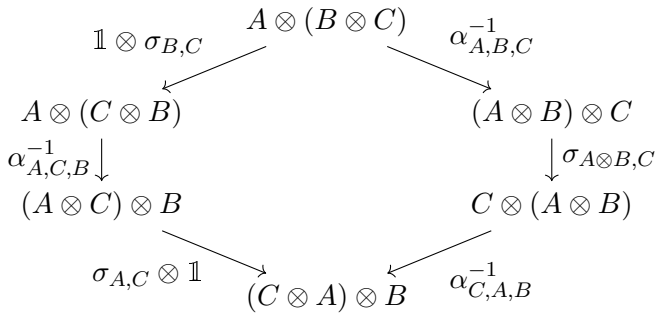


Figure 4.12: Hexagon diagram 2.

Property 4.9.18. The braiding $\sigma_{A,A}$ satisfies the Yang-Baxter equation. More generally the braiding σ satisfies the following equation for all objects $A, B, C \in \text{ob}(\mathbf{C})$:

$$(\sigma_{B,C} \otimes \mathbb{1}) \circ (\mathbb{1} \otimes \sigma_{A,C}) \circ (\sigma_{A,B} \otimes \mathbb{1}) = (\mathbb{1} \otimes \sigma_{A,B}) \circ (\sigma_{A,C} \otimes \mathbb{1}) \circ (\mathbb{1} \otimes \sigma_{B,C}) \tag{4.10}$$

Remark 4.9.19. When drawing the above equality using string diagrams one sees that the Yang-Baxter equation is equal to the invariance of string diagrams under a *Reidemeister III move*.

Definition 4.9.20 (Symmetric monoidal category). A braided monoidal category where the braiding σ satisfies:

$$\sigma_{X,Y} \circ \sigma_{Y,X} = \mathbb{1} \tag{4.11}$$

4.9.3 Duals

Definition 4.9.21 (Dual object). Let $(\mathbf{C}, \otimes, \mathbf{1})$ be a monoidal category and let $A \in \text{ob}(\mathbf{C})$. A left dual[20] A^* of A is an object in $\mathbf{C}$ together with two morphisms $\eta : \mathbf{1} \to A \otimes A^*$ and $\varepsilon : A^* \otimes A \to \mathbf{1}$, called the **unit** and **counit** morphisms[21], such that the diagrams 4.13 and 4.14 commute.

If the object A^* and the morphisms η, ε exist then A is said to be **dualizable**.

Property 4.9.22 (Braided categories). In a braided monoidal category the left and right duals of an object coincide.

[20] Analogously, A is called the **right dual** of A^*. The right dual of B is often denoted by ${}^*B$.
[21] Also called the **coevaluation** and **evaluation** morphisms.

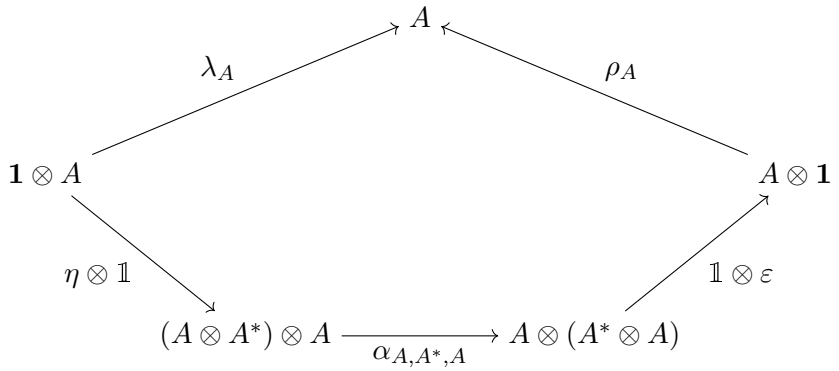


Figure 4.13: Dual object I.

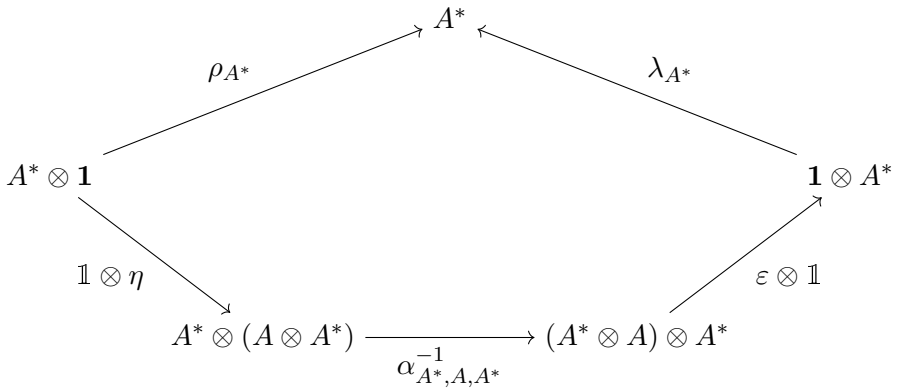


Figure 4.14: Dual object II.

Definition 4.9.23 (Rigid category[22]). A monoidal category in which all duals exist. If only left (resp. right) duals exist then the category is said to be left (resp. right) rigid.

Definition 4.9.24 (Compact closed category). A symmetric rigid category is also called a compact closed category.

Example 4.9.25 (FinVect). Consider the category of finite-dimensional vector spaces **FinVect** (assume that the base field is $\mathbb{R}$). The categorical dual of a vector space V is the algebraic dual V^*. The unit morphism is given by the *resolution of the identity*:

$$\eta : \mathbf{1} \to V \otimes V^* : 1 \mapsto \sum_{i=1}^{\dim(V)} e_i \otimes \phi^i \tag{4.12}$$

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

Definition 4.9.26 (Trace). Let $(\mathbf{C}, \otimes, \mathbf{1})$ be a rigid category. Let $f \in \mathrm{Hom}_C(A, A^{**})$. The left and right (categorical or quantum) traces of f are defined as the following morphisms in $\mathrm{End}_C(\mathbf{1})$:

$$\mathrm{tr}^L(f) : \varepsilon_{A^*} \circ (f \otimes \mathbb{1}) \circ \eta_A \tag{4.13}$$

$$\mathrm{tr}^R(f) : \varepsilon_{**A} \circ (\mathbb{1} \otimes f) \circ \eta_{*A} \tag{4.14}$$

Property 4.9.27. Following linear algebra-like properties hold for the categorical trace:

[22] Also called an **autonomous category**.

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

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

Definition 4.9.28 (Symmetric monoidal dagger category). A symmetric monoidal category $(\mathbf{C}, \otimes, \mathbf{1})$ which also carries the structure of a dagger category[23] such that:

$$(f \otimes g)^\dagger = f^\dagger \otimes g^\dagger \tag{4.15}$$

and such that the coherence and braiding morphisms are unitary.

Definition 4.9.29 (Dagger-compact category). A symmetric monoidal dagger category which is also a compact closed category such that the following diagram commutes:

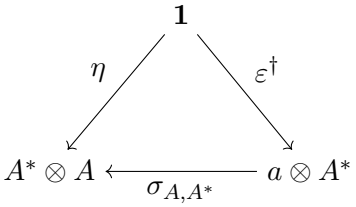


Figure 4.15: Dagger-compact category.

4.9.4 Fusion categories

Definition 4.9.30 (Pivotal category). Let $\mathbf{C}$ be a rigid monoidal category. A pivotal structure on $\mathbf{C}$ is a monoidal natural isomorphism $a_A : A \cong A^{**}$.

Definition 4.9.31 (Dimension). Let $(\mathbf{C}, a)$ be a pivotal category and consider an object $V \in \mathrm{ob}(\mathbf{C})$. The dimension of V is defined as follows:

$$\dim_a(V) := \mathrm{tr}^L(a_V) \tag{4.16}$$

Definition 4.9.32 (Spherical category). Let $(\mathbf{C}, a)$ be a pivotal category. If the left and right traces (with respect to a) coincide in $\mathbf{C}$, or equivalently if for all objects the left and right dimensions coincide, then the pivotal structure is said to be spherical.

Definition 4.9.33 (k-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 $\mathbf{Vect}_k$.

4.10 Internal structures

Definition 4.10.1 (Internal hom[24]). Let $(\mathbf{C}, \otimes, \mathbf{1})$ be a symmetric monoidal category. In this setting one can generalize the *currying* procedure, i.e. the identification of maps $X \times Y \to Z$ with maps $X \to (Y \to Z)$. If the object exists, the internal hom is defined as follows:

$$\hom(A \otimes B, C) = \hom(A, \underline{\hom}(B, C)) \tag{4.17}$$

Hence it is the right adjoint of the tensoring functor $- \otimes B$.

[23] See definition 4.2.4.

[24] Also called an **inner hom**.

Notation 4.10.2. A different, but frequently used, notation is $A \Rightarrow B$. However we will not use this as it might confuse with the notation for 2-morphisms.

Example 4.10.3 (Cartesian closed categories). In case of Cartesian categories, where the monoidal structure is given by the ordinary product, the internal hom object $\underline{\hom}(a, b)$ is given by the exponential object b^a.

Definition 4.10.4 (Internal category). Let C be a category. A category D internal to C consists of following objects:

- An object $D_0 \in \mathrm{ob}(\mathbf{C})$ of objects.
- An object $D_1 \in \mathrm{ob}(\mathbf{C})$ of morphisms.
- Source and target morphisms $s, t \in \hom_C(D_1, D_0)$.
- An identity-assigning morphism $e \in \hom_C(D_0, D_1)$ such that

$$s \circ e = \mathbb{1}_{D_0} \qquad t \circ e = \mathbb{1}_{D_0}$$

- A composition morphism $c : D_1 \times_{D_0} D_1 \to D_1$ such that the following equations hold:

$$\begin{aligned} s \circ c &= s \circ \pi_1 & t \circ c &= t \circ \pi_2 \\ \pi_1 &= c \circ (e \times_{D_0} \mathbb{1}) & c \circ (\mathbb{1} \times_{D_0} e) &= \pi_2 \\ & c \circ (c \times_{D_0} \mathbb{1}) = c \circ (\mathbb{1} \times_{D_0} c) \end{aligned}$$

where π_1, π_2 are the canonical projections associated with the pullback[25] $D_1 \times_{D_0} D_1$.

Remark 4.10.5. To make the above definition work it is required that all pullbacks of the source and target morphisms exist.

4.11 Higher category theory

4.11.1 n-Categories

Definition 4.11.1 (n-Category). A (strict) n-category consists of:

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

such that the composition of k-morphisms ($\forall k \leq n$) 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 an ∞-category.

Example 4.11.2 (2-Category). In a 2-category one can compose 2-morphisms in two different ways:

1. Horizontal composition: Consider the 2-morphisms $\alpha : f \Rightarrow g$ and $\beta : f' \Rightarrow g'$ where $f' \circ f, g' \circ g$ are well-defined. These 2-morphisms can be composed as:

$$\beta \circ \alpha : f' \circ f \Rightarrow g' \circ g$$

[25] See definition 4.6.12.

2. Vertical composition: Consider the 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:

$$\beta \cdot \alpha : f \Rightarrow h$$

Furthermore, the horizontal and vertical composition should satisfy an interchange law:

$$(\alpha \cdot \beta) \circ (\gamma \cdot \delta) = (\alpha \circ \gamma) \cdot (\beta \circ \delta) \tag{4.18}$$

Remark. n-morphisms are also often called **n-cells**.

Definition 4.11.3 ((n, r)-Category). A higher (∞-)category for which:

- All parallel k-morphisms with $k > n$ are equivalent.
- All k-morphisms with $k > r$ are invertible (or equivalences[26]).

Example 4.11.4. The classical example of a 1-category is **Set**, the classical example of a 2-category is **Cat**.

Definition 4.11.5 (Weak inverse). Let $\mathbf{C}$ be a 2-category. A 1-morphism $f : X \to Y$ is weakly invertible if there exists a 1-morphism $g : Y \to X$ and 2-isomorphisms $I : g \circ f \Rightarrow \mathbb{1}_X, J : f \circ g \Rightarrow \mathbb{1}_Y$.

Property 4.11.6 (Monoidal categories as 2-categories). Consider a monoidal category $(C, \otimes, \mathbf{1})$. From this monoidal category one can construct the so-called **delooping** $\mathbf{B}C$, which is a bicategory, in the following way:

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

4.12 Groupoids

Definition 4.12.1 (Categorical group). Let $(\mathbf{C}, \otimes, \mathbf{1})$ be a monoidal category. This category is called categorical group, **gr-category** or **weak 2-group**[27] if it satisfies the following conditions:

- All morphisms are invertible.
- Every object is weakly invertible with respect to the monoidal structure.

By property 4.11.6 above one can equivalently define a weak 2-group as 2-category with a single object, weakly invertible 1-morphisms and invertible 2-morphisms.

[26] In the fully weak ∞-sense.
[27] See also the end of this section.

Definition 4.12.2 (Groupoid). A (small) groupoid $\mathcal{G}$ is a (small) category in which all morphisms are invertible.

Definition 4.12.3 (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 4.12.4 (Orbit). Let $\mathcal{G}$ be a groupoid with O, M respectively the set of objects and morphisms. On O one can define an equivalence $a \sim b$ if there exists a morphism $\phi : a \to b$. The equivalence classes are called orbits and the set of orbits is denoted by O/M.

Definition 4.12.5 (Transitive component). Let $\mathcal{G}$ be a groupoid with O, M respectively the set of objects and morphisms. Let s, t denote the source and target maps on M. Given an orbit $o \in O/M$ one defines the transitive component of M associated to o as $s^{-1}(o)$ or equivalently $t^{-1}(o)$.

Property 4.12.6. Every groupoid is a (disjoint) union of its transitive components.

Definition 4.12.7 (Transitive groupoid). A groupoid $\mathcal{G}$ is said to be transitive if for all objects $x, y \in \mathrm{ob}(()\mathcal{G})$, where $x \neq y$, the set $\hom_{\mathcal{G}}(x, y) \neq \emptyset$.

Definition 4.12.8 (Lie groupoid[28]). A groupoid internal to **Diff**. As noted in the remark below definition 4.10.4 the pullbacks of the source and target morphisms should exist. In **Diff** this is equivalent to assuming that the source and target morphisms should be (surjective) submersions.

Remark 4.12.9. In the Ehresmannian approach one gives the manifold of composable morphisms $D_1 \times_{D_0} D_1$ as part of the data. Hence we do not have to assume anything about the source and target morphisms.

Definition 4.12.10 (2-Groupoid). A 2-groupoid is a 2-category in which all 1-morphisms are invertible and every 2-morphisms has a 'vertical' inverse. (The 'horizontal' inverse can be constructed from the other inverses.)

Definition 4.12.11 (Strict 2-group). A 2-group is defined as a 2-groupoid with only one object. From this it follows that the set of 1-morphisms forms a group and so does the set of 2-morphisms under horizontal composition. The 2-morphisms do not form a group under vertical composition[29].

4.13 Model categories

Definition 4.13.1 (Model structure). Let C be a category. A model structure (in the sense of Quillen[30]) on C consists of 3 classes of morphisms

1. Weak equivalences W
2. Fibrations Fib
3. Cofibrations Cof

which satisfy the following two conditions:

- W turns C into a category with weak equivalences (see 4.1.7).
- $(\mathrm{Cof}, \mathrm{Fib} \cap W)$ and $(\mathrm{Cof} \cap W, \mathrm{Fib})$ are weak factorization systems on C (see 4.7.16).

[28] Similarly one can define *topological groupoids, étalé groupoids, ...*

[29] Because the sources/targets may not match up.

[30] Also called a **closed model category**.

Definition 4.13.2 (Model category). A complete and cocomplete category equipped with a model structure.

Definition 4.13.3 (Acyclic). The morphisms in Fib$\cap W$ and Cof$\cap W$ are called **acyclic** or **trivial** fibrations and cofibrations.

Definition 4.13.4 (Fibrant). An object in a model category is said to be fibrant if the unique morphism to the terminal object is a fibration. Dually, an object in a model category is said to be cofibrant if the unique morphism from the initial object is a cofibration.

Definition 4.13.5 (Quillen equivalence). Let $\mathbf{C}, \mathbf{D}$ be two model categories. An adjuction

$$\mathbf{C} \underset{G}{\overset{F}{\underset{\longrightarrow}{\overset{\longleftarrow}{\perp}}}} \mathbf{D}$$

is called a **Quillen adjunction** if the left adjoint preserves cofibrations and acyclic cofibrations. It follows from the axioms that this is equivalent to requiring that the right adjoint preserves fibrations and acyclic fibrations. These adjoint functors are called (left and right) **Quillen functors**.

If (F, G) is a Quillen adjunction such that for all cofibrant objects A and fibrant objects B the morphism $FA \to B$ is a weak equivalence if and only if the adjunct morphism $A \to GB$ is a weak equivalence, then (F, G) is called a Quillen equivalence.

4.14 Operad theory

4.14.1 Operads

Definition 4.14.1 (Plain operad[31]). Let $\mathcal{O} = \{P(n)\}_{n\in\mathbb{N}}$ be a sequence of sets, called n**-ary operations** (n is the **arity**). The set $\mathcal{O}$ is called a plain operad if it satisfies following axioms:

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

$$\circ : P(n) \times P(k_1) \times \cdots \times P(k_n) \to P(k_1 + \cdots k_n) : (\psi, \theta_1, ..., \theta_n) \mapsto \psi \circ (\theta_1, ..., \theta_n) \quad (4.19)$$

that satisfies two additional axioms:

- Identity:

$$\theta \circ (\mathbb{1}, ..., \mathbb{1}) = \mathbb{1} \circ \theta = \theta \quad (4.20)$$

- Associativity:

$$\psi \circ \Big(\theta_1 \circ (\theta_{1,1}, ..., \theta_{1,k_1}), ..., \theta_n \circ (\theta_{n,1}, ..., \theta_{n,k_n})\Big) = \Big(\psi \circ (\theta_1, ..., \theta_n)\Big) \circ (\theta_{1,1}, ..., \theta_{1,k_1}, \theta_{2,1}, ..., \theta_{n,k_n}) \quad (4.21)$$

Remark 4.14.2. If one represents the operad 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 first glue the first two layers together or one can first glue the last two layers together.

[31] Also called a **non-symmetric operad** or **non-Σ operad**.

Example 4.14.3. 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 4.14.4 (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 4.14.5 (Categorical O-algebra). An O-algebra in the category Cat.

Chapter 5

Topos theory

5.1 Elementary topoi

Definition 5.1.1 (Subobject classifier). Consider a finitely complete category (in fact the existence of a terminal object suffices). A subobject classifier is a mono[1] $\text{true} : \mathbf{1} \hookrightarrow \Omega$ from the terminal object such that for every mono $\phi : a \hookrightarrow b$ there exists a unique morphism $\chi : b \to \Omega$ such that the following pullback square exists:

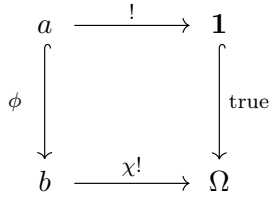


Figure 5.1: Subobject classifier.

Alternative Definition 5.1.2. Consider a well-powered category C. The assignment of subobjects $\text{Sub}(a)$ to an object $a \in \text{ob}(\mathbf{C})$ is a contravariant functor $\mathcal{S} : C \to \text{Set}$. A subobject classifier Ω is a representation of this functor, i.e. the following isomorphism is natural in a:

$$\mathcal{S}(a) \cong \hom(a, \Omega) \tag{5.1}$$

Example 5.1.3. The category Set has a subobject classifier, namely the 2-element set.

Property 5.1.4. Using Yoneda's lemma one obtains following statement: The subobject classifier Ω has a *generic* element $\mathfrak{t}$ such that for every object $a \in \text{ob}(\mathbf{C})$ and every element $b \in \text{Sub}(a)$ there exists a unique morphism $\chi : a \to \Omega$ such that $m = \chi^* \mathfrak{t}$.

Definition 5.1.5 (Elementary topos). An elementary topos is a cartesian closed category which has a subobject classifier and in which all finite limits exist. Equivalently one can define an elementary topos as finitely complete category in which all power objects.

5.2 Internal logic

In this subsection we consider categories with all finite limits and a subobject classifier (so they don't have to be a topos).

[1] The symbol for this morphism will become clear in subsection 5.2.

Definition 5.2.1 (Truth value). A global element of the subobject classifier, i.e. a morphism $\mathbf{1} \to \Omega$. The subobject classifier Ω is hence also sometimes called the **object of truth values**.

Property 5.2.2. The subobject classifier is an internal poset.

5.3 Geometric morphisms

Definition 5.3.1 (Logical morphism). Let $\mathcal{E}, \mathcal{F}$ be (elementary) toposes. A morphism $f : \mathcal{E} \to \mathcal{F}$ is called a logical morphism if it preserves finite limits, exponential objects and subobject classifiers.

Definition 5.3.2 (Geometric morphism). Let $\mathcal{E}, \mathcal{F}$ be toposes. A morphism $f : \mathcal{E} \to \mathcal{F}$ is called a geometric morphism if the left adjoint in the induced adjunction

$$\mathcal{E} \underset{f_*}{\overset{f^*}{\underset{\longrightarrow}{\overset{\longleftarrow}{\bot}}}} \mathcal{F}$$

preserves finite limits. The right adjoint f_* is called the **direct image** part of f and the left adjoint is called the **inverse image** part.

Example 5.3.3 (Topological spaces). Every continuous map $f : X \to Y$ induces a geometric morphism

$$\mathbf{Sh}(X) \underset{f_*}{\overset{f^*}{\underset{\longrightarrow}{\overset{\longleftarrow}{\bot}}}} \mathbf{Sh}(Y) \tag{5.2}$$

where the direct image functor f_* is defined as follows:

$$f_* F(U) = F(f^{-1} U) \tag{5.3}$$

for any sheaf $F \in \mathbf{Sh}(X)$ and any open subset $U \in \mathbf{Open}(Y)$.

By the previous example the global elements $* \to X$ of a topological space induce geometric morphisms of the form $\mathbf{Sh}(*) \to \mathbf{Sh}(X)$. By noting that $\mathbf{Sh}(*) = \mathbf{Set}$ we obtain the following generalization:

Definition 5.3.4 (Point). A point of a topos $\mathcal{E}$ is a geometric morphism $\mathbf{Set} \to \mathcal{E}$.

Notation 5.3.5 (Category of toposes). The category of elementary toposes and geometric morphisms forms a (2-)category which we will denote by **Topos**.

5.4 Sheaf topos

Property 5.4.1 (Presheaf topos). Consider the presheaf category $\mathbf{Psh}(X) = \widehat{\mathbf{Open}(X)}$ over a topological space X. This category is an elementary topos where the subobject classifier Ω is defined as follows:

$$\Omega(U) = \{V : V \text{ is an open subset of } U\} \tag{5.4}$$

Construction 5.4.2 (Sheafs and étalé bundles). Let X be a topological space. The functor

$$I : \mathbf{Open}(X) \to \mathbf{Top}/X : U \mapsto (U \hookrightarrow X)$$

induces the following adjunction:

$$\mathbf{Psh}(X) \underset{E}{\overset{\Gamma}{\underset{\longrightarrow}{\overset{\longleftarrow}{\bot}}}} \mathbf{Top}/X \tag{5.5}$$

The slice category on the right-hand side is equivalently the category of (topological) bundles[2] over X. Both directions of the adjunction have a clear interpretation. The left adjoint assigns to every bundle its sheaf of (global) sections. The right adjoint assigns to every presheaf its bundle of germs.

This adjunction restricts to an equivalence on the subcategories on which the unit and counit morphisms are isomorphisms. These subcategories are called the **sheaf** and **étalé bundle** categories respectively:

$$\mathbf{Sh}(X) \cong \mathbf{Et}(X) \tag{5.6}$$

Property 5.4.3 (Associated sheaf). The inclusion functor $\mathbf{Sh}(X) \hookrightarrow \mathbf{Psh}(X)$ admits a left adjoint. This is exactly the sheafification functor which assigns to every presheaf its associated sheaf. This functor is given by the composition $\Gamma \circ E$.[3]

The fact that the counit of adjunction 5.5 restricts to an isomorphism on the full subcategory $\mathbf{Sh}(X)$ is equivalent to the fact that the sheafification of a sheaf Γ is again Γ.

5.5 Grothendieck topos

Definition 5.5.1 (Discrete fibration). Let $F : \mathbf{A} \to \mathbf{B}$ be a functor. F is a discrete fibration if for every object $A \in \mathrm{ob}(\mathbf{A})$ and every morphism $f : B \to FA$ in $\mathbf{B}$ there exists a unique morphism $g : C \to A$ in $\mathbf{A}$ such that $F(g) = f$, where $B \in \mathrm{ob}(\mathbf{B}), C \in \mathrm{ob}(\mathbf{A})$.

Definition 5.5.2 (Sieve). Let $\mathbf{C}$ be a small category. A sieve S on $\mathbf{C}$ is a fully faithfull discrete fibration $S \hookrightarrow \mathbf{C}$.

A sieve S on an object $c \in \mathbf{C}$ is a sieve in the slice category $\mathbf{C}/c$. This means that S is a subset of $\mathrm{ob}(\mathbf{C})/c$ that is closed under *precomposition*, i.e. if $b \to c \in S$ and $a \to b \in \hom(\mathbf{C})$ then the composition $a \to b \to c \in S$.

All of this can be summarized by saying that a sieve on an object $c \in \mathrm{ob}(\mathbf{C})$ is a subfunctor of the hom-functor $\mathbf{C}(-, c)$.

Example 5.5.3 (Maximal sieve). Let C be a category. The maximal sieve on $c \in \mathrm{ob}(\mathbf{C})$ is the collection of all morphisms $\{f \in \hom(\mathbf{C}) : \mathrm{cod}(f) = c\}$.

The following definition is due to Giraud (the original definition used covers).

Definition 5.5.4 (Grothendieck topology). A Grothendieck topology on a category is a function J assigning to every object a collection of sieves satisfying the following conditions:

- For every object c the maximal sieve M_c is an element of $J(c)$.

[2] See chapter 29.

[3] This amounts to construction 10.2.10.

- If $S \in J(c)$ then $f^*S \in J(d)$ for every morphism $(f : d \to c) \in S$.
- Consider a sieve S on c. If there exists a sieve $R \in J(c)$ such that for every morphism $(f : d \to c) \in R$ the pullback sieve $f^*S \in J(d)$ then $S \in J(c)$.

The sieves in J are called (J-)**covering sieves**.

Example 5.5.5 (Topological spaces). These conditions have the following interpretation in the case of topological coverings:

- The collection of all open subsets covers a space U.
- If $\{U_i\}_{i\in I}$ covers U then $\{U_i \cap V\}_{i\in I}$ covers $U \cap V$.
- If $\{U_i\}_{i\in I}$ covers U and if for every $i \in I$ the collection $\{U_{ij}\}_{j\in J_i}$ covers U_i then $\{U_{ij}\}_{i\in I, j\in J_i}$ covers U.

The canonical Grothendieck topology on $\mathbf{Open}(X)$ is given by the sieves $S = \{U_i \hookrightarrow U\}_{i\in I}$ where $\bigcup_{i\in I} U_i = U$. This topology is denoted by $J_{\mathrm{Open}(X)}$.

Definition 5.5.6 (Site). A (small) category equipped with a Grothendieck topology J.

Definition 5.5.7 (Matching family). Consider a presheaf $F \in \mathrm{ob}(\widehat{\mathbf{C}})$ together with a sieve S on $c \in \mathrm{ob}(\mathbf{C})$. A matching family for S with respect to F is a natural transformation $\alpha : S \implies F$ between S regarded as a subfunctor of $\hom(-, c)$ and F.

More explicitly it is as assignment of an element $x_f \in Fd$ to every morphism $(f : d \to c) \in S$ such that:

$$Fg(x_f) = x_{f\circ g} \tag{5.7}$$

for all morphisms $g : e \to d$.

Given such a matching family one calls an element $z \in Fc$ an **amalgamation** if it satisfies

$$Ff(z) = x_f \tag{5.8}$$

for all morphisms $f \in S$.

Definition 5.5.8 (Sheaf). Consider a site $(\mathbf{C}, J)$. A presheaf F on $\mathbf{C}$ is called a J-sheaf if every matching family for any covering sieve (on any object in $\mathbf{C}$) in J admits a unique amalgamation.

From the natural transformation point of view this condition amounts to the existence of a unique extension of $\alpha : S \implies F$ along the inclusion $S \hookrightarrow \hom(-, c)$.

The category $\mathbf{Sh}(\mathbf{C}, J)$ of J-sheaves on the site $(\mathbf{C}, J)$ is the full subcategory of $\widehat{\mathbf{C}}$ on the presheaves which satisfy the above condition.

Example 5.5.9 (Topological spaces). The usual category of sheaves $\mathbf{Sh}(X)$ on a topological space X is obtained as the category of sheaves on the site $(\mathbf{Open}(X), J_{\mathrm{open}(X)})$.

Alternative Definition 5.5.10 (Grothendieck topos). A category[4] equivalent to a category of sheaves on a (small) site.

Property 5.5.11 (Balanced). All monics and epics in a Grothendiek topos are regular. Hence property 4.7.7 implies that every bimorphism is in fact an isomorphism and it follows that every Grothendieck topos is balanced[5].

[4] The fact that a Grothendieck topos is indeed an elementary topos follows from the fact that **Set** is a topos.
[5] See definition 4.7.3.

Property 5.5.12 (Epi-mono factorization). Every morphism $f : a \to b$ in a Grothendieck topos factorizes uniquely as an epi followed by a mono:

$$a \overset{e}{\twoheadrightarrow} c \overset{m}{\rightarrowtail} b \tag{5.9}$$

The mono is called the **image** of f.

Chapter 6

Homological algebra

A reference for this chapter is [1, 2].

6.1 Chain complexes

Definition 6.1.1 (Chain complex). Let $(A_k)_{k\in\mathbb{Z}}$ be a sequence of algebraic structures together with a sequence $\{\partial_k : A_k \to A_{k-1}\}_{k\in\mathbb{Z}}$ of morphisms, called the **boundary operators** or **differentials**, such that for all k:

$$\partial_k \circ \partial_{k+1} = 0 \tag{6.1}$$

This structure is called a chain complex[1]. Elements in $\mathrm{im}(\partial_k)$ are called **boundaries** and elements in $\ker(\partial_k)$ are called **cycles**. The chain complex $\{(A_k, \partial_k)\}_{k\in\mathbb{Z}}$ is often denotes by $A_\bullet$.

The morphisms of chain complexes are called **chain maps** and are defined as a collection of morphisms $\{f_k : A_k \to B_k\}_{k\in\mathbb{Z}}$ such that for all $k \in \mathbb{Z}$ the following equation holds:

$$\partial'_k \circ f_k = f_{k-1} \circ \partial_k \tag{6.2}$$

where ∂_k, ∂'_k are the boundary operators of respectively $A_\bullet$ and $B_\bullet$.

Remark 6.1.2. Given a chain (resp. cochain) complex one can easily construct a cochain (resp. chain) complex by setting $C'_k = C_{-k}$.

Definition 6.1.3 (Chain homology). Given a chain complex $C_\bullet$ one can define its homology groups $H_n(C_\bullet)$. Since $\partial^2 = 0$ the kernel $\ker(\partial_k)$ is a subgroup of the image $\mathrm{im}(\partial_{k+1})$. This way we can define the quotient groups:

$$H_n(C_\bullet) := \frac{\ker(\partial_k)}{\mathrm{im}(\partial_{k+1})} \tag{6.3}$$

The kernel in this definition is often called the group of **cycles** and is denoted by $Z_k(C_\bullet)$. The image in this definition is often called the group of **boundaries** and is denoted by $B_k(C_\bullet)$.

Definition 6.1.4 (Quasi-isomorphism). Consider a chain map $f_\bullet : C_\bullet \to D_\bullet$. If the induced maps on homology are isomorphisms then $f_\bullet$ is called a quasi-isomorphism.

[1] A **cochain complex** is constructed similarly. For this structure we consider an ascending order, i.e. $\partial_k : A_k \to A_{k+1}$.

Definition 6.1.5 (Chain homotopy). Two chain maps $f_\bullet, g_\bullet : C_\bullet \to D_\bullet$ are said to be chain homotopic if there exists a chain map $s_\bullet : C_\bullet \to D_\bullet$ such that the following equation is satisfied:

$$f - g = s \circ \partial_C + \partial_D \circ s \tag{6.4}$$

If both $f \circ g$ and $g \circ f$ are chain homotopic to the identity then $C_\bullet$ and $D_\bullet$ are said to be (chain) homotopy equivalent.

Property 6.1.6. Every (chain) homotopy equivalence is a quasi-isomorphism.

6.2 Resolutions

Consider some Abelian category $\mathbf{A}$. Let $\mathbf{Ch}(\mathbf{A})$ denote the category of chain complexes with objects in $\mathbf{A}$. In fact in this and coming sections we will only be interested in chain complex concentrated in positive degree, i.e. $C_k = 0$ for all $k < 0$, and henceforth we will assume that all chain complexes satisfy this condition.

Definition 6.2.1 (Acyclic complex). A chain complex $C_\bullet \in \mathbf{Ch}(\mathbf{A})$ is said to be acyclic if the sequence

$$\cdots \longrightarrow C_{k+1} \longrightarrow C_k \longrightarrow C_{k-1} \longrightarrow \cdots$$

is exact.

Remark. Some references, especially the older ones, use a slightly different definition of acyclicity. In their definition the sequence is exact except in degree 0, i.e. $H_0(C_\bullet) \neq 0$.

Definition 6.2.2 (Resolution). Consider an object $X \in \mathrm{ob}(\mathbf{A})$. A resolution of X is given by a chain complex $C_\bullet \in \mathbf{Ch}(\mathbf{A})$ which is acyclic except in degree 0 with the property that $H_0(C_\bullet) \cong X$. This definition is equivalent to stating that there exists an exact sequence of the form

$$\cdots \longrightarrow C_1 \longrightarrow C_0 \xrightarrow{\varepsilon} X \longrightarrow 0 \tag{6.5}$$

The morphism $\varepsilon : C_0 \to X$ is often called the **augmentation map**.

Often one can or should specialize the type of resolution. For example by considering chain complexes with only injective or projective objects (see figures 4.3a and 4.3b), one obtains injective or projective resolutions.

If every object $X \in \mathrm{ob}(\mathbf{A})$ admits a projective (resp. injective) resolution then $\mathbf{A}$ is said to have enough projectives (resp. injectives).

6.3 Derived functors

Given an additive functor (see definition 4.8.5) one can define its **prolongation** on the category of chain complexes:

Definition 6.3.1 (Prolongation). Let $F : \mathbf{A} \to \mathbf{A}'$ be an additive functor. Consider the categories of chain complexes $\mathbf{Ch}(\mathbf{A})$ and $\mathbf{Ch}(\mathbf{A}')$. The prolongation of F is a functor $\overline{F} : \mathbf{Ch}(\mathbf{A}) \to \mathbf{Ch}(\mathbf{A}')$ obtained by applying F to every object in a chain complex and to every diagram in the definition of a chain map.

To understand and unify the various long exact sequences in (co)homology and to give general statements about these theories we introduce the concept of derived functors.

Definition 6.3.2 (Left derived functor). Let $\mathbf{A}$ be an Abelian category with enough projectives and consider a right exact functor $F : \mathbf{A} \to \mathbf{A}'$. The left derived functors L_iF are defined in the following way. Pick an object $X \in \mathrm{ob}(\mathbf{A})$ and construct a projecive resolution $P_\bullet \xrightarrow{\varepsilon} X \to 0$. Apply the prolongation $\overline{F}$ to this resolution and construct its chain homology $H_\bullet$. The left derived functors are then given by:

$$L_iF(X) := H_i(\overline{F}P_\bullet) \tag{6.6}$$

Right derived functors (in the case of a left exact functor F) can be constructed dually by choosing an injective resolution, applying the prolongation and taking the cohomology of the resulting cochain complex. In the remainder of this section we will always work with left derived functor for simplicity.

Remark 6.3.3. The above construction is valid for covariant functors. For contravariant functors one has to change it slightly. For contravariant functors F one calculates the derived functors of the opposite functor F^{op}. This is equivalent to starting with an injective (resp. projective) resolution for the calculation of left (resp. right) derived functors since injective objects are projective in the opposite category and similarly homology becomes cohomology in the opposite category.

Property 6.3.4. If F is exact the above construction immediately implies that the derived functors vanish identically.

Property 6.3.5. Consider a right exact functor F together with its left derived functors L_iF. If an object P is projective then $L_iF(X) = 0$ for all $i \geq 1$. This can easily be proven by remarking that every projective object P admits a projective resolution of the form

$$\cdots \longrightarrow 0 \longrightarrow 0 \longrightarrow P \longrightarrow P \longrightarrow 0$$

Now of course we could wonder why the resolutions used in the construction of derived functors are required to be projective or injective. This seems to be a very strong requirement. The reason for this is that using our definition one obtains naturally isomorphic derived functors, i.e. the result is independent of the resolution used.

However for some cases one would like to work with a more general resolution. For example in the next section, when considering the tensor product it would be great to just work with *flat* modules. To this intent we introduce the following notion:

Definition 6.3.6 (Acyclic resolution). Consider a right exact functor F together with its left derived functors L_iF. An object X is said to be F**-acyclic** if $L_iF(X) = 0$ for all $i \geq 1$. A resolution of an object Y is said to be (F-)acyclic if all objects in the resolution are F-acyclic.

It can be proven that calculating the derived functors of F with respect to an F-acyclic resolution gives functors isomorphic to the ones obtained using the construction introduced above.

One of the motivating applications of derived functors is the relation with the exactness of a functor and the relation to various long exact sequences in (co)homology theories. All of these are a result of the following property:

Property 6.3.7. Let $F : \mathbf{A} \to \mathbf{A}'$ be a right exact (additive) functor (the left exact case proceeds in a similar way). Consider a short exact sequence in $\mathbf{A}$:

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0 \tag{6.7}$$

Now choose projective resolutions for A and C. By the *horseshoe lemma* we obtain a projective resolution for B which fits in a short (split) exact sequence of chain complexes:

$$0 \longrightarrow A_\bullet \longrightarrow B_\bullet \longrightarrow C_\bullet \longrightarrow 0 \tag{6.8}$$

Since F is additive and the above sequence is split exact the induced complex is also split exact, i.e. the following sequence is exact

$$0 \longrightarrow \overline{F}A_\bullet \longrightarrow \overline{F}B_\bullet \longrightarrow \overline{F}C_\bullet \longrightarrow 0 \tag{6.9}$$

is exact and so the *zig-zag lemma* is applicable. This theorem gives us the following long exact sequence in homology:

$$\cdots \longrightarrow H_i(\overline{F}B_\bullet) \longrightarrow H_i(\overline{F}C_\bullet) \longrightarrow H_{i-1}(\overline{F}A_\bullet) \longrightarrow H_{i-1}(\overline{F}B_\bullet) \longrightarrow \cdots \tag{6.10}$$

These homology groups are by definition the same as the left derived functors ($L_i = H_i \circ \overline{F}$) and accordingly we have a long exact sequence relating the different derived functors.

Corollary 6.3.8. By the above long exact sequence of derived functors we see that the first derived functor gives the obstruction to F being an exact functor. Since exact functors have vanishing derived functors we obtain the following result:

$$L_1F = 0 \implies L_iF = 0 \quad \forall i \geq 1 \tag{6.11}$$

6.3.1 Module categories

Let us first consider the tensor and hom bifunctors $- \otimes -$ and $\mathrm{Hom}(-,-)$. The tensor functor is right exact in both arguments while the hom functor is left exact in both arguments and hence we can form the associated left and right derived functors. For simplicity we will always work in the first argument. A proof that the derived functors are *balanced*, i.e. that one can use a projective resolution for either argument and obtain isomorphic results, can be found in the references cited in the beginning of the chapter.

Definition 6.3.9 (Tor-functor). Consider a finite group G and a G-module B. The Tor-functors $\mathrm{Tor}_G^n(-, B)$ are defined as the left derived functors of the tensor functor $- \otimes_{\mathbb{Z}G} B$.

Definition 6.3.10 (Ext-functor). Consider a finite group G and a G-module A. The Ext-functors $\mathrm{Ext}^G_n(-, A)$ are defined as the right derived functors of the hom functor $\mathrm{Hom}_{\mathbb{Z}G}(-, A)$.

For module categories over a ring one can use a simpler technique. For every module M one can construct a free resolution: Choose a free module which maps F_0 onto M; then choose a free module F_1 which maps onto the kernel of the map $p_0 : F_0 \to M$ and so forth. So without further ado we will always assume the resolutions to be free.

6.3.2 Group cohomology

In this section we will give an application of derived functors. In differen areas of mathematics and physics the concept of group cohomology pops up, e.g. group extensions, projective representations and symmetry-protected topological order. These applications however often start from an a priori ad hoc construction based on maps from a group G to a G-module (we will come back to this point further below).

For simplicity we will only consider finite groups and Abelian coefficients groups. Every (Abelian) G-module can be regarded as a module over the group ring $\mathbb{Z}G$, i.e. there exists an equivalences

of categories between **Ab** and $\mathbb{Z}G$-**Mod**. Assuming the axiom of choice, every module category over a ring has enough projectives and hence it makes sense to define group (co)homology[2] using derived functors. For groups we will give an explicit construction of a resolution which is not just $\mathbb{Z}G$-projective but even $\mathbb{Z}G$-free.

The homology and cohomology of a finite group G with coefficients in a G-module A is defined using the Ext- and Tor-functors defined above:

$$H^\bullet(G;A) := \mathrm{Ext}^\bullet_{\mathbb{Z}G}(\mathbb{Z},A) \tag{6.12}$$

$$H_\bullet(G;A) := \mathrm{Tor}^{\mathbb{Z}G}_\bullet(A,\mathbb{Z}) \tag{6.13}$$

where $\mathbb{Z}$ carries the trivial G-module structure. To explicitly calculate the (co)homology groups we have to find a $\mathbb{Z}G$-projective resolution of $\mathbb{Z}$:

Construction 6.3.11 (Normalized bar resolution). Let P'_n be the free $\mathbb{Z}G$-module on G^n. The boundary maps are defined as follows:

$$\partial_k(g_1,...,g_k) = g_1(g_2,...,g_k) + \sum_{i=1}^{k}(-1)^i(g_1,...,g_ig_{i+1},...,g_k) + (-1)^k(g_1,...,g_{k-1}) \tag{6.14}$$

To obtain the normalized bar[3] resolution (in inhomogeneous form) one has to quotient out the submodule of P'_n generated by tuples $(g_1,...,g_n)$ where one of the g_i's is the identity. It can be shown that the resulting quotient modules P_n form a $\mathbb{Z}G$-free (in particular $\mathbb{Z}G$-projective) resolution of $\mathbb{Z}$.

To explicitly calculate the cohomology groups $H^k(G;A) = H^k(\mathrm{Hom}_{\mathbb{Z}G}(P_\bullet,A))$ it is often easier to work with a more explicit description of the involved Hom-sets. Since P'_k is the free $\mathbb{Z}G$-module on G^k, it is isomorphic (as a module) to $\mathbb{Z}[G^{k+1}]$. This can be seen as follows. The generating set consists of all k-tuples of elements in G:

$$S = \{(g_1,...,g_k) : g_i \in G, \forall i \leq k\}$$

Since we generate a module over $\mathbb{Z}G$, we can write every element as formal combination over $\mathbb{Z}$ of elements of the form

$$g_0(g_1,...,g_k)$$

where the multiplication, regarded as the (left) G-action, between g_0 and the k-tuple is merely formal. We can now construct a morphism φ between this module and $\mathbb{Z}[G^{k+1}]$, which carries the diagonal G-action, in the following way. On the generating set S we define φ to be:

$$\varphi(g_1,...,g_k) = (e, g_1, g_1g_2, ..., g_1g_2\cdots g_k) \tag{6.15}$$

It is not hard to show that morphism is in fact an isomorphism (of G-modules) and hence we find that

$$H^k(G;A) = H^k(\mathrm{Hom}_{\mathbb{Z}G}(\mathbb{Z}[G^{k+1}],A)) \tag{6.16}$$

By a little more algebra it can also be shown that this Hom-set is isomorphic to the (set-theoretic) mapping space $\mathrm{Map}(G^k,A)$. This space can be given an Abelian group structure induced by the group structure of A. Combining these facts we get the following construction for the cohomology of groups:

[2] In fact most types of (co)homology for rings, algebras and modules can be defined this way.

[3] One of the possible explanations for this name is that the formal generating elements are often written as $[g_1|g_2|...|g_k]$.

Construction 6.3.12 (Group cohomology). Let G be a finite group and let A be a G-module. Let C^k be the free Abelian group generated by the set-theoretic functions $f : G^k \to A$ with the property that if any of its arguments is the identity then the result is 0. The boundary maps ∂^k, induced by the maps defined in equation 6.14, are given by:

$$(\partial^k f)(g_1, ..., g_{k+1}) = g_1 \cdot f(g_2, ..., g_{k+1}) + \sum_{i=1}^{k} (-1)^i f(..., g_i g_{i+1}, ...) + (-1)^{k+1} f(g_1, ..., g_k) \tag{6.17}$$

Property 6.3.13 (Finiteness). Let G be a finite group and let A be a G-module such that the underlying group is finitely generated. Since in this case the Hom-groups are finitely generated, the cohomology groups $G^k(G; A)$ with $k \geq 1$ are also finitely generated. Furthermore, they are annihilated by the order of G so in particular they are all torsion. It follows that all cohomology groups are finite.

6.4 Spectral sequences

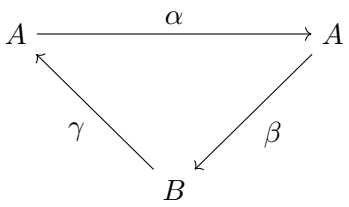
Definition 6.4.1 (Spectral sequence). Consider a sequence of differential objects $\{(E_i, d_i)\}$. This sequence is called a spectral sequence if it satisfies

$$H(E_i, d_i) = E_{i+1} \tag{6.18}$$

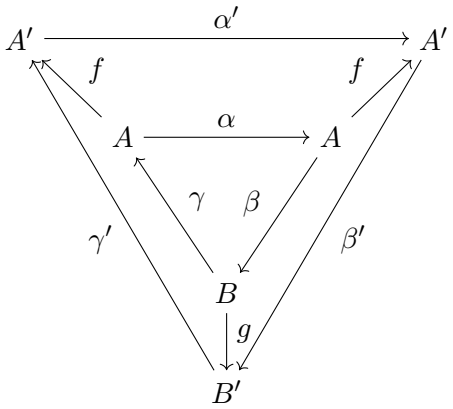
for every i. A morphism of spectral sequences is a collection of morphism $\{\varphi_i\}$ satisfying:

- $\varphi_i \circ d_i = d'_i \circ \varphi_i$
- $\varphi_{i+1} = H(\varphi_i)$

Definition 6.4.2 (Exact couple). An exact "couple" is a tuple $(A, B, \alpha, \beta, \gamma)$ such that the diagram 6.1a commutes.



(a) Exact couple.



(b) Morphism of exact couples.

Figure 6.1: Exact couples and their morphisms.

A morphism of exact couples is a pair of morphisms $(f, g) : (A, B) \to (A', B')$ that make the diagram 6.1b commute.

Part III

Topology

Chapter 7

General Topology

7.1 Topological spaces

Definition 7.1.1 (Topology). Let Ω be a set. Let $\tau \subseteq 2^{\Omega}$. The set τ is a topology on Ω if it satisfies following axioms:

1. $\emptyset \in \tau$ and $\Omega \in \tau$
2. $\forall\ \mathcal{F} \subseteq \tau : \bigcup_{V \in \mathcal{F}} V \in \tau$
3. $\forall\ U, V \in \tau : U \cap V \in \tau$

Furthermore we call the elements of τ open sets and the couple (Ω, τ) a topological space.

Remark. On topological spaces the open sets are thus defined by axioms.

Property 7.1.2. Consider a topological space (X, τ). Let $U \subseteq V \in \tau$. The inclusion maps $U \hookrightarrow V$ are morphisms. The set of these morphisms together with the topology τ form a (small) category $\mathrm{Open}(X)$.

Definition 7.1.3 (Relative topology[1]). Let (X, τ_X) be a topological space and Y a subset of X. We can turn Y into a topological space by equipping it with the following topology:

$$\tau_{\mathrm{rel}} = \{U_i \cap Y : U_i \in \tau_X\} \tag{7.1}$$

Definition 7.1.4 (Disjoint union). Let $\{X_i\}_{i \in I}$ be a family of topological spaces. Now consider the disjoint union

$$X = \bigsqcup_{i \in I} X_i \tag{7.2}$$

together with the canonical inclusion maps $\phi_i : X_i \to X : x_i \mapsto (x_i, i)$. We can turn X into a topological space by equipping it with the following topology:

$$\tau_X = \{U \subseteq X | \forall i \in I : \phi_i^{-1}(U) \text{ is open in } X_i\} \tag{7.3}$$

Definition 7.1.5 (Quotient space). Let X be a topological space and let $\sim$ be an equivalence relation defined on X. The set $X/_\sim$ can be turned into a topological space by equipping it with the following topology:

$$\tau_\sim = \{U \subseteq X/_\sim | \pi^{-1}(U) \text{ is open in } X\} \tag{7.4}$$

where π is the canonical surjective map from X to $X/_\sim$.

[1] Sometimes called the **subspace topology**.

Example 7.1.6 (Discrete topology). The discrete topology is the topology such that every subset is open (and thus also closed).

Example 7.1.7 (Product topology). First consider the case where the index set I is finite. The product space $X = \prod_{i\in I} X_i$ can be turned into a topological space by equipping it with the topology generated by the following basis:

$$\mathcal{B} = \left\{ \prod_{i\in I} U_i \;\middle|\; U_i \in \tau_i \right\} \tag{7.5}$$

For general cases (countably infinite and uncountable index sets) the topology can be defined using the canonical projections $\pi_i : X \to X_i$. The general product topology (**Tychonoff topology**) is the coarsest (finest) topology such that all projections π_i are continuous.

Definition 7.1.8 (Topological group). A topological group is a group G equipped with a topology such that both the multiplication and inversion map are continuous.

Definition 7.1.9 (Pointed topological space). Let $x_0 \in X$. The triple (X, τ, x_0) is called a pointed topological space with base point x_0.

Construction 7.1.10 (Suspension). Let X be a topological space. The suspension of X is defined as the following quotient space:

$$SX = (X \times [0,1])/\{(x,0) \sim (y,0) \text{ and } (x,1) \sim (y,1) | x, y \in X\} \tag{7.6}$$

Construction 7.1.11 (Attaching space). Let X, Y be two topological spaces and consider a subspace $A \subseteq Y$. For every continuous map $f : A \to X$, called the **attaching map**, we can construct the attaching space[2] $X \cup_f Y$ in the following way:

$$X \cup_f Y = (X \sqcup Y)/\{A \sim f(A)\} \tag{7.7}$$

Construction 7.1.12 (Join). Let $\{A_i\}_{i\leq n}$ be a collection of topological spaces. The join, denoted by $A = A_1 \circ \cdots \circ A_n$, is defined as follows: Every point of A is defined by an n-tuple of non-negative numbers $\{t_i\}_{i\leq n}$ satisfying[3] $\sum_i t_i = 1$ and for each index i such that $t_i \neq 0$ a point $a_i \in A_i$. This point in A is then denoted by $t_1 a_1 \oplus \cdots \oplus t_n a_n$.

In the case of two spaces one has a more intuitive (but equivalent) construction: Let A, B be two topological spaces. The join $A \circ B$ is defined as the quotient space $(A \times B \times [0,1])/\sim$ where the relation $\sim$ is defined as follows:

- For all $a \in A$ and $b, b' \in B$: $(a, b, 0) \sim (a, b', 0)$
- For all $a, a' \in A$ and $b \in B$: $(a, b, 1) \sim (a', b, 1)$

which can be viewed as collapsing one end of the cilinder $(A \times B) \times [0,1]$ to A and the other end to B.

Property 7.1.13. The join induces a monoidal structure on the category **Top** where the tensor unit is given by the empty space $\emptyset$.

7.2 Neighbourhoods

7.2.1 Neighbourhoods

Definition 7.2.1 (Neighbourhood). A set $V \subseteq \Omega$ is a neighbourhood of a point $a \in \Omega$ if there exists an open set $U \in \tau$ such that $a \in U \subseteq V$.

[2] Sometimes called the **adjunction space**.

[3] Hence an element of an n-simplex. (See definition 9.2.1.)

Definition 7.2.2 (Basis). Let $\mathcal{B} \subseteq \tau$ be a family of open sets. The family $\mathcal{B}$ is a basis for the topological space (Ω, τ) if every $U \in \tau$ can be written as:

$$U = \bigcup_{V \in \mathcal{F}} V \tag{7.8}$$

where $\mathcal{F} \subseteq \mathcal{B}$.

Definition 7.2.3 (Local basis). Let $\mathcal{B}_x$ be a family of open neighbourhoods of a point $x \in \Omega$. $\mathcal{B}_x$ is a local basis of x if every neighbourhood of x contains at least one element in $\mathcal{B}_x$.

Definition 7.2.4 (First-countability). A topological space (Ω, τ) is first-countable if for every point $x \in \Omega$ there exists a countable local basis.

Property 7.2.5 (Decreasing basis). Let $x \in \Omega$. If there exists a countable local basis for x then there also exists a countable decreasing local basis for x.

Definition 7.2.6 (Second-countability). A topological space (Ω, τ) is second-countable if there exists a countable global basis.

Property 7.2.7. Let X be a topological space. The closure of a subset V is given by:

$$\overline{V} = \{x \in X | \exists \text{ a net } (x_\lambda)_{\lambda \in I} \text{ in } X : x_\lambda \to x\} \tag{7.9}$$

This implies that the topology on X is completely determined by the convergence of nets[4].

Corollary 7.2.8. In first-countable spaces we only have to consider the convergence of sequences.

Definition 7.2.9 (Germ). Let X be a topological space and let Y be a set. Consider two functions $f, g : X \to Y$. If there exists a neighbourhood U of a point $x \in X$ such that

$$f(u) = g(u) \qquad \forall u \in U$$

then this property defines an equivalence relation denoted by $f \sim_x g$ and the equivalence classes are called **germs**.

Property 7.2.10. Let the set Y in the previous definition be the set of reals $\mathbb{R}$. Then the germs at a point $p \in X$ satisfy following closure/linearity relations:

- $[f] + [g] = [f + g]$
- $\lambda[f] = [\lambda f]$
- $[f][g] = [fg]$

where $[f], [g]$ are two germs at p and $\lambda \in \mathbb{R}$ is a scalar.

7.2.2 Separation axioms

Definition 7.2.11 (T_0 axiom[5]). A topological space is T_0 if for every two distinct points x, y at least one of them has a neighbourhood not containing the other. The points are said to be topologically distinguishable.

Definition 7.2.12 (T_1 axiom[6]). A topological space is T_1 if for every two distinct points x, y there exists a neighbourhood U of x such that $y \notin U$. The points are said to be separated.

[4] See definition 2.3.9.

[5] T_0 spaces are also said to carry the **Kolmogorov topology**.

[6] T_1 spaces are also said to carry the **Fréchet topology**.

Definition 7.2.13 (Hausdorff space). A topological space is a Hausdorff space or T_2 space if it satisfies the following axiom:

$$(\forall x, y \in \Omega)(\exists \text{ neighbourhoods } U, V)(x \in U, y \in V, U \cap V = \emptyset) \tag{7.10}$$

This axiom is called the **Hausdorff separation axiom** or T_2 axiom. The points are said to be separated by neighbourhoods.

Property 7.2.14. Every singleton (and thus also every finite set) is closed in a Hausdorff space.

Definition 7.2.15 (Urysohn space[7]). A topological space is an Urysohn space if every two distinct points are separated by closed neighbourhoods.

Definition 7.2.16 (Regular space). A topological space is said to be regular if for every closed subset F and every point $x \notin F$ there exist disjoint open subsets U, V such that $x \in U$ and $F \subset V$.

Definition 7.2.17 (T_3 axiom). A space that is both regular and T_0 is T_3.

Definition 7.2.18 (Normal space). A topological space is said to be normal if every two closed subsets have disjoint neighbourhoods.

Definition 7.2.19 (T_4 axiom). A space that is both normal and T_1 is T_4.

Property 7.2.20. A space satisfying the separation axiom T_k also satisfies all separation axioms $T_{i \leq k}$.

7.3 Morphisms

7.3.1 Convergence

Definition 7.3.1 (Convergence). A sequence $(x_n)_{n \in \mathbb{N}}$ in X is said to converge to a point $a \in X$ if:

$$(\forall \text{ neighbourhoods } U \text{ of } a)(\exists N > 0)(\forall n > N)(x_n \in U) \tag{7.11}$$

Property 7.3.2. Every subsequence of a converging sequence converges to the same point[8].

Property 7.3.3. Let X be a Hausdorff space. The limit of a converging sequence in X is unique.

7.3.2 Continuity

Definition 7.3.4 (Continuity). A function $f : X \to Y$ is continuous if the inverse image $f^{-1}(U)$ of every open set U is also open.

Theorem 7.3.5. *Let X be a first-countable space. Consider a function $f : X \to Y$. The following statements are equivalent:*

- *f is continuous*
- *The sequence $(f(x_n))_{n \in \mathbb{N}}$ converges to $f(a) \in Y$ whenever the sequence $(x_n)_{n \in \mathbb{N}}$ converges to $a \in X$.*

[7] Sometimes called a $T_{2 1/2}$ space.

[8] This limit does not have to be unique. See the next property for more information.

Corollary 7.3.6. If the space Y in the previous theorem is Hausdorff then the limit $f(a)$ does not need to be known because the limit is unique (see 7.3.3).

Remark 7.3.7. If the space X is not first-countable, we have to consider the convergence of nets 2.3.9.

Theorem 7.3.8 (Urysohn's lemma). *A topological space X is normal*[9] *if and only if every two closed disjoint subsets $A, B \subset X$ can be separated by a continuous function $f : X \to [0, 1]$ i.e.*

$$f(a) = 0, \forall a \in A \qquad f(b) = 1, \forall b \in B \tag{7.12}$$

Theorem 7.3.9 (Tietze extension theorem). *Let X be a normal space and let $A \subset X$ be a closed subset. Consider a continuous function $f : A \to \mathbb{R}$. There exists a continuous function $F : X \to \mathbb{R}$ such that $\forall a \in A : F(a) = f(a)$. Furthermore, if the function f is bounded then F can be chosen to be bounded by the same number.*

Remark. The Tietze extension theorem is equivalent to Urysohn's lemma.

7.3.3 Homeomorphisms

Definition 7.3.10 (Homeomorphism). A map f is called a homeomorphism if both f and f^{-1} are continuous and bijective.

Definition 7.3.11 (Diffeomorphism). A homeomorphism, differentiable of class C^k, is called a C^k-diffeomorphism.

Definition 7.3.12 (Embedding). A continuous map is an embedding if it is a homeomorphism onto its image.

Definition 7.3.13 (Local homeomorphism). A continuous map $f : X \to Y$ is a local homeomorphism if for every point $x \in X$ there exists an open neighbourhood U such that $f(U)$ is open and such that f_U is an embedding.

Definition 7.3.14 (Mapping cylinder). Let $f : X \to Y$ be a continuous function. The mapping cylinder M_f is defined as follows:

$$M_f = \left([0, 1] \times X \bigsqcup Y\right) / \sim_f \tag{7.13}$$

where the equivalence relation $\sim_f$ is generated by the relations $(0, x) \sim f(x)$. From this definition it follows that the "top" of the cylinder is homeomorphic to X and the "base" is homeomorphic to $f(X) \subseteq Y$.

Definition 7.3.15 (Covering space). Consider two topological spaces X, C and a continuous surjective map $\phi : C \to X$, called the **covering map**. C is said to be a covering space of X if for all points $x \in X$ there exists a neighbourhood U of x such that $\phi^{-1}(U)$ can be written as a disjoint union $\bigsqcup_i C_i$ of open sets in C such that every set C_i is mapped homeomorphically onto U. The neighbourhoods U are said to be **evenly covered**.

Definition 7.3.16 (Universal covering space). A covering space C is said to be universal if it is simply-connected[10].

Universal property 7.3.17. Let X be a topological space and let C_X be the universal covering space of X, every other covering space C of X is also covered by C_X.

[9] See definition 7.2.18.
[10] See definition 9.1.13.

Definition 7.3.18 (Deck transformation). Let $p : C \to X$ be a covering map of X. The group of deck transformations[11] of (C, p) is given by all homeomorphisms φ satisfying $p \circ \varphi = p$.

Definition 7.3.19 (Étalé space). Let X be a topological space. A topological space Y is an étalé space over X if there exists a continuous surjective map $\pi : Y \to X$ such that π is a local homeomorphism. The preimage $\pi^{-1}(x)$ of a point $x \in X$ is called the **stalk** of π over x.

Property 7.3.20. Every covering space is an étalé space.

Definition 7.3.21 (Section). A section of an étalé space $\pi : Y \to X$ over an open set $U \subseteq X$ is a continuous map f such that $\pi \circ f = \mathbb{1}_U$.

Definition 7.3.22 (Pseudogroup). Let X be a topological space. A collection $\mathcal{G}$ of homeomorphisms $\phi : U \subseteq M \to M$ such that:

- $\mathbb{1}_M \in \mathcal{G}$
- If $\phi \in \mathcal{G}$ then $\phi^{-1} \in \mathcal{G}$
- If $V \subset U$ is open then $\phi|_V \in \mathcal{G}$
- If $U = \bigcup_{i \in I} U_i$ and $\phi_i : U_i \to$ is an element of $\mathcal{G}$ for all $i \in I$ then $\phi \in \mathcal{G}$
- If $\phi : U \to V$ and $\psi : U' \to V'$ are elements of $\mathcal{G}$ and $V \cap U' \neq \emptyset$ then $\psi \circ \phi|_{\phi^{-1}(V \cap U')} \in \mathcal{G}$

7.4 Connected spaces

Definition 7.4.1 (Connected space). A topological space X is connected if it cannot be written as the disjoint union of two non-empty open sets. Equivalently, X is connected if the only clopen sets are X and $\emptyset$.

Property 7.4.2. Let X be a connected space. Let f be a function on X. If f is locally constant, i.e. for every $x \in X$ there exists a neighbourhood U on which f is constant, then f is constant on all of X.

Theorem 7.4.3 (Intermediate value theorem). *Let X be a connected space. Let $f : X \to \mathbb{R}$ be a continuous function. If $a, b \in f(X)$ then for every $c \in]a, b[$ we have that $c \in f(X)$.*

Definition 7.4.4 (Path-connected space[12]). Let X be a topological space. If for every two points $x, y \in X$ there exists a continuous function $\varphi : [0, 1] \to X$ (i.e. a **path**) such that $\varphi(0) = x$ and $\varphi(1) = y$ then the space is said to be path-connected.

Property 7.4.5. Every path-connected space is connected.

The converse does not hold. There exists however the following (stronger) relation:

Property 7.4.6. A connected and locally path-connected space is path-connected.

Remark 7.4.7. The notions of connectedness and path-connectedness define equivalence relations on the space X. The equivalence classes are closed in X and form a cover of X.

[11] In fact this group forms the automorphism group of (C, p) in the category of covering spaces of X.

[12] A similar notion is that of **arcwise-connectedness** where the function φ is required to be a homeomorphism.

7.5 Compact spaces

7.5.1 Compactness

Definition 7.5.1 (Sequentially compact). A topological space is sequentially compact if every sequence[13] has a convergent subsequence.

Definition 7.5.2 (Finite intersection property). A family $\mathcal{F} \subseteq 2^X$ of subsets has the finite intersection property[14] if every finite subfamily has a non-zero intersection:

$$\bigcap_{i \in I} V_i \neq \emptyset \tag{7.14}$$

for all finite index sets I.

Definition 7.5.3 (Locally finite cover). An open cover of a topological space X is said to be locally finite if every $x \in X$ has a neighbourhood that intersects only finitely many sets in the cover of X.

Property 7.5.4. A first-countable space is sequentially compact if and only if every countable open cover has a finite subcover.

Definition 7.5.5 (Lindelöf space). A space for which every open cover has a countable subcover.

Property 7.5.6. Every second-countable space is also a Lindelöf space.

Definition 7.5.7 (Compact space). A topological space X is compact if every open cover of X has a finite subcover.

Theorem 7.5.8 (Heine-Borel[15]). *If a topological space X is sequentially compact and second-countable then every open cover has a finite subcover. This implies that X is compact.*

Theorem 7.5.9 (Heine-Borel on real numbers). *A subset of $\mathbb{R}^n$ is compact if and only if it is closed and bounded.*

Theorem 7.5.10 (Tychonoff's theorem). *Any product[16] of compact topological spaces is again compact when equipped with the (Tychonoff) product topology 7.1.7.*

Definition 7.5.11 (Relatively compact). A topological space is called relatively compact if its closure is compact.

Definition 7.5.12 (Locally compact). A topological space is locally compact if every point $x \in X$ has a compact neighbourhood.

Theorem 7.5.13 (Dini). *Let (X, τ) be a compact space. Let $(f_n)_{n \in \mathbb{N}}$ be an increasing sequence of continuous functions $f_n : X \to \mathbb{R}$. If $(f_n)_n \to f$ pointwise to a continuous function f then the convergence is uniform.*

Definition 7.5.14 (Paracompact space). A topological space is paracompact if every open cover has a locally finite open refinement.

Property 7.5.15. A paracompact Hausdorff space is normal.

[13] The sequence itself does not have to converge.
[14] The family is then called a FIP-family.
[15] Also Borel-Lebesgue.
[16] Finite, countably infinite or even uncountably infinite.

Definition 7.5.16 (ω-boundedness). Let X be a topological space. X is said to be ω-bounded if the closure of every countable subset is compact.

Definition 7.5.17 (Partition of unity). Let $\{\varphi_i : X \to [0,1]\}_i$ be a collection of continuous functions such that for every $x \in X$:

- For every neighbourhood U of x, the set $\{f_i : \text{supp} f_i \cap U \neq \emptyset\}$ is finite.
- $\sum_i f_i = 1$

Definition 7.5.18 (Subordinate). Consider an open cover $\{V_i\}_{i\in I}$ of X, indexed by a set I. If there exists a partition of unity, also indexed by I, such that $\text{supp}(\varphi_i) \subseteq U_i$, then this partition of unity is said to be **subordinate** to the open cover.

Property 7.5.19. A paracompact space is Hausdorff if and only if it admits a partition of unity subordinate to any open cover.

Definition 7.5.20 (Numerable open cover). An open cover $\{U_i\}_{i\in I}$ of a space X is said to be numerable if X admits a partition of unity subordinate to $\{U_i\}_{i\in I}$.

7.5.2 Compactifications

Definition 7.5.21 (Dense). A subset $V \subseteq X$ is dense in a topological space X if $\overline{V} = X$.

Definition 7.5.22 (Separable space). A topological space is separable if it contains a countable dense subset.

Property 7.5.23. Every second-countable space is separable.

Definition 7.5.24 (Compactification). A compact topological space (X', τ') is a compactification of a topological space (X, τ) if X is a dense subspace of X'.

Example 7.5.25. Standard examples of compactifications are the extended real line $\mathbb{R} \cup \{-\infty, +\infty\}$ and the extended complex plane $\mathbb{C} \cup \{\infty\}$ for the real line and the complex plane respectively.

Remark. It is important to note that compactifications are not necessarily unique.

Definition 7.5.26 (One-point compactification). Let X be a Hausdorff space. A one-point compactification or **Alexandrov compactification** is a compactification X' such that $X' \setminus X$ is a singleton.

7.6 Locales

Property 7.6.1. Consider the poset $\mathbf{Open}(X)$ of opens of a topological space X. This set is closed under finite intersections (limits) and arbitrary unions (colimits). Furthermore, arbitrary unions distribute over finite intersections:

$$V \cap \left(\bigcup_{i\in I} U_i \right) = \bigcup_{i\in I} (V \cap U_i) \tag{7.15}$$

This implies that the poset $\mathbf{Open}(X)$ is a frame[17].

[17] See definition 2.3.19.

Definition 7.6.2 (Locale). The previous property can be used to generalize the notion of topological space to include *pointless spaces*. Let **Frame** denote the category of frames together with frame homomorphisms. The category of locales is defined as the opposite category:

$$\mathbf{Loc} = \mathbf{Frame}^{op}$$

Construction 7.6.3 (From locale to topological space). There exists an adjunction

$$\mathbf{Loc} \underset{\text{Point}}{\overset{\iota}{\underset{\longrightarrow}{\overset{\longleftarrow}{\bot}}}} \mathbf{Top}$$

where the right adjoint is defined as follows:

Let L be a locale. For a topological space the points are given by continuous maps $* \to X$ and hence by frame morphisms $\mathbf{Open}(X) \to \Omega_{\text{Frame}} = \{0, 1\}$. Generalizing this to locales one defines the set of points of L as the Ω_{Loc}-elements:

$$\text{Point}(L) = \mathbf{Loc}(\Omega_{\text{Loc}}, L)$$

This set can be given a topology by declaring for every $U \in L$ the set $\{p \in \text{Point}(L) : p^{-1}(U) = 1\}$ to be open.

Chapter 8

Metric Spaces

8.1 General definitions

Definition 8.1.1 (Metric). A metric (or distance) on a set M is a map $d : M \times M \to \mathbb{R}^+$ that satisfies the following properties:

- Non-degeneracy: $d(x, y) = 0 \iff x = y$
- Symmetry: $d(x, y) = d(y, x)$
- Triangle inequality: $d(x, z) \leq d(x, y) + d(y, z) \quad , \forall x, y, z \in M$

Definition 8.1.2 (Metric space). A set M equipped with a metric d is called a metric space and is denoted by (M, d).

Definition 8.1.3 (Diameter). The diamater of a subset $U \subset M$ is defined as

$$\operatorname{diam}(U) = \sup_{x,y \in U} d(x, y) \tag{8.1}$$

Definition 8.1.4 (Bounded). A subset $U \subseteq M$ is bounded if $\operatorname{diam}(U) < +\infty$.

Property 8.1.5. Every metric space is a topological space[1].

Multiple topological notions can be reformulated in terms of a metric. The most important of them are given below:

Definition 8.1.6 (Open ball). An open ball centered on a point $x_0 \in M$ with radius $R > 0$ is defined as the set:

$$\boxed{B(x_0, R) = \{x \in M : d(x, x_0) < R\}} \tag{8.2}$$

Definition 8.1.7 (Closed ball). The closed ball $\overline{B}(x_0, R)$ is defined as the union of the open ball $B(x_0, R)$ and its boundary, i.e. $\overline{B}(x_0, R) = \{x \in M : d(x, x_0) \leq R\}$.

Definition 8.1.8 (Interior point/neighbourhood). Let N be a subset of M. A point $x \in N$ is said to be an interior point of N if there exists an $R > 0$ such that $B(x, R) \subset M$. Furthermore, N is said to be a neighbourhood of x.

Definition 8.1.9 (Open set). A subset $N \subset M$ is said to be open if every point $x \in N$ is an interior point of N.

[1] See next chapter.

Definition 8.1.10 (Closed set). A subset $V \subset M$ is said to be closed if its complement is open.

Definition 8.1.11 (Limit point). Let S be a subset of X. A point $x \in X$ is called a limit point of S if every neighbourhood of x contains at least one point of S different from x.

Definition 8.1.12 (Accumulation point). Let $x \in X$ be a limit point of S. Then x is an accumulation point of S if every open neighbourhood of x contains infinitely many points of S.

Definition 8.1.13 (Convergence). A sequence $(x_n)_{n\in\mathbb{N}} : \mathbb{N} \to M$ in a metric space (M, d) is said to be convergent to a point $a \in M$ if:

$$\forall \varepsilon > 0 : \exists N_0 \in \mathbb{N} : \forall n \geq N_0 : d(x_n, a) < \varepsilon \tag{8.3}$$

Definition 8.1.14 (Continuity). Let (M, d) and (M', d') be two metric spaces. A function $f : M \to M'$ is said to be continuous at a point $a \in \mathrm{dom}(f)$ if:

$$\forall \varepsilon > 0 : \exists \delta_\varepsilon : \forall x \in \mathrm{dom}(f) : d(a, x) < \delta_\varepsilon \implies d'(f(a), f(x)) < \varepsilon \tag{8.4}$$

Property 8.1.15. Let (M, d) be a metric space. The distance function $d : M \times M \to \mathbb{R}$ is a continuous function.

Definition 8.1.16 (Uniform continuity). Let (M, d) and (M', d') be two metric spaces. A function $f : M \to M'$ is said to be uniformly continuous if:

$$\forall \varepsilon > 0 : \exists \delta_\varepsilon : \forall x, y \in \mathrm{dom}(f) : d(x, y) < \delta_\varepsilon \implies d'(f(x), f(y)) < \varepsilon \tag{8.5}$$

This is clearly a stronger notion than that of continuity as the number ε is equal for all points $y \in \mathrm{dom}(f)$.

8.2 Examples of metrics

Definition 8.2.1 (Product space). Consider the cartesian product

$$M = M_1 \times M_2 \times ... \times M_n$$

with $\forall n : (M_n, d_n)$ a metric space. If equipped with the distance function $d(x, y) = \max\limits_{1\leq i\leq n} d_i(x_i, y_i)$ this product is also a metric space. It is called the product metric space.

Property 8.2.2. The projection associated with the set M_j is defined as:

$$\mathrm{pr}_j : M \to M_j : (a_1, ..., a_n) \mapsto a_j \tag{8.6}$$

A sequence in a product metric space M converges if and only if every component $(\mathrm{pr}_j(x_m))_{m\in\mathbb{N}}$ converges in (M_j, d_j).

Example 8.2.3 (Supremum distance). Let $K \subset \mathbb{R}^n$ be a compact set. Denote the set of continuous functions $f : K \to \mathbb{C}$ by $\mathcal{C}(K, \mathbb{C})$. The following map defines a metric on $\mathcal{C}(K, \mathbb{C})$:

$$d_\infty(f, g) = \sup_{x\in K} |f(x) - g(x)| \tag{8.7}$$

Example 8.2.4 (p-metric). We can define following set of metrics on $\mathbb{R}^n$:

$$\boxed{d_p(x, y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p\right)^{1/p}} \tag{8.8}$$

Example 8.2.5 (Chebyshev distance).

$$d_\infty(x,y) = \max_{1\le i\le n} |x_i - y_i| \tag{8.9}$$

It is also called the **maximum metric** or L_∞ metric.

Remark 8.2.6. This metric is also an example of a product metric defined on the Euclidean product space $\mathbb{R}^n$. The notation d_∞, which is also used for the supremum distance, can be justified if the space $\mathbb{R}^n$ is identified with the set of maps $\{1,...,n\} \to \mathbb{R}$ equipped with the supremum distance. Another justification is the following relation:

$$d_\infty(x,y) = \lim_{p\to\infty} d_p(x,y) \tag{8.10}$$

which is also the origin of the name L_∞ metric.

8.3 Metrizable spaces

Definition 8.3.1 (Metrizable space). A topological space X is metrizable if it is homeomorphic to a metric space M or equivalently if there exists a metric function $d : X \times X \to \mathbb{R}$ such that it induces the topology on X.

Theorem 8.3.2 (Urysohn's metrization theorem). *Every second-countable T_3 space is metrizable.*

8.4 Compactness in metric spaces

Theorem 8.4.1 (Stone). *Every metric space is paracompact.*

Definition 8.4.2 (Totally bounded). A metric space M is said to be totally bounded if it satisfies the following equivalent statements:

- For every $\varepsilon > 0$ there exists a finite cover $\mathcal{F}$ of M with $\forall F \in \mathcal{F} : \operatorname{diam}(F) \le \varepsilon$.
- For every $\varepsilon > 0$ there exists a finite subset $E \subset M$ such that $M \subseteq \bigcup_{x\in E} B(x,\varepsilon)$.

Property 8.4.3. Every totally bounded set is bounded and every subset of a totally bounded set is also totally bounded. Furthermore, every totally bounded space is second-countable.

The following theorem is a generalization of the statement "*a set is compact if and only if it is closed and bounded*" known from Euclidean space $\mathbb{R}^n$.

Theorem 8.4.4. *For a metric space M the following statements are equivalent:*

- *M is compact.*
- *M is sequentially compact.*
- *M is complete and totally bounded.*

Theorem 8.4.5 (Heine-Cantor). *Let M, M' be two metric spaces with M being compact. Every continuous function $f : M \to M'$ is also uniformly continuous.*

Definition 8.4.6 (Equicontinuity). Let X be a topological space and let M be a metric space. A collection $\mathcal{F}$ of maps $X \to M$ is equicontinuous in $a \in X$ if for all neighbourhoods U of a:

$$(\forall f \in \mathcal{F})(\forall x \in U)(d(f(x), f(a)) \leq \varepsilon) \tag{8.11}$$

for all $\varepsilon \geq 0$.

Property 8.4.7. Let $I \subseteq \mathbb{R}$ be an open interval. Let $\mathcal{F}$ be a collection of differentiable functions such that $\{f'(t) : f \in \mathcal{F}, t \in I\}$ is bounded. Then $\mathcal{F}$ is equicontinuous.

Theorem 8.4.8 (Arzelà-Ascoli). *Let K be a compact topological space and let M be a complete metric space. The following statements are equivalent for any collection $\mathcal{F} \subseteq C(K, M)$:*

- *$\mathcal{F}$ is compact with respect to the supremum distance[2].*
- *$\mathcal{F}$ is equicontinuous, closed under uniform convergence and $\{f(x) : f \in \mathcal{F}\}$ is totally bounded for every $x \in K$.*

8.5 Complete metric spaces

Definition 8.5.1 (Cauchy sequence). A sequence $(x_n)_{n\in\mathbb{N}}$ in a metric space (M, d) is Cauchy (or has the Cauchy property) if

$$\boxed{(\forall \varepsilon > 0)(\exists N \in \mathbb{N})(\forall m, n \geq N)(d(x_m, x_n) < \varepsilon)} \tag{8.12}$$

Property 8.5.2 (Cauchy criterion). A metric space (M, d) satisfies the Cauchy criterion if a sequence converges to a point $a \in M$ if and only if it is Cauchy.

Definition 8.5.3 (Complete metric space). A metric space is complete if it satisfies the Cauchy criterion.

Property 8.5.4.

- Every closed subset of a complete metric space is complete.
- Every complete subset of a metric space is closed.

8.5.1 Injective metric spaces

Definition 8.5.5 (Metric retraction). Let (M, d) be a metric space. A function $f : X \to X$ is said to be a retraction of metric spaces if:

- f is idempotent
- f is non-expansive, i.e. the following relation holds for all $x, y \in M$:

$$d(f(x), f(y)) \leq d(x, y) \tag{8.13}$$

The image of f is called a (metric) retract of M.

Definition 8.5.6 (Injective metric space). A metric space M is said to be injective if whenever M is isometric to a subspace Y of a metric space X then Y is a metric retract of X.

Property 8.5.7. Every injective metric space is complete.

[2] See formula 8.7.

8.5.2 Convex metric spaces

Definition 8.5.8 (Convex space). A metric space (M, d) is said to be convex if for every two points $x, y \in M$ there exists a third point $z \in M$ such that:

$$d(x, z) = d(x, y) + d(y, z) \tag{8.14}$$

Property 8.5.9. A closed subset of Euclidean space is a convex metric space if and only if it is a convex set.

Definition 8.5.10 (Hyperconvex space). A convex space for which the set of closed balls has the Helly property[3] is called a hyperconvex space.

Theorem 8.5.11 (Aronszajn & Panitchpakdi). *A metric space is injective if and only if it is hyperconvex.*

[3]See definition 2.1.6.

Chapter 9

Algebraic Topology

References for this chapter are [3, 4].

9.1 Homotopy theory

9.1.1 Homotopy

Definition 9.1.1 (Retraction). Let X be a topological space and let $A \subseteq X$ be a subspace. A continuous function $f : X \to A$ is called a retraction (and A is called a **retract** of X) if it satisfies $f(a) = a$ for all $a \in A$.

Definition 9.1.2 (Homotopy). Let $f, g \in \mathcal{C}(X, Y)$ where X, Y are topological spaces. If there exists a continuous function $H : X \times [0, 1] \to Y$ such that $f(x) = H(x, 0)$ and $g(x) = H(x, 1)$ then f and g are said to be homotopic. This relation also induces an equivalence relation on $\mathcal{C}(X, Y)$.

Definition 9.1.3 (Deformation retraction). Let X be a topological space and let $A \subseteq X$ be a subspace. A is called a deformation retract if there exists a homotopy between the identity map on A and a retraction $f : X \to A$.

Definition 9.1.4 (Homotopy type). Let X, Y be two topological spaces. X and Y are said to be homotopy equivalent, or of the same homotopy type, if there exist functions $f : X \to Y$ and $g : Y \to X$ such that $f \circ g$ is homotopic to $\mathbb{1}_Y$ and $g \circ f$ is homotopic to $\mathbb{1}_X$. The maps f, g are called **homotopy equivalences**.

Property 9.1.5. Every homeomorphism is a homotopy equivalence.

Definition 9.1.6 (Null-homotopic). A continuous function is null-homotopic if it is homotopic to a constant function.

Definition 9.1.7 (Contractible space). A topological space X is said to be contractible if the identity map $\mathbb{1}_X$ is null-homotopic. Equivalently, a space is called contractible if it is homotopy-equivalent to a point.

9.1.2 Homotopy groups

In this subsection we will always assume to be working with pointed spaces 7.1.9. The base point will be denoted by x_0.

Definition 9.1.8 (Loop space). The set of all **loops** in a pointed topological space $(X, *)$, i.e. all continuous maps[1] $\delta : (S^1, t_0) \to (X, *)$ for which $\delta(t_0) = *$. It is denoted by ΩX. This set can be equipped with a multiplication operation corresponding to the concatenation of loops[2].

When one drops the requirement of based loops, i.e. one considers the space of all continuous maps $S^1 \to X$, the resulting space is called the **free loop space** on X. This object is denoted by LX.

Definition 9.1.9 (Fundamental group). The fundamental group $\pi_1(X, x_0)$ based at $x_0 \in X$ is defined as the loop space (with base x_0) modulo homotopy. As the name implies the fundamental group can be given the structure of a multiplicative group where the operation is inherited from that of the loop space.

Remark 9.1.10. In general, as the notation implies, the fundamental group depends on the base point x_0. However when the space X is path-connected, the fundamental groups belonging to different base points are isomorphic. It follows that we can speak of "the" fundamental group in the case of path-connected spaces.

Definition 9.1.11 (Fundamental groupoid). Let X be a topological space. The fundamental groupoid $\Pi_1(X)$ is the groupoid which has the points of X as objects and the basepoint-preserving homotopy classes of continuous function $f : X \to X$ as morphisms. The fundamental group $\pi_1(X, a)$ is then defined as the automorphism group of $a \in \mathrm{ob}(\Pi_1(X))$.

Property 9.1.12 (Universal cover). Consider a topological space X and let $\widetilde{X}$ be its universal covering space. The group of deck transformations[3] $\mathrm{Aut}(\widetilde{X})$ is isomorphic to the fundamental group $\pi_1(X)$. Hence we obtain:

$$\widetilde{X} \cong X/\pi_1(X) \tag{9.1}$$

Definition 9.1.13 (Simply-connected space). A topological space is said to be simply-connected if it is path-connected and if the fundamental group is trivial.

The definition of a fundamental group can be generalized to arbitrary dimensions in the following way[4]:

Definition 9.1.14 (Homotopy group). The homotopy group $\pi_n(X, x_0)$ is defined as the set of homotopy classes of continuous maps $f : S^n \to X$ based at $x_0 \in X$. The set $\pi_0(X, x_0)$ is defined as the set of path-connected components of X.

Property 9.1.15. For $n \geq 1$ the sets $\pi_n(X, x_0)$ are groups.

Property 9.1.16. For $n \geq 2$ the homotopy groups $\pi_n(X, x_0)$ are Abelian.

Property 9.1.17. If X is path-connected, then the homotopy groups $\pi_n(X, x_0)$ and $\pi_n(X, x_1)$ are isomorphic for all $x_0, x_1 \in X$ and all $n \in \mathbb{N}$.

Property 9.1.18. Homeomorphic spaces have the same homotopy groups π_n.

Formula 9.1.19. Let (X, x_0) and (Y, y_0) be pointed topological spaces with homotopy groups $\pi_n(X, x_0)$ and $\pi_n(Y, y_0)$. The homotopy groups of their product is given by:

$$\pi_n(X \times Y, (x_0, y_0)) = \pi_n(X, x_0) \otimes \pi_n(Y, y_0) \tag{9.2}$$

where $\otimes$ denotes the direct product of groups 3.1.35.

[1] The mapping space is equipped with the compact-open topology.

[2] It should be noted that the speed at which the concatenated loops are traversed is doubled because the parameter t should remain an element of $S^1 \cong [0, 1]/_{0\sim 1}$.

[3] See definition 7.3.18.

[4] Note however that we replace the interval $[0, 1]$ by the sphere S^1. This is non-restrictive as we can construct S^n by identifying the boundary of $[0, 1]^n$ with the basepoint x_0.

Definition 9.1.20 (n-connected space). A topological space is said to be n-connected if its first n homotopy groups are trivial.

Definition 9.1.21 (Weak homotopy equivalence). A continuous map which induces isomorphisms on all homotopy groups.

Property 9.1.22 (Homotopy category). The homotopy category hTop has as objects the topological spaces and as morphisms the homotopy classes of continuous maps. It is immediately clear that there exists a functor $F : \mathbf{Top} \to \mathbf{hTop}$ that maps topological spaces to themselves and continuous maps to their homotopy classes.

In fact the above definition is often to restrictive. Quillen gave a more general construction: The homotopy category (in the sense of Quillen) is obtained as the localization[5] of **Top** with respect to the collection of weak homotopy equivalences.

Definition 9.1.23 (Eilenberg-MacLane space). Let G be a group and choose a positive integer $n \in \mathbb{N}_0$. The Eilenberg-MacLane space $K(G, n)$ is a topological space with the following property:

$$\pi_i\Big(K(G,n)\Big) = \begin{cases} G & i = n \\ 0 & i \neq n \end{cases} \tag{9.3}$$

It follows from property 9.1.16 above that for $n > 1$ the group G has to be Abelian.

Property 9.1.24. For every G and n the space $K(G, n)$ is unique up to weak homotopy equivalence.

Property 9.1.25. The loop space $\Omega K(G, n)$ is homotopy equivalent to $K(G, n - 1)$.

9.1.3 CW complexes

Definition 9.1.26 (n-cell). An open n-cell is a subset of a topological space homeomorphic to the n-dimensional open ball. A closed n-cell is the image of an n-dimensional closed ball under an attaching map[6].

Definition 9.1.27 (CW complex). A CW complex is a Hausdorff space X together with a partition of X in open cells satsifying following conditions:

- A subset of X is closed if and only if it meets the closure of each cell in a closed et.
- For each open n-cell C in the partition there exists an attaching map $f : \overline{B}_n \to X$ such that:
  - $f|_{B_n}$ is homeomorphic to C.
  - $f(\partial\overline{B}_n)$ is covered by a finite number of open cells in the partition, each having dimension smaller than n.

where $\overline{B}_n$ denotes the closed n-dimensional ball.

Definition 9.1.28 (Regular CW complex). A CW complex is called regular if for every open cell C the attaching map f is a homeomorphism onto the closure $\overline{C}$.

Construction 9.1.29. Every CW complex can, up to isomorphism, be constructed inductively:

First choose a discrete space X_0, i.e. a topological space equipped with the discrete topology. This space forms a 0-cell. Then we can add 1-cells C_1 using appropriate attaching maps

[5] See definition 4.7.17.
[6] See definition 7.1.11.

$f : \partial\overline{B}_1 \to X_0$. This way we obtain a 1-dimensional CW complex X_1. Inductively one obtains a sequence of nested n-dimensional CW complex $X_0 \subset X_1 \subset \cdots \subset X_n$.

The spaces X_i are also called i-**skeletons**.

Remark 9.1.30. Infinite-dimensional CW complexes can be obtained by taking the direct limit[7] of the sequence above.

Theorem 9.1.31 (Whitehead). *A continuous map between CW-complexes is a homotopy equivalence if and only if it is a weak homotopy equivalence.*

9.1.4 Fibrations

Definition 9.1.32 (Homotopy lifting property). Consider a continuous map $\pi : E \to B$ between topological spaces. The map π is said to have the homotopy lifting property with respect to a topological space X if for every homotopy $f : X \times [0,1] \to B$ and lifting $\widetilde{f}_0 : X \to E$ of $f_0 = f|_{X\times\{0\}}$ there exists a homotopy $\widetilde{f} : X \times [0,1]$ lifting f such that the following diagram commutes:

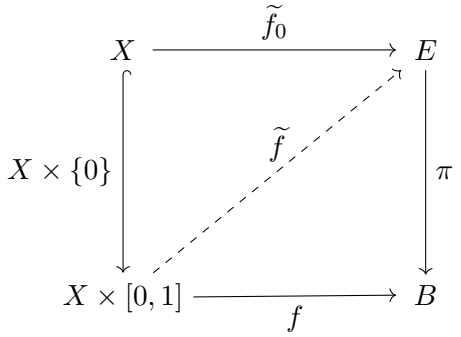


Figure 9.1: Homotopy lifting property.

where $\widetilde{f}$ denotes the lifting of f, i.e. $f = \pi \circ \widetilde{f}$.

Definition 9.1.33 (Hurewicz fibration). A map π satisfying the homotopy lifting property with respect to every topological space X is called a (Hurewicz) fibration.[8]

Property 9.1.34. Consider a fibration $\pi : E \to B$ with B path-connected. All fibres, i.e. sets $\pi^{-1}(\{b\})$ where $b \in B$, are homotopy equivalent. Therefore a fibration is often denoted by the diagram $F \hookrightarrow E \to B$.

We give two important examples:

Example 9.1.35 (Hopf fibration). The Hopf fibration is given by

$$S^1 \hookrightarrow S^3 \to S^2 \tag{9.4}$$

Adam's theorem states that this fibration can be generalized to higher dimensions as $S^n \hookrightarrow S^{2n+1} \to S^{2n}$ only for $n \in \{0, 1, 3, 7\}$.

Example 9.1.36. For all $n \in \mathbb{N}$ the following sequence forms a fibration:

$$\mathrm{SO}(n) \hookrightarrow \mathrm{SO}(n+1) \to S^n \tag{9.5}$$

[7] See definition 3.34.

[8] If the homotopy lifting property only holds with respect to CW complexes (see definition 9.1.27) then it is called a **Serre fibration**.

9.1.5 Rational homotopy theory

Definition 9.1.37 (Rational space). A simply connected topological space X such that the homotopy groups $\pi_n(X)$ are rational vector spaces.

Definition 9.1.38 (Rational homotopy equivalence). A continuous function $f : X \to Y$ for which the induced maps on rational homotopy groups

$$\pi_n(f) \otimes \mathbb{Q} : \pi_n(X) \otimes \mathbb{Q} \to \pi_n(Y) \otimes \mathbb{Q} \tag{9.6}$$

is an isomorphism for all $n \in \mathbb{N}$.

Definition 9.1.39 (Rational homotopy category). Consider the category **Top** of topological spaces. The rational homotopy category is obtained as the localization of **Top** with respect to the collection of rational homotopy equivalences.

9.2 Simplicial homology

9.2.1 Simplices

Definition 9.2.1 (Simplex). A k-simplex $\sigma^k = [t_0, ..., t_k]$ is defined as the following set:

$$\sigma^k = \left\{ \sum_{i=0}^{k} \lambda_i t_i \,\middle|\, \sum_{i=0}^{k} \lambda_i = 1 \text{ and } \lambda_i \geq 0 \right\} \tag{9.7}$$

where the points (**vertices**) $t_i \in \mathbb{R}^n$ are **affinely independent**, i.e. the vectors $t_i - t_0$ are linearly independent. Equivalently, a simplicial k-simplex is the convex hull of the $k+1$ vertices $\{t_0, ..., t_k\}$.

Remark 9.2.2 (Barycentric coordinates). The coordinates λ_i from previous definition are called barycentric coordinates. This follows from the fact that the point $\sum_{i=0}^{k} \lambda_i t_i$ represents the barycenter of a gravitational system consisting of masses λ_i placed at the points t_i.

Notation 9.2.3 (Face). Consider a simplicial k-simplex $[v_0, ..., v_k]$. The face opposite to the vertex v_i is the simplicial $(k-1)$-simplex $[v_0, ..., \hat{v}_i, ..., v_k]$ obtained by removing the vertex v_i.

Definition 9.2.4 (Simplicial complex). A simplicial complex $\mathcal{K}$ is a set of simplices satisfying following conditions:

- If σ is a simplex in $\mathcal{K}$ then so are its faces.
- If $\sigma_1, \sigma_2 \in \mathcal{K}$ then either $\sigma_1 \cap \sigma_2 = \emptyset$ or $\sigma_1 \cap \sigma_2$ is a face of both σ_1 and σ_2.

A simplicial k-complex is a simplicial complex where every simplex has dimension at most k.

Definition 9.2.5 (Path-connectedness). Let $\mathcal{K}$ be a simplicial complex. $\mathcal{K}$ is said to be path-connected if every two vertices in $\mathcal{K}$ are connected by edges in $\mathcal{K}$.

Definition 9.2.6 (Polyhedron). Let $\mathcal{K}$ be a simplicial complex. The polyhedron associated with $\mathcal{K}$ is the topological space constructed by equipping $\mathcal{K}$ with the Euclidean subspace topology.

Definition 9.2.7 (Triangulable spaces). Let X be a topological space and let $\mathcal{K}$ be a polyhedron. If there exists a homeomorphism $\varphi : \mathcal{K} \to X$ then we say that X is triangulable and we call $\mathcal{K}$ a **triangulation** of X.

Theorem 9.2.8. *Let $\mathcal{K}$ be a path-connected polyhedron with basepoint a_0. Let $\mathcal{C} \subset \mathcal{K}$ be a contractible 1-dimensional subpolyhedron containing all vertices of $\mathcal{K}$. Let G be the free group generated by the elements g_{ij} corresponding to the ordered 1-simplices $[v_i, v_j] \in \mathcal{C}$.*

The group G is isomorphic to the fundamental group $\pi_1(\mathcal{K}, a_0)$ if the generators g_{ij} satisfy following two relations:

- *$g_{ij}g_{jk} = g_{ik}$ for every ordered 2-simplex $[v_i, v_j, v_k] \in \mathcal{K}\backslash\mathcal{C}$*
- *$g_{ij} = e$ if $[v_i, v_j] \in \mathcal{C}$.*

Corollary 9.2.9. From the theorem that homeomorphic spaces have the same homotopy groups it follows that the fundamental group of a triangulable space can be computed by looking at its triangulations.

Remark 9.2.10 (Hauptvermutung). Although the homology invariants (which we will define below) do not depend on the choice of triangulation we should give a remark about the existence of non-equivalent triangulations. Before the construction of a counterexample it was believed, hence the terms *Hauptvermutung* in German or *main conjecture* in English, that every two triangulations of a topological space allowed a common refinement and hence where equivalent for many constructions. However it was shown that this conjecture is generally false, e.g. for topological manifolds in dimensions 5 and higher there exist an infinite number of non-equivalent triangulations. In dimensions below 4 it was proven by Radó and Moise that the Hauptvermutung holds for topological manifolds (see theorem 27.1.9).

Definition 9.2.11 (Delta-complex). Let $\{\sigma_\alpha : \Delta^n \to X\}$ be a collection of morphisms from simplicices to X where the dimension n may depend on the subscript α. This collection forms a Delta complex (sometimes Δ-complex) on X if it satisfies the following conditions:

- The restrictions $\sigma_\alpha|_{\overset{\circ}{\Delta}{}^n}$ are injective and every point in X lies in the image of exactly one such restriction.
- The restriction of a morphism σ_α to the any one of the faces of Δ^n is equal to some other σ_β.
- A set in X is open if and only if it is open in all the inverse images σ_α^{-1}.

Similar to a CW-complex or cellular complex these conditions imply that every Delta complex can be constructed inductively from a (discrete) set of vertices by gluing and identifying edges.

9.2.2 Simplicial homology

Definition 9.2.12 (Chain group). Let $\mathcal{K}$ be a simplicial n-complex. The k^{th} chain group $C_k(\mathcal{K})$ is defined as the free Abelian group generated by the k-simplices in $\mathcal{K}$:

$$C_k(\mathcal{K}) = \left\{ \sum_i a_i\sigma_i \;\middle|\; \sigma_i \text{ is a } k\text{-simplex in } \mathcal{K} \text{ and } a_i \in \mathbb{Z} \right\} \tag{9.8}$$

For $k > n$ we define $C_k(\mathcal{K})$ to be $\{0\}$.

Definition 9.2.13 (Boundary operator). The boundary operator $\partial_k : C_k(\mathcal{K}) \to C_{k-1}(\mathcal{K})$ is the group morphism defined by following properties:

- Linearity:

$$\partial_k \left(\sum_i a_i\sigma_i \right) = \sum_i a_i \partial_k \sigma_i \tag{9.9}$$

- For every oriented k-simplex $[v_0, ..., v_k]$:

$$\partial_k[v_0, ..., v_k] = \sum_{i=0}^{k}(-1)^i[v_0, ..., \hat{v}_i, ..., v_k] \tag{9.10}$$

- The boundary of every 0-chain is the identity 0.

Remark 9.2.14. The alternating sum comes down the fact that we want the *oriented* boundary.

Property 9.2.15. The boundary operators satisfy following relation:

$$\partial_k \circ \partial_{k+1} = 0 \tag{9.11}$$

This property turns the system (C_k, ∂_k) into a chain complex[9].

Definition 9.2.16 (Cycle group). The k^{th} cycle group $Z_k(\mathcal{K})$ is defined as the set of k-chains σ_k such that $\partial_k\sigma_k = 0$. These chains are called **cycles**.

Definition 9.2.17 (Boundary group). The k^{th} boundary group $B_k(\mathcal{K})$ is defined as the set of k-chains σ_k for which there exists a $(k+1)$-chain N such that $\partial_{k+1}N = \sigma_k$. These chains are called **boundaries**.

Definition 9.2.18 (Homology group). From property 9.11 it follows that $B_k(\mathcal{K}) \subset Z_k(\mathcal{K})$ is a subgroup. We can thus define the k^{th} homology group $H_k(\mathcal{K})$ as the following quotient group:

$$H_k(\mathcal{K}) = Z_k(\mathcal{K})/B_k(\mathcal{K}) \tag{9.12}$$

Theorem 3.1.50 tells us that we can write $H_k(\mathcal{K})$ as $G_k \oplus T_k$. Both of these groups tell us something about $\mathcal{K}$. The rank of G_k, denoted by $R_k(\mathcal{K})$, is equal to the number of $(k+1)$-dimensional holes in $\mathcal{K}$. The torsion subgroup T_k tells us how the space $\mathcal{K}$ is twisted.

Property 9.2.19. If two topological spaces have the same homotopy type then they have isomorphic homology groups. It follows that homeomorphic spaces have isomorphic homology groups.

Corollary 9.2.20. As was the case for the fundamental group, it follows from the definition of a triangulation that we can construct the homology groups for a given triangulable space by looking at its triangulations.

Definition 9.2.21 (Betti numbers). The ranks $R_k(\mathcal{K})$ are called the Betti numbers of $\mathcal{K}$.

Formula 9.2.22 (Euler characteristic). The Euler characteristic of a triangulable space X is defined as follows[10]:

$$\boxed{\chi(X) = \sum_i (-1)^i R_i(X)} \tag{9.13}$$

Construction 9.2.23. The definition of homology groups can be generalized by letting the (formal) linear combinations used in the definition of the chain group (see 9.2.12) be of the following form:

$$c^k = \sum_i g_i \sigma_i^k \tag{9.14}$$

where $G = \{g_i\}$ is an Abelian group and σ_i^k are k-simplices. The k^{th} homology group of X with coefficients in G is denoted by $H_k(X; G)$.

[9] See definition 6.1.1.

[10] This formula is sometimes called the *Poincaré* or *Euler-Poincaré* formula.

Property 9.2.24. When G is a field, such as $\mathbb{Q}$ or $\mathbb{R}$, the torsion subgroups T_k vanish. The relation between integral homology and homology with coefficients in a group is given by the *Universal coefficient theorem.*

Formula 9.2.25 (Künneth formula). Let X, Y be two triangulable spaces. The homology groups of the Cartesian product $X \times Y$ with coefficients in a field F is given by:

$$H_k(X \times Y; F) = \bigoplus_{k=i+j} H_i(X; F) \otimes H_j(Y; F) \tag{9.15}$$

Remark. When the requirement of F being a field is relaxed to it merely being a group, the torsion subgroups have to be taken into account. This will not be done here.

9.2.3 Relative homology

In this section we use a simplicial complex K and a subcomplex L.

Definition 9.2.26 (Relative chain group). The k-chain group of K modulo L is defined as the following quotient group:

$$C_k(K, L) = C_k(K)/C_k(L) \tag{9.16}$$

Definition 9.2.27 (Relative boundary operator). The relative boundary operator $\overline{\partial}_k$ is defined as follows:

$$\overline{\partial}_k(c_k + C_k(L)) = \partial_k c_k + C_{k-1}(L) \tag{9.17}$$

where $c_k \in C_k(K)$. This operator is a group morphism, just like the ordinary boundary operator ∂_k.

Definition 9.2.28 (Relative homology groups). The relative cycle and relative boundary groups are defined analogous to their ordinary counterparts. The relative homology groups are then defined as follows:

$$H_k(K, L) = \frac{\ker(\overline{\partial}_k)}{\mathrm{im}(\overline{\partial}_{k+1})} \tag{9.18}$$

Elements $h_k \in H_k(K, L)$ can thus be written as $h_k = z_k + C_k(L)$ where z_k does not have to be a relative k-cycle but merely a chain in $C_{k-1}(L)$.

Definition 9.2.29 (Homology sequence). Using the relative homology groups we obtain following (long) exact sequence:

$$\cdots \to H_k(L) \xrightarrow{i_*} H_k(K) \xrightarrow{j_*} H_k(K, L) \xrightarrow{\partial_k} H_{k-1}(L) \to \cdots \tag{9.19}$$

where i_* and j_* are the homology morphisms induced by the inclusions $i : L \to K$ and $j : K \to (K, L)$.

Theorem 9.2.30 (Excision theorem). *Let U, V and X be triangulable spaces such that $U \subset V \subset X$. If the closure $\overline{U}$ is contained in the interior V° then:*

$$H_k(X, V) = H_k(X \backslash U, V \backslash U) \tag{9.20}$$

9.2.4 Examples

Example 9.2.31. Let X be a contractible space.

$$H_k(X) = \begin{cases} \mathbb{Z} & k = 0 \\ \{0\} & k > 0 \end{cases} \tag{9.21}$$

Example 9.2.32. Let P be a path-connected polyhedron (or path-connected triangulable space):

$$H_0(P) = \mathbb{Z} \tag{9.22}$$

Furthermore, every point $p \in P$ determines a generator $\langle p \rangle \in H_0(P)$.

Example 9.2.33. The homology groups of the n-sphere S^n are given by:

$$H_k(S^n) = \begin{cases} \mathbb{Z} & k = 0 \text{ or } k = n \\ \{0\} & \text{otherwise} \end{cases} \tag{9.23}$$

Definition 9.2.34 (Homology sphere). A n-dimensional manifold having the same homology groups as the n-sphere.

Definition 9.2.35 (Degree). From the example above we know that $H_n(S^n) = \mathbb{Z}$. Given a map $f : S^n \to S^n$ the induced map f_* on homology is an endomorphism of $\mathbb{Z}$ and hence is of the form $f(x) = dx$ where $d \in \mathbb{Z}$. This coefficient is called the degree of f.

Property 9.2.36. Two maps $f : S^n \to S^n$ have the same degree if and only if they are homotopic.

Example 9.2.37. Consider a closed, connected and orientable manifold M with $\dim(M) = n$.

$$H_n(M) = \mathbb{Z} \tag{9.24}$$

Corollary 9.2.38 (Orientation). A choice of orientation of M coincides with a choice of generator for $H_n(M)$. This generator is called the **fundamental class**. In the case M is disconnected, the fundamental class equals the direct sum of the generators of the connected components[11].

9.3 Singular homology

Definition 9.3.1 (Singular simplex). Consider the **standard** k-simplex Δ^k:

$$\Delta^k = \left\{ (x_0, ..., x_k) \in \mathbb{R}^{k+1} \middle| \sum_i x_i = 1 \text{ and } x_i \geq 0 \right\} \tag{9.25}$$

A singular k-simplex in a topological space X is defined as a continuous map $\sigma^k : \Delta^k \to X$.

Remark. The name singular comes from the fact that the maps σ^k need not be injective.

Definition 9.3.2 (Singular chain group). The singular chain group $S_k(X)$ with coefficients in a group G is defined as the set of formal linear combinations $\sum_i g_i \sigma_i^k$. The basis of this free group is in most cases infinite as there are multiple ways to map Δ^k to X.

[11] Following the idea of the additivity axiom (see 9.4.1).

Before continuing we first need to introduce an important concept in the context of simplicial complexes:

Definition 9.3.3 (Face map). The face maps are morphisms $\varepsilon_i^k : \Delta^{k-1} \to \Delta^k$ that map Δ^{k-1} onto the ith face of Δ^k. They are explicitly given by:

$$\varepsilon_i^k(s_0, ..., s_{k-1}) = (s_0, ..., s_{i-1}, 0, s_i, ..., s_{k-1}) \tag{9.26}$$

Their defining property is the following relation:

$$\varepsilon_i^k \circ \varepsilon_j^{k-1} = \varepsilon_j^k \circ \varepsilon_{i-1}^{k-1} \tag{9.27}$$

where $j \leq i$.

Remark. Some authors (for example the authors at nLab) call these maps *degeneracy maps* and call what in these notes are called *degeneracy maps* face maps.

Definition 9.3.4 (Singular boundary operator). The singular boundary operator ∂ (we use the same notation as for simplicial boundary operators) is defined by its linear action on the singular chain group $S_k(X)$. It follows that we only have to know the action on the singular simplices σ^k.

The action of the boundary operator on the singular simplex σ^k is then given by:

$$\partial_k \sigma^k = \sum_{i=0}^{k} (-1)^i \sigma^k \circ \varepsilon_i^k \tag{9.28}$$

where the ε_i^k are the face maps defined above. The singular boundary operators satisfy the same relation as the the simplicial boundary operators:

$$\partial_{k-1} \circ \partial_k = 0 \tag{9.29}$$

and hence define a chain complex.

Definition 9.3.5 (Singular homology group). The singular homology groups are defined as follows:

$$H_k(X; G) = \frac{\ker(\partial_k)}{\mathrm{im}(\partial_{k+1})} \tag{9.30}$$

Theorem 9.3.6. *For triangulable spaces the singular homology is equivalent to simplicial homology.*

Remark 9.3.7. When X is not triangulable the previous theorem is not valid. The singular approach to homology is a more general construction, but it is often more difficult to compute the homology groups (even in the case of triangulable spaces).

Property 9.3.8 (Induced morphism). Consider a continuous map $f : X \to Y$ between topological spaces. This induces a map $f_k^S : S^k(X) \to S^k(Y)$ on the chain groups as follows:

$$f_k^S \left(\sum_\sigma c_\sigma \sigma \right) = \sum_\sigma c_\sigma f \circ \sigma \tag{9.31}$$

This map takes cycle (resp. boundary) groups to (subgroups of) cycle (resp. boundary) groups and hence induces a morphism of homology groups[12]:

$$f_* : H_k(X) \to H_k(Y) : \langle h \rangle \mapsto \langle f_k^S(h) \rangle \tag{9.32}$$

[12]These induced morphisms are also called **pushforwards**.

Corollary 9.3.9. H_k is a functor Top $\to$ Ab that maps topological spaces to their homology groups and continuous maps f to their pushforward f_*.

Theorem 9.3.10 (Hurewicz). *Let X be path-connected. Let $[\cdot]$ and $\langle\cdot\rangle$ denote the equivalence classes in the homotopy and homology groups respectively. Then the map*[13] *$h : \pi(X) \to H_1(X) : [\gamma] \mapsto \langle\gamma\rangle$ defines a group morphism. Furthermore, this map induces an isomorphism $h' : \pi(X)/[\pi(X), \pi(X)] \to H_1(X)$.*

More generally for every topological space Y and every $n \in \mathbb{N}$ there exists a morphism $h_* : \pi_n(Y) \to H_n(Y)$. If Y is $(n-1)$-connected then for every $k \leq n$ this morphism is in fact an isomorphism.

Property 9.3.11. Let X be a CW-complex. There exists an isomorphism $[X, K(G,n)] \to H^n(X;G)$ between the homotopy classes of maps $X \to K(G,n)$ and the n^{th} singular cohomology of X with coefficients in G.

This morphism takes a map f to the pullback $f^*\psi$ where ψ is the inverse of the Hurewicz isomorphism $\pi_n(K) \to H_n(K;\mathbb{Z})$ and using the isomorphism $G = H^n(K;G) \cong \mathrm{Hom}(H_n(K;\mathbb{Z}), G)$.

9.4 Axiomatic approach

Definition 9.4.1 (Eilenberg-Steenrod axioms). All homology theories have a set of properties in common. By treating these properties as axioms we can construct homology theories as a sequence of functors $H_k : \mathrm{Top} \times \mathrm{Top} \to \mathrm{Ab}$. The axioms are as follows:

1. **Homotopy**: If f, g are homotopic maps then their induced homology maps are the same, i.e.
$$f \cong g \implies H_k(f) = H_k(g), \forall k \in \mathbb{N}$$
2. **Excision**[14]: If $U \subset V \subset X$ and $\overline{U} \subset V^\circ$ then $H_k(X,V) \cong H_k(X\backslash U, V\backslash U)$
3. **Dimension**: If X is a singleton then $H_k(X) = \{0\}$ for all $k \geq 1$. The group $H_0(X)$ is called the **coefficient group** and gives the coefficients used in the linear combinations of the chain group.
4. **Additivity**: If $X = \bigsqcup_i X_i$ then $H_k(X) \cong \bigoplus_i H_k(X_i)$
5. **Exactness**: Each pair (X, A), where $A \subset X$, induces a long exact sequence
$$\cdots \to H_k(A) \xrightarrow{i_*} H_k(X) \xrightarrow{j_*} H_k(X,A) \xrightarrow{\partial_k} H_{k-1}(A) \to \cdots \tag{9.33}$$
where i_* and j_* are the homology morphisms induced by the inclusions $i : A \to X$ and $j : X \to (X, A)$.

Remark 9.4.2. If the dimension axiom is removed from the set of axioms then we obtain a so-called *extraordinary homology theory.*

9.5 Equivariant cohomology

In this section we will consider topological spaces equipped with a continuous action of a topological group G. These spaces will be called topological G-spaces or just G-spaces.

[13] Every path (and hence loop) is essentially a singular 1-cycle.
[14] See also theorem 9.2.30.

Definition 9.5.1 (Equivariant cohomology). Let X be a topological G-space for which the G-action is free. The equivariant cohomology of X is defined as:

$$H^*_G(X) = H^*(X/G) \tag{9.34}$$

where the X/G is the orbit space with respect to the action of G on X.

Chapter 10

Sheaf Theory

A reference for this chapter is [5].

10.1 Presheafs

Definition 10.1.1 (Presheaf). Let (X, τ) be a topological space. A presheaf over X consists of an algebraic structure $\mathcal{F}(U)$ for every open set $U \in \tau$ and a morphism $\Phi_V^U : \mathcal{F}(U) \to \mathcal{F}(V)$ for every two open sets $U, V \in \tau$ with $V \subseteq U$ such that the following conditions are satisfied:

1. $\Phi_U^U = \mathrm{Id}$
2. If $W \subseteq V \subseteq U$ then $\Phi_W^U = \Phi_W^V \circ \Phi_V^U$.

The set $\mathcal{F}(U)$ is called the set of **sections** over U and the morphisms Φ_V^U are called the **restriction maps**.

Definition 10.1.2 (Morphism of presheaves). Let $\mathcal{F}, \mathcal{F}'$ be two presheaves over a space X. A morphism $\mathcal{F} \to \mathcal{F}'$ is a set of morphisms $\Psi_U : \mathcal{F}(U) \to \mathcal{F}'(U)$ that commute with the restriction maps Φ_V^U.

Alternative Definition 10.1.3 (Category theory). Using the language of category theory one can more easily introduce presheaves: Let $\mathbf{C}$ be a category and let X be a topological space. A $\mathbf{C}$-valued presheaf on X is a contravariant functor $\mathcal{F} : \mathrm{Open}(X) \to \mathbf{C}$. A morphism of presheafs is accordingly a natural transformation between such functors.

As such the category of presheaves on a topological space X is in fact $\mathbf{Set}^{\mathbf{Open}(X)^{op}}$, the presheaf topos on $\mathbf{Open}(X)$.

Example 10.1.4 (Constant presheaf). Let S be any set. The constant presheaf over X with target S is defined by:

$$\mathcal{F}(U) = S$$

for every open set $U \subseteq X$.

10.2 Sheafs

Definition 10.2.1 (Sheaf). Let (X, τ) be a topological space. A sheaf $\mathcal{O}_X$ over X is a presheaf $\mathcal{F}$ satisfying the following additional conditions:

1. Locality[1]: Let $\{U_i \in \tau\}$ be an open cover of $U \subseteq X$ and consider sections $s, t \in \mathcal{F}(U)$. If $\forall i : s|_{U_i} = t|_{U_i}$ then $s = t$.
2. Gluing: Let $\{U_i \in \tau\}$ be an open cover of $U \subseteq X$ and let $\{s_i \in \mathcal{F}(U_i)\}$ be a collection of sections. If $\forall i, j : s_i|_{U_i \cap U_j} = s_j|_{U_i \cap U_j}$ then there exists a section $s \in \mathcal{F}(U)$ such that $\forall i : s|_{U_i} = s_i$.

Property 10.2.2. Let X be a topological space and let $\mathcal{F}$ be a presheaf over X. $\mathcal{F}$ is a sheaf over X if an y open $U \subseteq X$ and every open cover $\{U_i\}_{i \in I}$ of U the following diagram is an equalizer diagram:

$$\mathcal{F}(U) \to \prod_{i \in I} \mathcal{F}(U_i) \rightrightarrows \prod_{i,j \in I} \mathcal{F}(U_i \cap U_j) \tag{10.1}$$

The two morphisms on the right are obtained by combining the projection morphisms of the direct product and the restriction morphisms $\Phi^{U_i}_{U_i \cap U_j}$ and $\Phi^{U_j}_{U_i \cap U_j}$.

Definition 10.2.3 (Stalk). Let $x \in X$ and consider the set of all neighbourhoods of x. This set can be turned into a directed set[2] by equipping it with the (partial) order relation

$$U \subseteq V \implies U \geq V$$

This turns the sheaf $\mathcal{F}$ over X into a directed system. The stalk over x is then defined as the following direct limit[3]:

$$\mathcal{F}_x = \varinjlim_{U \ni x} \mathcal{F}(U) \tag{10.2}$$

The equivalence class of a section $s \in \mathcal{F}(U)$ in $\mathcal{F}_x$ is called the **germ** of s at x. Two sections belong to the same germ at x if there exists a neighbourhood of x on which they coincide.

Notation 10.2.4. Similar to the notation of the restriction morphisms we denote the morphism that maps every section to its germ at x by Φ^U_x.

Property 10.2.5. Two subsheaves of a sheaf $\mathcal{G}$ over X are equal if and only if their stalks are equal as subsets of $\mathcal{G}_x$ for all points $x \in X$.

Construction 10.2.6 (Associated sheaf). Consider a presheaf $\mathcal{F}$ over a topological space X. From this presheaf one can construct a sheaf $\overline{\mathcal{F}}$, called the **sheafification** or associated sheaf of $\mathcal{F}$, in the following way:

First we define a presheaf $\mathcal{G}$ such that:[4]

$$\mathcal{G}(U) = \left\{ (s_x)_{x \in U} \in \prod_{x \in U} \mathcal{F}_x \,\middle|\, \forall x \in U : \exists \text{ open } V \ni x, t \in \mathcal{F}(V) : \forall v \in V : s_v = \Phi^V_v(t) \right\} \tag{10.3}$$

The restriction maps ρ^U_V are defined by:

$$\rho^U_V : (s_x)_{x \in U} \mapsto (s_x)_{x \in V} \tag{10.4}$$

It is easily proven that this presheaf is in fact a sheaf and hence we obtain the sheafification of $\mathcal{F}$ by setting $\overline{\mathcal{F}} = \mathcal{G}$. This construction also gives a canonical morphism $\varphi : \mathcal{F} \to \overline{\mathcal{F}}$ defined by:

$$\varphi(s) : U \to \prod_{x \in U} \mathcal{F}_x : x \mapsto s_x := \Phi^U_x(s) \tag{10.5}$$

where $s \in \mathcal{F}(U)$ and $x \in U$.

[1] This is in fact a corollary of the second axiom.
[2] See definition 2.3.8.
[3] See definition 3.34.
[4] Sections in this sheaf are said to be *continuous*. This statement can be made formal using the concept of an étalé space. (See construction 10.2.10.)

Remark. We see that the sections in $\overline{\mathcal{F}}$ arise from (locally) resticting sections in $\mathcal{F}$ to the stalks $\mathcal{F}_x$.

Universal property 10.2.7. Let $\mathcal{F}$ be a presheaf over X with associated sheaf $\overline{\mathcal{F}}$. For any sheaf $\mathcal{G}$ over X and any morphism $\rho : \mathcal{F} \to \mathcal{G}$ there exists a unique morphism $\sigma : \overline{\mathcal{F}} \to \mathcal{G}$ such that $\rho = \sigma \circ \varphi$ where φ is the canonical morphism $\mathcal{F} \to \overline{\mathcal{F}}$.

Property 10.2.8. Let $\mathcal{F}$ be a presheaf over X with associated sheaf $\overline{\mathcal{F}}$. The morphism $\varphi : \mathcal{F} \to \overline{\mathcal{F}}$ induces an isomorphism $\varphi_x : \mathcal{F}_x \to \overline{\mathcal{F}}_x$ for all $x \in X$.

Property 10.2.9. Let $\mathcal{F}$ be a sheaf over X with associated sheaf $\overline{\mathcal{F}}$. The morphism $\varphi : \mathcal{F} \to \overline{\mathcal{F}}$ is an isomorphism.

There exists another, more topological, construction of the associated sheaf:

Construction 10.2.10 (Étalé spaces). Let $\mathcal{F}$ be a presheaf over X. Consider the disjoint union

$$\overline{\mathcal{F}} = \bigsqcup_{x \in X} \mathcal{F}_x \tag{10.6}$$

Define for every local section $s \in \mathcal{F}(U)$ a function $\overline{s} : U \to \overline{\mathcal{F}} : x \mapsto s_x \in \mathcal{F}_x$ The union $\overline{\mathcal{F}}$ can be turned into an étalé space[5] over X by equipping it with the topology with basis

$$\{\overline{s}(U) \mid U \text{ open in } X, s \in \mathcal{F}(U)\} \tag{10.7}$$

The projection map π is given by $\pi : s_x \in \mathcal{F}_x \mapsto x$. The sheafification $\mathcal{F}^*$ is then given by the sheaf of sections of $\overline{\mathcal{F}}$.

Example 10.2.11 (Constant sheaf). Consider the constant presheaf over X with target S (see example 10.1.4). The constant sheaf, denoted by S_X, is defined as the associated sheaf of this presheaf. The stalks over every point $x \in X$ can be identified with S. The continuous sections $S_X(U)$ are the locally constant functions $f : U \to S$.

Notation 10.2.12 (Category of sheaves). Similar to the case of presheaves one can define a morphism of sheaves as a morphism commuting with the restriction maps.[6] The sheafs and sheaf morphisms over a space X form a full subcategory of the category of presheafs, denoted by $\mathrm{Sh}(X)$.

Property 10.2.13. The category of sheafs $\mathrm{Sh}(X)$ is in fact an elementary topos, called the sheaf topos over X.

Definition 10.2.14 (Global sections functor). Let X be a topological space. The global sections functor $\Gamma(X, -)$ is defined as the functor $\Gamma(X, -) : \mathrm{Sh}(X) \to \mathrm{Set} : \mathcal{F} \to \mathcal{F}(X)$.

Property 10.2.15. The global sections functor is only left exact.

10.3 Resolutions and cohomology

In this section we will only use Abelian sheafs, i.e. sheafs with values in Ab, unless stated otherwise.

Property 10.3.1. Every Abelian sheaf admits an injective resolution.

[5] See definition 7.3.19.

[6] When treating a (pre)sheaf as a functor, the (pre)sheaf morphism is a natural transformation.

Definition 10.3.2 (Sheaf cohomology group). Let $\mathcal{F}$ be a sheaf over X. Given an injective resolution[7] I of $\mathcal{F}$ one defines the sheaf cohomology groups of $\mathcal{F}$ over X as the cohomology groups of the complex

$$\cdots \longrightarrow \Gamma(X, I^i) \longrightarrow \Gamma(X, I^{i+1}) \longrightarrow \cdots \tag{10.8}$$

The cohomology group $H^0(X, \mathcal{F})$ is equal to $\Gamma(X, \mathcal{F})$.

Definition 10.3.3 (Acyclic sheaf). A sheaf is called acyclic if its higher cohomology groups vanish.

Theorem 10.3.4 (de Rham & Weil). *Let $\mathcal{F}$ be a sheaf. There exists an isomorphism between the sheaf cohomology groups defined above and the ones obtained by using an acyclic resolution of $\mathcal{F}$.*

Definition 10.3.5 (Image and kernel). Given a morphism of sheafs $\phi : \mathcal{F} \to \mathcal{G}$ over a space X one can define the kernel/image presheafs which assign to every open subset $U \subseteq X$ the image/kernel of ϕ_U. The kernel presheaf is already a sheaf and will be denoted by $\ker(\phi)$. The sheafification of the image presheaf will be denoted by $\mathrm{im}(\phi)$.

Definition 10.3.6 (Cohomology sheafs). Let $\mathcal{F}^\bullet$ be a complex of sheafs over X. The cohomology sheafs $\underline{H}^i(X, \mathcal{F}^\bullet)$ assign to every open subset $U \subseteq X$ the quotient group $\ker(d^i_U)/\mathrm{im}(d^{i-1}_U)$.

10.4 Ringed spaces

Definition 10.4.1 (Ringed space). A ringed space is a topological space X equipped with a sheaf of rings $\mathcal{O}_X$.

Definition 10.4.2 (Locally ringed space). A ringed space $(X, \mathcal{O}_X)$ is said to be locally ringed if the stalk over every point $x \in X$ is a local ring[8].

[7] In fact this construction is independent of the chosen injective resolution.
[8] See definition 3.2.18.

Chapter 11

Algebraic Geometry

11.1 Polynomials and Galois theory

11.1.1 Polynomials

Definition 11.1.1 (Polynomial ring). Let R be a (commutative) ring. The polynomial ring on the indeterminates $X = \{x_i\}_{i \in I}$ is the free commutative R-algebra on X.

Definition 11.1.2 (Degree). The exponent of the highest order power in x. It is often denoted by $\deg(f)$.

Definition 11.1.3 (Monic polynomial). A polynomial for which the highest order term has coefficient 1.

Theorem 11.1.4 (Fundamental theorem of algebra). *Let $f(x) \in K[x]$ with $\deg(f) \geq 1$. Then $f(x)$ has at least 1 root in $\mathbb{C}$.*

Corollary 11.1.5. If $f(x) \in \mathbb{C}[x]$ is a monic polynomial with $\deg(f) \geq 1$, we can write:

$$f(x) = \prod_{i=1}^{k} (x - a_i)^{n_i}$$

Where $a_1, ..., a_k \in \mathbb{C}$ and $n_1, ..., n_k \in \mathbb{N}$.

Definition 11.1.6 (Transcendental element). Consider a base field k and a field extension L/k. An element $x \in L$ for which there exists no non-trivial polynomial p over k such that $p(x) = 0$ is said to be transcendental. Otherwise it is said to be **algebraic**.

Definition 11.1.7 (Algebraic independence). Consider a base field k and a field extension L/k. A subset $S \subseteq L$ is said to be algebraically independent over k if the elements of S do not satisfy any non-trivial polynomial over k.

11.2 Schemes

11.2.1 Spectrum of a ring

Definition 11.2.1 (Spectrum). Let R be a commutative ring. The spectrum $\mathrm{Spec}(R)$ is defined as the set of prime ideals of R. This set can be turned into a topological space by equipping it with the **Zariski topology**: Let V_I be the set of prime ideals containing the ideal I. The collection of closed sets, inducing the Zariski topology, is given by $\{V_I\}_{I \text{ ideal of } R}$.

Remark 11.2.2. A basis for the above topology is given by the sets $D_f = \{I_p \not\ni f : f \in R, I_p \text{ is a prime ideal}\}$.

Property 11.2.3. $\mathrm{Spec}(R)$ is a compact T_0 space.

Definition 11.2.4 (Structure sheaf). Given a spectrum $X = \mathrm{Spec}(R)$, equipped with its Zariski topology, we can define a sheaf[1] $\mathcal{O}_X$ by setting $\forall f \in R : \Gamma(D_f, \mathcal{O}_X) = R_f^*$, where R_f^* is the localization of R with respect to the monoid of powers of f.

Property 11.2.5. The spectrum $\mathrm{Spec}(R)$ together with its structure sheaf forms a ringed space.

11.2.2 Affine schemes

Definition 11.2.6 (Affine scheme). A ringed space, isomorphic to the spectrum $\mathrm{Spec}(R)$ for some commutative ring R, is called an affine scheme.

11.2.3 Zariski tangent space

Definition 11.2.7 (Zariski tangent space). Consider a variety X with structure sheaf $\mathcal{O}_X$. At every point $x \in X$ the ring $\mathcal{O}_{X,x}$ is a local ring and hence we obtain a maximal ideal $\mathfrak{m}_x$. The quotient $\mathfrak{m}_x/\mathfrak{m}_x^2$ is a vector space over the residue field $\mathcal{O}_{X,x}/\mathfrak{m}_x$. It is called the Zariski cotangent space at $x \in X$. Its algebraic dual is called the Zariski tangent space at $x \in X$.

[1] In fact this is merely a *B-sheaf* as it is only defined on the basis of the topology. However, every B-sheaf can be extended to a sheaf by taking the appropriate limits.

Part IV

Calculus

Chapter 12

Calculus

12.1 Sequences

Definition 12.1.1 (Limit superior). Let $(x_i)_{i\in\mathbb{N}}$ be a sequence of real numbers. The limit superior is defined as follows:

$$\limsup_{i\to+\infty} x_i = \inf_{i\geq 1}\left\{\sup_{k\geq i} x_k\right\} \tag{12.1}$$

Definition 12.1.2 (Limit inferior). Let $(x_i)_{i\in\mathbb{N}}$ be a sequence of real numbers. The limit superior is defined as follows:

$$\liminf_{i\to+\infty} x_i = \sup_{i\geq 1}\left\{\inf_{k\geq i} x_k\right\} \tag{12.2}$$

Theorem 12.1.3. *A sequence $(x_i)_{i\in\mathbb{N}}$ converges pointwise if and only if $\limsup_{i\to+\infty} x_i = \liminf_{i\to+\infty} x_i$.*

12.2 Continuity

Definition 12.2.1 (Lipschitz continuity). A function $f : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous if there exists a constant $C > 0$ such that

$$|f(x) - f(x')| \leq C|x - x'| \tag{12.3}$$

for all $x, x' \in \mathbb{R}$.

Theorem 12.2.2 (Darboux's theorem). *Let f be a differentiable function on a closed interval I. Then f' has the intermediate value property*[1]*.*

Remark 12.2.3 (Darboux function). Functions that have the intermediate value property are called Darboux functions.

Corollary 12.2.4 (Bolzano). If $f(a) < 0$ and $f(b) > 0$ (or vice versa) then there exists at least one point x_0 where $f(x_0) = 0$.

Corollary 12.2.5. The image of a compact set is also a compact set.

Theorem 12.2.6 (Weierstrass' extreme value theorem). *Let $I = [a, b] \subset \mathbb{R}$ be a compact interval. Let f be a continuous function defined on I. Then f attains a minimum and maximum at least once on I.*

[1] This means that the function satisfies the conclusion of the intermediate value theorem 7.4.3.

12.3 Convergence

Definition 12.3.1 (Pointwise convergence). Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions. The sequence is said to converge pointwise to a limit function $f(x)$ if

$$\forall x \in \mathrm{dom}(f_n) : \lim_{n\to+\infty} f_n(x) = f(x) \tag{12.4}$$

Definition 12.3.2 (Uniform convergence). Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions. The sequence is said to converge uniformly to a limit function $f(x)$ if

$$\sup_{x\in\mathrm{dom}(f_n)} \left\{ |\lim_{n\to+\infty} f_n(x) - f(x)| \right\} = 0 \tag{12.5}$$

12.4 Derivative

12.4.1 Single variable

Formula 12.4.1 (Derivative).

$$\boxed{f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}} \tag{12.6}$$

Theorem 12.4.2 (Mean value theorem). *Let f be continuous on the closed interval $[a, b]$ and differentiable on the open interval $]a, b[$. Then there exists a point $c \in]a, b[$ such that:*

$$f'(c) = \frac{f(b) - f(a)}{b - a} \tag{12.7}$$

Definition 12.4.3 (Differentiablity class). Let I be a set. Let f be a function defined on I. If f is n times continuously differentiable on I (i.e. $f^{(i)}$ exists and is continuous for $i = 1, \dots, n$) then f is said to be of class $\mathbf{C}^n(I)$.

Definition 12.4.4 (Smooth function). A function f is said to be smooth if it is of class $\mathbf{C}^\infty$.

Theorem 12.4.5 (Boman). *Consider a function $f : \mathbb{R}^d \to \mathbb{R}$. If for every smooth function $g : \mathbb{R} \to \mathbb{R}^d$ the composition $f \circ g$ is also smooth then the function f is smooth too.*[2]

Definition 12.4.6 (Analytic function). A function f is said to be analytic if it is smooth and if its Taylor series expansion around any point x_0 converges to f in some neighbourhood of x_0. The set of analytic functions defined on V is denoted by $\mathrm{C}^\omega(V)$.

Theorem 12.4.7 (Schwarz' theorem[3]). *Let $f \in C^2(\mathbb{R}^n, \mathbb{R})$. The mixed partial derivatives of f coincide:*

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right) = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right) \tag{12.8}$$

for all indices $i, j \leq n$.

[2] In the context of Part VI: *Differential Geometry*, this theorem can be generalized to smooth manifolds by replacing $\mathbb{R}^d$ with a smooth manifold M.

[3] Also called **Clairaut's theorem**.

Method 12.4.8 (Derivative of $f(x)^{g(x)}$). Let us consider a function of the form $u(x) = f(x)^{g(x)}$. To find the derivative of this function we can use the derivative of the natural logarithm:

$$\ln[u(x)] = g(x)\ln[f(x)]$$

Taking the derivative gives us:

$$\frac{d\ln[u(x)]}{dx} = \frac{d}{dx}\Big(g(x)\ln[f(x)]\Big) = \frac{dg(x)}{dx}\ln[f(x)] + \frac{g(x)}{f(x)}\frac{df(x)}{dx}$$

The derivative of a logarithm also satisfies:

$$\frac{d\ln[u(x)]}{dx} = \frac{1}{u(x)}\frac{du}{dx}$$

Combining these two equations gives:

$$\boxed{\frac{d}{dx}\Big[f(x)^{g(x)}\Big] = f(x)^{g(x)}\left[\frac{dg}{dx}(x)\ln[f(x)] + \frac{g(x)}{f(x)}\frac{df}{dx}(x)\right]} \tag{12.9}$$

Theorem 12.4.9 (Euler). *Let f be a homogeneous function, i.e.*

$$f(ax_1, ..., ax_n) = a^n f(x_1, ..., x_n)$$

This function satisfies following equality:

$$\sum_k x_k \frac{\partial f}{\partial x_k} = nf(x_1, ..., x_n) \tag{12.10}$$

12.5 Riemann integral

Definition 12.5.1 (Improper Riemann integral).

$$\boxed{\int_{-\infty}^{+\infty} f(x)dx = \lim_{\substack{a\to-\infty \\ b\to+\infty}} \int_a^b f(x)dx} \tag{12.11}$$

12.6 Fundamental theorems

Theorem 12.6.1 (First fundamental theorem of calculus). *Let f be a continuous function defined on the open interval I. Let $c \in I$. The following theorem establishes a link between integration and differentiation:*

$$\boxed{\exists F(x) = \int_c^x f(x')dx' : F'(x) = f(x)} \tag{12.12}$$

Furthermore this function $F(x)$ is uniformly continuous on I.

Remark 12.6.2. The function $F(x)$ in the previous theorem is called a **primitive function** of $f(x)$. Remark that $F(x)$ is just 'a' primitive function as adding a constant to $F(x)$ does not change anything because the derivative of a constant is zero.

Theorem 12.6.3 (Second fundamental theorem of calculus). *Let $f(x)$ be a function defined on the interval $[a, b]$. Furthermore, let $f(x) \in C^1[a, b]$. We then find the following important theorem:*

$$\boxed{\int_a^b f'(x)dx = f(b) - f(a)} \tag{12.13}$$

Theorem 12.6.4 (Differentiation under the integral sign[4]).

$$\boxed{\frac{d}{dx}\int_{a(x)}^{b(x)} f(x,y)dy = f(x,b(x)) \cdot b'(x) - f(x,a(x)) \cdot a'(x) + \int_{a(x)}^{b(x)} \frac{\partial f(x,y)}{\partial x}dy} \tag{12.14}$$

12.7 Taylor expansion

Formula 12.7.1 (Exponential function).

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \tag{12.15}$$

12.8 Euler integrals

12.8.1 Euler integral of the first kind

Formula 12.8.1 (Beta function).

$$\boxed{B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}dt} \tag{12.16}$$

12.8.2 Euler integral of the second kind

Formula 12.8.2 (Gamma function).

$$\boxed{\Gamma(x) = \int_0^{+\infty} t^{x-1}e^{-t}dt} \tag{12.17}$$

Formula 12.8.3. $\mathrm{B(x,\,y)} = \dfrac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$

Formula 12.8.4 (Recursion formula). Let $n!$ denote the factorial for integer numbers.

$$\Gamma(n) = (n-1)! \tag{12.18}$$

Formula 12.8.5 (Stirling).

$$\ln(n!) \approx n\ln n - n \tag{12.19}$$

[4]This is a more general version of the so called 'Leibnitz integral rule'.

12.9 Gaussian integrals

Formula 12.9.1 (n-dimensional Gaussian integral). A general Gaussian integral is of the form

$$I(A,\vec{\boldsymbol{b}}) = \int_{-\infty}^{+\infty} d^n x \exp\left(-\frac{1}{2}\vec{\boldsymbol{x}}\cdot A\vec{\boldsymbol{x}} + \vec{\boldsymbol{b}}\cdot\vec{\boldsymbol{x}}\right) \tag{12.20}$$

where A is a real symmetric matrix. By performing the transformation $\vec{\boldsymbol{x}} \to A^{-1}\vec{\boldsymbol{b}} - \vec{\boldsymbol{x}}$ and diagonalising A one can obtain the following result:

$$I(A,\vec{\boldsymbol{b}}) = (2\pi)^{n/2}\det(A)^{-1/2}\exp\left(\frac{1}{2}\vec{\boldsymbol{b}}\cdot A^{-1}\vec{\boldsymbol{b}}\right) \tag{12.21}$$

Corollary 12.9.2. A functional generalisation is given by:

$$\begin{aligned} I(iA,iJ) &= \int [d\varphi]\exp\left(-i\int d^n x d^n y\ \varphi(x)A(x,y)\varphi(y) + i\int d^n x\ \varphi(x)J(x)\right) \\ &= C\det(A)^{-1/2}\exp\left(\frac{i}{2}\int d^n x d^n y\ J(x)A^{-1}(x,y)J(y)\right) \end{aligned} \tag{12.22}$$

where we used the analytic continuation $I(iA,iJ)$ of equation 12.21. One should pay attention to the normalisation factor C which is infinite in general.

12.10 Series

12.10.1 Convergence tests

Theorem 12.10.1. *A series $\sum_{i=1}^{+\infty} a_i$ can only converge if $\lim_{i\to+\infty} a_i = 0$.*

Property 12.10.2 (Absolute/conditional convergence). If $S' = \sum_{i=1}^{+\infty}|a_i|$ converges then so does the series $S = \sum_{i=1}^{+\infty} a_i$ and S is said to be absolutely convergent. If S converges but S' does not, then S is said to be conditionally convergent.

Definition 12.10.3 (Majorizing series). Let $S_a = \sum_{i=1}^{+\infty} a_i$ and $S_b = \sum_{i=1}^{+\infty} b_i$ be two series. The series S_a is said to majorize S_b if for every $k > 0$ the partial sum $S_{a,k} \geq S_{b,k}$.

Method 12.10.4 (Comparison test). Let S_a, S_b be two series such that S_a majorizes S_b. We have the following cases:

- If S_b diverges, then S_a diverges.
- If S_a converges, then S_b converges.
- If S_b converges, nothing can be said about S_a.
- If S_a diverges, nothing can be said about S_b.

Method 12.10.5 (MacLaurin-Cauchy integral test). Let f be a non-negative continuous monotone decreasing function on the interval $[n,+\infty[$. If $\int_n^{+\infty} f(x)dx$ is convergent then so is $\sum_{k=n}^{+\infty} f(k)$. On the other hand, if the integral is divergent, so is the series.

Remark 12.10.6. The function does not have to be non-negative and decreasing on the complete interval. As long as it does on the interval $[N,+\infty[$ for some $N \geq n$. This can be seen by writing $\sum_{k=n}^{+\infty} f(k) = \sum_{k=n}^{N} f(k) + \sum_{k=N}^{+\infty} f(k)$ and noting that the first term is always finite (the same argument applies for the integral).

Property 12.10.7. If the integral in the previous theorem converges, then the series has following lower and upper bounds:

$$\int_n^{+\infty} f(x)dx \leq \sum_{i=n}^{+\infty} a_i \leq f(n) + \int_n^{+\infty} f(x)dx \tag{12.23}$$

Method 12.10.8 (d'Alembert's ratio test).

$$R = \lim_{n\to+\infty} \left|\frac{a_{n+1}}{a_n}\right| \tag{12.24}$$

Following cases arise:

- $R < 1$: the series converges absolutely
- $R > 1$: the series does not converge
- $R = 1$: the test is inconclusive

Method 12.10.9 (Cauchy's root test).

$$R = \limsup_{n\to+\infty} \sqrt[n]{|a_n|} \tag{12.25}$$

We have the following cases:

- $R < 1$: the series converges absolutely
- $R > 1$: the series does not converge
- $R = 1$ and the limit approaches strictly from above: the series diverges
- $R = 1$: the test is inconclusive

Definition 12.10.10 (Radius of convergences). The number $\frac{1}{R}$ is called the radius of convergence.

Remark 12.10.11. The root test is stronger than the ratio test. Whenever the ratio test determines the convergence/divergence of a series, the radius of convergence of both tests will coincide.

Method 12.10.12 (Gauss's test). If $u_n > 0$ for all n then we can write the ratio of successive terms as follows:

$$\left|\frac{u_n}{u_{n+1}}\right| = 1 + \frac{h}{n} + \frac{B(n)}{n^k} \tag{12.26}$$

where $k > 1$ and $B(n)$ is a bounded function when $n \to \infty$. The series converges if $h > 1$ and diverges otherwise.

12.10.2 Asymptotic expansions

Definition 12.10.13 (Asysmptotic expansion). Let $f(x)$ be a continuous function. A series expansion of order N is called an asymptotic expansion of $f(x)$ if it satisfies:

$$f(x) - \sum_{n=0}^{N} = O(x^{N+1}) \tag{12.27}$$

Method 12.10.14 (Borel transform†). Define the function $F(x) = \sum_{n=0}^{+\infty} \frac{a_n}{n!} x^n$. If the integral

$$\int_0^{+\infty} e^{-t} F(xt) dt < +\infty \tag{12.28}$$

for all $x \in \mathbb{R}$ then $F(x)$ is called the Borel transform of $f(x)$. Furthermore the integral will give a convergent expression for $f(x)$.

Theorem 12.10.15 (Watson). *The uniqueness of the function $F(x)$ is guaranteed if the function $f(x)$ is holomorphic on the domain $\{z \in \mathbb{C} : |\arg(z)| < \frac{\pi}{2} + \varepsilon\}$.*

Chapter 13

Complex Analysis

13.1 Complex algebra

The set of complex numbers $\mathbb{C}$ forms a 2-dimensional vector space over the field of real numbers. Furthermore the operations of complex addition and complex multiplication also turn the complex numbers into a field.

Definition 13.1.1 (Complex conjugate). The complex conjugate $\overline{z} : a + bi \mapsto a - bi$ is an involution, i.e. $\overline{\overline{z}} = z$. It is sometimes denoted by z^* instead of $\overline{z}$.

Formula 13.1.2 (Real/imaginary part). A complex number z can also be written as $\mathrm{Re}(z) + i\mathrm{Im}(z)$ where

$$\mathrm{Re}(z) = \frac{z + \overline{z}}{2} \tag{13.1}$$

$$\mathrm{Im}(z) = \frac{z - \overline{z}}{2i} \tag{13.2}$$

Definition 13.1.3 (Argument). Let z be a complex number parametrized as $z = re^{i\theta}$. The number θ is called the argument of z and it is denoted by $\arg(z)$.

Definition 13.1.4 (Riemann sphere). Consider the one-point compactification[1] $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$. This set is called the Riemann sphere or **extended complex plane**. The standard operations on $\mathbb{C}$ can be generalized to $\overline{\mathbb{C}}$ in the following way:

$$\begin{aligned} z + \infty &= \infty \\ z * \infty &= \infty \\ \frac{z}{\infty} &= 0 \end{aligned} \tag{13.3}$$

for all non-zero $z \neq \infty$. As there exists no multiplicative inverse for ∞ the Riemann sphere does not form a field.

13.2 Complex maps

13.2.1 Holomorphic maps

Definition 13.2.1 (Holomorphic). A function f is holomorphic on an open set U if it is complex differentiable at every point $z_0 \in U$.

[1] See definition 7.5.26.

Definition 13.2.2 (Biholomorphic). A complex function f is said to be biholomorphic if both f and f^{-1} are holomorphic.

Definition 13.2.3 (Entire). A function holomorphic at every point $z \in \mathbb{C}$.

Property 13.2.4 (Cauchy-Riemann conditions). A holomorphic function $f(z)$ satisfies the following conditions:

$$\boxed{\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \qquad \text{and} \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}} \tag{13.4}$$

or equivalently:

$$\boxed{\frac{\partial f}{\partial \overline{z}} = 0} \tag{13.5}$$

Theorem 13.2.5 (Looman-Menchoff[2]). *Let $f(z)$ be a continuous complex-valued function defined on a subset $U \in \mathbb{C}$. If the partial derivatives of the real and imaginary part exist and if f satisfies the Cauchy-Riemann conditions then f is holomorphic on U.*

Property 13.2.6. Functions u, v satisfying the CR-conditions are harmonic functions, i.e. they satisfy Laplace's equation.

Property 13.2.7. Functions u, v satisfying the CR-conditions have orthogonal level curves 2.14.

13.2.2 Conformal maps

Property 13.2.8. Every analytic map is also conformal.

13.3 Complex integrals

In this and further sections, all contours have been chosen to be evaluated counterclockwise (by convention). To obtain results concerning clockwise evaluation, most of the time adding a minus sign is sufficient.

Definition 13.3.1 (Contour). A contour is a curve $z(t)$ that can be parametrized by

$$\left.\begin{array}{l} x = x(t) \\ y = y(t) \end{array}\right\} \rightarrow z(t) = x(t) + iy(t) \tag{13.6}$$

Formula 13.3.2 (Complex contour integral). The complex contour integral of a function $f(z) = u(z) + iv(z)$ is defined as the following line integral:

$$\int_{z_1}^{z_2} f(z)dz = \int_{(x_1,y_1)}^{(x_2,y_2)} [u(x,y) + iv(x,y)](dx + idy) \tag{13.7}$$

Theorem 13.3.3 (Cauchy's Integral Theorem[3]). *Let Ω be a simply-connected subset of $\mathbb{C}$ and let f be a holomorphic function on Ω. Then for every closed rectifiable[4] contour C in Ω:*

$$\boxed{\oint_C f(z)dz = 0} \tag{13.8}$$

[2]This is the strongest (most general) theorem on the holomorphy of continuous functions as it generalizes the original results by Riemann and Cauchy-Goursat.

[3]Also called the *Cauchy-Goursat theorem.*

[4]A contour with finite length.

Corollary 13.3.4. The contour integral of a holomorphic function depends only on the limits of integration and not on the contour connecting them.

Formula 13.3.5 (Cauchy's Integral Formula). Let Ω be a connected subset of $\mathbb{C}$ and let f be a holomorphic function on Ω. Let C be a contour in Ω. For every point z_0 inside C we find:

$$\boxed{f(z_0) = \frac{1}{2\pi i} \oint_C \frac{f(z)}{z - z_0} dz} \tag{13.9}$$

Corollary 13.3.6 (Analytic function). Let Ω be a connected subset of $\mathbb{C}$ and C a closed contour in Ω. If f is holomorphic on Ω then f is analytic[5] on Ω and:

$$\boxed{f^{(n)}(z_0) = \frac{1}{2\pi i} \oint_C f(z) \frac{n!}{(z - z_0)^{n+1}} dz} \tag{13.10}$$

Furthermore, the derivatives are also holomorphic on Ω.

Theorem 13.3.7 (Morera). *If f is continuous on a connected open set Ω and $\oint_C f(z)dz = 0$ for every closed contour C in Ω, then f is holomorphic on Ω.*

Theorem 13.3.8 (Liouville). *Every bounded entire function is constant.*

Theorem 13.3.9 (Sokhotski-Plemelj[6]). *Let $f(x)$ be a continuous complex-valued function defined on the real line and let $a < 0 < b$.*

$$\lim_{\varepsilon \to 0^+} \int_a^b \frac{f(x)}{x \pm i\varepsilon} dx = \mp i\pi f(0) + \mathcal{P} \int_a^b \frac{f(x)}{x} dx \tag{13.11}$$

where $\mathcal{P}$ denotes the Cauchy principal value.

13.4 Laurent series

Definition 13.4.1 (Laurent series). If f is a function, analytic on an annulus A, then f can be expanded as the following series:

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n \qquad \text{with} \qquad a_n = \frac{1}{2\pi i} \oint \frac{f(z')}{(z' - z_0)^{n+1}} dz' \tag{13.12}$$

Remark 13.4.2. The Laurent series of an analytic function f converges uniformly to f in the ring shaped region (**annulus**) $R_1 < |z - z_0| < R_2$, with R_1 and R_2 the distances from z_0 to the two closest poles.

Definition 13.4.3 (Principal part). The principal part of a Laurent series at the point z_0 is defined as the sum:

$$\sum_{n=-\infty}^{-1} a_n (z - z_0)^n \tag{13.13}$$

[5] See definition 12.4.6.

[6] See for example [6], page 104.

13.5 Singularities

13.5.1 Poles

Definition 13.5.1 (Pole). A function $f(z)$ has a pole of order $m > 0$ at a point z_0 if its Laurent series at z_0 satisfies $\forall n < -m : a_n = 0$ and $a_{-m} \neq 0$.

Definition 13.5.2 (Meromorphic). A function f is called meromorphic when it is analytic on the whole complex plane with exception of isolated poles and removable singularities.

Definition 13.5.3 (Essential singularity). A function $f(z)$ has an essential singularity at a point z_0 if its Laurent series at z_0 satisfies $\forall n \in \mathbb{N} : a_{-n} \neq 0$, i.e. its Laurent series has infinitely many negative degree terms.

Theorem 13.5.4 (Picard's great theorem). *Let $f(z)$ be an analytic function with an essential singularity at z_0. On every punctured neighbourhood of z_0, $f(z)$ takes on all possible complex values, with at most a single exception, infinitely many times.*

Method 13.5.5 (Frobenius transformation). To study the behaviour of a function $f(z)$ at $z \to \infty$, one should apply the Frobenius transformation $h = 1/z$ and study the limit $\lim_{h\to 0} f(h)$.

13.5.2 Branch cuts

Formula 13.5.6 (Roots). Let $z \in \mathbb{C}$. The n^{th} roots[7] of $z = re^{i\theta}$ are given by:

$$z^{1/n} = \sqrt[n]{r} \exp\left(i\frac{\theta + 2\pi k}{n}\right) \tag{13.14}$$

where $k \in \{0, 1, ..., n\}$.

Formula 13.5.7 (Complex logarithm). We parametrize z as $z = re^{i\theta}$.

$$\mathrm{LN}(z) = \ln(r) + i(\theta + 2\pi k) \tag{13.15}$$

Definition 13.5.8 (Branch). From these two formulas it is clear that the complex roots and logarithms are multi-valued functions. To get an unambiguous image it is necessary to fix a value of the parameter k. By doing so there will arise curves in the complex plane where the function is discontinuous. These are the branch cuts. A **branch** is then defined as a particular choice of the parameter k. For the logarithm the choice for $\arg(\mathrm{LN}) \in]\alpha, \alpha + 2\pi]$ is often denoted by LN_α or $\log_\alpha$.

Definition 13.5.9 (Branch point). Let $f(z)$ be a complex valued function. A point z_0 such that there exists no neighbourhood $|z - z_0| < \varepsilon$ where $f(z)$ is single valued is called a branch point.

Definition 13.5.10 (Branch cut). A line connecting exactly two branch points is called a branch cut. One of the branch points can be at infinity. In case of multiple branch cuts, they do not cross.

Example 13.5.11. Consider the complex function

$$f(z) = \frac{1}{\sqrt{(z - z_1)...(z - z_n)}}$$

[7] Also see the *fundamental theorem of algebra* 11.1.4.

This function has singularities at $z_1, ..., z_n$. If n is even, this function will have n (finite) branch points. This implies that the points can be grouped in pairs connected by non-intersecting branch cuts. If n is odd, this function will have n (finite) branch points and one branch point at infinity. The finite branch points will be grouped in pairs connected by non-intersecting branch cuts and the remaining branch point will be joined to infinity by a branch cut which does not intersect the others.(See [7] for the proof.)

Definition 13.5.12 (Principal value). The principal value of a multi-valued complex function is defined as the choice of branch such that $\arg(f) \in]-\pi, \pi]$.

13.5.3 Residue theorem

Definition 13.5.13 (Residue). By applying formula 13.7 to a polynomial function we find:

$$\int_C (z - z_0)^n dz = 2\pi i \delta_{n,-1} \tag{13.16}$$

where C is a circular contour around the pole $z = z_0$. This means that integrating a Laurent series around a pole isolates the coefficient a_{-1}. This coefficient is therefore called the residue of the function at the given pole.

Notation 13.5.14. The residue of a complex function $f(z)$ at a pole z_0 is denoted by:

$$\text{Res}[f(z)]_{z=z_0}$$

Formula 13.5.15. For a pole of order m, the residue is calculated as follows:

$$\text{Res}\left[f(z)\right]_{z=z_j} = a_{-1} = \lim_{z \to z_0} \frac{1}{(m-1)!} \left(\frac{\partial}{\partial z}\right)^{m-1} (f(z)(z - z_0)) \tag{13.17}$$

For essential singularities the residue can be found by writing out the Laurent series explicitly.

Theorem 13.5.16 (Residue theorem). *If $f(z)$ is a meromorphic function in Ω and if C is a closed contour in Ω which contains the poles z_j of $f(z)$, then:*

$$\boxed{\oint_C f(z)dz = 2\pi i \sum_j \text{Res}\left[f(z)\right]_{z=z_j}} \tag{13.18}$$

Remark 13.5.17. For poles on the contour C, only half of the residue contributes to the integral.

Formula 13.5.18 (Argument principle). Let $f(z)$ be a meromorphic function. Let Z_f, P_f be respectively the number of zeroes and poles of $f(z)$ inside the contour C. From the residue theorem we can derive the following formula:

$$\frac{1}{2\pi i} \oint_C \frac{f(z)}{f'(z)} dz = Z_f - P_f \tag{13.19}$$

Formula 13.5.19 (Winding number). Let $f(z)$ be a meromorphic function and let C be a simple closed contour. For all $a \notin f(C)$ the winding number, also called the **index**, of a with respect to the function f is defined as:

$$\text{Ind}_f(a) = \frac{1}{2\pi i} \oint_C \frac{f'(z)}{f(z) - a} dz \tag{13.20}$$

This number will always be an integer.

13.6 Limit theorems

Theorem 13.6.1 (Small limit theorem). *Let f be a function that is holomorphic almost every where on $\mathbb{C}$. Let the contour C be a circular segment with radius ε and central angle α. If z is parametrized as $z = \varepsilon e^{i\theta}$ then*

$$\int_C f(z)dz = i\alpha A$$

with

$$A = \lim_{\varepsilon \to 0} f(z)$$

Theorem 13.6.2 (Great limit theorem). *Let f be a function that is holomorphic almost every where on $\mathbb{C}$. Let the contour C be a circular segment with radius R and central angle α. If z is parametrized as $z = Re^{i\theta}$ then*

$$\int_C f(z)dz = i\alpha B$$

with

$$B = \lim_{R \to +\infty} f(z)$$

Theorem 13.6.3 (Jordan's lemma). *Let g be a continuous function with $g(z) = f(z)e^{bz}$. Let the contour C be a semicircle lying in the half-plane bounded by the real axis and oriented away of the point $\overline{b}i$. If z is parametrized as $z = Re^{i\theta}$ and*

$$\lim_{R \to \infty} f(z) = 0$$

then

$$\int_C g(z)dz = 0$$

13.7 Analytic continuation

Theorem 13.7.1 (Schwarz' reflection principle). *Let $f(z)$ be analytic on the upper half plane. If $f(z)$ is real when z is real then*

$$f(\overline{z}) = \overline{f(z)} \tag{13.21}$$

Chapter 14

Measure Theory and Lebesgue Integration

14.1 Measure

14.1.1 General definitions

Definition 14.1.1 (Measure). Let X be a set. Let Σ be a σ-algebra over X. A function $\mu : \Sigma \to \overline{\mathbb{R}}$ is called a measure if it satisfies the following conditions:

1. Non-negativity: $\forall E \in \Sigma : \mu(E) \geq 0$
2. Null empty set: $\mu(\emptyset) = 0$
3. Countable-additivity[1] : $\forall i \neq j : E_i \cap E_j = \emptyset \implies \mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i)$

Definition 14.1.2 (Measure space). The pair (X, Σ) is called a measurable space. The elements $E \in \Sigma$ are called measurable sets. The triplet (X, Σ, μ) is called a measure space.

Method 14.1.3. To show that two measures coincide on a σ-algebra, it suffices to show that they coincide on the generating sets and apply the monotone class theorem 2.4.13.

Definition 14.1.4 (Almost everywhere[2]). Let (X, Σ, μ) be a measure space. A property P is said to hold on X almost everywhere (a.e.) if it satisfies the following equation:

$$\mu\left(\{x \in X : \neg P(x)\}\right) = 0 \tag{14.1}$$

Definition 14.1.5 (Complete measure space). The measure space (X, Σ, μ) is said to be complete if for every $E \in \Sigma$ with $\mu(E) = 0$ the following property holds for all $A \subset E$:

$$A \in \Sigma \quad \text{and} \quad \mu(A) = 0 \tag{14.2}$$

Definition 14.1.6 (Completion). Let $\mathcal{F}, \mathcal{G}$ be σ-algebras over a set X. $\mathcal{G}$ is said to be the completion of $\mathcal{F}$ if it is the smallest σ-algebra such that the measure space $(X, \mathcal{G}, \mu)$ is complete.

Definition 14.1.7 (Regular Borel measure). Let μ be a non-negative countably additive set function defined on $\mathcal{B}$. μ is called a regular Borel measure if it satisifes following equations for every Borel set B:

$$\mu(B) = \inf\{\mu(O) : O \text{ open}, O \supset B\} \tag{14.3}$$

$$\mu(B) = \sup\{\mu(F) : F \text{ closed}, F \subset B\} \tag{14.4}$$

[1] also called σ-additivity

[2] In probability theory this is foten often called **almost surely**.

Definition 14.1.8 (σ-finite measure). Let $(\Omega, \mathcal{F}, P)$ be a measure space. The measure P is said to be σ-finite if there exists a sequence $(A_i)_{i\in\mathbb{N}}$ of measurable sets such that $\bigcup_{i=1}^{+\infty} A_i = \Omega$ with $\forall A_i : P(A_i) < +\infty$.

Definition 14.1.9 (Measure-preserving map). Let (X, Ω) be a measure space. Consider a map $T : X \to X$. T is said to be measure-preserving if it satisfies the following equation:

$$\mu\left(T^{-1}(A)\right) = \mu(A) \tag{14.5}$$

for all $A \in \Omega$. This equation can also be written using a pullback notation: $T_*\mu = \mu$.

Definition 14.1.10 (Ergodic map). Let (X, Ω) be a measure space. Consider a measure-preserving map $T : X \to X$. T is said to be ergodic if the following conditions is satisfied:

$$T(A) = A \implies \mu(A) = 0 \vee \mu(X \backslash A) = 0 \tag{14.6}$$

14.1.2 Lebesgue measure

Formula 14.1.11 (Length of an interval). The length of an open interval $I = (a, b)$ is defined as:

$$l\left(I\right) = b - a \tag{14.7}$$

Definition 14.1.12 (Null set). A set $A \subset \mathbb{R}$ is called a null set if it can be covered by a sequence of intervals of arbitrarily small length: $\forall \varepsilon > 0$ there exists a sequence $(I_n)_{n\in\mathbb{N}}$ such that

$$A \subseteq \bigcup_{n=1}^{+\infty} I_n \tag{14.8}$$

with

$$\sum_{i=1}^{+\infty} l(I_n) < \varepsilon \tag{14.9}$$

Theorem 14.1.13. *Let $(E_i)_{i\in\mathbb{N}}$ be a sequence of null sets. The union $\bigcup_{i=1}^{+\infty} E_i$ is also null.*

Corollary 14.1.14. Any countable set is null.

Definition 14.1.15 (Outer measure). Let $X \subseteq \mathbb{R}$ be an open set. The (Lebesgue) outer measure is defined as:

$$\boxed{m^*(X) = \inf\left\{\sum_{i=1}^{+\infty} l(I_i) \text{ with } (I_i)_{i\in\mathbb{N}} \text{ a sequence of open intervals that covers } X\right\}} \tag{14.10}$$

Property 14.1.16. Let I be an interval. The outer measure equals the length: $m^*(I) = l(I)$.

Property 14.1.17. The outer measure is translation invariant: $m^*(A + t) = m^*(A) \quad , \forall A, t$

Property 14.1.18. $m^*(A) = 0$ if and only if A is null.

Property 14.1.19. If $A \subset B$ then $m^*(A) \leq m^*(B)$.

Property 14.1.20 (Countable subadditivity). For every sequence of sets $(E_i)_{i\in\mathbb{N}}$ the following inequality holds:

$$m^*\left(\bigcup_{i=1}^{+\infty} E_i\right) \leq \sum_{i=1}^{+\infty} m^*(E_i) \tag{14.11}$$

Theorem 14.1.21 (Carathéodory's criterion / Lebesgue measure). *Let X be a set. If X satisfies the following equation, it is said to be Lebesgue measurable:*

$$\forall E \subseteq \mathbb{R} : m^*(E) = m^*(E \cap X) + m^*(E \cap X^c) \tag{14.12}$$

This is denoted by $X \in \mathcal{M}$ and the outer measure $m^*(X)$ is called the Lebesgue measure of X denoted by $m(X)$.

Property 14.1.22. All null sets and intervals are measurable.

Property 14.1.23 (Countable additivity). For every sequence $(E_i)_{i\in\mathbb{N}}$ with $E_i \in \mathcal{M}$ satisfying $i \neq j : E_i \cap E_j = \emptyset$ the following equation holds:

$$\boxed{m\left(\bigcup_{i=1}^{+\infty} E_i\right) = \sum_{i=1}^{+\infty} m(E_i)} \tag{14.13}$$

Remark. Previous property, together with the properties of the outer measure, implies that the Lebesgue measure is indeed a proper measure as defined in 14.1.1.

Property 14.1.24. $\mathcal{M}$ is a σ-algebra[3] over $\mathbb{R}$.

Theorem 14.1.25. *For every $A \subset \mathbb{R}$ there exists a sequence $(O_i)_{i\in\mathbb{N}}$ of open sets such that:*

$$A \subset \bigcap_i O_i \qquad \text{and} \qquad m\left(\bigcap_i O_i\right) = m^*(A) \tag{14.14}$$

Theorem 14.1.26. *For every $E \in \mathcal{M}$ there exists a sequence $(F_i)_{i\in\mathbb{N}}$ of closed sets such that:*

$$\bigcup_i F_i \subset E \qquad \text{and} \qquad m\left(\bigcup_i F_i\right) = m(E) \tag{14.15}$$

Remark. The previous 2 theorems imply that the Lebesgue measure is a regular Borel measure 14.3.

Theorem 14.1.27. *Let $E \subset \mathbb{R}$. $E \in \mathcal{M}$ if and only if for every $\varepsilon > 0$ there exist an open set $O \supset E$ and a closed set $F \subset E$ such that $m^*(O \backslash E) < \varepsilon$ and $m^*(E \backslash F) < \varepsilon$.*

Property 14.1.28. Let $(A_i)_{i\in\mathbb{N}}$ be a sequence of sets with $\forall i : A_i \in \mathcal{M}$. The following two properties apply:

$$\forall i : A_i \subseteq A_{i+1} \implies m\left(\bigcup_{i=1}^{+\infty} A_i\right) = \lim_{i\to+\infty} m(A_i) \tag{14.16}$$

$$\forall i : A_i \supseteq A_{i+1} \wedge m(A_1) < +\infty \implies m\left(\bigcap_{i=1}^{+\infty} A_i\right) = \lim_{i\to+\infty} m(A_i) \tag{14.17}$$

Remark 14.1.29. This property is not only valid for the Lebesgue measure but for every countably additive set function.

Property 14.1.30. The Lebesgue measure $m(X)$ is continuous at $\emptyset$, i.e. if $(A_i)_{i\in\mathbb{N}} \to \emptyset$ then $\lim_{i\to+\infty} m(A_i) = 0$.

Theorem 14.1.31. *$\mathcal{M}$ is the completion of $\mathcal{B}$.*

Corollary 14.1.32. $\mathcal{B} \subset \mathcal{M} \subset \mathcal{F}_{\mathbb{R}}$

Definition 14.1.33 (Restricted Lebesgue measure). Let $B \subset \mathbb{R}$ be a measurable set with measure $m(B) > 0$. The restriction of the Lebesgue measure to the set B is defined as follows:

$$\mathcal{M}_B = \{A \cap B : A \in \mathcal{M}\} \qquad \text{and} \qquad \forall E \in \mathcal{M}_B : m_B(E) = m(E) \tag{14.18}$$

Furthermore, the measure space $(B, \mathcal{M}_B, m_B)$ is complete.

[3] See definition 2.4.2.

14.1.3 Measurable functions

Definition 14.1.34 (Measurable function). A function f is (Lebesgue) measurable if for every interval $I \subset \mathbb{R} : f^{-1}(I) \in \mathcal{M}$.

Definition 14.1.35 (Borel measurable function). A function f is called Borel measurable[4] if for every interval $I \subset \mathbb{R} : f^{-1}(I) \in \mathcal{B}$.

Remark 14.1.36. Inclusion 14.1.32 implies that every Borel function is also Lebesgue measurable.

Theorem 14.1.37. *The class of Lebesgue measurable[5] functions defined on $E \in \mathcal{M}$ is closed under multiplication and it forms a vector space.*

Property 14.1.38. Following types of functions are measurable:

- monotone functions
- continuous functions
- indicator functions

Corollary 14.1.39. Let f, g be measurable functions. Let $F : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a continuous function. The composition $F(f(x), g(x))$ is also measurable.

Property 14.1.40. Let f be a measurable function. The set[6] $\{x : f(x) = a\}$ is also measurable for all $a \in \mathbb{R}$.

Theorem 14.1.41. *Define following functions, which are measurable if f is measurable as a result of previous properties:*

$$f^+(x) = \begin{cases} f(x) & \text{if } f(x) > 0 \\ 0 & \text{if } f(x) \leq 0 \end{cases} = \max(f, 0) \tag{14.19}$$

$$f^-(x) = \begin{cases} 0 & \text{if } f(x) > 0 \\ -f(x) & \text{if } f(x) \leq 0 \end{cases} = \max(-f, 0) \tag{14.20}$$

The function $f : E \to \mathbb{R}$ is measurable if and only if both f^+ and f^- are measurable. Furthermore f is measurable if $|f|$ is measurable, the converse is false.

14.1.4 Limit operations

Property 14.1.42. Let $(f_i)_{i \in \mathbb{N}}$ be a sequence of measurable[7] functions. The following operations are measurable:

- $\min\limits_{i \leq k} f_i$ and $\max\limits_{i \leq k} f_i$
- $\inf\limits_{i \in \mathbb{N}} f_i$ and $\sup\limits_{i \in \mathbb{N}} f_i$
- $\liminf\limits_{i \to +\infty} f_i$ and $\limsup\limits_{i \to +\infty} f_i$

Remark. The measurability of the limit inferior and limit superior follows from their definitions and from the measurability of the inf / sup and min / max.

[4] These functions are often simply called 'Borel functions'.
[5] This property is also valid for Borel functions.
[6] This set is called the 'level set' of f.
[7] This property is also valid for Borel functions.

Property 14.1.43. Let f be a measurable function. Let g be a function such that $f = g$ almost everywhere. The function g is measurable.

Corollary 14.1.44. A result of the previous two properties is the following: if a sequence of measurable functions converges pointwise a.e. then the limit is also a measurable function.

Definition 14.1.45 (Essential supremum).

$$\text{ess sup} f = \sup\{z : f \geq z \text{ a.e.}\} \tag{14.21}$$

Definition 14.1.46 (Essential infimum).

$$\text{ess inf} f = \inf\{z : f \leq z \text{ a.e.}\} \tag{14.22}$$

Property 14.1.47. Let f be a measurable function. $f \leq \text{ess sup} f$ a.e. and $f \geq \text{ess inf} f$ a.e. We also have that: $\text{ess sup} f \leq \sup f$ and $\text{ess inf} f \geq \inf f$, furthermore this last pair of inequalities becomes a pair of equalities if f is continuous.

Property 14.1.48. Let f, g be measurable functions. $\text{ess sup}(f+g) \leq \text{ess sup} f + \text{ess sup} g$. An analogous inequality holds for the essential infimum.

14.2 Lebesgue integral

14.2.1 Simple functions

Definition 14.2.1 (Indicator function). An important function when working with sets is the following one:

$$\boxed{\mathbb{1}_A(x) = \begin{cases} 1 & \text{if} \quad x \in A \\ 0 & \text{if} \quad x \notin A \end{cases}} \tag{14.23}$$

Definition 14.2.2 (Simple function). Let f be a function that takes on a finite number of non-negative values $\{a_i\}$ with for every $i \neq j : f^{-1}(a_i) \cap f^{-1}(a_j) = \emptyset$. f is called a simple function if it can be expanded in the following way:

$$f(x) = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}(x) \tag{14.24}$$

with $A_i = f^{-1}(a_i) \in \mathcal{M}$

Remark 14.2.3 (Step function). If the sets A_i are intervals, the above function is often called a step function.

Formula 14.2.4 (Lebesgue integral of simple functions). Let φ be a simple function as defined in equation 14.24. Let $\mu : \mathcal{M} \to \mathbb{R}$ be a Lebesgue measure and let E be a measurable set. The Lebesgue integral of φ over a E with respect to μ is given by:

$$\int_E \varphi d\mu = \sum_{i=1}^{n} a_i \mu(E \cap A_i) \tag{14.25}$$

Example 14.2.5. Let $\mathbb{1}_{\mathbb{Q}}$ be the indicator function of the set of rational numbers. This function is clearly a simple function. Previous formula makes it possible to integrate the rational indicator function over the real line, which is not possible in the sense of Riemann:

$$\int_{\mathbb{R}} \mathbb{1}_{\mathbb{Q}} dm = 1 \times m(\mathbb{Q}) + 0 \times m(\mathbb{R} \backslash \mathbb{Q}) = 0 \tag{14.26}$$

where the measure of the rational numbers is 0 because it is a countable set (see corollary 14.1.14.

14.2.2 Measurable functions

Formula 14.2.6 (Lebesgue integral). Let f be a non-negative measurable function. Let A be measurable set. The Lebesgue integral of f over E is defined as:

$$\int_E f dm = \sup\left\{\int_E \varphi dm : \varphi \text{ a simple function such that } \varphi \leq f\right\} \tag{14.27}$$

Property 14.2.7. The Lebesgue integral $\int_E f dm$ of a measurable function f is always non-negative.

Notation 14.2.8. The following notation is frequently used (both in the sense of Riemann and Lebesgue):

$$\int f dm = \int_{\mathbb{R}} f dm \tag{14.28}$$

Formula 14.2.9. The following equality is easily proved as we know that for every set $A \subseteq \mathbb{R}$: $A \cup A^c = \mathbb{R}$.

$$\int_A f dm = \int f \mathbb{1}_A dm \tag{14.29}$$

Theorem 14.2.10. *Let f be a non-negative measurable function. Then $f = 0$ a.e. if and only if $\int_{\mathbb{R}} f dm = 0$.*

Property 14.2.11. The Lebesgue integral over a null set is 0.

Property 14.2.12. Let f, g me measurable functions. The Lebesgue integral has the following properties:

- $f \leq g$ a.e. implies $\int f dm \leq \int g dm$.
- Let A be a measurable set. Let $B \subset A$. Then $\int_B f dm \leq \int_A f dm$.
- The Lebesgue integral is linear.
- For every two disjoint measurable sets A and B we have that $\int_{A\cup B} f dm = \int_A f dm + \int_B f dm$.

Property 14.2.13 (Mean value theorem). If $a \leq f(x) \leq b$, then $am(A) \leq \int_A f dm \leq bm(A)$.

Theorem 14.2.14. *Let f be a non-negative measurable function. There exists an increasing sequence $(\varphi_i)_{i\in\mathbb{N}}$ of simple functions such that $\varphi_i \nearrow f$.*

Theorem 14.2.15. *Let f be a bounded measurable function defined on the interval $[a, b]$. For every $\varepsilon > 0$ there exists a step function[8] h such that $\int_a^b |f - h| dm < \varepsilon$.*

14.2.3 Integrable functions

Definition 14.2.16 (Integrable function). Let $E \in \mathcal{M}$. A measurable function f is said to be integrable over E if both $\int_E f^+ dm$ and $\int_E f^- dm$ are finite. The Lebesgue integral of f over E is defined as:

$$\int_E f dm = \int_E f^+ dm - \int_E f^- dm \tag{14.30}$$

[8] See remark 14.2.3.

Remark. The difference between the integral 14.27 and the integral of an integrable function is that with the latter f does not have to be non-negative.

Theorem 14.2.17. *f is integrable if and only if $|f|$ is integrable. Furthermore, $\int_E |f|dm = \int_E f^+dm + \int_E f^-dm$.*

Property 14.2.18. Let f, g be integrable functions. The following important properties apply:

- $f + g$ is also integrable.
- $\forall E \in \mathcal{M}, \int_E f dm \leq \int_E g dm \implies f \leq g$ a.e.
- Let $c \in \mathbb{R}$. $\int_E (cf)dm = c\int_E f dm$.
- f is finite a.e.
- $|\int f dm| \leq \int |f| dm$
- $f \geq 0 \wedge \int f dm = 0 \implies f = 0$ a.e.

Theorem 14.2.19. *The set of functions integrable over a set $E \in \mathcal{M}$ forms a vector space. It is denoted by $\mathcal{L}^1(E)$.*

Property 14.2.20. Let $f \in \mathcal{L}^1$ and $\varepsilon > 0$. There exists a continuous function g, vanishing outside some finite interval, such that $\int |f - g| dm < \varepsilon$.

Property 14.2.21. Let $f \geq 0$. The mapping $E \mapsto \int_E f dm$ is a measure on E (if it exists, hence if f is integrable). Furthermore, this measure is said to be **absolutely continuous**.

Remark. See section 14.6 for further information.

14.2.4 Convergence theorems

Theorem 14.2.22 (Fatou's lemma). *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of non-negative measurable functions.*

$$\int_E \left(\liminf_{n\to\infty} f_n\right) dm \leq \liminf_{n\to\infty} \int_E f_n dm \tag{14.31}$$

Theorem 14.2.23 (Monotone convergence theorem). *Let $E \in \mathcal{M}$. Let $(f_n)_{n\in\mathbb{N}}$ be an increasing sequence of non-negative measurable functions such that $f_n \nearrow f$ pointwise a.e. We have the following powerful equality:*

$$\boxed{\int_E f dm = \lim_{n\to\infty} \int_E f_n(x) dm} \tag{14.32}$$

Method 14.2.24. To prove 'linear' results concerning integrable functions in spaces such as $\mathcal{L}^1(E)$ we proceed according to the following steps:

1. Verify that the property holds for indicator functions. (This often follows by definition.)
2. Use the linearity to extend the property to simple functions.
3. Apply the monotone convergence theorem to show that the property holds for all non-negative measurable functions.
4. Extend the property to all integrable functions by writing $f = f^+ - f^-$ and applying the linearity again.

Theorem 14.2.25 (Dominated convergence theorem). *Let $E \in \mathcal{M}$. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of measurable functions with $\forall n : |f_n| \leq g$ a.e. for a function $g \in \mathcal{L}^1(E)$. If $f_n \to f$ pointwise a.e. then f is integrable over E and*

$$\int_E f dm = \lim_{n\to\infty} \int_E f_n(x) dm \tag{14.33}$$

Property 14.2.26. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of non-negative measurable functions. The following equality applies:

$$\int \sum_{n=1}^{+\infty} f_n(x) dm = \sum_{n=1}^{+\infty} \int f_n(x) dm \tag{14.34}$$

We cannot conclude that the right-hand side is finite a.e., so the series on the left-hand side need not be integrable.

Theorem 14.2.27 (Beppo-Levi). *Suppose that*

$$\sum_{i=1}^{\infty} \int |f_n|(x) dm \text{ is finite.}$$

The series $\sum_{i=1}^{\infty} f_n(x)$ converges a.e. Furthermore, the series is integrable and

$$\int \sum_{i=1}^{\infty} f_n(x) dm = \sum_{i=1}^{\infty} \int f_n(x) dm \tag{14.35}$$

Theorem 14.2.28 (Riemann-Lebesgue lemma). *Let $f \in \mathcal{L}^1$. The sequences*

$$s_k = \int_{-\infty}^{+\infty} f(x) \sin(kx) dx$$

and

$$c_k = \int_{-\infty}^{+\infty} f(x) \cos(kx) dx$$

both converge to 0.

Remark. This theorem is useful in Fourier analysis.

14.2.5 Relation to the Riemann integral

Theorem 14.2.29 (Fundamental theorem of calculus). *If $f : [a, b] \to \mathbb{R}$ is continuous then f is integrable and the function $F : x \mapsto \int_a^x f dm$ is differentiable for $x \in]a, b[$ such that $F' = f$.*

Theorem 14.2.30. *Let $f : [a, b] \to \mathbb{R}$ be a bounded function.*

- *f is Riemann-integrable if and only if f is continuous a.e. with respect to the Lebesgue measure on $[a, b]$.*
- *Riemann-integrable functions on $[a, b]$ are integrable with respect to the Lebesgue measure on $[a, b]$ and the integrals coincide.*

Theorem 14.2.31. *If $f \geq 0$ and the improper Riemann integral 12.11 exists, then the Lebesgue integral $\int f dm$ exists and the two integrals coincide.*

14.3 Examples

Definition 14.3.1 (Dirac measure[9]). We define the Dirac measure as follows:

$$\delta_a(X) = \begin{cases} 1 & \text{if } a \in X \\ 0 & \text{if } a \notin X \end{cases} \tag{14.36}$$

The integration with respect to the Dirac measure has the following nice property[10]:

$$\int g(x) d\delta_a = g(a) \tag{14.37}$$

Example 14.3.2. Let $\mu = \delta_2, X = (-4; 1)$ and $Y = (-2; 17)$. The following two integrals are easily computed:

$$\int_X d\mu = 0 \qquad\qquad \int_Y d\mu = 1$$

14.4 Space of integrable functions

14.4.1 Distance

To define a distance between functions, we first have to define some notion of length of a function. Normally this would not be a problem, because we know to integrate integrable functions, however the fact that two functions differing on a null set have the same integral carries problems with it, i.e. a non-zero function could have a zero length. Therefore we will define the 'length' function on a restricted vector space:

Define the following set of equivalence classes $L^1(E) = \mathcal{L}^1(E)_{/\equiv}$ by introducing the equivalence relation: $f \equiv g$ if and only if $f = g$ a.e.

Property 14.4.1. $L^1(E)$ is a Banach space[11].

Formula 14.4.2. A norm on $L^1(E)$ is given by:

$$||f||_1 = \int_E |f| dm \tag{14.38}$$

14.4.2 Hilbert space L^2

Property 14.4.3. L^2 is a Hilbert space[12].

Formula 14.4.4. A norm on $L^2(E)$ is given by:

$$||f||_2 = \left(\int_E |f|^2 dm \right)^{\frac{1}{2}} \tag{14.39}$$

This norm is induced by the following inner product:

$$\boxed{\langle f|g\rangle = \int_E f\overline{g} dm} \tag{14.40}$$

[9] Compare to definition 15.2.3.
[10] This equality can be proved by applying formula 40.14 with $X \equiv a$.
[11] See definition 21.1.6.
[12] See definition 21.2.1.

Now instead of deriving L^2 from $\mathcal{L}^2$ we do the opposite. We define $\mathcal{L}^2$ as the set of measurable functions for which equation 14.39 is finite.

Definition 14.4.5 (Orthogonality). As L^2 is a Hilbert space and thus has an inner product $\langle\cdot|\cdot\rangle$, it is possible to introduce the concept of orthogonality of functions in the following way:

$$\langle f|g\rangle = 0 \implies \text{f and g are orthogonal} \tag{14.41}$$

Furthermore it is also possible to introduce the angle between functions in the same way as equation 19.41.

Formula 14.4.6 (Cauchy-Schwarz inequality). Let $f, g \in L^2(E, \mathbb{C})$. We have that $fg \in L^1(E\mathbb{C})$ and:

$$\boxed{\left|\int_E f\overline{g}dm\right| \leq ||fg||_1 \leq ||f||_2||g||_2} \tag{14.42}$$

Remark. This follows immediately from formula 14.44.

Property 14.4.7. If E has finite Lebesgue measure then $L^2(E) \subset L^1(E)$.

14.4.3 L^p spaces

Generalizing the previous two Lebesgue function classes leads us to the notion of L^p spaces with the following norm:

Property 14.4.8. For all $1 \leq p \leq +\infty$ $L^p(E)$ is a Banach space with a norm given by:

$$||f||_p = \left(\int_E |f|^p \, dm\right)^{\frac{1}{p}} \tag{14.43}$$

Remark 14.4.9. Note that L^2 is the only L^p space that is also a Hilbert space. The other L^p spaces do not have a norm induced by an inner product.

Formula 14.4.10 (Hölder's inequality). Let $\frac{1}{p} + \frac{1}{q} = 1$ with $p \geq 1$. For every $f \in L^p(E)$ and $g \in L^q(E)$ we have that $fg \in L^1(E)$ and:

$$||fg||_1 \leq ||f||_p||g||_q \tag{14.44}$$

Formula 14.4.11 (Minkowski's inequality). For every $p \geq 1$ and $f, g \in L^p(E)$ we have

$$||f + g||_p \leq ||f||_p + ||g||_p \tag{14.45}$$

Property 14.4.12. If E has finite Lebesgue measure then $L^q(E) \subset L^p(E)$ when $1 \leq p \leq q < +\infty$.

14.4.4 L^∞ space of essentially bounded measurable functions

Definition 14.4.13 (Essentially bounded function). Let f be a measurable function satisfying ess $\sup|f| < +\infty$. The function f is said to be essentially bounded and the set of all such functions is denoted by $L^\infty(E)$.

Formula 14.4.14. A norm on L^∞ is given by:

$$||f||_\infty = \text{ess sup}|f| \tag{14.46}$$

This norm is called the **supremum norm** and it induces the supremum metric 8.7.

Property 14.4.15. L^∞ is a Banach space.

14.5 Product measures

14.5.1 Real hyperspace $\mathbb{R}^n$

The notions of intervals and lengths from the one dimensional case can be generalized to more dimensions in the following way:

Definition 14.5.1 (Hypercube). Let $I_1, ..., I_n$ be a sequence of intervals.

$$\mathbf{I} = I_1 \times ... \times I_n \tag{14.47}$$

Definition 14.5.2 (Generalized length). Let $\mathbf{I}$ be a hypercube induced by the sequence of intervals $I_1, ..., I_n$. The length of $\mathbf{I}$ is given by:

$$l(\mathbf{I}) = \prod_{i=1}^{n} l(I_i) \tag{14.48}$$

14.5.2 Construction of the product measure

Property 14.5.3 (General condition). The general condition for multi-dimensional Lebesgue measures is given by following equation which should hold for all $A_1 \in \mathcal{F}_1$ and $A_2 \in \mathcal{F}_2$:

$$\boxed{P(A_1 \times A_2) = P_1(A_1)P_2(A_2)} \tag{14.49}$$

Definition 14.5.4 (Section). Let $A = A_1 \times A_2$. The following two sets are called sections:

$$A_{\omega_1} = \{\omega_2 \in \Omega_2 : (\omega_1, \omega_2) \in A\} \subset \Omega_2$$

$$A_{\omega_2} = \{\omega_1 \in \Omega_1 : (\omega_1, \omega_2) \in A\} \subset \Omega_1$$

Property 14.5.5. Let $\mathcal{F} = \mathcal{F}_1 \times \mathcal{F}_2$. If $A \in \mathcal{F}$ then for each ω_1, $A_{\omega_1} \in \mathcal{F}_2$ and for each ω_2, $A_{\omega_2} \in \mathcal{F}_1$. Equivalently the sets $\mathcal{G}_1 = \{A \in \mathcal{F} : \forall\omega_1, A_{\omega_1} \in \mathcal{F}_2\}$ and $\mathcal{G}_2 = \{A \in \mathcal{F} : \forall\omega_2, A_{\omega_2} \in \mathcal{F}_1\}$ coincide with the product σ-algebra $\mathcal{F}$.

Property 14.5.6. The function $A_{\omega_2} \mapsto P(A_{\omega_2})$ is a step function:

$$P(A_{\omega_2}) = \begin{cases} P_1(A_1) & \text{if} \quad \omega_2 \in A_2 \\ 0 & \text{if} \quad \omega_2 \notin A_2 \end{cases}$$

Formula 14.5.7 (Product measure). From previous property it follows that we can write the product measure $P(A)$ in the following way:

$$\boxed{P(A) = \int_{\Omega_2} P_1(A_{\omega_2}) dP_2(\omega_2)} \tag{14.50}$$

Property 14.5.8. Let P_1, P_2 be finite. If $A \in \mathcal{F}$ then the functions

$$\omega_1 \mapsto P_2(A_{\omega_1}) \qquad\qquad \omega_2 \mapsto P_1(A_{\omega_2})$$

are measurable with respect to $\mathcal{F}_1$ and $\mathcal{F}_2$ respectively and

$$\boxed{\int_{\Omega_2} P_1(A_{\omega_2}) dP_2(\omega_2) = \int_{\Omega_1} P_2(A_{\omega_1}) dP_1(\omega_1)} \tag{14.51}$$

Furthermore the set function P is countably additive and if any other product measure coincides with P on all rectangles, it is equal to P on the whole product σ-algebra.

14.5.3 Fubini's theorem

Property 14.5.9. Let $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$ be a non-negtaive function. If f is measurable with respect to $\mathcal{F}_1 \times \mathcal{F}_2$ then for each $\omega_1 \in \Omega_1$ the function $\omega_2 \mapsto f(\omega_1, \omega_2)$ is measurable with respect to $\mathcal{F}_2$ (and vice versa). There integrals with respect to P_1 and P_2 respectively are also measurable.

Definition 14.5.10 (Section of a function). The functions $\omega_1 \mapsto f(\omega_1, \omega_2), \omega_2 \mapsto f(\omega_1, \omega_2)$ are called sections of f.

Theorem 14.5.11 (Tonelli's theorem). *Let $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$ be a non-negative function. The following equalities apply:*

$$\begin{aligned}\int_{\Omega_1 \times \Omega_2} f(\omega_1, \omega_2) d(P_1 \times P_2)(\omega_1, \omega_2) &= \int_{\Omega_1} \left(\int_{\Omega_2} f(\omega_1, \omega_2) dP_2(\omega_2) \right) dP_1(\omega_1) \\ &= \int_{\Omega_2} \left(\int_{\Omega_1} f(\omega_1, \omega_2) dP_1(\omega_1) \right) dP_2(\omega_2)\end{aligned} \tag{14.52}$$

Corollary 14.5.12 (Fubini's theorem). Let $f \in L^1(\Omega_1 \times \Omega_2)$. The sections are integrable in the appropriate spaces. Furthermore the functions $\omega_1 \mapsto \int_{\Omega_2} f dP_2$ and $\omega_2 \mapsto \int_{\Omega_1} f dP_1$ are in $L^1(\Omega_1)$ and $L^1(\Omega_2)$ respectively and equality 14.52 holds.

Remark 14.5.13. The previous construction and theorems also apply for higher dimensional product spaces. These thereoms provide a way to construct higher-dimensional Lebesgue measures m_n by defining them as the completion of the product of n one-dimensional Lebesgue measures.

14.6 Radon-Nikodym theorem

Definition 14.6.1. Let $(\Omega, \mathcal{F})$ be a measurable space. Let μ, ν be two measures defined on this space. ν is said to be **absolutely continuous with respect to** μ if

$$\forall A \in \mathcal{F} : \mu(A) = 0 \implies \nu(A) = 0 \tag{14.53}$$

Notation 14.6.2. This relation is denoted by $\nu \ll \mu$.

Theorem 14.6.3 (Absolute continuity). *Let μ, ν be finite measures on a measurable space $(\Omega, \mathcal{F})$. Then $\nu \ll \mu$ if and only if*

$$\forall \varepsilon > 0 : \exists \delta > 0 : \forall A \in \mathcal{F} : \mu(A) < \delta \implies \nu(A) < \varepsilon \tag{14.54}$$

Property 14.2.21 can be generalized to arbitrary measure spaces as follows:

Property 14.6.4. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Let $f : \Omega \to \mathbb{R}$ be a measurable function such that $\int f d\mu$ exists. Then $\nu(f) = \int_F f d\mu$ defines a measure $\nu \ll \mu$.

Definition 14.6.5 (Dominated measure). Let μ, ν be two measures. μ is said to **dominate** ν if $0 \leq \nu(F) \leq \mu(F)$ for every $F \in \mathcal{F}$.

Theorem 14.6.6 (Radon-Nikodym theorem for dominated measures).
Let μ be a measure such that $\mu(\Omega) = 1$. Let ν be a measure dominated by μ. There exists a non-negative $\mathcal{F}$-measurable function h such that $\nu(F) = \int_F h d\mu$ for all $F \in \mathcal{F}$.

Remark. The assumption $\mu(\Omega) = 1$ is non-restrictive as every other finite measure ϕ can be normalized by putting $\mu = \frac{\phi}{\phi(\Omega)}$.

Definition 14.6.7 (Radon-Nikodym derivative). The function h as defined in previous theorem is called the Radon-Nikodym derivative of ν with respect to μ and we denote it by $\frac{d\nu}{d\mu}$.

Theorem 14.6.8 (Radon-Nikodym theorem). *Let $(\Omega, \mathcal{F})$ be a measurable space. Let μ, ν be two σ-finite measures defined on this space such that $\nu \ll \mu$. There exists a non-negative measurable function $g : \Omega \to \mathbb{R}$ such that $\nu(F) = \int_F g d\mu$ for all $F \in \mathcal{F}$.*

Remark 14.6.9. The function g in the previous theorem is unique up to a μ-null (or ν-null) set.

Property 14.6.10. Let μ, ν be finite measures such that μ dominates ν. Let $h_\nu = \frac{d\nu}{d\mu}$ be the associated Radon-Nikodym derivative. For every non-negative $\mathcal{F}$-measurable function f we have

$$\int_\Omega f d\nu = \int_\Omega f h_\nu d\mu \tag{14.55}$$

Remark 14.6.11. This property also holds for all functions $f \in L^1(\mu)$.

Property 14.6.12. Let λ, ν, μ be σ-finite measures. If $\lambda \ll \mu$ and $\nu \ll \mu$ then we have:

- $\frac{d(\lambda + \nu)}{d\mu} = \frac{d\lambda}{d\mu} + \frac{d\lambda}{d\mu}$ a.e.
- Chain rule: if $\lambda \ll \nu$ then $\frac{d\lambda}{d\mu} = \frac{d\lambda}{d\nu}\frac{d\nu}{d\mu}$ a.e.

14.7 Lebesgue-Stieltjes measure

Chapter 15

Distributions

15.1 Generalized function

Definition 15.1.1 (Schwartz space). The Schwartz space or **space of rapidly decreasing functions**[1] $S(\mathbb{R})$ is defined as:

$$S(\mathbb{R}) = \left\{ f(x) \in C^\infty(\mathbb{R}) : \forall i, j \in \mathbb{N} : \forall x \in \mathbb{R} : |x^i f^{(j)}(x)| < +\infty \right\} \tag{15.1}$$

Remark 15.1.2. This definition can be generalized to functions in $C^\infty(\mathbb{R}^n)$ or functions $f : \mathbb{R}^n \to \mathbb{C}$. The Schwartz space is then denoted by $S(\mathbb{R}^n, \mathbb{C})$.

Definition 15.1.3 (Functions of slow growth). The set of functions of slow growth $N(\mathbb{R})$ is defined as:

$$N(\mathbb{R}) = \{ f(x) \in C^\infty(\mathbb{R}) : \forall i \in \mathbb{N}, \exists M_i > 0 : |f^{(i)}(x)| = O(|x|^i) \text{ for } |x| \to +\infty \} \tag{15.2}$$

Remark. It is clear that all polynomials belong to $N(\mathbb{R})$ but not to $S(\mathbb{R})$.

Property 15.1.4. If $f(x) \in S(\mathbb{R})$ and $a(x) \in N(\mathbb{R})$ then $a(x)f(x) \in S(\mathbb{R})$.

Definition 15.1.5 (Generalized function). Let $g(x) \in S(\mathbb{R})$ be a test function. Let $\{f_n(x) \in S(\mathbb{R})\}, \{h_n(x) \in S(\mathbb{R})\}$ be sequences such that

$$\lim_{n \to +\infty} \langle f_n(x) | g(x) \rangle = \lim_{n \to +\infty} \int_{-\infty}^{+\infty} f_n(x) g(x) dx$$

and similarly for h_n. Define the equivalence relation $\{f_n(x) \in S(\mathbb{R})\} \sim \{h_n(x) \in S(\mathbb{R})\}$ by saying that the two sequences, satisfying the previous condition, are equivalent if and only if

$$\lim_{n \to +\infty} \langle f_n(x) | g(x) \rangle = \lim_{n \to +\infty} \langle h_n(x) | g(x) \rangle$$

A generalized function is defined as a complete equivalence class under previous relation.

Notation 15.1.6. Let ψ be a generalized function. Let $f \in S(\mathbb{R})$. The inner product 14.40 is generalized by following functional:

$$\langle \psi | f \rangle = \lim_{n \to +\infty} \int_{-\infty}^{+\infty} \psi_n(x) f(x) dx \tag{15.3}$$

[1] These functions are said to be rapidly decreasing because every derivative $f^{(j)}(x)$ decays faster than any polynomial x^i for $x \to +\infty$.

Property 15.1.7. Let ψ be a generalized function. Let $f(x) \in S(\mathbb{R})$. The previous functional has following properties:

- $\forall i \in \mathbb{N} : \langle \psi^{(i)} | f \rangle = (-1)^i \langle \psi | f^{(i)} \rangle$
- $\forall a, b \in \mathbb{R}, a \neq 0 : \langle \psi(ax+b) | f(x) \rangle = |a|^{-1} \langle \psi(x) | f(x - b/a) \rangle$
- $\forall a(x) \in N(\mathbb{R}) : \langle a\psi | f \rangle = \langle \psi | af \rangle$

Property 15.1.8 (Ordinary function as generalized function). Let $f : \mathbb{R} \to \mathbb{C}$ be a function such that $\exists M \geq 0 : (1+x^2)^{-M} |f(x)| \in L(\mathbb{R}, \mathbb{C})$[2]. There exists a generalized function $\psi \sim \{f_n(x) \in S(\mathbb{R}, \mathbb{C})\}$ such that for every $g(x) \in S(\mathbb{R}, \mathbb{C})$:

$$\langle \psi | g \rangle = \langle f | g \rangle$$

Furthermore if $f(x)$ is continuous on an interval, then $\lim_{n \to +\infty} f_n(x) = f(x)$ converges pointwise on that interval.

15.2 Dirac Delta distribution

Definition 15.2.1 (Heaviside function). Define the generalized function $H \sim \{H_n(x) \in S(\mathbb{R})\}$ as:

$$H(x) = \begin{cases} 0 & if \quad x < 0 \\ 1 & if \quad x \geq 0 \end{cases} \tag{15.4}$$

From this definition it follows that for every $f \in S(\mathbb{R})$:

$$\langle H | f \rangle = \int_0^{+\infty} f(x) dx \tag{15.5}$$

Remark 15.2.2. For the above integral to exist, $f(x)$ does not need to be an element of $S(\mathbb{R})$. It is a sufficient condition, but not a necessary one.

Definition 15.2.3 (Generalized delta function). The Dirac delta function is defined as a representant of the equivalence class of generalized functions $\{H_n'(x) \in S(\mathbb{R})\}$. By equations 15.1.7 and 15.5 we have for every $f \in S(\mathbb{R})$:

$$\begin{aligned} \langle \delta | f \rangle &= \langle H' | f \rangle \\ &= -\langle H | h \rangle \\ &= -\int_0^{+\infty} f'(x) dx \\ &= f(0) \end{aligned}$$

Property 15.2.4 (Sampling property). The result from previous definition can be generalized in the following way:

$$\boxed{f(x_0) = \int_{\mathbb{R}} f(x) \delta(x - x_0) dx} \tag{15.6}$$

Example 15.2.5 (Dirac comb).

$$III_b(x) = \sum_n \delta(x - nb) \tag{15.7}$$

[2]The space of Lebesgue integrable functions 14.2.19.

Property 15.2.6. Let $f(x) \in C^1(\mathbb{R})$ be a function with roots at $x_1, x_2, ..., x_n$ such that $f'(x_i) \neq 0$. The Dirac delta distribution has the following property:

$$\delta[f(x)] = \sum_{i=1}^{n} \frac{1}{|f'(x_i)|}\delta(x - x_i) \tag{15.8}$$

Property 15.2.7 (Convolution with delta function). Let $f(x) \in S(\mathbb{R})$. Let $\otimes$ denote the convolution.

$$\delta(x) \otimes f(x) = \int_{-\infty}^{+\infty} \delta(x - \alpha) f(\alpha) d\alpha = f(x) \tag{15.9}$$

Formula 15.2.8 (Differentiation across discontinuities). Let $f(x)$ be a piecewise continuous function with discontinuities at $x_1, ..., x_n$. Let f satisfy the conditions to be a generalized function. Define $\sigma_i = f^+(x_i) - f^-(x_i)$ which are the jumps of f at its discontinuities. Next, define the (continuous) function

$$f_c(x) = f(x) - \sum_{i=1}^{n} \sigma_i H(x - x_i)$$

Differentiation gives

$$f'(x) = f_c'(x) + \sum_{i=1}^{n} \sigma_i \delta(x - x_i)$$

It follows that the derivative in the generalized sense of a piecewise continuous function equals the derivative in the classical sense plus a summation of delta functions at every jump discontinuity.

15.3 Fourier series

Definition 15.3.1 (Dirichlet kernel). The Dirichlet kernel is the collection of functions of the form

$$D_n(x) = \frac{1}{2\pi} \sum_{k=-n}^{n} e^{ikx} \tag{15.10}$$

Formula 15.3.2 (Sieve property). If $f \in C^1[-\pi, \pi]$ then

$$\lim_{n \to +\infty} \int_{-\pi}^{\pi} f(x) D_n(x) dx = 0 \tag{15.11}$$

Formula 15.3.3. For 2π-periodic functions, the n^{th} degree Fourier approximation is given by the following convolution:

$$s_n(x) = \sum_{k=-n}^{n} \widetilde{f}(k) e^{ikx} = (D_n * f)(x) \tag{15.12}$$

Theorem 15.3.4 (Convergence of the Fourier series). *Let $f : \mathbb{R} \to \mathbb{R}$ be a function with period 2π. If $f(x)$ is piecewise C^1 on $[-\pi, \pi]$ the the limit $\lim_{n \to +\infty}(D_n * f)(x)$ converges to $\frac{f(x+)+f(x-)}{2}$ for all $x \in \mathbb{R}$.*

Formula 15.3.5 (Generalized Fourier series). Let $f(x) \in \mathcal{L}^2[-l, l]$ be a $2l$-periodic function. This function can be approximated by the following series:

$$\boxed{f(x) = \sum_{n=-\infty}^{+\infty} \left(\frac{1}{2l} \int_{-l}^{l} e^{-i\frac{n\pi x'}{l}} f(x') dx' \right) e^{i\frac{n\pi x}{l}}} \tag{15.13}$$

Formula 15.3.6 (Fourier coefficients). As seen in the general formula, the Fourier coefficient $\widetilde{f}(n)$ can be calculated by taking the inner product 21.1 of $f(x)$ and the n-th eigenfunction e_n:

$$\widetilde{f}(n) = \langle e_n | f \rangle = \int_{-l}^{l} e_n^*(x) f(x) dx \qquad \text{with} \qquad e_n = \sqrt{\frac{1}{2l}} e^{i\frac{n\pi x}{l}} \tag{15.14}$$

Definition 15.3.7 (Periodic extension). Let $f(x)$ be piecewise C^1 on $[-L, L]$. The periodic extension $f^L(x)$ is defined by repeating the restriction of $f(x)$ to $[-L, L]$ every $2L$. The **normalized periodic extension** is defined as

$$f^{L,\nu}(x) = \frac{f^L(x+) + f^L(x-)}{2} \tag{15.15}$$

Theorem 15.3.8. *If $f : \mathbb{R} \to \mathbb{R}$ is piecewise C^1 on $[-L, L]$ then the Fourier series approximation of $f(x)$ converges to $f^{L,\nu}(x)$ for all $x \in \mathbb{R}$.*

15.4 Fourier transform

The Fourier series can be used to expand a $2l$-periodic function as an infinite series of exponentials. For expanding a non-periodic function we need the Fourier integral:

$$\boxed{\mathcal{F}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt} \tag{15.16}$$

$$f(t) = \mathcal{F}^{-1}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathcal{F}(\omega) e^{i\omega t} d\omega \tag{15.17}$$

Equation 15.16 is called the (forward) Fourier transform of $f(t)$ and equation 15.17 is called the inverse Fourier transform.

Notation 15.4.1. The Fourier transform of a function $f(t)$, as seen in equation 15.16, is often denoted by $\widetilde{f}(\omega)$.

Theorem 15.4.2 (Convergence of the Fourier integral). *If $f : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous (see 12.3) and if $\int_{-\infty}^{+\infty} |f(x)| dx$ is convergent then the Fourier integral converges to $f(x)$ for all $x \in \mathbb{R}$.*

Theorem 15.4.3 (Fourier inversion theorem). *If both $f(t), \mathcal{F}(\omega) \in \mathcal{L}^1(\mathbb{R})$ are continuous then the Cauchy principal value in 15.17 can be replaced by a normal integral.*

Remark 15.4.4. Schwartz functions are continuous elements of $\mathcal{L}^1(\mathbb{R})$ and as such the Fourier inversion theorem also holds for these functions. This is interesting because checking the conditions for Schwartz functions is often easier then checking the more general conditions of the theorem.

Property 15.4.5. From the Riemann-Lebesgue lemma 14.2.28 it follows that

$$\mathcal{F}(\omega) \to 0 \qquad \text{if} \qquad |\omega| \to 0 \tag{15.18}$$

Property 15.4.6 (Parceval's theorem). Let $(f, \widetilde{f})$ and $(g, \widetilde{g})$ be two Fourier transform pairs.

$$\int_{-\infty}^{+\infty} f(x) g(x) dx = \int_{-\infty}^{+\infty} \widetilde{f}(k) \widetilde{g}(k) dk \tag{15.19}$$

Corollary 15.4.7 (Plancherel theorem). The integral of the square (of the modulus) of a Fourier transform is equal to the integral of the square (of the modulus) of the original function:

$$\int_{-\infty}^{+\infty} |f(x)|^2 dx = \int_{-\infty}^{+\infty} |\widetilde{f}(k)|^2 dk \tag{15.20}$$

Theorem 15.4.8. *Let $f(x), F(k)$ be a Fourier transform pair. If $f(x) \in S(\mathbb{R}, \mathbb{C})$, then $F(k) \in S(\mathbb{R}, \mathbb{C})$. It follows that for a sequence $\{f_n(x) \in S(\mathbb{R}, \mathbb{C})\}$ the sequence of Fourier transformed functions $\{F_n(x) \in S(\mathbb{R}, \mathbb{C})\}$ is also a subset of the Schwartz space. Furthermore Parceval's theorem 15.19 gives*

$$\int_{-\infty}^{+\infty} f_n(x) g(x) dx = \int_{-\infty}^{+\infty} F_n(x) G(x) dx \in \mathbb{R}$$

where $g(x) \in S(\mathbb{R}, \mathbb{C})$. From these two properties it follows that the Fourier transform of a generalized function is also a generalized function.

Property 15.4.9. Let ψ be a generalized function with Fourier transform Ψ. Let $f(x) \in S(\mathbb{R}, \mathbb{C})$ with Fourier transform $F(k)$. We have the following equality:

$$\langle \psi | F \rangle = \langle \Psi | f \rangle \tag{15.21}$$

15.4.1 Convolution

Formula 15.4.10 (Convolution).

$$\boxed{(f * g)(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(\tau) g(t - \tau) d\tau} \tag{15.22}$$

Property 15.4.11 (Commutativity).

$$f * g = g * f \tag{15.23}$$

Theorem 15.4.12 (Convolution theorem).

$$\widetilde{f * g} = \widetilde{g}\widetilde{f} \tag{15.24}$$

15.5 Laplace transform

Formula 15.5.1 (Laplace transform).

$$\mathcal{L}\{F(t)\}_{(s)} = \int_0^{\infty} f(t) e^{-st} dt \tag{15.25}$$

Formula 15.5.2 (Bromwich integral).

$$f(t) = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} \mathcal{L}\{F(t)\}_{(s)} e^{st} ds \tag{15.26}$$

Notation 15.5.3. The Laplace transform as defined in equation 15.25 is sometimes denoted by $f(s)$.

15.6 Mellin transform

Formula 15.6.1 (Mellin transform).

$$\mathcal{M}\{f(x)\}(s) = \int_0^{+\infty} x^{s-1} f(x) dx \tag{15.27}$$

Formula 15.6.2 (Inverse Mellin transform).

$$f(x) = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} \mathcal{M}\{f(x)\}_{(s)} x^{-s} ds \tag{15.28}$$

15.7 Integral representations

Formula 15.7.1 (Heaviside step function).

$$\theta(x) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{e^{ikx}}{x - i\varepsilon} dk \tag{15.29}$$

Formula 15.7.2 (Dirac delta function).

$$\delta^{(n)}(\vec{\boldsymbol{x}}) = \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} e^{i\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{x}}} d^n k \tag{15.30}$$

Chapter 16

Ordinary differential equations

16.1 Boundary conditions

Unigue solutions of a differential equation are obtained by supplying additional conditions. These are called boundary conditions.

16.1.1 Periodic boundary conditions

Periodic boundary conditions are conditions of the following form:

$$y(x) = y(x + \varphi) \tag{16.1}$$

By induction it follows that for every n:

$$y(x) = y(x + n\varphi) \tag{16.2}$$

16.1.2 Dirichlet boundary conditions

Dirichlet boundary conditions are conditions of the following form:

$$y(x) = f(x) \qquad , \qquad x \in \partial\Omega \tag{16.3}$$

where Ω is the domain of the problem.

Remark 16.1.1. When y is a function of multiple variables, α can be a function as well. For example (in spherical coordinates: ρ, ϕ, θ):

$$y(x, \phi, \theta) = \alpha(\phi, \theta) \tag{16.4}$$

16.1.3 Neumann boundary conditions

Neumann boundary conditions are conditions of the following form:

$$y'(a) = \alpha \tag{16.5}$$

Remark 16.1.2. When y is a function of multiple variables, we obtain the following form (where S is the boundary of the domain and $\hat{n}$ a normal vector to this boundary):

$$\frac{\partial y}{\partial \hat{n}}(\vec{x}) = f(\vec{x}) \qquad , \qquad \vec{x} \in S \tag{16.6}$$

16.2 First order ODE's

Formula 16.2.1 (First order ODE).

$$\boxed{y'(t) + a(t)y(t) = R(t)} \tag{16.7}$$

If the function $R(t)$ is identically zero, then the ODE is said to be **homogenous**.

Theorem 16.2.2. *Let $U \subseteq \mathbb{R}$ be an open set. Let the functions $a(t), R(t) : U \to \mathbb{R}$ be continuous. The solutions $\varphi(t) : U \to \mathbb{R}$ of equation 16.7 are given by:*

$$\boxed{\varphi(t) = e^{-\int a(t)dt}\left(c + \int R(t)e^{\int a(t)dt}dt\right)} \tag{16.8}$$

where c is a constant.

16.3 Second order ODE's

Formula 16.3.1 (Second order ODE).

$$\boxed{y''(t) + a(t)y'(t) + b(t)y(t) = R(t)} \tag{16.9}$$

Formula 16.3.2 (Homogeneous second order ODE).

$$y''(t) + a(t)y'(t) + b(t)y(t) = 0 \tag{16.10}$$

16.3.1 General solution

Formula 16.3.3. Let $\varphi : U \to \mathbb{R}$ be a nowhere zero solution of the homogeneous equation 16.10. The general solution of equation 16.9 is then given by:

$$\boxed{y(t) = c_1\ \varphi + c_2\ \varphi \int \frac{e^{-\int a}}{\varphi^2} + \psi_0} \tag{16.11}$$

where ψ_0 is a particular solution of equation 16.9.

Theorem 16.3.4. *Let ψ_0 be a solution of equation 16.9. The set of all solutions is given by the affine space:*

$$\{\psi_0 + \chi\ :\ \chi \text{ is a solution of the homogeneous equation 16.10}\} \tag{16.12}$$

Theorem 16.3.5. *Two solutions of the homogeneous equation 16.10 are independent if the wronskian is nonzero:*

$$W\left(\varphi_1(x), \varphi_2(x)\right) = \begin{vmatrix} \varphi_1(x) & \varphi_2(x) \\ \varphi_1'(x) & \varphi_2'(x) \end{vmatrix} \neq 0 \tag{16.13}$$

Formula 16.3.6 (Abel's identity). An explicit formula for the wronskian is given by:

$$\boxed{W(x) = W(x_0)\exp\left(-\int_{x_0}^{x} a(x')dx'\right)} \tag{16.14}$$

16.3.2 Constant coefficients

Theorem 16.3.7. *A map $\varphi : U \to \mathbb{C}$ is a complex solution of equation 16.10 if and only if $Re\{\varphi\}$ and $Im\{\varphi\}$ are real solutions of equation 16.10.*

Formula 16.3.8 (Characteristic equation). When having an ODE of the form[1] :

$$y''(t) + py'(t) + qy(t) = 0 \tag{16.15}$$

where p and q are constants, we define the characteristic equation as follows:

$$\lambda^2 + p\lambda + q = 0 \tag{16.16}$$

This polynomial equation generally[2] has two distinct (complex) roots λ_1 and λ_2. From these roots we can derive the solutions of equation 16.15 using the following rules (c_1 and c_2 are constants):

- $\lambda_1 \neq \lambda_2$, λ_1 and $\lambda_2 \in \mathbb{R}$: $y(t) = c_1\ e^{\lambda_1 t} + c_2\ e^{\lambda_2 t}$
- $\lambda_1 = \lambda_2$: $y(t) = c_1\ e^{\lambda t} + c_2\ te^{\lambda t}$
- $\lambda_1 = \lambda_2^*$, where $\lambda_1 = a + ib$: $y(t) = c_1\ e^{at}\cos(bt) + c_2\ e^{at}\sin(bt)$

16.3.3 Method of Frobenius

Formula 16.3.9 (Method of Frobenius). To find a solution of the homogeneous equation 16.10 we assert a solution of the form:

$$\boxed{y(x) = \sum_{i=0}^{\infty} a_i (x - x_0)^{i+k}} \tag{16.17}$$

where k is a constant.

Definition 16.3.10 (Indicial equation). After inserting the solution 16.17 into the homogeneous equation 16.10 we obtain[3] an equation of the form $\sum_{i=n}^{\infty} H_i(k)x^i = 0$ where $n \in \mathbb{R}$ and $H_i(k)$ is a polynomial in k. This means that for every i we obtain an equation of the form $H_i(k) = 0$, due to the independence of polynomial terms. The equation for the lowest power will be quadratic in k and it is called the indicial equation.

Theorem 16.3.11. *The indicial equation generally has two roots k_1, k_2. The following possibilities arise:*

- *$k_1 = k_2$: Only one solution will be found with the method of Frobenius (another one can be found as in the second term of equation 16.11)*
- *$k_1 - k_2 \in \mathbb{Z}$: A second independent solution might be obtained using this method. If not, then a second solution can be found as mentioned in the previous case.*
- *$k_1 - k_2 \notin \mathbb{Z}$: Two independent solutions can be found using this method.*

Theorem 16.3.12 (Fuch's theorem). *If $a(z)$ and $b(z)$ are analytic at $z = z_0$ then the general solution $y(z)$ can be expressed as a Frobenius' series.*

[1] Any other form of homogeneous second order ODE's with constant coefficients can be rewritten in this form.
[2] See theorem 11.1.4 ("Fundamental theorem of algebra").
[3] It is important to 'sync' the power of all terms in order to obtain one 'large' coefficient.

16.4 Sturm-Liouville theory

Definition 16.4.1 (Sturm-Liouville boundary value problem). The following ODE, subject to mixed boundary conditions, is called a Sturm-Liouville boundary value problem:

$$\boxed{\frac{d}{dx}\left[p(x)\frac{dy}{dx}\right] + [g(x) + \lambda r(x)]\, y(x) = 0} \tag{16.18}$$

where $p(x), q(x)$ and $r(x)$ are continuous on $a \leq x \leq b$. $p(x) \in C^1(a,b)$ with $p(x) < 0$ or $p(x) > 0$ for $a \leq x \leq b$. $r(x) \geq 0$ or $r(x) \leq 0$ for $a \leq x \leq b$ and $r(x)$ is not identically zero on any subinterval.

The boundary conditions are given by

$$\begin{cases} \alpha_1 y(a) + \beta_1 y'(a) = 0 \\ \alpha_2 y(b) + \beta_2 y'(b) = 0 \end{cases} \tag{16.19}$$

where at least one of the constants $\alpha_1, \alpha_2, \beta_1$ or β_2 is non-zero.

Formula 16.4.2. The solutions are of the form

$$y(x) = c_1 u_1(\lambda; x) + c_2 u_2(\lambda; x)$$

Only for certain values of λ will these solutions (u_1, u_2) be non-trivial. The values of λ for which the solutions are non-trivial are called **eigenvalues** and the associated solutions are called **eigenfunctions**. Substituting this form in the boundary conditions gives the following determinant condition for non-trivial solutions, which is also the defining equation of the eigenvalues λ:

$$\begin{vmatrix} \alpha_1 u_1(a;\lambda) + \beta_1 u_1'(a;\lambda) & \alpha_1 u_2(a;\lambda) + \beta_1 u_2'(a;\lambda) \\ \alpha_1 u_1(b;\lambda) + \beta_1 u_1'(b;\lambda) & \alpha_1 u_2(b;\lambda) + \beta_1 u_2'(b;\lambda) \end{vmatrix} = 0 \tag{16.20}$$

The independent eigenfunctions can be found by substituting the found eigenvalues in the ODE 16.18.

Definition 16.4.3 (Self-adjoint form). The SL-problem can be rewritten as

$$\left[\hat{\mathcal{L}} + \lambda r(x)\right] y(x) = 0$$

The operator $\hat{\mathcal{L}} = \frac{d}{dx}\left[p(x)\frac{d}{dx} + g(x)\right]$ is called the self-adjoint form. Now consider the general linear ODE

$$\left[a_2(x)\frac{d^2}{dx^2} + a_1(x)\frac{d}{dx} + a_0(x)\right] y(x) = 0 \tag{16.21}$$

This equation can be rewritten in a self-adjoint form by setting:

$$p(x) = e^{\int \frac{a_1}{a_2} dx} \qquad \text{and} \qquad g(x) = \frac{a_0}{a_2} e^{\int \frac{a_1}{a_2} dx} \tag{16.22}$$

Property 16.4.4. The eigenfunctions corresponding to distinct eigenvalues are orthogonal with respect to the weight function $r(x)$.

Theorem 16.4.5 (Oscillation theorem). *Let f_n be the n^{th} eigenfunction of a Sturm-Liouville boundary condition problem. Then f_n has precisely $n-1$ roots.*

Chapter 17

Partial differential equations

17.1 General linear equations

Formula 17.1.1 (Cramer's rule). Let $Ax = b$ be a system of linear equations where the matrix A has a nonzero determinant. Then Cramer's rule gives a unique solution where the unknowns are given by;

$$x_i = \frac{\det(A_i)}{\det(A)} \tag{17.1}$$

where A_i is the matrix obtained by replacing the i^{th} column of A by the column matrix b.

Definition 17.1.2 (Characteristic curve). Curve along which the highest order partial derivatives are not uniquely defined.

17.2 First order PDE

Formula 17.2.1 (First order quasilinear PDE).

$$\boxed{P(x,y,z)\frac{\partial z}{\partial x} + Q(x,y,z)\frac{\partial z}{\partial y} = R(x,y,z)} \tag{17.2}$$

Formula 17.2.2 (Characteristic curve). The PDE will have no unique solution if

$$\begin{vmatrix} P & Q \\ dx & dy \end{vmatrix} = 0 \tag{17.3}$$

and will have a non-unique solution if

$$\begin{vmatrix} P & R \\ dx & dz \end{vmatrix} = 0 \tag{17.4}$$

The characteristic curves are thus defined by $\frac{dx}{P} = \frac{dy}{Q}$ and along thse curves the condition $\frac{dx}{P} = \frac{dz}{R}$ should hold to ensure a solution.

Theorem 17.2.3. *The general solution of 17.2 is implicitly given by $F(\xi,\eta) = 0$ with $F(\xi,\eta)$ an arbitrary differentiable function where $\xi(x,y,z) = c_1$ and $\eta(x,y,z) = c_2$ are solutions of the equation*

$$\frac{dx}{P} = \frac{dy}{Q} = \frac{dz}{R} \tag{17.5}$$

where c_1, c_2 are constants which are fixed by boundary conditions.

Remark 17.2.4. Looking at the defining equations of the characteristic curve, it is clear that these fix the general solution of the PDE.

17.3 Characteristics

Formula 17.3.1 (Second order quasilinear PDE). Consider the following pseudolinear differential equation for the function $u(x, y)$:

$$R(x,y)u_{xx} + S(x,y)u_{xy} + T(x,y)u_{yy} = W(x,y,u,p,q) \tag{17.6}$$

where $p = u_x$ and $q = u_y$.

Formula 17.3.2 (Equation of characteristics). Consider the following differential equations:

$$\begin{cases} u_{xx}dx + u_{xy}dy = dp \\ u_{xy}dx + u_{yy}dy = dq \end{cases} \tag{17.7}$$

According to Cramer's rule 17.1 these equations, together with the PDE 17.6, give the following condition for the characteristic curves:

$$\begin{vmatrix} R(x,y) & S(x,y) & T(x,y) \\ dx & dy & 0 \\ 0 & dx & dy \end{vmatrix} = 0 \tag{17.8}$$

which is equivalent to following equation:

$$\boxed{R\left(\frac{dy}{dx}\right)^2 - S\left(\frac{dy}{dx}\right) + T = 0} \tag{17.9}$$

Definition 17.3.3 (Types of characteristics). Equation 17.9 is quadratic in $\frac{dy}{dx}$. If this equation has two distinct real roots then the PDE is said to be **hyperbolic**. If the equation has only one root, the PDE is said to be **parabolic**. In the remaining case, where the equation has two distinct complex roots, the PDE is said to be **elliptic**.

Formula 17.3.4 (Canonical form). Consider the general change of variables $\xi = \xi(x, y)$, $\eta = \eta(x, y)$ and $z = \zeta$. With this change, the PDE 17.6 becomes:

$$A(\xi_x, \xi_y)\frac{\partial^2\zeta}{\partial\xi^2} + 2B(\xi_x, \xi_y, \eta_x, \eta_y)\frac{\partial^2\zeta}{\partial\xi\partial\eta} + A(\eta_x, \eta_y)\frac{\partial^2\zeta}{\partial\eta^2} = F(\xi, \eta, \zeta, \zeta_\xi, \zeta_\eta) \tag{17.10}$$

where $A(a, b) = Ra^2 + Sab + Tb^2$ and $B = R\xi_x\eta_x + \frac{1}{2}S(\xi_x\xi_y + \eta_x\eta_y) + Tbd$. Solving the quadratic equation 17.9 will lead to the following three canonical forms:

- **hyperbolic PDE**: With the solutions $\lambda_1(x, y)$ and $\lambda_2(x, y)$ the defining equation can be separated into two ODE's

  $$\left(\frac{dy}{dx} + \lambda_1(x,y)\right)\left(\frac{dy}{dx} + \lambda_2(x,y)\right) = 0$$

  It is clear that the solutions of these ODE's are also roots of the $A(a, b)$ coefficients such that the change of variables $\xi = f_1(x, y)$ and $\eta = f_2(x, y)$ gives the canonical hyperbolic form

  $$\boxed{\frac{\partial^2\zeta}{\partial\xi\partial\eta} = H(\xi, \eta, \zeta, \zeta_\xi, \zeta_\eta)} \tag{17.11}$$

  where $H = \frac{F}{2B}$.

- **parabolic PDE**: As in the hyperbolic case we perform the change of variable $\xi = f(x, y)$, however there is only one root of the defining equation so the second variable can be chosen randomly, yet indepedent of $f_1(x, y)$. From the condition $S^2 + 4RT = 0$ it is alos possible to derive the condition that $B(\xi_x, \xi_y\eta_x\eta_y) = 0$ and $A(\eta_x, \eta_y) \neq 0$. This gives the parabolic canonical form

$$\boxed{\frac{\partial^2\zeta}{\partial\eta^2} = G(\xi, \eta, \zeta, \zeta_\xi, \zeta_\eta)} \tag{17.12}$$

  where $G = \dfrac{F}{A(\eta_x, \eta_y)}$.

- **elliptic PDE**: Again there are two (complex) roots, so the A coefficients will dissappear. Writing $\xi = \alpha + i\beta$ and $\eta = \alpha - i\beta$ gives the following (real) equation

$$\frac{\partial^2\zeta}{\partial\xi\partial\eta} = \frac{1}{4}\left(\frac{\partial^2\zeta}{\partial\alpha^2} + \frac{\partial^2\zeta}{\partial\beta^2}\right)$$

  Substituting this in the hyperbolic case results in the following elliptic canonical form

$$\boxed{\frac{\partial^2\zeta}{\partial\alpha^2} + \frac{\partial^2\zeta}{\partial\beta^2} = K(\alpha, \beta, \zeta, \zeta_\alpha, \zeta_\beta)} \tag{17.13}$$

Theorem 17.3.5 (Maximum principle). *Consider a PDE of the parabolic or elliptic type. The maximum of the solution on a domain is to be found on the boundary of that domain.*

17.3.1 D'Alemberts method

Consider the wave equation

$$\frac{\partial^2 u}{\partial x^2}(x, t) = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2}(x, t) \tag{17.14}$$

By applying the method from previous subsection, it is clear that the characteristics are given by

$$\xi = x + ct \qquad \text{and} \qquad \eta = x - ct \tag{17.15}$$

Furthermore, it follows that the wave equation is a hyperbolic equation which can be rewritten in the canonical form:

$$\frac{\partial^2 u}{\partial\xi\partial\eta}(\xi, \eta) = 0 \tag{17.16}$$

Integration with respect to ξ and η and rewriting the solution in terms of x and t gives

$$u(x, t) = f(x + ct) + g(x - ct) \tag{17.17}$$

where f, g are arbitrary functions. This solution represents a superposition of a left-moving wave and a right-moving wave.

Now consider the wave equation subject to the general conditions

$$u(x, 0) = v(x) \qquad \text{and} \qquad \frac{\partial u}{\partial t}(x, 0) = q(x) \tag{17.18}$$

By applying these conditions to the general solution 17.17 it can be shown that the general solution subject to the given boundary conditions is given by:

$$\boxed{u(x, t) = \frac{1}{2}\left[v(x + ct) + v(x - ct)\right] + \frac{1}{2c}\int_{x-ct}^{x+ct} q(z)dz} \tag{17.19}$$

Remark 17.3.6. Because x is not bounded, this solution is only valid for infinite strings.

17.4 Separation of variables

Remark. We begin this section with the remark that solutions obtained by this method are generalized Fourier series, which tend to converge rather slowly. For numerical purposes, other techniques are recommended. However, the series solutions often give a good insight in the properties of the obtained solutions.

17.4.1 Cartesian coordinates

Method 17.4.1 (Separation of variables). Let $\hat{\mathcal{L}}$ be the operator associated with a partial diferential equation such that $\hat{\mathcal{L}}u(\vec{x}) = 0$ where $\vec{x} = (x_1, ..., x_n)$ is the set of variables. A useful method is to propose a solution of the form

$$u(\vec{x}) = \prod_{i=1}^{n} u_i(x_i)$$

By substituting this form in the PDE and using (basic) algebra it is sometimes (!!) possible to reduce the partial differential equation to a system of n ordinary differential equations.

Example 17.4.2. Consider following PDE:

$$\frac{\partial u}{\partial t} - a\frac{\partial^2 u}{\partial x^2} = 0 \tag{17.20}$$

Substituting a solution of the form $u(x,t) = X(x)T(t)$ gives

$$X(x)\frac{dT(t)}{dt} - aT(t)\frac{d^2X(x)}{dx^2} = 0$$

which can be rewritten as (the arguments are dropped for convenience)

$$\frac{1}{aT}\frac{dT}{dt} = \frac{1}{X}\frac{d^2X}{dx^2}$$

As both sides are independent, it is clear that they are equal to a constant, say λ. This results in the following system of ordinary differential equations:

$$\begin{cases} X''(x) &= \lambda X(x) \\ T'(t) &= a\lambda T(t) \end{cases}$$

17.4.2 Dirichlet problem

The (interior) Dirichlet problem[1] is the problem of finding a solution to a PDE in a finite region, given the value of the function on the boundary of the region. The uniqueness of this solution can be proven with the maximum principle 17.3.5 if the PDE is of the elliptic kind (!!) such as the Laplace equation[2].

Proof. Let ϕ, ψ be two solutions of the interior Dirichlet problem. Due to the linearity both $\psi - \phi$ and $\phi - \psi$ are solutions too (without applying the boundary conditions). According to the maximum principle, these solutions achieve their maximum on the boundary of the domain. Furthermore, due to the Dirichlet boundary conditions, $\phi(x) = \psi(x)$ for all $x \in \partial\Omega$. Combining these two facts gives $\max(\psi - \phi) = \max(\phi - \psi) = 0$ or alternatively $\psi \leq \phi$ and $\phi \leq \psi$ in the complete domain. Which means that $\phi = \psi$ in the complete domain. □

There is also an exterior Dirichlet problem, where one has to find the solution of the PDE, given the boundary conditions, outside of the boundary.

[2]Think of the Dirichlet boundary condition 16.3.

[2]The Dirichlet boundary problem originated with the Laplace equation.

17.5 Non-homogeneous boundary conditions

Formula 17.5.1 (Non-homogeneous boundary condition).

$$\alpha u(a,t) + \beta \frac{\partial u}{\partial x}(a,t) = h(t) \tag{17.21}$$

When $h(t)$ is identically zero, the boundary condition becomes homogeneous.

Method 17.5.2 (Steady-state solution). Assume that the function $h(t)$ is constant. In this case it is useful to rewite the solution as

$$u(x,t) = v(x) + w(x,t)$$

The 'time'-independent function is called the steady-state solution and the function $w(x,t)$ represents the deviation of this steady-state scenario.
As the PDE is linear, we require the partial solutions $v(x)$ and $w(x,t)$ to individually satisfy the equation. Furthermore we require the function $v(x)$ to also satisfy the given non-homogeneous boundary conditions. This results in $w(x,t)$ being the solution of a homogeneous PDE with homogeneous boundary conditions. This can be seen in the following proof:

Proof. Assume a boundary condition of the form $\alpha u(a,t) + \beta \frac{\partial u}{\partial x}(a,t) = u_0$. Due to the requirements, we also have $\alpha v(a) + \beta \frac{\partial v}{\partial x}(a) = u_0$. Combining these two conditions gives

$$\alpha \left[v(a) + w(a,t)\right] + \beta \left[\frac{\partial v}{\partial x}(a) + \frac{\partial w}{\partial x}(a,t)\right] = \alpha v(a) + \beta \frac{\partial v}{\partial x}(a)$$

which can be reduced to

$$\alpha w(a,t) + \beta \frac{\partial w}{\partial x}(a,t) = 0$$

The steady-state deviation $w(x,t)$ thus satisfies homogeneous boundary conditions. □

Method 17.5.3. If the function $h(t)$ is not a constant, we use a different method. Rewrite the solution as $u(x,t) = v(x,t) + w(x,t)$ where we only require $v(x,t)$ to be some function that satisfies the boundary conditions (and not the PDE)[3]. This will lead to $w(x,t)$ satisfying the homogeneous boundary conditions as in the previous method. After substituting the function $v(x,t)$ in the PDE, we obtain a differential equation for $w(x,t)$ but it can be non-homogeneous.

Method 17.5.4. A third, sometimes useful, method is the following. If the problem consists of 3 homogeneous and 1 non-homogeneous boundary condition then the problem can be solved by first applying the homogenous conditions to restrict the values of the separation constant and obtain a series expansion. Afterwards the obtained series can be fitted to the non-homogeneous condition to obtain the final remaining coefficients.

If there is more than 1 non-homogeneous boundary condition, the method can be extended. Let there be j boundary conditions. Rewrite the general solution as $u(x,t) = \sum_{i=1}^{j} v_j(x,t)$ where $v_j(x,t)$ satisfies the j^{th} non-homogeneous condition and the homogeneous versions of the other conditions. This way the general solution still satisfies all conditions and the first part of the method can be applied to all functions $v_j(x,t)$ to obtain a series expansion.

Method 17.5.5 (Non-homogeneous PDE). A possible way to solve non-homogeneous second order partial differential equations of the form

$$\hat{\mathcal{L}} u(x,t) = f(x,t)$$

[3] As there are infinitely many possible functions that satisfy the boundary conditions, the best choice for $v(x,t)$ is the one that makes the equation for $w(x,t)$ as simple as possible.

given a set of homogeneous boundary conditions and inital value conditions $w(x,0) = \psi(x)$, is the following, where we assume all involved functions to be expandable as a generalized Fourier series:

1. Solve the homogeneous version of the PDE, which will result in a series expansion $\sum_n w_n(t)e_n(x)$, where $e_n(x)$ are a complete set of eigenfunctions in the variable x. This solution should satisfy the (homogeneous[4]) boundary conditions.
2. Expand the function $f(x,t)$ in the same way as $u(x,t)$. The coefficients f_n can be found by using the orthogonality realtions of the functions $e_n(x)$.
3. Inserting these expansions in the original PDE and rewriting the equation will lead to a summation of the form:
$$\sum_n \left[\left(\hat{D}w_n(t)\right)e_n(x)\right] = 0$$
where $\hat{D}$ is a linear first order differential operator. As all terms are independent, this gives n first order ODE's to obtain the functions $w_n(t)$. These can be generally solved by using formula 16.8.
4. Initial value conditions for the functions $w_n(t)$ are applied by setting $t = 0$ in the series expansion of $u(x,t)$ and equating it with the series expansion of $\psi(x)$. This results in $w_n() = \Psi_n$.
5. The obtained ODE's together with the found boundary conditions $w_n(0) = \Psi_n$ will give the solutions of $w_n(t)$.
6. Entering these solutions in the series expansion of $u(x,t)$ will give the general solution of the non-homogeneuous PDE.

Remark 17.5.6. It is clear that the requirement that all involved functions are expandable as a generalized Fourier series is restricting. Not all non-homogeneous PDE's are solvable with this method.

17.6 Higher dimensions

17.6.1 Symbols

Definition 17.6.1 (Symbol). Consider a general k^{th}-order differential operator

$$\hat{P} = \sum_{|\alpha|\leq k} c_\alpha(x)D^\alpha \tag{17.22}$$

where we used multi-indices α. The symbol of this operator is defined by replacing the partial derivatives by unknowns $\{\xi^i\}$:

$$p(\hat{P},\xi) = \sum_{|\alpha|\leq k} c_\alpha(x)\xi^\alpha \tag{17.23}$$

Definition 17.6.2 (Principal symbol). The principal symbol of a k^{th}-order differential operator $\hat{P}$ is defined as the highest degree component of $p(\hat{P})$:

$$\sigma_P(\xi) = \sum_{|\alpha|=k} c_\alpha(x)\xi^\alpha \tag{17.24}$$

[4]Non-homogeneous boundary conditions can be turned into homogeneous ones by the previous two methods.

Property 17.6.3. The principal symbol of a differential operator transforms as a tensor.

Property 17.6.4. A PDE

$$\hat{P}f(x) = 0$$

is elliptic if and only if σ_P is invertible.

Chapter 18

Bessel functions

18.1 Bessel's differential equation (BDE)

A Bessel's differential equation is an ordinary differential equation of the following form:

$$\boxed{z^2y'' + zy' + (z^2 - n^2)y = 0} \tag{18.1}$$

The solutions of this ODE are the Bessel functions of the first and second kind (also called respectivelly Bessel and Neumann functions).

$$J_n(z) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!(m+n)!} \left(\frac{z}{2}\right)^{2m+n} \tag{18.2}$$

$$N_n(z) = \lim_{\nu \to n} \frac{cos(\nu\pi)J_N(z) - J_{-n}(z)}{sin(\nu\pi)} \tag{18.3}$$

Remark. Solution 18.2 can be found by using Frobenius' method.

Property 18.1.1. For $n \notin \mathbb{N}$ the solutions $J_n(z)$ and $J_{-n}(z)$ are independent.

Remark 18.1.2. For $n \notin \mathbb{N}$ the limit operation in function 18.3 is not necessary as $\sin(n\pi)$ will never become 0 in this case.

18.2 Generating function

Define the following function:

$$g(x,t) = exp\left[\frac{x}{2}\left(t - \frac{1}{t}\right)\right] \tag{18.4}$$

If we expand this function as a Laurent series, we obtain the following formula:

$$g(x,t) = \sum_{n=-\infty}^{+\infty} J_n(x)t^n \tag{18.5}$$

By applying the residue theorem 13.5.16, we can express the functions $J_n(x)$ as follows:

$$J_n(x) = \frac{1}{2\pi i} \oint_C \frac{g(x,t)}{t^{n+1}} dt \tag{18.6}$$

The function $g(x,t)$ is called the generating function of the Bessel functions.

18.3 Applications

18.3.1 Laplace equation

When solving the Laplace equation in cylindrical coordinates we obtain a BDE with integer n, which has the **cylindrical Bessel functions** 18.2 and 18.3 as solutions.

18.3.2 Helmholtz equation

When solving the Helmholtz equation in spherical coordinates we obtain a variant of the BDE for the radial part:

$$z^2y'' + 2zy' + [z^2 - n(n+1)]y = 0 \tag{18.7}$$

where n is an integer. The solutions, called **spherical Bessel functions**, are related to the cylindrical Bessel functions in the following way:

$$j_n(r) = \sqrt{\frac{\pi}{2x}} J_{n+\frac{1}{2}}(r) \tag{18.8}$$

and similarly for the Neumann functions.

Part V

Linear Algebra

Chapter 19

Linear Algebra

19.1 General

19.2 Vector spaces

In this and coming sections all vector spaces can be both finite- or infinite-dimensional. If necessary, the dimension will be specified.

Definition 19.2.1 (K-vector space). Let K be a field. A K-vector space V is a set equipped with two operations, vector addition $V \times V \to V$ and scalar multiplication $K \times V \to V$, that satisfy the following 8 axioms:

1. V forms an Abelian group under vector addition
2. $a(b\vec{v}) = (ab)\vec{v}$
3. $1_K\vec{v} = \vec{v}$ where 1_K is the identity element of the field K
4. Distributivity of scalar multiplication with respect to vector addition: $a(\vec{v} + \vec{w}) = a\vec{v} + a\vec{w}$

19.2.1 Linear independence

Definition 19.2.2 (Linear combination). The vector w is a linear combination of elements in the set $\{v_n\}$ if it can be written as:

$$w = \sum_n \lambda_n v_n \tag{19.1}$$

for some subset $\{\lambda_n\}$ of the field K.

Definition 19.2.3 (Linear independence). A set finite $\{v_n\}_{n \leq N}$ is said to be linearly independent if the following relation holds:

$$\sum_{n=0}^{N} \lambda_n v_n = 0 \iff \forall n : \lambda_n = 0 \tag{19.2}$$

A general set $\{w_i\}_{i \in I}$ is linearly independent if every finite subset of it is linearly independent.

Definition 19.2.4 (Span). A set of vectors $\{v_n\}$ is said to span V if every vector $v \in V$ can be written as a linear combination of $\{v_n\}$.

Definition 19.2.5 (Frame). A k-frame is an ordered set of k linearly independent vectors.

Definition 19.2.6 (Stiefel manifold). Let V be an inner product space over a field K (real, complex or quaternionic numbers). The set of orthonormal k-frames can be embedded in $K^{n\times k}$. It becomes a compact embedded submanifold, called the Stiefel manifold of k-frames over V.

19.2.2 Bases

Definition 19.2.7 (Basis). A set $\{v_n\}$ is said to be a basis of V if $\{v_n\}$ is linearly independent and if $\{v_n\}$ spans V.

Corollary 19.2.8. Every set T that spans V contains a basis of V.

Remark 19.2.9. In the previous definition we implicitly used the concept of a *Hamel* basis, which is based on two conditions:

- The basis is linearly independent.
- Every element in the vector space can be written as a linear combination of a finite subset of the basis.

Hence for finite-dimensional spaces we do not have to worry. In infinite-dimensional spaces however we have to keep this in mind. An alternative construction, which allows combinations of a countably infinite number of elements is given by the *Schauder basis*.

We now continue by constructing a Hamel basis:

Construction 19.2.10 (Hamel basis). Let V be a vector space over a field K. Consider the set of all linearly independent subsets of V. Under the relation of inclusion this set becomes a partially ordered set[1]. Zorn's lemma 2.3.7 tells us that there exists at least one maximal linearly independent set.

Now we will show that this maximal subset S is also a generating set of V. Let us choose a vector $v \in V$ that is not already in S. From the maximality of S it follows that $S \cup v$ is linearly dependent and hence there exists a finite sequence of numbers $(a^1, ..., a^n, b)$ in K and a finite sequence of elements $(e_1, ..., e_n)$ in S such that:

$$\sum_{i=0}^{n} a^i e_i + bv = 0 \tag{19.3}$$

where not all scalars are zero. This then implies that $b \neq 0$ because else the set $\{e_i\}_{i\leq n}$ and hence S would be linearly dependent. It follows that we can write v as[2]:

$$v = -\frac{1}{b}\sum_{i=0}^{n} a^i e_i \tag{19.4}$$

Because v was randomly chosen we conclude that S is a generating set for V. It is called a Hamel basis of V.

Remark 19.2.11. This construction clearly assumes the ZFC axioms of set theory, only ZF does not suffice. It can even be shown that the existence of a Hamel basis for every vector space[3] is equivalent to the axiom of choice or Zorn's lemma.

[1] See definition 2.3.2.

[2] It is this step that requires R to be a division ring in property 3.2.23 because else we would not generally be able to divide by $b \in R$.

[3] This would turn a vector space into a free object in the category of vector spaces.

Definition 19.2.12 (Dimension). Let V be a finite-dimensional K-vector space. Let $\{v_n\}$ be a basis for V that contains n elements. We then define the dimension of V as following:

$$\boxed{\dim(V) = n} \tag{19.5}$$

Property 19.2.13. Let V be a finite-dimensional K-vector space. Every basis of V has the same number of elements.[4]

19.2.3 Subspaces

Definition 19.2.14 (Subspace). Let V be a K-vector space. A subset W of V is a subspace if W itself is a K-vector space under the operations of V. Alternatively we can write this as:

$$W \leq V \iff \forall w_1, w_2 \in W, \forall \lambda, \mu \in K : \lambda w_1 + \mu w_2 \in W \tag{19.6}$$

Definition 19.2.15 (Grassmannian). Let V be a K-vector space. The set of all subspaces of dimension k is the Grassmannian $\mathrm{Gr}(k, V)$.

Property 19.2.16. $\mathrm{GL}(V)$ acts transitively[5] on all k-dimensional subspaces of V. From property 3.1.65 it follows that the coset space $\mathrm{GL}(V)/H_W$ for any stabilizer H_W of some $W \in \mathrm{Gr}(k, V)$ is isomorphic (as a set) to $\mathrm{Gr}(k, V)$.

Definition 19.2.17 (Flag). Let V be a finite-dimensional vector space. A sequence of proper subspaces $V_1 \leq ... \leq V_n$ is called a flag of V. The sequence $(\dim V_1, ..., \dim V_n)$ is called the **signature** of the flag. If for all i, $\dim V_i = i$ then the flag is called **complete**.

Definition 19.2.18 (Flag variety). The set of all flags of a given signature over a vector space V forms a homogeneous space, called the (generalized) flag variety (of that signature). If the underlying field is the field of real (or complex) numbers then the flag variety is a smooth (or complex) manifold, called the **flag manifold**.

19.2.4 Sum and direct sum

Definition 19.2.19 (Sum). Let V be a K-vector space. Let $W_1, ..., W_k$ be subspaces of V. The sum of the subspaces $W_1, ..., W_k$ is defined as follows:

$$W_1 + ... + W_k := \left\{ \sum_{i=1}^{k} w_i : w_i \in W_i \right\} \tag{19.7}$$

Definition 19.2.20 (Direct sum). If every element v of the sum as defined above can be written as a unique linear combination, then the sum is called a direct sum.

Notation 19.2.21 (Direct sum). The direct sum of vector spaces is in general written in the following way:

$$W_1 \oplus ... \oplus W_k = \bigoplus_{i=1}^{k} W_i$$

Formula 19.2.22. Let V be a finite-dimensional K-vector space. Let W_1, W_2 be two subspaces of V. Then the following relation holds:

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2) \tag{19.8}$$

[4] This theorem can be generalized to infinite-dimensional spaces by stating that all bases have the same *cardinality*.

[5] See definition 3.1.63

Property 19.2.23. Let V be a K-vector space. Let W be decomposed as $W = W_1 \oplus W_2$. If $\mathcal{B}_1$ is a basis of W_1 and if $\mathcal{B}_2$ is a basis of W_2, then $\mathcal{B}_1 \cup \mathcal{B}_2$ is a basis of W.

Definition 19.2.24 (Complement). Let V be a K-vector space. Let W be a subspace of V. A subspace W' of V is called a complement of W if $V = W \oplus W'$.

Property 19.2.25. Let V be a K-vector space. Let U, W be two subspaces of V. If $V = U + W$, then there exists a subspace $Y \leq U$ such that $V = W \oplus Y$. Furthermore every subset W of V has a complement in V.

19.2.5 Algebras

Definition 19.2.26 (Algebra). Let V be a K-vector space. Let V be equipped with the binary operation $\star : V \times V \to V$. $(V, \star)$ is called an algebra over K if it satisfies the following conditions[6]:

1. Right distributivity: $(\vec{x} + \vec{y}) \star \vec{z} = \vec{x} \star \vec{z} + \vec{y} \star \vec{z}$
2. Left distributivity: $\vec{x} \star (\vec{y} + \vec{z}) = \vec{x} \star \vec{y} + \vec{x} \star \vec{z}$
3. Compatibility with scalars: $(a\vec{x}) \star (b\vec{y}) = (ab)(\vec{x} \star \vec{y})$

These conditions say that the binary operation is bilinear.

Definition 19.2.27 (Unital algebra). An algebra V is said to be unital if it contains an identity element with respect to the bilinear map $\star$.

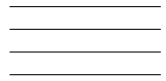
Remark 19.2.28. More generally one can define an algebra over a commutative unital ring R. The defining conditions remain the same except that we require V to be an R-module instead of a K-vector space.

Definition 19.2.29 (Temperley-Lieb algebra). Let R be a commutative unital ring and fix an element $\delta \in R$. The Temperley-Lieb algebra $\mathrm{TL}_n(\delta)$ is the unital R-algebra with generators $\{U_i\}_{i<n}$ that satisfy the **Jones relations**:

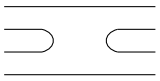
- $U_i^2 = \delta U_i$
- $U_i U_j = U_j U_i$ if $|i - j| \neq 1$
- $U_i U_j U_i = U_i$ if $|i - j| = 1$

One can represent the elements of a Temperley-Lieb algebra diagrammatically. All elements of $\mathrm{TL}_n(\delta)$ are represented as diagrams with n inputs and n outputs.

The unit is given by the diagram where all inputs are connected to the outputs directly across the diagram. The generators $\{U_i\}_{i<n}$ are constructed by connecting the i^{th} input (resp. output) to the $i + 1^{th}$ input (resp. output) and all other inputs are connected to the output directly across the diagram. Multiplication in $\mathrm{TL}_n(\delta)$ is performed diagrammatically by placing two diagrams side by side. Closed loops are replaced by a factor δ.



(a) Unit in $\mathrm{TL}_4(\delta)$.



(b) Generator U_2 in $\mathrm{TL}_4(\delta)$.

[6] These conditions imply that the binary operation is a bilinear map.

19.2.6 Graded vector spaces

Similar to definition 3.2.29 we can define the following:

Definition 19.2.30 (Graded vector space). Let V_n be a vector space for all $n \in \mathbb{N}$. The vector space

$$V = \bigoplus_{n \in \mathbb{N}} V_n \tag{19.9}$$

is called a graded vector space. In fact one can replace $\mathbb{N}$ by any countable (finite or infinite) index set. For most operations however one requires the index set to be closed under addition operations. The index n is often called the **degree** of the subspace V_n in V.

Definition 19.2.31 (Graded algebra). Let V be a graded vector space with the additional structure of an algebra given by the multiplication $\star$. Then V is a graded algebra if $\star$ maps $V^k \times V^l$ to V^{k+l}.

Example 19.2.32 (Superalgebra). A $\mathbb{Z}_2$-graded algebra, i.e. there exists a decomposition

$$A = A_0 \oplus A_1 \tag{19.10}$$

such that for all $i, j \mod 2$:

$$A_i \star A_j \subseteq A_{i+j} \tag{19.11}$$

19.3 Linear maps[7]

Definition 19.3.1 (Injective). A map $f : A \to B$ is called injective if the following condition is satisfied:

$$\forall a, a' \in A : f(a) = f(a') \implies a = a' \tag{19.12}$$

Notation 19.3.2 (Injective map).

$$f : A \hookrightarrow B$$

Definition 19.3.3 (Surjective). A map $f : A \to B$ is called surjective if the following condition is satisfied:

$$\forall b \in B, \exists a \in A : f(a) = b \tag{19.13}$$

Notation 19.3.4 (Surjective map).

$$f : A \twoheadrightarrow B$$

Definition 19.3.5 (Bijective). A map is called bijective if it is both injective and surjective.

Notation 19.3.6 (Bijective map).

$$f : A \xrightarrow{\sim} B \qquad \text{or} \qquad f : A \cong B$$

Notation 19.3.7 (Isomorphic). If two K-vector spaces V, W are isomorphic we denote this by

$$V \cong W$$

Property 19.3.8. Let V be finite-dimensional K-vector space. Let $f : V \to V$ be a linear map. The following statements are equivalent:

[7] Other names are **linear mapping** and **linear transformation**.

- f is injective
- f is surjective
- f is bijective

Definition 19.3.9 (Automorphism). An isomorphism from V to V is called an automorphism[8]. The set of all automorphisms on V, which is in fact a group, is denoted by $\mathrm{Aut}(V)$.

Definition 19.3.10 (General linear group[9]). The set of all automorphisms $f : V \to V$ is called the general linear group and denoted by $\mathrm{GL}_K(V)$ or $\mathrm{GL}(V)$ when the base field is obvious.

Definition 19.3.11 (Rank). The dimension of the image of a linear map is called the rank.

Definition 19.3.12 (Kernel). The kernel of a linear map $f : V \to W$ is the following subset of V:

$$\ker(f) = \{v \in V \mid f(v) = 0\} \tag{19.14}$$

Definition 19.3.13 (Nullity). The dimension of the kernel is called the nullity.

Theorem 19.3.14. *A linear map $f : V \to W$ is injective if and only if $ker(f) = \{0\}$.*

Property 19.3.15. Let $f : V \to W$ be a linear map. Let $U \leq V$. We have the following two properties of the restriction $f|_U$ of f to U:

- $\ker(f|_U) = \ker(f) \cap U$
- $\mathrm{im}(f|_U) \leq \mathrm{im}(f)$

19.3.1 Dimension

Theorem 19.3.16 (Dimension theorem[10]). *Let $f : V \to W$ be a linear map.*

$$\dim(\mathrm{im}(f)) + \dim(\ker(f)) = \dim(\mathrm{V}) \tag{19.15}$$

Property 19.3.17. Two K-vector spaces are isomorphic if and only if they have the same dimension.

19.3.2 Homomorphisms

Definition 19.3.18 (Homomorphism space). Let V, W be two K-vector spaces. The set of all linear maps between V and W is called the homomorphism space from V to W, or shorter the `hom-space` from V to W.

$$\mathrm{Hom}_K(V, W) = \{f : V \to W \mid \text{f is linear}\} \tag{19.16}$$

Formula 19.3.19. If V, W are two finite-dimensional K-vector spaces we have:

$$\dim(\mathrm{Hom}_K(V, W)) = \dim(V) \cdot \dim(W) \tag{19.17}$$

Definition 19.3.20 (Endomorphism ring). The space $\mathrm{Hom}_K(V, V)$ with the composition as multiplication forms a ring, the endomorphism ring. It is denoted by $\mathrm{End}_K(V)$ or $\mathrm{End}(V)$.

[8] In some case also called a **linear operator**, but this terminology is also often used for a general linear map in *operator theory*.

[9] This group is isomorphic to the general linear group of invertible matrices, hence the similar name and notation. (See definition 19.5.7)

[10] Also called the **rank-nullity theorem**.

Property 19.3.21. The endomormphism ring $\text{End}(V)$ forms a Lie algebra[11] when equipped with the commutator $[A, B] = A \circ B - B \circ A$.

Property 19.3.22 (Jordan-Chevalley decomposition). Every endomorphism A can be decomposed as follows:

$$A = A_{ss} + A_n \tag{19.18}$$

where

- A_{ss} is **semisimple**, i.e. for every invariant subspace of A_{ss} there exists an invariant complementary subspace.
- A_n is **nilpotent**, i.e. $\exists k \in \mathbb{N} : A_n^k = 0$.

Furthermore, this decomposition is unique and the endomorphisms A_{ss}, A_n can be written as polynomials in A.

Definition 19.3.23 (Minimal polynomial). Let $f \in \text{End}(V)$ with V a finite-dimensional K-vector space. The monic polynomial $\mu_f(x)$ of lowest order such that $\mu_f(f) = 0$ is called the minimal polynomial of f.

Property 19.3.24. Let $f \in \text{End}(V)$. Let $\mu_f(x)$ be the minimal polynomial of f and let $\varphi(x) \in K[x]$. If $\varphi(f) = 0$, then the minimal polynomial $\mu_f(x)$ divides $\varphi(x)$.

19.3.3 Dual space

Definition 19.3.25 (Dual space). Let V be a K-vector space. The (algebraic) dual V^* of V is the following vector space:

$$V^* := \text{Hom}_K(V, K) = \{f : V \to K \ : \ f \text{ is a linear map}\} \tag{19.19}$$

Definition 19.3.26 (Linear form). The elements of V^* are called linear forms.

Property 19.3.27. From theorem 19.3.19 it follows that $\dim(V^*) = \dim(V)$.

Remark 19.3.28. If V is infinite-dimensional, theorem 19.3.27 is never valid. In the infinite-dimensional case we **always** have $|V^*| > |V|$ (where we now use the cardinality instead of the dimension).

Definition 19.3.29 (Dual basis). Let $\mathcal{B} = \{e_1, e_2, ..., e_n\}$ be a basis for a finite-dimensional K-vector space V. We can define a basis $\mathcal{B}^* = \{\varepsilon_1, \varepsilon_2, ..., \varepsilon_n\}$ for V^*, called the dual basis of $\mathcal{B}$, as follows:

$$\boxed{\varepsilon_i : V \to K : \sum_{j=1}^{n} a_i e_i \mapsto a_i} \tag{19.20}$$

The relation between the basis and dual basis can also be written as:

$$\varepsilon^i(e_j) = \delta^i_j \tag{19.21}$$

Definition 19.3.30 (Dual map). Let $f : V \to W$ be a linear map. The linear map $f^* : W^* \to V^* : \varphi \to \varphi \circ f$ is called the dual map or **transpose** of f.

Notation 19.3.31 (Transpose). When $V = W$ the dual map f^* is often denoted by f^T.

Definition 19.3.32 (Natural pairing). The natural pairing of V and its dual V^* is defined as the following bilinear map:

$$\langle v, v^* \rangle = v^*(v) \tag{19.22}$$

[11] See also 28.2.22.

19.3.4 Convex functions

Definition 19.3.33 (Convex function). Let X be a convex subset of V, i.e. if $x, y \in X$ then $tx = (1-t)y \in X$ for all $t \in [0,1]$. A function $f : X \to \mathbb{R}$ is convex if for all $x, y \in X$ and $t \in [0,1]$:

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) \tag{19.23}$$

Remark 19.3.34. For the definition of a **concave** function we have to turn the inequality around.

Corollary 19.3.35. A linear map $f : X \to \mathbb{R}$ is both convex and concave.

Theorem 19.3.36 (Karamata's inequality). *Let $I \subset \mathbb{R}$ be an interval and let $f : I \to \mathbb{R}$ be a convex function. If $(x_1, ..., x_n)$ is a tuple that majorizes $(y_1, ..., y_n)$, i.e. $\forall k \leq n$*

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \tag{19.24}$$

$$x_{(1)} + ... + x_{(k)} \geq y_{(1)} + ... + y_{(k)} \tag{19.25}$$

where $x_{(i)}$ denotes the ordering[12] *of the tuple $(x_1, ..., x_n)$, then*

$$\sum_{i=1}^{n} f(x_i) \geq \sum_{i=1}^{n} f(y_i) \tag{19.26}$$

19.4 Inner product

In the following section all vector spaces V will be defined over $\mathbb{R}$ or $\mathbb{C}$.

Notation 19.4.1 (Inner product). Let v, w be two vectors in V. The map $\langle \cdot | \cdot \rangle : V \times V \to K$ is called an inner product on V if it satisfies the following properties:

1. **Conjugate symmetry:** $\langle v|w\rangle = \langle w|v\rangle^*$
2. **Linearity in the first argument:** $\langle \lambda u + v|w\rangle = \lambda\langle u|w\rangle + \langle v|w\rangle$
3. **Non-degeneracy:** $\langle v|v\rangle = 0 \iff v = 0$
4. **Positive-definiteness** $\langle v|v\rangle \geq 0$

Remark 19.4.2. Inner products are special cases of **non-degenerate Hermitian forms** which do not satisfy the positive-definiteness property.

Corollary 19.4.3. The first two properties have the result of conjugate linearity in the second argument:

$$\langle f|\lambda g + \mu h\rangle = \overline{\lambda}\langle f|g\rangle + \overline{\mu}\langle f|h\rangle \tag{19.27}$$

19.4.1 Inner product space

Definition 19.4.4 (Inner product space[13]**).** A vector space equipped with an inner product $\langle \cdot | \cdot \rangle$ is called an inner product space.

[12] In decreasing order: $x_{(1)} \geq ... \geq x_{(n)}$.

[13] Sometimes called a **pre-Hilbert space**.

Definition 19.4.5 (Metric dual[14]). Using the inner product (or any other non-degenerate Hermitian form) one can define the metric dual of a vector v by the following map:

$$L : V \to V^* : v \mapsto \langle v|\cdot\rangle \tag{19.28}$$

Definition 19.4.6 (Adjoint operator). Let A be a linear operator on V. The **Hermitian adjoint** of A is defined as the linear operator $A^\dagger$ that satisfies

$$\langle A^\dagger v, w\rangle = \langle v, Aw\rangle \tag{19.29}$$

for all $v, w \in V$. Alternatively one can define the adjoint using the transpose and metric dual as follows:

$$\boxed{A^\dagger = L^{-1} \circ A^T \circ L} \tag{19.30}$$

If $A = A^\dagger$ then A is said to be **Hermitian** or **self-adjoint**.

Corollary 19.4.7. The Hermitian adjoint of a complex matrix $A \in \mathbb{C}^{m\times n}$ is given by:

$$A^\dagger = \overline{A}^T \tag{19.31}$$

where $\overline{A}$ denotes the complex conjugate of A and A^T the transpose of A.

Example 19.4.8. The Lie algebra associated with the group of isometries $\mathrm{Isom}(V)$ of a non-degenerate Hermitian form satisfies following condition:

$$\langle Xv, w\rangle = -\langle v, Xw\rangle \tag{19.32}$$

for all Lie algebra elements X. It follows that the Lie algebra consists of all anti-Hermitian operators.

19.4.2 Orthogonality

Definition 19.4.9 (Orthogonal). Let $v, w \in V$. The vectors v and w are said to be orthogonal, denoted by $v \perp w$, if they obey the following relation:

$$\langle v|w\rangle = 0 \tag{19.33}$$

An orthogonal **system** is a set of vectors, none of them equal to 0, that are mutually orthogonal.

Property 19.4.10. Orthogonal systems are linearly independent.

Definition 19.4.11 (Orthonormal). A set of vectors $\{v_n\}$ is said to be orthonormal if it is orthogonal and if all the elements v_n obey the following relation:

$$\langle v|v\rangle = 1 \tag{19.34}$$

Definition 19.4.12 (Orthogonal complement[15]). Let W be a subspace of V. The following subspace is called the orthogonal complement of W:

$$W^\perp = \{v \in V \mid \forall w \in W : \langle v|w\rangle = 0\} \tag{19.35}$$

Property 19.4.13. Let V be a finite-dimensional K-vector space. The orthogonal complement $W^\perp$ is a complementary subspace[16] to W, i.e. $W \oplus W^\perp = V$.

[14] See also definition 33.1.
[15] $W^\perp$ is pronounced as 'W-perp'.
[16] hence the name

Corollary 19.4.14. Let $W \leq V$ where V is a finite-dimensional K-vector space. We have the following relation:

$$(W^{\perp})^{\perp} = W \tag{19.36}$$

Definition 19.4.15 (Orthogonal projection). Let V be a finite-dimensional K-vector space. Let $W \leq V$. Let $w \in W$ and let $\{w_1, ..., w_k\}$ be an orthonormal basis of W. We define the projection of $v \in V$ on W and $w \in W$ as follows:

$$\text{proj}_W(v) = \sum_{i=1}^{k} \langle v|w_i\rangle w_i \tag{19.37}$$

$$\text{proj}_w(v) = \frac{\langle v|w\rangle}{\langle w|w\rangle} w \tag{19.38}$$

Property 19.4.16.

1. $\forall w \in W : \text{proj}_W(w) = w$
2. $\forall u \in W^{\perp} : \text{proj}_W(u) = 0$

Method 19.4.17 (Gram-Schmidt orthonormalisation). Let $\{u_n\}$ be a set of linearly independent vectors. We can construct an orthonormal set $\{e_n\}$ out of $\{u_n\}$ in the following way:

$$\begin{aligned} w_1 &= u_1 & e_1 &= \frac{w_1}{||w_1||} \\ w_2 &= u_2 - \frac{\langle u_2|w_1\rangle}{||u_2||^2} w_1 & e_2 &= \frac{w_2}{||w_2||} \\ &\vdots & &\vdots \\ w_n &= u_n - \sum_{k=1}^{n-1} \frac{\langle u_n|w_k\rangle}{||u_n||^2} w_k & e_n &= \frac{w_n}{||w_n||} \end{aligned} \tag{19.39}$$

Definition 19.4.18 (Householder transformation). Let v be an element of an inner product space V. The Householder transformation generated by v is given by the linear map

$$\sigma_v : w \mapsto w - 2\frac{\langle w|v\rangle}{\langle v|v\rangle} v \tag{19.40}$$

This transformation amounts to a reflection in the hyperplane orthogonal to v.

19.4.3 Angle

Definition 19.4.19 (Angle). Let v, w be elements of an inner product space V. The angle θ between v and w is defined as:

$$\boxed{\cos\theta = \frac{\langle v|w\rangle}{||v||||w||}} \tag{19.41}$$

19.5 Matrices

Notation 19.5.1. The set of all $m \times n$-matrices defined over the field K is denoted by $M_{m,n}(K)$. If $m = n$, the set is denoted by $M_n(K)$.

Property 19.5.2 (Dimension). The dimension of $M_{m,n}(K)$ is mn.

Definition 19.5.3 (Trace). Let $A = (a_{ij}) \in M_n(K)$. We define the trace of A as follows:

$$\boxed{\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}} \tag{19.42}$$

Property 19.5.4. Let $A, B \in M_n(K)$. We have the following properties of the trace:

1. $\mathrm{tr} : M_n(K) \to K$ is a linear map
2. $\mathrm{tr}(AB) = \mathrm{tr}(BA)$
3. $\mathrm{tr}(AB) \neq \mathrm{tr}(A)\mathrm{tr}(B)$
4. $\mathrm{tr}(A^T) = \mathrm{tr}(A)$

Formula 19.5.5 (Hilbert-Schmidt norm[17]). The Hilbert-Schmidt matrix norm is given by following formula:

$$||A||_{HS}^2 = \sum_{i,j} |A_{ij}|^2 = \mathrm{tr}(A^\dagger A) \tag{19.43}$$

If one identifies $M_n(\mathbb{C})$ with $\mathbb{C}^{2n}$ then this norm equals the standard Hermitian norm.

Formula 19.5.6 (Hadamard product). The Hadamard product of two matrices $A, B \in M_{m\times n}(K)$ is defined as the entry-wise product:

$$(A \circ B)_{ij} = A_{ij}B_{ij} \tag{19.44}$$

Definition 19.5.7 (General linear group). The set of invertible matrices is called the general linear group and is denoted by $\mathrm{GL}_n(K)$.

Property 19.5.8. For all $A \in \mathrm{GL}_n(K)$ we have:

- $A^T \in \mathrm{GL}_n(K)$
- $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$

Property 19.5.9. Let $A \in M_{m,n}(K)$. Denote the set of columns as $\{A_1, A_2, ..., A_n\}$ and the set of rows as $\{R_1, R_2, ..., R_m\}$. The set of columns is a subset of K^m and the set of rows is a subset of K^n. Furthermore we have:

$$\dim(\mathrm{span}(A_1, ..., A_n)) = \dim(\mathrm{span}(R_1, ..., R_m))$$

Definition 19.5.10 (Rank of a matrix). We can define the rank of matrix $A \in M_{m,n}(K)$ as follows:

$$\mathrm{rk}(A) := \dim(\mathrm{span}(A_1, ..., A_n)) \overset{19.5.9}{=} \dim(\mathrm{span}(R_1, ..., R_m)) \tag{19.45}$$

Property 19.5.11. The rank of a matrix has the folowing properties:

1. Let $A \in M_{m,n}(K)$ and $B \in M_{n,r}(K)$. We have $\mathrm{rk}(AB) \leq \mathrm{rk}(A)$ and $\mathrm{rk}(AB) \leq \mathrm{rk}(A)$.
2. Let $A \in \mathrm{GL}_n(K)$ and $B \in M_{n,r}(K)$. We have $\mathrm{rk}(AB) = \mathrm{rk}(B)$.
3. Let $A \in \mathrm{GL}_n(K)$ and $B \in M_{r,n}(K)$. We have $\mathrm{rk}(BA) = \mathrm{rk}(B)$.

Property 19.5.12. Let $A \in M_{m,n}(K)$. First define the following linear map:

$$\boxed{L_A : K^n \to K^m : v \mapsto Av} \tag{19.46}$$

This map has the following properties:

1. $\mathrm{im}(L_A) = \mathrm{span}(A_1, ..., A_n)$
2. $\dim(\mathrm{im}(L_A)) = \mathrm{rk}(A)$

Remark. The second property is a direct consequence of the first one and definition 19.45.

[17] Also called the **Frobenius norm**.

19.5.1 System of equations

Theorem 19.5.13. *Let $AX = w$ with $A \in M_{m,n}(K)$, $w \in K^m$ and $X \in K^n$ be a system of m equations in n variables. Let L_A be the linear map as defined in equation 19.46. We then have the following properties:*

1. *The system is false if and only if $w \notin im(L_A)$.*
2. *If the system is not false, the solution set is an affine space. If $v_0 \in K^n$ is a solution, then the solution set is given by: $L_A^{-1}(w) = v_0 + ker(L_A)$.*
3. *If the system is homogeneous ($AX = 0$), then the solution set is equal to $ker(L_A)$.*

Theorem 19.5.14 (Uniqueness). *Let $AX = w$ with $A \in M_n(K)$ be a system of n equations in n variables. If $rk(A) = n$, then the system has a unique solution.*

19.5.2 Coordinates and matrix representations

Definition 19.5.15 (Coordinate vector). Let $\mathcal{B} = \{b_1, ..., b_n\}$ be a basis of V. Let $v \in V$ such that $v = \sum_{i=1}^{n} \lambda_i b_i$. We define the coordinate vector of v with respect to $\mathcal{B}$ as $(\lambda_1, ..., \lambda_n)^T$. The λ_i's are called the **coordinates** of v with respect to $\mathcal{B}$.

Definition 19.5.16 (Coordinate isomorphism). With the previous definition in mind we can define the coordinate isomorphism of v with respect to $\mathcal{B}$ as follows:

$$\boxed{\beta : V \to K^n : \sum_{i=1}^{n} \lambda_i b_i \mapsto (\lambda_1, ..., \lambda_n)^T} \tag{19.47}$$

Definition 19.5.17 (Matrix representation). Let V be an n-dimensional K-vector space and W an m-dimensional K-vector space. Let $f : V \to W$ be a linear map. Let $\mathcal{B} = \{b_1, ..., b_n\}, \mathcal{C} = \{c_1, ..., c_m\}$ be a basis for V, respectively W. The matrix representation of f with respect to $\mathcal{B}$ and $\mathcal{C}$ can be derived as follows: For every $j \in \{1, ..., n\}$ we can write $f(b_j) = \sum_{i=1}^{m} a_{ij} c_i$, so with this in mind we can define the matrix $(a_{ij}) \in M_{m,n}(K)$ as the matrix represenation of f.

Notation 19.5.18 (Matrix representation of a linear map). The matrix representation of f with respect to $\mathcal{B}$ and $\mathcal{C}$ is denoted by $A_{f,\mathcal{B},\mathcal{C}}$.

Method 19.5.19 (Construction of a matrix representation). From definition 19.5.17 we can see that j-th column of $A_{f,\mathcal{B},\mathcal{C}}$ coincides with the coordinate vector of $f(b_j)$ with respect to $\mathcal{C}$. We use this relation to construct $A_{f,\mathcal{B},\mathcal{C}}$ by writing for every $j \in \{1, ..., n\}$ the coordinate vector of $f(b_j)$ in the j-th column.

Theorem 19.5.20. *Let $(\lambda_1, ..., \lambda_n)^T$ be the coordinate vector of $v \in V$ with respect to $\mathcal{B}$. Let $(\mu_1, ..., \mu_m)^T$ be the coordinate vector of $f(v)$ with respect to $\mathcal{C}$. Then the following relation holds:*

$$\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \end{pmatrix} = A_{f,\mathcal{B},\mathcal{C}} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \tag{19.48}$$

Theorem 19.5.21. *For every matrix $A \in M_{m,n}(K)$ there exists a linear map $f : V \to W$ such that $A_{f,\mathcal{B},\mathcal{C}} = A$.*

On the other hand we also have the following theorem:

Theorem 19.5.22. *Let $f : K^n \to K^m$ be a linear map. There exists a matrix $A \in M_{m,n}(K)$ such that $f = L_A$.*

Theorem 19.5.23. *Let β and γ be the coordinate isomorphisms with respect to respectively $\mathcal{B}$ and $\mathcal{C}$. From theorem 19.5.20 it follows that:*

$$\gamma(f(v)) = A_f \cdot \beta(v) \tag{19.49}$$

or alternatively

$$\gamma \circ f = L_{A_f} \circ \beta \tag{19.50}$$

Theorem 19.5.24. *The map $Hom_K(V,W) \to M_{m,n}(K) : f \mapsto A_f$ is an isomorphism and for every $f \in Hom_K(V,W)$ and $g \in Hom_K(W,U)$ we have:*

$$A_{g\circ f} = A_g A_f \tag{19.51}$$

Theorem 19.5.25. *The map $End_K(V) \to M_n(K) : f \mapsto A_{f,\mathcal{B},\mathcal{B}}$ is an isomorphism and for every $f, g \in End_K(V)$ we have:*

$$A_{g\circ f} = A_g A_f \tag{19.52}$$

Theorem 19.5.26. *Let $f \in End_K(V)$. Let A_f be the corresponding matrix representation. The linear map f is invertible if and only if A_f is invertible. Furthermore, if A_f is invertible, we have that*

$$(A_f)^{-1} = A_{f^{-1}}$$

In other words, the following map is an isomorphism[18]*:*

$$GL_K(V) \to GL_n(K) : f \mapsto A_f \tag{19.53}$$

Remark 19.5.27. The sets $\mathrm{GL}_K(V)$ and $\mathrm{GL}_N(K)$ are groups. So the previous theorem states that the map $f \mapsto A_f$ is a group isomorphism.

Theorem 19.5.28. *Let $V = K^n$. Let $f \in V^*$. From construction 19.5.19 it follows that $A_f = (f(e_1), ..., f(e_n)) \in M_{1,n}(K)$ with respect to the standard basis of V. This combined with theorem 19.5.20 gives:*

$$f(\lambda_1, ..., \lambda_n)^T = (f(e_1), ..., f(e_n))(\lambda_1, ..., \lambda_n)^T = \sum_{i=1}^{n} f(e_i)\lambda_i \tag{19.54}$$

or alternatively with $\{\varepsilon_1, ..., \varepsilon_n\}$ the dual basis to the standard basis of V:

$$\boxed{f = \sum_{i=1}^{n} f(e_i)\varepsilon_i} \tag{19.55}$$

Theorem 19.5.29. *Let $f : V \to W$ be a linear map. Let $f^* : W^* \to V^*$ be the corresponding dual map. If A_f is the matrix representation of f with respect to $\mathcal{B}$ and $\mathcal{C}$, then the transpose A_f^T is the matrix representation of f^* with respect to the dual basis of $\mathcal{C}$ and the dual basis of $\mathcal{B}$.*

[18] Follows from theorem 19.5.25.

19.5.3 Coordinate transforms

Definition 19.5.30 (Transition matrix). Let $\mathcal{B} = \{b_1, ..., b_n\}$ and $\mathcal{B}' = \{b'_1, ..., b'_n\}$ be two bases of V. Every element of $\mathcal{B}'$ can be written as a linear combination of elements in $\mathcal{B}$:

$$b'_j = q_{1j}b_1 + ... + q_{nj}b_n \tag{19.56}$$

The matrix $Q = (q_{ij}) \in M_n(K)$ is called the transition matrix from the 'old' basis $\mathcal{B}$ to the 'new' basis $\mathcal{B}'$.

Theorem 19.5.31. *Let $\mathcal{B}, \mathcal{B}'$ be two basis of V. Let Q be the transition matrix from $\mathcal{B}$ to $\mathcal{B}'$. We find the following statements:*

1. *Let $\mathcal{C}$ be an arbitrary basis of V with γ the corresponding coordinate isomorphism. Define the following matrices:*

   $$B = (\gamma(b_1), ..., \gamma(b_n)) \quad \textit{and} \quad B' = (\gamma(b'_1), ..., \gamma(b'_n))$$

   Then $BQ = B'$.

2. *$Q \in GL_n(K)$ and Q^{-1} is the transition matrix from $\mathcal{B}'$ to $\mathcal{B}$.*

3. *Let $v \in V$ with $(\lambda_1, ..., \lambda_n)^T$ the coordinate vector with respect to $\mathcal{B}$ and $(\lambda'_1, ..., \lambda'_n)^T$ the coordinate vector with respect to $\mathcal{B}'$. Then:*

   $$Q \begin{pmatrix} \lambda'_1 \\ \vdots \\ \lambda'_n \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \quad \textit{and} \quad \begin{pmatrix} \lambda'_1 \\ \vdots \\ \lambda'_n \end{pmatrix} = Q^{-1} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

Theorem 19.5.32. *Let V, W be two finite-dimensional K-vector spaces. Let $\mathcal{B}, \mathcal{B}'$ be two bases of V and $\mathcal{C}, \mathcal{C}'$ two bases of W. Let Q, P be the transition matrices from $\mathcal{B}$ to $\mathcal{B}'$ and from $\mathcal{C}$ to $\mathcal{C}'$ respectively. Let $A = A_{f,\mathcal{B},\mathcal{C}}$ and $A' = A_{f,\mathcal{B}',\mathcal{C}'}$. Then:*

$$A' = P^{-1}AQ \tag{19.57}$$

Corollary 19.5.33. Let $f \in \mathrm{End}_K(V)$ and let Q be the transition matrix. From theorem 19.5.32 it follows that:

$$A' = Q^{-1}AQ \tag{19.58}$$

Definition 19.5.34 (Matrix conjugation). Let $A \in M_n(K)$. The set

$$\{Q^{-1}AQ \mid Q \in \mathrm{GL}_n(K)\} \tag{19.59}$$

is called the conjugacy class[19] of A. Another name for conjugation is **similarity transformation**.

Remark 19.5.35. If A is a matrix representation of a linear operator f, then the conjugacy class of A consists out of every possible matrix representation of f.

Property 19.5.36. From property 19.5.4 it follows that the trace of a matrix is invariant under similarity transformations:

$$\boxed{\mathrm{tr}(Q^{-1}AQ) = \mathrm{tr}(A)} \tag{19.60}$$

[19] This is the general definition of conjugacy classes for groups. Furthermore, these classes induce a partitioning of the group.

Definition 19.5.37 (Matrix congruence). Let $A, B \in M_n(K)$. If there exists a matrix P such that

$$A = P^T BP \tag{19.61}$$

then the matrices are said to be congruent.

Property 19.5.38. Every matrix congruent to a symmetric matrix is also symmetric.

Theorem 19.5.39. *Let $(V, \langle .|. \rangle)$ be an inner-product space defined over $\mathbb{R}$ (or $\mathbb{C}$). Let $\mathcal{B}, \mathcal{B}'$ be two orthonormal bases of V and let Q be the transition matrix. Then Q is orthogonal:*

$$Q^T Q = \mathbb{1}_n \tag{19.62}$$

19.5.4 Determinant

Definition 19.5.40 (Minor). The (i, j)-th minor of A is defined as:

$$\det(A_{ij})$$

where $A_{ij} \in M_{n-1}(K)$ is the matrix obtained by removing the i-th row and the j-th column from A.

Definition 19.5.41 (Cofactor). The cofactor α_{ij} of the matrix element a_{ij} is equal to:

$$(-1)^{i+j} \det(A_{ij})$$

where $\det(A_{ij})$ is the minor as previously defined.

Definition 19.5.42 (Adjugate matrix). The adjugate matrix of $A \in M_n(K)$ is defined as follows:

$$\text{adj}(A) := \begin{pmatrix} \alpha_{11} & \alpha_{21} & \cdots & \alpha_{n1} \\ \alpha_{12} & \alpha_{22} & \cdots & \alpha_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_{1n} & \alpha_{2n} & \cdots & \alpha_{nn} \end{pmatrix} \tag{19.63}$$

or shorter: $\text{adj}(A) = (\alpha_{ij})^T$.

Remark. It is important to notice that we have to transpose the matrix after the elements have been replaced by their cofactor.

Property 19.5.43. Let $A, B \in M_n(K)$. Denote the columns of A as $A_1, \ldots, A_n$. We have the following properties of the determinant:

1. $\det(A^T) = \det(A)$
2. $\det(AB) = \det(BA) = \det(A)\det(B)$
3. $\det(A_1, \ldots, A_i + \lambda A_i', \ldots, A_n) = \det(A_1, \ldots, A_i, \ldots, A_n) + \lambda \det(A_1, \ldots, A_i', \ldots, A_n)$ for all $A_i, A_i' \in M_{n,1}(K)$.
4. If two columns of A are equal then $\det(A) = 0$.
5. $\det(A_{\sigma(1)}, \ldots, A_{\sigma(n)}) = \text{sgn}(\sigma) \det(A_1, \ldots, A_n)$
6. The determinant can be evaluated as follows:

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+k} a_{ik} \det(A_{ik}) \tag{19.64}$$

Theorem 19.5.44. *Let $A \in M_n(K)$, the following statements are equivalent:*

1. $\det(A) \neq 0$
2. $rk(A) = n$
3. $A \in GL_n(K)$

Theorem 19.5.45. *For all $A \in M_n(K)$ we find $Aadj(A) = adj(A)A = \det(A)I_n$.*

Formula 19.5.46. For all $A \in \mathrm{GL}_n(K)$ we find:

$$A^{-1} = \det(A)^{-1} \operatorname{adj}(A) \tag{19.65}$$

An alternative definition of a $k \times k$-minor is:

Definition 19.5.47. Let $A \in M_{m,n}(K)$ and $k \leq \min(m,n)$. A $k \times k$-minor of A is the determinant of a $k \times k$-partial matrix obtained by removing $m-k$ rows and $n-k$ columns from A.

Theorem 19.5.48. *Let $A \in M_{m,n}(K)$ and $k \leq \min(m,n)$. We find that $rk(A) \geq k$ if and only if A contains a non-zero $k \times k$-minor.*

Theorem 19.5.49. *Let $f \in \mathrm{End}_K(V)$. The determinant of the matrix representation of f is invariant under basis transformations.*

Definition 19.5.50 (Determinant of a linear operator). The previous theorem allows us to unambiguously define the determinant of $f \in \mathrm{End}_K(V)$ as follows:

$$\det(f) := \det(A)$$

where A is some matrix representation of f.

19.5.5 Characteristic polynomial

Definition 19.5.51 (Characteristic polynomial[20]). Let V be a finite-dimensional K-vector space. Let $f \in \mathrm{End}_K(V)$ be a linear operator with the matrix representation A (with respect to some arbitrary basis). We then find:

$$\chi_f(x) := \det(x\mathbb{1}_n - A) \in K[x] \tag{19.66}$$

is a monic polynomial of degree n in the variable x and the polynomial does not depend on the choice of basis.

Definition 19.5.52 (Characteristic equation[21]). The following equation is called the characteristic equation of f:

$$\boxed{\chi_f(x) = 0} \tag{19.67}$$

Formula 19.5.53. Let $A = (a_{ij}) \in M_n(K)$ with characteristic polynomial:

$$\chi_A(x) = x^n + c_{n-1}x^{n-1} + \ldots + c_1 x + c_0$$

We then have the following result:

$$\begin{cases} c_0 = (-1)^n \det(A) \\ c_{n-1} = -\mathrm{tr}(A) \end{cases} \tag{19.68}$$

[20] This polynomial can also be used directly for a matrix A as theorem 19.5.21 matches every matrix A with some linear operator f.

[21] This equation is sometimes called the **secular equation**.

Theorem 19.5.54 (Cayley-Hamilton).

1. *Let $A \in M_n(K)$ with characteristic polynomial $\chi_A(x)$. We find the following relation:*

$$\chi_A(A) = A^n + \sum_{i=1}^{n-1} c_i A^i = 0 \tag{19.69}$$

2. *Let $f \in \mathrm{End}_K(V)$ with characteristic polynomial $\chi_f(x)$. We find that*

$$\chi_f(f) = f^n + \sum_{i=1}^{n-1} c_i f^i = 0 \tag{19.70}$$

Corollary 19.5.55. From theorem 19.3.24 and the Cayley-Hamilton theorem it follows that the minimal polynomial $\mu_f(x)$ is a divisor of the characteristic polynomial $\chi_f(x)$.

19.5.6 Linear groups

Definition 19.5.56 (Elementary matrix). An elementary matrix is a matrix of the following form:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & c_{ij} & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \vdots & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & c_{ij} & \ddots & \vdots \\ 0 & 0 & \vdots & 1 \end{pmatrix}, \cdots$$

i.e. equal to the sum of an identity matrix and a multiple of a matrix unit $U_{ij}, i \neq j$.

Notation 19.5.57 (Elementary matrix). $E_{ij}(c)$ is the elementary matrix with element c on the i,j-th position.

Property 19.5.58. We have the following property:

$$\det(E_{ij}(c)) = 1 \tag{19.71}$$

which implies that $E_{ij}(c) \in \mathrm{GL}_n(K)$.

Property 19.5.59. We find the following results concerning the multiplication by an elementary matrix:

1. Left multiplication by an elementary matrix $E_{ij}(c)$ comes down to replacing the i-th row of the matrix with the i-th row plus c times the j-th row.
2. Right multiplication by an elementary matrix $E_{ij}(c)$ comes down to replacing the j-th column of the matrix with the j-th column plus c times the i-th column.

Theorem 19.5.60. *Every matrix $A \in GL_n(K)$ can be written in the following way:*

$$A = SD$$

where S is a product of elementary matrices and $D = diag(1, \ldots, 1, \det(A))$.

Definition 19.5.61 (Special linear group). The following subset of $\mathrm{GL}_n(K)$ is called the special linear group:

$$\mathrm{SL}_n(K) = \{A \in \mathrm{GL}_n(K) \mid \det(A) = 1\} \tag{19.72}$$

Theorem 19.5.62. *Every $A \in SL_n(K)$ can be written as a product of elementary matrices.*[22]

Definition 19.5.63 (Orthogonal group). The orthogonal and special orthogonal group are defined as follows:

$$\begin{aligned} \mathrm{O}_n(K) &= \{A \in \mathrm{GL}_n(K) \mid AA^T = A^T A = I_n\} \\ \mathrm{SO}_n(K) &= \mathrm{O}_n(K) \cap \mathrm{SL}_n(K) \end{aligned}$$

Property 19.5.64. For orthogonal matrices, conjugacy 19.5.34 and congruency 19.5.37 are equivalent.

Definition 19.5.65 (Unitary group). The unitary and special unitary group are defined as follows:

$$\begin{aligned} \mathrm{U}_n(K,\sigma) &= \{A \in \mathrm{GL}_n(K) \mid A\overline{A}^T = \overline{A}^T A = I_n\} \\ \mathrm{SU}_n(K,\sigma) &= \mathrm{U}_n(K) \cap \mathrm{SL}_n(K) \end{aligned}$$

where σ denotes the *involution*[23] $a^\sigma \equiv \overline{a}$.

Remark. If $K = \mathbb{C}$ where the involution is taken to be the complex conjugate, the σ is often ommited in the definition: $\mathrm{U}_n(K)$ and $\mathrm{SU}_n(K)$.

Definition 19.5.66 (Unitary equivalence). Let A, B be two matrices in $\mathrm{M}_n(K)$. If there is a unitary matrix U such that

$$A = U^\dagger B U$$

then the matrices A and B are said to be **unitarily equivalent**.

Definition 19.5.67 (Symplectic group). Consider a vector space V with an antisymmetric nonsingular matrix Ω. The symplectic group $\mathrm{Sp}_n(V,\Omega)$ is defined as follows:

$$\mathrm{Sp}(V,\Omega) = \{A \in \mathrm{GL}(V) \mid A^T \Omega A = \Omega\} \tag{19.73}$$

On the real or complex numbers one can define the canonical **symplectic** matrix

$$\Omega_{st} = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}$$

The group of matrices that preserve this matrix are often denoted by $\mathrm{Sp}_n(\mathbb{R})$ or $\mathrm{Sp}_n(\mathbb{C})$.

Property 19.5.68. Symplectic groups can only be defined on even-dimensional spaces because the defining matrix Ω can only be nonsingular if n is even.

Definition 19.5.69 (Compact symplectic group). The compact symplectic group is defined as follows:

$$\mathrm{Sp}(n) = \mathrm{Sp}_{2n}(\mathbb{C}) \cap \mathrm{U}(2n) \tag{19.74}$$

This is in fact isomorphic to the *quaternionic unitary group* in n quaternionic dimensions.

Example 19.5.70. For $n = 1$ we find $\mathrm{Sp}(1) \cong \mathrm{SU}(2)$.

[22] This follows readily from theorem 19.5.60.

[23] An involution is an operator that is its own inverse: $f(f(x)) = x$.

19.5.7 Matrix decomposition

Method 19.5.71 (QR Decomposition). Every square complex matrix M can be decomposed as:

$$M = QR \tag{19.75}$$

where Q is unitary and R is upper-triangular. The easiest (but not the most numerically stable) way to do this is by applying the Gram-Schmidt orthonormalisation process:

Let $\{v_i\}_{i\leq n}$ be a basis for the column space of M. By applying the Gram-Schmidt process to this basis one obtains a new orthonormal basis $\{e_i\}_{i\leq n}$. The matrix M can then be written as QR where:

- R is an upper-triangular matrix with entries $R_{ij} = \langle e_i | \mathrm{col}_j(M)\rangle$ where $\mathrm{col}_j(M)$ denotes the j^{th} column of M.
- $Q = (a_1 \cdots a_n)$ is the unitary matrix constructed by setting the i^{th} column equal to the i^{th} basis vector a_i

Property 19.5.72. If M is invertible and if the diagonal elements of R are required to have positive norm then the QR-decomposition is unique.

19.6 Eigenvectors

Definition 19.6.1 (Eigenvector). A vector $v \in V \setminus \{0\}$ is called an **eigenvector** of the linear operator $f : V \to V$ if it satisfies the following equation:

$$f(v) = \lambda v \tag{19.76}$$

Where $\lambda \in K$ is the **eigenvalue** belonging to v.

Definition 19.6.2 (Eigenspace). The subspace of V consisting of the zero vector and the eigenvectors of an operator is called the eigenspace associated with that operator. It is given by:

$$\ker(\lambda \mathbf{1}_V - f) \tag{19.77}$$

Theorem 19.6.3 (Characteristic equation[24]). *Let $f \in End_K(V)$ be a linear operator. A scalar $\lambda \in K$ is an eigenvalue of f if and only if it satisfies the characteristic equation 19.67.*

Theorem 19.6.4. *A linear operator $f \in End_K(V)$ defined over an n-dimensional K-vector space V has at most n different eigenvalues.*[25]

Method 19.6.5 (Finding the eigenvectors of a matrix). To calculate the eigenvectors of a matrix one should perform the following steps:

1. First we find the eigenvalues λ_i of $\mathbf{A}$ by applying theorem 19.6.3.
2. Then we find the eigenvector v_i belonging to the eigenvalue λ_i by using the following equation:

$$(\mathbf{A} - \lambda_i \mathbf{1}_V)\, v_i = 0 \tag{19.78}$$

[24] This theorem also holds for the eigenvalues of a matrix $A \in M_n(K)$.
[25] This theorem also holds for a matrix $A \in M_n(K)$.

19.6.1 Diagonalization

Definition 19.6.6 (Diagonalizable operator). Let V be a finite-dimensional K-vector space. An operator $f \in \mathrm{End}_K(V)$ is diagonalizable if there exists a matrix representation $A \in M_n(K)$ of f such that A is a diagonal matrix.

Property 19.6.7. Every diagonalizable operator is semisimple[26]. Conversely, in finite dimensions a semisimple operator, over an algebraically closed field, is diagonalizable.

Theorem 19.6.8. *A linear operator f defined on a finite-dimensional K-vector space V is diagonalizable if and only if the set of eigenvectors of f forms a basis of V.*

Theorem 19.6.9. *A matrix $A \in M_n(K)$ is diagonalizable if and only if there exists a matrix $P \in GL_n(K)$ such that $P^{-1}AP$ is diagonal.*

Corollary 19.6.10. Using the fact that the trace of a linear operator is invariant under similarity transformations (see property 19.60) we get following useful formula:

$$\boxed{\mathrm{tr}(f) = \sum_i \lambda_i} \tag{19.79}$$

where $\{\lambda_i\}_{0 \leq i \leq n}$ are the eigenvalues of f.

Property 19.6.11. Let V be an n-dimensional K-vector space. Let $f \in \mathrm{End}_K(V)$ be a linear operator. We find the following properties of the eigenvectors/eigenvalues of f:

1. The eigenvectors of f belonging to different eigenvalues are linearly independent.
2. If f has exactly n eigenvalues, f is diagonalizable.
3. If f is diagonalizable, V is the direct sum of the eigenspaces of f belonging to the different eigenvalues of f.

19.6.2 Multiplicity

Definition 19.6.12 (Multiplicity). Let V be a K-vector space. Let $f \in \mathrm{End}_K(V)$ be a linear operator with characteristic polynomial[27]:

$$\chi_f(x) = \prod_{i=1}^{n} (x - \lambda_i)^{n_i} \tag{19.80}$$

We can define the following multiplicities:

1. The *algebraic multiplicity* of an eigenvalue λ_i is equal to n_i.
2. The *geometric multiplicity* of an eigenvalue λ_i is equal to the dimension of the eigenspace belonging to that eigenvalue.

Remark 19.6.13. The geometric multiplicity is always at least 1.

Property 19.6.14. The algebraic multiplicity is always greater than or equal to the geometric multiplicity.

[26] See 19.3.22.

[27] We assume that the characteristic polynomial can be written in this form. This depends on the possibility to completely factorize the polynomial in K (i.e. it has 'enough' roots in K). If not, f cannot even be diagonalized. However, there always exists a field F containing K, called a *splitting field*, where the polynomial has 'enough' roots.

Theorem 19.6.15. *Let $f \in End_K(V)$ be a linear operator. f is diagonalizable if and only if for every eigenvalue the algebraic multiplicity is equal to the geometric multiplicity.*

Property 19.6.16. Every Hermitian operator $f \in \mathrm{End}_K(\mathbb{C}^n)$ has the following properties:

1. All the eigenvalues of f are real.
2. Eigenvectors belonging to different eigenvalues are orthogonal.
3. f is diagonalizable and there always exists an orthonormal basis of eigenvectors of f.[28]

Property 19.6.17. Let $A, B \in \mathrm{End}_K(V)$ be two linear operators. If the commutator $[A, B] = 0$, then the two operators have a common eigenbasis.

Theorem 19.6.18 (Sylvester's law of inertia). *Let S be a symmetric matrix. The number of positive and negative eigenvalues is invariant with respect to similarity transformations*[29]*.*

19.7 Euclidean space $\mathbb{R}^n$

A finite-dimensional $\mathbb{R}$-vector space is called a **Euclidean space**.

19.7.1 Angle

Definition 19.7.1 (Angle). Let $(V, \langle .|. \rangle)$ be a real inner-product space. For every $u, v \in V \setminus \{0\}$ we can define the angle between them as[30]:

$$\sphericalangle(u, v) = \mathrm{acos} \frac{\langle u|v \rangle}{||u|| \cdot ||v||} \tag{19.81}$$

where we set the range of acos as $[0, \pi]$.

Notation 19.7.2. When working in a Euclidean space the inner product $\langle v|w \rangle$ is often written as $v \cdot w$ or even vw.

19.7.2 Vector product

Definition 19.7.3 (Orientation). Let $\mathcal{B}, \mathcal{B}'$ be two ordered bases of $\mathbb{R}^n$. Let Q be the transition matrix from $\mathcal{B}$ to $\mathcal{B}'$. If $\det(Q) > 0$ then the bases are said to have the same orientation (or be *consistently oriented*). If $\det(Q) < 0$ then the bases are said to have an opposite orientation.

Corollary 19.7.4 (Positive orientation). The previous definition imposes an equivalence relation on the set of bases of $\mathbb{R}^n$. The set of bases consists out of two equivalence classes. Take one class and call the bases in it *positively* or *directly* oriented. The bases in the other class are then said to be *negatively* or *indirectly* oriented.

Remark 19.7.5. It is convenient to take the standard basis $(e_1, \ldots, e_n)$ to be positively oriented.

[28] This implies that the matrix P diagonalizing the Hermitian operator is unitary, i.e. $P^{-1} = P^\dagger$.

[29] Also with respect to conjugation, which are equivalent to similarity transformations according to property 19.5.64.

[30] This formula follows readily from the Cauchy-Schwarz inequality (see theorem 21.2.6).

Formula 19.7.6 (Cross product).

$$\boxed{(v \times w)_i = \varepsilon_{ijk} v_j w_k} \tag{19.82}$$

where ε_{ijk} is the 3-dimensional Levi-Civita symbol.

Remark 19.7.7. It is important to note that the previous construction is only valid in 3 dimenensions.

Chapter 20

Vector & Tensor Calculus

References for this chapter are [8].

20.1 Nabla-operator

Definition 20.1.1 (Nabla).

$$\nabla \equiv \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \tag{20.1}$$

Following formulas can be found by using basic properties of (vector) calculus.

Formula 20.1.2 (Gradient).

$$\nabla V = \left(\frac{\partial V_x}{\partial x}, \frac{\partial V_y}{\partial y}, \frac{\partial V_z}{\partial z} \right) \tag{20.2}$$

Formula 20.1.3. Let $\varphi(\vec{\boldsymbol{x}})$ be a scalar field. The total differential $d\varphi$ can be rewritten as

$$d\varphi = \nabla\varphi \cdot d\vec{r} \tag{20.3}$$

Property 20.1.4. The gradient of a scalar function V is perpendicular to the level sets 2.14 of V.

Definition 20.1.5 (Directional derivative). Let $\vec{a}$ be a unit vector. The directional derivative $\nabla_{\vec{\boldsymbol{a}}} V$ is defined as the change of the function V in the direction of $\vec{a}$:

$$\nabla_{\vec{\boldsymbol{a}}} V \equiv (\vec{a} \cdot \nabla) V \tag{20.4}$$

Example 20.1.6. Let $\varphi(\vec{\boldsymbol{x}})$ be a scalar field. Let $\vec{t}$ denote the tangent vector to a curve $\vec{r}(s)$ with s natural parameter. The variation of the scalar field $\varphi(\vec{\boldsymbol{x}})$ along $\vec{r}(s)$ is given by

$$\frac{\partial\varphi}{\partial s} = \frac{d\vec{r}}{ds} \cdot \nabla\varphi \tag{20.5}$$

Definition 20.1.7 (Conservative vector field). A vector field obtained as the gradient of a scalar function.

Property 20.1.8. A vector field is conservative if and only if its line integral is path independent.

Formula 20.1.9 (Gradient of tensor). Let T be a tensor field with coordinates x^i. Let $\vec{\boldsymbol{e}}^i(x^1, x^2, x^3)$ be a curvilinear orthogonal frame[1]. The gradient of T is defined as follows:

$$\nabla T = \frac{\partial T}{\partial x^i} \otimes \vec{\boldsymbol{e}}^i \tag{20.6}$$

Formula 20.1.10 (Divergence).

$$\nabla \cdot \vec{\boldsymbol{A}} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \tag{20.7}$$

Definition 20.1.11 (Solenoidal vector field). A vector field $\vec{\boldsymbol{V}}(\vec{\boldsymbol{x}})$ is said to be solenoidal if it satisfies:

$$\nabla \cdot \vec{\boldsymbol{V}} = 0 \tag{20.8}$$

It is also known as a **divergence free vector field**.

Formula 20.1.12 (Rotor / curl).

$$\nabla \times \vec{\boldsymbol{A}} = \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}, \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}, \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \tag{20.9}$$

Definition 20.1.13 (Irrotational vector field). A vector field $\vec{\boldsymbol{V}}(\vec{\boldsymbol{x}})$ is said to be irrotational if it satisfies:

$$\nabla \times \vec{\boldsymbol{V}} = 0 \tag{20.10}$$

Remark 20.1.14. All conservative vector fields are irrotational but irrotational vector fields are only conservative if the domain is simply-connected[2]

20.1.1 Laplacian

Definition 20.1.15 (Laplacian).

$$\triangle V \equiv \nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} \tag{20.11}$$

$$\nabla^2 \vec{\boldsymbol{A}} = \nabla \left(\nabla \cdot \vec{\boldsymbol{A}} \right) - \nabla \times \left(\nabla \times \vec{\boldsymbol{A}} \right) \tag{20.12}$$

Remark 20.1.16. Equation 20.12 is called the **vector laplacian**.

Formula 20.1.17 (Laplacian in different coordinate systems).

- Cylindrical coordinates (ρ, ϕ, z):

$$\frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2}{\partial \phi^2} + \frac{\partial^2}{\partial z^2} \tag{20.13}$$

- Spherical coordinates (r, ϕ, θ):

$$\frac{1}{r^2} \left[\frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) \right] \tag{20.14}$$

[1] See definition 30.2.17.
[2] See definition 9.1.13.

20.1.2 Mixed properties[3]

$$\nabla \times (\nabla V) = 0 \tag{20.15}$$

$$\nabla \cdot \left(\nabla \times \vec{\boldsymbol{V}} \right) = 0 \tag{20.16}$$

In Cartesian coordinates equation 20.12 can be rewritten as follows:

$$\nabla^2 \vec{\boldsymbol{A}} = (\triangle A_x, \triangle A_y, \triangle A_z) \tag{20.17}$$

20.1.3 Helmholtz decomposition

Formula 20.1.18 (Helmholtz decomposition). Let $\vec{\boldsymbol{P}}$ be a vector field that decays rapidly (more than $1/r$) when $r \to \infty$. $\vec{\boldsymbol{P}}$ can be written as follows:

$$\vec{\boldsymbol{P}} = \nabla \times \vec{\boldsymbol{A}} + \nabla V \tag{20.18}$$

20.2 Line integrals

Formula 20.2.1 (Line integral of a continuous function). Let $f : \mathbb{R}^3 \to \mathbb{R}$ be a continuous function. Let Γ be a piecewise smooth curve with parametrization $\vec{\boldsymbol{\varphi}}(t), t \in [a, b]$. We define the line integral of f over Γ as follows:

$$\int_\Gamma f(s)ds = \int_a^b f(\vec{\boldsymbol{\varphi}}(t))||\vec{\boldsymbol{\varphi}}'(t)||dt \tag{20.19}$$

Formula 20.2.2 (Line integral of a continuous vector field). Let $\vec{\boldsymbol{F}}$ be a continuous vector field. Let Γ be a piecewise smooth curve with parametrization $\vec{\boldsymbol{\varphi}}(t), t \in [a, b]$. We define the line integral of F over Γ as follows:

$$\int_\Gamma \vec{\boldsymbol{F}}(\vec{\boldsymbol{s}}) \cdot d\vec{\boldsymbol{s}} = \int_a^b \vec{\boldsymbol{F}}(\vec{\boldsymbol{\varphi}}(t)) \cdot \vec{\boldsymbol{\varphi}}'(t)dt \tag{20.20}$$

20.3 Integral theorems[4]

Theorem 20.3.1 (Fundamental theorem of calculus for line integrals).
Let $\vec{\Gamma} : \mathbb{R} \to \mathbb{R}^3$ be a smooth curve.

$$\int_{\Gamma(a)}^{\Gamma(b)} \nabla f(\vec{\boldsymbol{r}}) \cdot d\vec{\boldsymbol{r}} = \varphi(\Gamma(b)) - \varphi(\Gamma(a)) \tag{20.21}$$

Theorem 20.3.2 (Kelvin-Stokes' theorem).

$$\oint_{\partial S} \vec{\boldsymbol{A}} \cdot d\vec{\boldsymbol{l}} = \iint_S \left(\nabla \times \vec{\boldsymbol{A}} \right) dS \tag{20.22}$$

[3] See remark 30.4.16 for a differential geometric approach.
[4] These theorems follow from the more general Stokes' theorem 32.6.

Theorem 20.3.3 (Divergence theorem[5]).

$$\oiint_{\partial V} \vec{\boldsymbol{A}} \cdot d\vec{\boldsymbol{S}} = \iiint_V \left(\nabla \cdot \vec{\boldsymbol{A}}\right) dV \tag{20.23}$$

Corollary 20.3.4 (Green's identity).

$$\oiint_{\partial V} (\psi \nabla\phi - \phi\nabla\psi) \cdot d\vec{\boldsymbol{S}} = \iiint_V \left(\psi\nabla^2\phi - \phi\nabla^2\psi\right) dV \tag{20.24}$$

20.4 Curvilinear coordinates

In this section the differential operators are generalized to curvilinear coordinates. To do this we need the scale factors as formally defined in equation 26.14. Also there is no Einstein summation used, all summations are written explicitly.

Formula 20.4.1 (Unit vectors).

$$\frac{\partial \vec{r}}{\partial q^i} = h_i \hat{e}_i \tag{20.25}$$

Formula 20.4.2 (Gradient).

$$\nabla V = \sum_{i=1}^{3} \frac{1}{h_i} \frac{\partial V}{\partial q^i} \hat{e}_i \tag{20.26}$$

Formula 20.4.3 (Divergence).

$$\nabla \cdot \vec{\boldsymbol{A}} = \frac{1}{h_1 h_2 h_3} \left(\frac{\partial}{\partial q^1}(A_1 h_2 h_3) + \frac{\partial}{\partial q^2}(A_2 h_3 h_1) + \frac{\partial}{\partial q^3}(A_3 h_1 h_2) \right) \tag{20.27}$$

Formula 20.4.4 (Rotor).

$$(\nabla \times \vec{\boldsymbol{A}})_i = \frac{1}{h_j h_k} \left(\frac{\partial}{\partial q^j}(A_k h_k) - \frac{\partial}{\partial q^k}(A_j h_j) \right) \tag{20.28}$$

where $i \neq j \neq k$.

20.5 Tensor product

20.5.1 Tensor product

There are two possible ways to introduce the components of a tensor (on finite dimensional spaces). One way is to interpret tensors as multilinears maps another way is to interpret the components as expansion coefficients with respect to the tensor space basis.

Definition 20.5.1. The tensor product of vector spaces V and W is defined as[6] the set of multilinear maps on the Cartesian product $V^* \times W^*$. Let v, w be vectors in respectively V and W. Let g, h be vectors in the corresponding dual spaces. The tensor product of v and w is then defined as:

$$\boxed{(v \otimes w)(g, h) = v(g)w(h)} \tag{20.29}$$

[5] Also known as *Gauss's theorem* or the *Gauss-Ostrogradsky theorem*.

[6] *isomorphic to* would be a better terminology. See the "universal property" 20.5.3. For a complete proof and explanation, see [9].

Definition 20.5.2 (Tensor component). One way to define the tensor components is as follows: Let $\mathbf{T}$ be a tensor that takes r vectors and s covectors as input and returns a scalar. The different components are given by $\mathbf{T}(e_i, ..., e_j, e^k, ..., e^l) = T_{i...j}{}^{k...l}$.

The following property can also be seen as the defining property of a tensor product (also in the case of infinite-dimensional spaces):

Universal property 20.5.3. Let Z be a vector space. For every bilinear map $T : V \times W \to Z$ there exists a linear map $f : V \otimes W \to Z$ such that $T = f \circ \varphi$, where φ is the bilinear map $V \times W \to V \otimes W$.

Corollary 20.5.4. The tensor product is unique up to a linear isomorphism. This results in the commutativity of the tensor product:

$$V \otimes W \cong W \otimes V \tag{20.30}$$

where the isomorphism is explicitly given by:

$$v(f) \equiv f(v) \tag{20.31}$$

for all $v \in V$ and $f \in V^*$.

Notation 20.5.5 (Tensor power).

$$V^{\otimes n} = \underbrace{V \otimes ... \otimes V}_{n \text{ copies}} \tag{20.32}$$

Remark 20.5.6. More generally, the tensor product of r copies of V and s copies of V^* is the vector space $\mathcal{T}^r_s(V) = V^{\otimes r} \otimes V^{*\otimes s}$. These tensors are said to be of **type** (r, s).

Remark 20.5.7. Generally the space $\mathcal{T}^1_1 V$ is only isomorphic to the space $\mathrm{End}(V^*)$. The isomorphism is given by the map $\hat{T} : V^* \to V^* : \omega \mapsto \mathbf{T}(\cdot, \omega)$ for every $\mathbf{T} \in \mathcal{T}^1_1 V$. Furthermore the spaces $\mathcal{T}^0_1 V$ and V^* are isomorphic.

For finite-dimensional vector spaces the space $\mathcal{T}^1_1 V$ is also isomorphic to $\mathrm{End}(V)$ (see property 19.3.27) and the space $\mathcal{T}^1_0 V$ will also be isomorphic to V itself.

Definition 20.5.8. The scalars (elements of the base field K) are by definition the $(0, 0)$ tensors.

Alternative Definition 20.5.9. The tensor space $\mathcal{T}^r_s(V)$ is spanned by the elements

$$\underbrace{e_i \otimes ... \otimes e_j}_{r \text{ basis vector}} \otimes \underbrace{\varepsilon^k \otimes ... \otimes \varepsilon^l}_{s \text{ dual basis vectors}}$$

where the operation $\otimes$ satisfies following properties:

1. Associativity: $u \otimes (v \otimes w) = u \otimes v \otimes w$
2. Multilinearity: $a(v \otimes w) = (av) \otimes w = v \otimes (aw)$ and $v \otimes (u + w) = v \otimes u + v \otimes w$

The expansion coefficients in this basis are written as $T^{i...j}{}_{k...l}$

Property 20.5.10 (Dimension of tensor product). From the previous construction it follows that the dimension of $\mathcal{T}^r_s(V)$ is equal to rs.

We now have to proof that the values of the tensor operating on r basis vectors and s basis covectors are equal to the corresponding expansion coefficients:

Proof. Let $\mathbf{T} = T_{i...j}{}^{k...l} e^i \otimes ... \otimes e^j \otimes e_k \otimes ... \otimes e_l$. Applying 20.5.1 and using the definition of the dual vectors 19.21 we have:

$$\begin{aligned} \mathbf{T}(e_a, ..., e_b, \varepsilon^m, ..., \varepsilon^n) &= T_{i...j}{}^{k...l} e^i(e_a)...e^j(e_b)e_k(e^m)...e_l(e^n) \\ &= T_{i...j}{}^{k...l} \delta^i_a ... \delta^j_b \delta^m_k ... \delta^n_l \\ &= T_{a...b}{}^{m...n} \end{aligned}$$

This is exactly the same result as the one we get by applying the first definition. □

Definition 20.5.11 (Tensor algebra). The tensor algebra over a vector space V is defined as follows:

$$T(V) = \bigoplus_{k \geq 0} V^{\otimes k} \tag{20.33}$$

20.5.2 Quotient space

On infinite-dimensional spaces there exists a more general definition (that coincides with the previous one on finite-dimensional spaces[7]):

Construction 20.5.12 (Tensor product). Consider two vector spaces V, W over a field K. First construct the free vector space $F(V \times W)$ over K. Then construct the subspace N of $F(V \times W)$ spanned by the following elements:

- $(v + v', w) - (v, w) - (v', w)$
- $(v, w + w') - (v, w) - (v, w')$
- $(kv, w) - k(v, w)$
- $(v, lw) - l(v, w)$

where $v \in V, w \in W$ and $k, l \in K$. The tensor product $V \otimes W$ is then given by the quotient $F(V \times W)/N$.

20.6 Transformation rules

Let the basis for V transform as $e'_i = A^j{}_i e_j$ and $e_i = B^j{}_i e'_j$. Because the basis transformation should be well-defined, the operators A and B are each other's inverses: $B = A^{-1}$.

Definition 20.6.1 (Contravariant). A tensor component that transforms by the following rule is called contravariant:

$$v^i = A^i{}_j v'^j \tag{20.34}$$

Definition 20.6.2 (Covariant). A tensor component that transforms by the following rule is called covariant:

$$p_i = B^j{}_i \, p'_j \tag{20.35}$$

Example 20.6.3 (Mixed tensor). As an example of a mixed tensor we give the transformation formula for the mixed third-order tensor T^k_{ij}:

$$T^k_{ij} = A^k{}_w B^u{}_i B^v{}_j T'^w{}_{uv}$$

[7] This can be checked using the universal property.

Theorem 20.6.4 (Quotient rule). *Assume we have an equation such as* $K_i A^{jk} = B_i{}^{jk}$ *or* $K_i{}^j A_{jl}{}^k = B_{il}{}^k$ *with* A *and* B *two known tensors*[8]*. The quotient rule asserts the following: "If the equation of interest holds under all transformations, then* K *is a tensor of the indicated rank and covariant/contravariant character".*

Remark. This rule is a useful substitute for the "illegal" division of tensors.

20.7 Tensor operations

20.7.1 General operations

Definition 20.7.1 (Contraction). Let A be a tensor of type (n, m). Setting a sub- and superscript equal and summing over this index gives a new tensor of type $(n-1, m-1)$. This operation is called the contraction of A. It is given by the evaluation map

$$V \otimes V^* : e_i \otimes e^j \mapsto e^j(e_i) \tag{20.36}$$

Definition 20.7.2 (Direct product). Let A and B be two random tensors (both rank and co-/contravariancy). The tensor constructed by the componentwise multiplication of A and B is called the direct product of A and B.

Example 20.7.3. Let $A^i{}_k$ and $B^j{}_{lm}$ be two tensors. The direct product is equal to:

$$C^{i\ j}_{\ k\ lm} = A^i{}_k B^j{}_{lm}$$

Formula 20.7.4 (Operator product). It is also possible to combine operators working on different vector spaces so to make them work on the tensor product space. To do this we use following definition:

$$\boxed{(\hat{A} \otimes \hat{B})(v \otimes w) = (\hat{A}v) \otimes (\hat{B}w)} \tag{20.37}$$

Remark. Consider an operator $\hat{A}$ working on a space V_1. When working with a combined space $V_1 \otimes V_2$ the corresponding operator is in fact $\hat{A} \otimes \mathbb{1}$ but it is often still denoted by $\hat{A}$ in physics.

Notation 20.7.5. Consider a tensor with two indices T_{ij}. The antisymmetric part is written as follows:

$$T_{[ij]} = \frac{1}{2}(T_{ij} - T_{ji}) \tag{20.38}$$

20.7.2 Determinant

Definition 20.7.6 (Form). An n-form is a totally antisymmetric element $\omega \in \mathcal{T}^0_n V$.

Definition 20.7.7 (Volume form). A form of rank $\dim V$ is also called a **top form** or **volume form**.

Definition 20.7.8 (Determinant). Let V be finite-dimensional with basis $\{e_i\}_{i \leq n}$. Let φ be a tensor in $\mathcal{T}^1_1 V \cong \mathrm{End}(V)$ and let ω be a volume form. The determinant of φ is then defined as:

$$\det \varphi = \frac{\omega(\varphi(e_1), ..., \varphi(e_n))}{\omega(e_1, ..., e_n)} \tag{20.39}$$

This definition is well-defined, i.e. it is independent of the choice of volume form and basis. Furthermore it coincides with definition 19.5.50.

[8] This rule does not necessarily hold when $B = 0$ as transformations rules are not defined for the null-tensor.

One should note that the determinant is only well-defined for $(1,1)$-tensors. Although other types of tensors can also be represented as matrices, definition 19.5.50 would not be independent of a choice of basis anymore. An alternative concept can be defined using principal bundles and more precisely frame bundles (see section 31).

Corollary 20.7.9.

$$\omega(e_1, ..., e_{i-1}, X, e_{i+1}, ..., e_n) = X_i \tag{20.40}$$

20.7.3 Differentiation

Property 20.7.10.

$$\nabla \cdot (\vec{\boldsymbol{A}} \otimes \vec{\boldsymbol{B}}) = (\nabla \cdot \vec{\boldsymbol{A}})\vec{\boldsymbol{B}} + (\vec{\boldsymbol{A}} \cdot \nabla)\vec{\boldsymbol{B}} \tag{20.41}$$

20.7.4 Levi-Civita tensor

Definition 20.7.11 (Levi-Civita tensor). Let e^i be the dual vector to e_i. In n dimensions, we define the Levi-Civita tensor as follows:

$$\boldsymbol{\varepsilon} = \varepsilon_{12...n} e^1 \otimes e^2 \otimes ... \otimes e^n \tag{20.42}$$

where

$$\varepsilon_{i...n} = \begin{cases} 1 & \text{if } (i...n) \text{ is an even permutation of } (12...n) \\ -1 & \text{if } (i...n) \text{ is an odd permutation of } (12...n) \\ 0 & \text{if any of the indices occurs more than once} \end{cases}$$

Remark 20.7.12. The Levi-Civita symbol is not a tensor, but a pseudotensor. This means that the sign changes under reflections (or any transformation with determinant -1). To turn it into a proper tensor one should multiply it by a factor $\sqrt{g}$ where g is the determinant of the metric.

Formula 20.7.13 (Cross product). By using the Levi-Civita symbol, we can define the i-th component of the cross product[9] as follows:

$$\boxed{(\vec{\boldsymbol{v}} \times \vec{\boldsymbol{w}})_i = \varepsilon_{ijk} v_j w_k} \tag{20.43}$$

20.7.5 Complexification

Definition 20.7.14 (Complexification). Let V be a real vector space. The complexification of V is defined as the following tensor product:

$$V^{\mathbb{C}} = V \otimes \mathbb{C} \tag{20.44}$$

As such this is still a real vector space. However we can turn this space into a complex vector space by generalizing the scalar product as follows:

$$\alpha(v \otimes \beta) = v \otimes (\alpha\beta) \tag{20.45}$$

for all $\alpha, \beta \in \mathbb{C}$.

Property 20.7.15. By noting that every element $v_{\mathbb{C}} \in V^{\mathbb{C}}$ can be written as

$$v_{\mathbb{C}} = (v_1 \otimes 1) + (v_2 \otimes i)$$

we can decompose the complexification as follows:

$$V^{\mathbb{C}} \cong V \oplus iV \tag{20.46}$$

[9] Following from remark 20.7.12 we can see that the cross product is in fact not a vector, but a pseudovector.

20.8 (Anti)symmetric tensors

20.8.1 Symmetric tensors

Notation 20.8.1. The space of symmetric $(n, 0)$ tensors is denoted by $S^n(V)$. The space of symmetric $(0, n)$ tensors is denoted by $S^n(V^*)$.

20.8.2 Antisymmetric tensors

Definition 20.8.2 (Antisymmetric tensor). Tensors that change sign under the interchange of any two indices.

Notation 20.8.3. The space of antisymmetric $(0, n)$ tensors is denoted by $\Lambda^n(V^*)$. The space of antisymmetric $(n, 0)$ tensors is denoted by $\Lambda^n(V)$.

Remark. Elements of $\Lambda^2(V)$ are also known as **bivectors**. Elements of $\Lambda^k(V)$ are generally known as k-**blades**.

Property 20.8.4. Let $n = \dim(V)$. $\Lambda^r(V)$ equals the null-space for all $r \geq n$.

20.8.3 Wedge product

Formula 20.8.5 (Antisymmetrization). Let $\{P_i\}_i$ be the set of all permutations of the sequence $(1, ..., k)$.

$$\mathrm{Alt}(e_1 \otimes ... \otimes e_k) = \sum_i \mathrm{sgn}(P_i) e_{P_i(1)} \otimes ... \otimes e_{P_i(k)} \tag{20.47}$$

Definition 20.8.6 (Wedge product). Let $\{e_i\}_{1 \leq i \leq \dim(V)}$ be a basis for V.

$$e_1 \wedge ... \wedge e_k = \mathrm{Alt}(e_1 \otimes ... \otimes e_k) \tag{20.48}$$

From this definition it immediately follows that the wedge product is (totally) antisymmetric.

Construction 20.8.7. Let $\{e_i\}_{1 \leq i \leq \dim(V)}$ be a basis for V. It is clear from the definition 20.48 that a basis for $\Lambda^r(V)$ is given by

$$\{e_{i_1} \wedge ... \wedge e_{i_r} \ : \ \forall k : 1 \leq i_k \leq \dim(V)\}$$

The dimension of this space is given by:

$$\dim \Lambda^r(V) = \binom{n}{r} \tag{20.49}$$

From the antisymmetry it follows that for $r > \dim(V)$ the spaces $\Lambda^r(V)$ are zero.

Remark 20.8.8. For $k = 0$, the above construction is not useful, so we just define $\Lambda^0(V) = \mathbb{R}$.

Formula 20.8.9. Let $v \in \Lambda^r(V)$ and $w \in \Lambda^m(V)$.

$$v \wedge w = \frac{1}{r!m!} \mathrm{Alt}(v \otimes w) \tag{20.50}$$

where the antisymmetrization operator Alt is defined in equation 20.47.

Formula 20.8.10 (Levi-Civita symbol). The Levi-Civita tensor in n dimensions as introduced in 20.42 can now be rewritten more concisely as:

$$\boldsymbol{\varepsilon} = e_1 \wedge ... \wedge e_n \tag{20.51}$$

Formula 20.8.11. In 3 dimensions there exists an important isomorphism $J : \Lambda^2(\mathbb{R}^3) \to \mathbb{R}^3$:

$$J(\lambda)^i = \frac{1}{2}\varepsilon^i{}_{jk}\lambda^{jk} \tag{20.52}$$

where $\lambda \in \Lambda^2(\mathbb{R}^3)$.

Looking at the definition of the cross product 19.82, we can see that $\vec{\boldsymbol{v}} \times \vec{\boldsymbol{w}}$ is actually the same as $J(\vec{\boldsymbol{v}} \wedge \vec{\boldsymbol{w}})$. One can thus use the wedge product to generalize the cross product to higher dimensions.

Example 20.8.12. Let A, B and C be three vectors in V. Now consider following expression:

$$(C \wedge B)(L(A), \cdot)$$

where $L(A)$ is the metric dual of A (see 19.28). Evaluating this formula using the properties of the wedge and tensor products leads to the well known BAC-CAB rule of triple cross products:

$$(C \cdot A)B - (B \cdot A)C$$

20.8.4 Exterior algebra

Definition 20.8.13 (Exterior power). In the theory of exterior algebras, the space $\Lambda^k(V)$ is often called the k^{th} exterior power of V. Its elements are called (exterior) k**-forms**.

Definition 20.8.14 (Exterior algebra). We can define a graded vector space[10] $\Lambda^*(V)$ as follows:

$$\Lambda^*(V) = \bigoplus_{k \geq 0} \Lambda^k(V)$$

Then we can turn this graded vector space into a graded algebra by taking the wedge product as the multiplication:

$$\wedge : \Lambda^k(V) \times \Lambda^l(V) \to \Lambda^{k+l}(V)$$

This algebra is called the exterior algebra or **Grassmann algebra** over V.

Alternative Definition 20.8.15 (†). Let $T(V)$ be the tensor algebra over the vector space V, i.e.

$$T(V) = \bigoplus_{k \geq 0} V^{\otimes k} \tag{20.53}$$

The exterior algebra over V is generally defined as the quotient of $T(V)$ by the two-sided ideal I generated by $\{v \otimes v : v \in V\}$.

Property 20.8.16. The exterior algebra is both a unital associative algebra with unit element $1 \in \mathbb{R}$ and a coalgebra. Furthermore it is also commutative in the graded sense (see 3.33).

Property 20.8.17. The graded commutativity implies that the wedge product of any odd exterior form with itself is identically 0. The wedge product of an even exterior form with itself vanishes if and only if the form can be decomposed as a product of 1-forms.

[10] See definition 19.9.

20.8.5 Hodge star

It follows from equation 20.49 that the spaces $\Lambda^k(V)$ and $\Lambda^{n-k}(V)$ have the same dimension, so there exists an isomorphism between them. This map is given by the Hodge star $*$. However this map can only be defined independent of the choice of (ordered) basis if we restrict ourselves to vector spaces equipped with a non-degenerate Hermitian form 19.4.2.

Definition 20.8.18 (Volume element). Let V be an n-dimensional vector space. A volume element on V is a distinct choice of top-dimensional exterior form on V, i.e. a distinct choice of $\text{Vol} \in \Lambda^n(V)$.

When equipped with an inner product and hence an orthonormal basis $\{e_i\}$, there exists a canonical volume form given by $\text{Vol} = e_1 \wedge ... \wedge e_n$. This will be the convention adopted in the remainder of this section.

Definition 20.8.19 (Orientation). Let Vol be the choice of volume form on the vector space V. From the definition of a volume form it follows that every other $\dim(V)$-blade is a scalar multiple of Vol. Hence the choice of volume form induces an orientation on V: if the scalar $r > 0$ then the orientation is said to be **positive**, if $r < 0$ then the orientation is **negative**.

Formula 20.8.20 (Inner product). Let V be equipped with an inner product $\langle \cdot, \cdot \rangle$. Then we can define an inner product on $\Lambda^k(V)$ by:

$$\boxed{\langle v_1 \wedge ... \wedge v_k | w_1 \wedge ... \wedge w_k \rangle_k = \det(\langle v_i, w_j \rangle)} \tag{20.54}$$

For an orthogonal basis, this formula factorises into:

$$\langle v_1 \wedge ... \wedge v_k | w_1 \wedge ... \wedge w_k \rangle_k = \langle v_1 | w_1 \rangle \cdots \langle v_k | w_k \rangle \tag{20.55}$$

Definition 20.8.21 (Hodge star). The Hodge star $* : \Lambda^k(V) \to \Lambda^{n-k}(V)$ is defined as the isomorphism such that for all $\omega \in \Lambda^k(V)$ and $\rho \in \Lambda^{n-k}(V)$ we have the following equality:

$$\omega \wedge \rho = \langle *\omega, \rho \rangle_{n-k} \text{Vol}(V) \tag{20.56}$$

where $\langle \cdot, \cdot \rangle$ is the inner product 20.54 on $\Lambda^{n-k}(V)$. Furthermore, this isomorphism is unique.

Proof. Because $\omega \wedge \rho$ is an element of $\Lambda^n(V)$ it is a scalar multiple of $\text{Vol}(V)$. This implies that it can be written as

$$c(\rho)\text{Vol}(V)$$

The map $c : \Lambda^{n-k}(V) \to \mathbb{R} : \rho \mapsto c(\rho)$ is a linear map and thus a continuous map, so we can apply Riesz' representation theorem to identify c with a unique element $*\omega \in \Lambda^{n-k}(V)$ such that

$$c(\rho) = \langle *\omega, \rho \rangle_{n-k}$$

□

Formula 20.8.22. Let $\{e_i\}_{i \leq n}$ be a positively oriented ordered orthonormal (possibly in a Lorentzian signature) basis for V. An explicit formula for the Hodge star is given by the following construction. Let $\{i_1, ..., i_k\}$ and $\{j_1, ..., j_{n-k}\}$ be two complementary index sets with increasing subindices. Let $\omega = e_{i_1} \wedge ... \wedge e_{i_k}$.

$$\boxed{*\omega = \text{sgn}(\tau) \prod_{m=1}^{n-k} \langle e_{j_m} | e_{j_m} \rangle e_{j_1} \wedge ... \wedge e_{j_{n-k}}} \tag{20.57}$$

where τ is the permutation that maps $e_{i_1} \wedge ... \wedge e_{i_k} \wedge e_{j_1} \wedge ... \wedge e_{j_{n-k}}$ to $\text{Vol}(V)$.

Corollary 20.8.23. Consider three vectors $u, v, w \in \mathbb{R}^3$.

$$*(v \wedge w) = v \times w \tag{20.58}$$
$$*(v \times w) = v \wedge w \tag{20.59}$$
$$*(u \wedge v \wedge w) = u \cdot (v \times w) \tag{20.60}$$

Remark 20.8.24. Formula 20.52 is an explicit evaluation of the first equation 20.58.

Proof. The sign $\text{sgn}(\tau)$ can be written using the Levi-Civita symbol ε_{ijk} as defined in 20.42. The factor $\frac{1}{2}$ is introduced to correct for the double counting due to the contraction over both the indices j and k.

Property 20.8.25. Consider an inner product space V, then

$$\boxed{** \ \omega = (-1)^{k(n-k)}\omega} \tag{20.61}$$

In $n = 4$ this leads to $**\omega = \omega$ which means that the Hodge star is an involution in 4-dimensional inner product spaces.

Definition 20.8.26 (Self-dual). Let V be a 4-dimensional inner product space. Consider $\omega \in \Lambda^2(V)$. Then ω is said to be self-dual if $*\omega = \omega$. Furthermore every $v \in \Lambda^2(V)$ can be uniquely decomposed as the sum of a self-dual and an anti-self-dual 2-form.

20.8.6 Grassmann numbers

Although this section does not really belong to the chapter about tensors, we have included it here as it is an application of the concept of exterior algebras. The concept of Grassmann numbers (or variables) is used in QFT when performing calculations in the fermionic sector.

Definition 20.8.27 (Grassmann numbers). Let V be a complex vector space spanned by a set of generators θ_i. The Grassmann algebra with Grassmann variables θ_i is the exterior algebra over V. The wedge symbol of Grassmann variables is often ommitted when writing the product: $\theta_i \wedge \theta_j \equiv \theta_i\theta_j$.

Remark 20.8.28. Furthermore, from the anti-commutativity it follows that we can regard the Grassmann variables as being non-zero square-roots of zero.

Property 20.8.29. Consider a one-dimensional Grassmann algebra. When constructing the polynomial ring $\mathbb{C}[\theta]$ generated by θ, we see that, due to the anti-commutativity, $\mathbb{C}[\theta]$ is spanned only by 1 and θ. All higher degree terms vanish because $\theta^2 = 0$. This implies that the most general polynomial over a one-dimenisonal Grassmann algebra can be written as

$$p(\theta) = a + b\theta \tag{20.62}$$

Definition 20.8.30. We can equip the exterior algebra Λ with Grassmann variables θ_i with an involution similar to that on $\mathbb{C}$:

$$(\theta_i\theta_j...\theta_k)^* = \theta_k...\theta_j\theta_i \tag{20.63}$$

Elements $z \in \Lambda$ such that $z^* = z$ are called **(super)real**, elements such that $z^* = -z$ are called **(super)imaginary**. This convention is called the *DeWitt* convention.

Chapter 21

Normed Spaces

In this chapter the term "linear operator", which was previously reserved for vector space automorphisms, is now used instead of "linear map". This was done to keep the vocabulary in track with that of the standard literature on Banach spaces and operator spaces.

For a revision of inner product spaces see section 19.4.

21.1 Banach spaces

Definition 21.1.1 (Topological vector space). A topological vector space (TVS) over a base field K is a vector space for which the addition and scalar multiplication over K are continuous.

Definition 21.1.2 (Weak topology). The initial topology on a TVS with respect to its dual.

Definition 21.1.3 (Norm). Let V be a TVS over a field K. A function $||\vec{v}|| : V \to [0, +\infty[$ is called a norm if it satisfies following conditions:

- **Non-degeneracy:** $||\vec{v}|| = 0 \iff \vec{v} = 0$
- **Homogeneity:** $||a\vec{v}|| = |a|||\vec{v}||$ for all scalars $a \in K$
- **Triangle equality (subadditivity):** $||\vec{v} + \vec{w}|| \leq ||\vec{v}|| + ||\vec{w}||$

Remark 21.1.4. A norm $|| \cdot ||$ clearly induces a metric[1] by setting $d(x, y) = ||x - y||$.

Definition 21.1.5 (Normed vector space). A TVS equipped with a norm $|| \cdot ||$.

Definition 21.1.6 (Banach space). A normed vector space that is complete[2] in the norm-topology.

Property 21.1.7. The topological (continuous) dual of a Banach space is also a Banach space.

Definition 21.1.8 (Reflexive space). A Banach space V that coincides with its double (topological) dual, i.e. $V = (V^*)^*$.

Property 21.1.9. Every finite-dimensional Banach spaces is reflexive. This follows from property 19.3.27.

[1] See definition 8.1.1.
[2] See condition 8.12.

Property 21.1.10. Let (x_n) be a Cauchy sequence in a normed space V. Then $(||x_n||)$ is a convergent sequence in $\mathbb{R}$. This implies that every Cauchy sequence in a normed space is bounded.

Property 21.1.11. Let X be a TVS. Every linear map $\varphi : \mathbb{K}^n \to X$ is continuous.

Property 21.1.12. Let X be a finite-dimensional normed vector space. Every linear bijection $\varphi : \mathbb{K}^n \to X$ is a homeomorphism.

Corollary 21.1.13. Two finite-dimensional normed vector spaces with the same dimension are homeomorphic. It follows that all metrics on a finite-dimensional normed vector space are equivalent.

Theorem 21.1.14 (Open mapping theorem[3]). *Let $f : V \to W$ be a continuous linear operator between two Banach spaces. If f is surjective then it is also open.*

Theorem 21.1.15 (Hahn-Banach theorem). *Let V be a Banach space. Let $f : V \to \mathbb{R}$ be a sublinear map, i.e. a map that is both subadditive and positive-homogeneous, and let $\phi : U \to \mathbb{R}$ be a linear map dominated by f, defined on a linear subspace $U \subset V$. Then there exists a linear extension $\psi : V \to \mathbb{R}$ of ϕ that is dominated by f on all of V.*

21.2 Hilbert space

Definition 21.2.1 (Hilbert space). A vector space that is both a Banach space and an inner product space (where the norm is induced by the inner product).

Example 21.2.2. Let $f, g \in \mathcal{L}^2([a,b], \mathbb{C})$, the inner product of f and g is defined as:

$$\boxed{\langle f|g\rangle = \int_a^b f^*(x)\overline{g(x)}dx} \tag{21.1}$$

Remark 21.2.3. See section 14.4.2 for a more formal treatment of this subject.

Formula 21.2.4. It is also possible to define an inner product with respect to a weight function $\phi(x)$:

$$\int_a^b f^*(x)g(x)\phi(x)dx \tag{21.2}$$

Using this formula it is possible to define orthogonality with respect to a weight function.

21.2.1 Inner products and norms

Formula 21.2.5. Let V be an inner product space. A norm on V can be induced by the inner product in the following way:

$$||v||^2 = \langle v|v\rangle \tag{21.3}$$

However not every norm induces an inner product. Only norms that satisfy the parallellogram law 21.5 induce an inner product. This inner product can be recovered through the polarization identity 21.6 (see below).

Property 21.2.6 (Cauchy-Schwarz inequality).

$$\boxed{|\langle v|w\rangle| \leq ||v||\ ||w||} \tag{21.4}$$

where the equality holds if and only if v and w are linearly dependent.

[3]Sometimes called the *Banach-Schauder* theorem.

Corollary 21.2.7. The Cauchy-Schwarz inequality can be used to prove the triangle inequality. Together with the properties of an inner product this implies that an inner product space is also a normed space.

Formula 21.2.8 (Parallellogram law).

$$||v+w||^2 + ||v-w||^2 = 2(||v||^2 + ||w||^2) \tag{21.5}$$

Formula 21.2.9 (Polarization identity).

$$4\langle v|w\rangle = ||v+w||^2 - ||v-w||^2 + i\left(||v+iw||^2 - ||v-iw||^2\right) \tag{21.6}$$

Formula 21.2.10 (Pythagorean theorem). In an inner product space the triangle equality reduces to the well-known Pythagorean theorem for orthogonal vectors v, w:

$$||v+w||^2 = ||v||^2 + ||w||^2 \tag{21.7}$$

This formula can be extended to any set of orthogonal vectors $x_1, ..., x_n$:

$$\boxed{\left\|\sum_{i=1}^{n} x_i\right\|^2 = \sum_{i=1}^{n} ||x_i||^2} \tag{21.8}$$

21.2.2 Generalized Fourier series

Property 21.2.11 (Bessel's inequality). First of all we have following general equality for orthonormal vectors $x_1, ..., x_n$ and complex scalars $a_1, ..., a_n$:

$$\left\|x - \sum_{i=1}^{n} a_i x_i\right\|^2 = ||x||^2 - \sum_{i=1}^{n} |\langle x, x_i\rangle|^2 + \sum_{i=1}^{n} |\langle x, x_i\rangle - a_i|^2 \tag{21.9}$$

This expression becomes minimal for $a_i = \langle x, x_i\rangle$ (last term becomes 0). This leads to Bessel's inequality:

$$\boxed{\sum_{i=1}^{n} |\langle x, x_i\rangle|^2 \leq ||x||^2} \tag{21.10}$$

Corollary 21.2.12. The sum in 21.10 is bounded for all n, so the series $\sum_{i=1}^{+\infty}$ converges for all x. This implies that the sequences $(\langle x, x_n\rangle)$ belongs to the space l^2 of square-summable sequences.

This result does however not imply that the generalized Fourier series $\sum_{i=1}^{+\infty}\langle x, x_i\rangle x_i$ converges to x. The following theorem gives a necessary and sufficient condition for the convergence.

Theorem 21.2.13. *Let $\mathcal{H}$ be a Hilbert space. Let (x_n) be an orthonormal sequence in $\mathcal{H}$ and let (a_n) be a sequence in $\mathbb{C}$. The expansion $\sum_{i=1}^{+\infty} a_i x_i$ converges in $\mathcal{H}$ if and only if $(a_n) \in l^2$. Furthermore the expansion satisfies following equality:*

$$\left\|\sum_{i=1}^{+\infty} a_i x_i\right\|^2 = \sum_{i=1}^{+\infty} |a_i|^2 \tag{21.11}$$

As we noted the sequence $(\langle x, x_n\rangle)$ belongs to l^2 so the generalized Fourier series converges of $x \in \mathcal{H}$ converges in $\mathcal{H}$.

Remark 21.2.14. Although the convergence of the generalized Fourier series of $x \in \mathcal{H}$ can be established using previous theorem, it does not follow that the expansion converges to x itself. We can merely say that the Fourier expansion is the best approximation of x with respect to the norm on $\mathcal{H}$.

21.2.3 Complete sets

Definition 21.2.15 (Complete set). Let $\{e_i\}_{i\in I}$ be a set (possibly a sequence) of orthonormal vectors in an inner product space V. This set is said to be complete if every vector $x \in V$ can be expressed as follows:

$$x = \sum_{i\in I} \langle x, x_i\rangle x_i \tag{21.12}$$

This implies that a complete set is a basis for the vector space.

Another characterization is the following.

Alternative Definition 21.2.16. A complete set of orthonormal vectors is a set $S \subset V$ such that we cannot add another vector w to it satisfying:

$$\forall v_i \in S : \langle v_i, w\rangle = 0 \qquad \wedge \qquad w \neq 0 \tag{21.13}$$

Property 21.2.17. For complete sequences (x_n) the inequality of Bessel 21.10 becomes an equality. Furthermore, the generalized Fourier series with respect to the complete sequence is unique.

Using previous property we can prove the following theorem due to Parceval.

Theorem 21.2.18 (Parceval). *Let (x_n) be a complete sequence in a Hilbert space $\mathcal{H}$. Every vector $x \in \mathcal{H}$ has a unique Fourier series representation $\sum_{i=1}^{+\infty} a_i x_i$ where the Fourier coefficients (a_i) belong to l^2 and the inequality of Bessel is an equality.*

Conversely if the inequality of Bessel becomes an equality for every $x \in \mathcal{H}$ then the sequence (x_n) is complete.

Property 21.2.19. A sequence (x_n) in a Hilbert space $\mathcal{H}$ is complete if and only if $\langle x, x_i\rangle = 0$ for all x_i implies that $x = 0$.

21.2.4 Orthogonality and projections

The basic notions on orthogonality in inner product space can be found in section 19.4.2.

Property 21.2.20. Let S be a subset (not necessarily a subspace) of a Hilbert space $\mathcal{H}$. The orthogonal complement $S^\perp$ is closed in $\mathcal{H}$.

Corollary 21.2.21. The previous property implies that the orthogonal complemement of some arbitrary subset of a Hilbert space is a Hilbert space itself.

Theorem 21.2.22 (Projection theorem). *Let H be a Hilbert space and $K \leq H$ a complete subspace. For every $h \in H$ there exists a unique $h' \in K$ such that $h - h'$ is orthogonal to every $k \in K$, i.e $h - h' \in K^\perp$.*

Remark 21.2.23. An equivalent definition for the unique $h' \in K$ is $||h - h'|| = \inf\{||h - k|| : k \in K\}$.

Corollary 21.2.24. It follows that given a complete (or closed) subspace S the Hilbert space $\mathcal{H}$ can be decomposed as $\mathcal{H} = S \oplus S^\perp$.

Definition 21.2.25 (Trace). Let $\mathcal{H}$ be a Hilbert space wih orthogonal basis e_k. Given a bounded linear operator $S \in \mathcal{B}(\mathcal{H})$ we define its trace by the following formula:

$$\mathrm{tr}(S) = \sum_k \langle |S|e_k, e_k\rangle \tag{21.14}$$

21.2.5 Separable Hilbert spaces

The definition of separable spaces in the sense of point-set topology is given in 7.5.22. An equivalent definition for Hilbert spaces is the following.

Alternative Definition 21.2.26 (Separable Hilbert space). A Hilbert space is separable if it contains a complete sequence of orthonormal vectors.

Corollary 21.2.27. Using the Gram-Schmidt method it follows from previous definition that every finite-dimensional Hilbert space is separable.

The following theorem shows that (up to an isomorphism) there are only 2 distinct types of separable Hilbert spaces.

Theorem 21.2.28. *Let $\mathcal{H}$ be separable. If $\mathcal{H}$ is finite-dimensional with dimension n then it is isometrically isomorphic to $\mathbb{C}^n$. If $\mathcal{H}$ is infinite-dimensional then it is isometrically isomorphic to l^2.*

Property 21.2.29. Every orthogonal subset of a separable Hilbert space is countable.

21.2.6 Linear functionals

Property 21.2.30. Let f be a continuous linear functional. Then two possibilities arise:

1. $\dim(\ker f)^\perp = 0$ if $f \equiv 0$.
2. $\dim(\ker f)^\perp = 1$

Theorem 21.2.31 (Riesz' representation theorem). *Let $\mathcal{H}$ be a Hilbert space. For every continuous linear functional $\rho \in \mathcal{H}^*$ there exists a unique element $x_0 \in \mathcal{H}$ such that*

$$\rho(h) = \langle h, x_0 \rangle \tag{21.15}$$

for all $h \in \mathcal{H}$. This implies that $\mathcal{H}$ and $\mathcal{H}^*$ are isometrically isomorphic. Furthermore the operator norm of ρ is equal to the norm of x_0.

Remark 21.2.32. This theorem justifies the bra-ket notation used in quantum mechanics where one associates to every ket $|\psi\rangle \in \mathcal{H}$ a bra $\langle\psi| \in \mathcal{H}^*$.

21.3 Seminorms

Definition 21.3.1 (Seminorm). Let V be a K-vector space. A function $p : V \to [0, +\infty[$ is called a seminorm if it satisfies following conditions:

- **Homogeneity:** $||a\vec{v}|| = |a|||\vec{v}||$ for all scalars $a \in K$
- **Triangle equality (subadditivity):** $||\vec{v} + \vec{w}|| \leq ||\vec{v}|| + ||\vec{w}||$

Theorem 21.3.2 (Hahn-Banach). *Let X be a TVS. If f is a continuous linear functional on X such that $|f(y)| \leq p(y)$ on a subspace $Y \leq X$ for some seminorm p defined on X then there exists a linear extension F on X such that:*

- $F(y) = f(y), \forall y \in Y$.
- $|F(x)| \leq p(x), \forall x \in X$.

21.3.1 Topology

In this subsection we denote by $\mathscr{P}$ a family of seminorms defined on a TVS X. By I we denote the index family of $\mathscr{P}$.

Definition 21.3.3 ($\mathscr{P}$-open ball). A $\mathscr{P}$-open ball centered on x_0 is a subset $Y \subseteq X$ such that all points $y \in Y$ satisfy the following condition for a finite number of seminorms $p_i \in \mathscr{P}, i \in I$:

$$p_i(y - x_0) \leq \varepsilon_i \tag{21.16}$$

where $\varepsilon_i > 0$.

Property 21.3.4. The set of $\mathscr{P}$-open balls generates a topology on X. This topology is often called the $\mathscr{P}$-topology.

Definition 21.3.5 (Separated family). A family of seminorms $\mathscr{P}$ is said to be separated if for every point $x \in X$ there exists a seminorm $p \in \mathscr{P}$ such that $p(x) \neq 0$.

Property 21.3.6. A separated family of seminorms generates a Hausdorff topology on X. If the family is finite or countable the topology is also metrizable. In the case it is finite, the metric is induced by the norm $\sum_{i \in I} p_i$.

Chapter 22

Clifford Algebra

22.1 Clifford algebra

Definition 22.1.1 (Clifford algebra). Let V be unital associative algebra. The Clifford algebra over V with quadratic form $Q : V \to K$ is the free algebra generated by V under the following condition:

$$v \cdot v = Q(v)1 \tag{22.1}$$

where 1 is the unit element in V. This condition implies that the square of a vector is a scalar.

Notation 22.1.2. The Clifford algebra corresponding to V and Q is denoted by $C\ell(V, Q)$.

Construction 22.1.3. The previous definition can be given an explicit construction. First we construct the tensor algebra of V:

$$T(V) = \bigoplus_{k \in \mathbb{N}} V^{\otimes k} \tag{22.2}$$

Then we construct a two-sided ideal I of V generated[1] by $\{v \otimes v - Q(v)1_V \mid v \in V\}$. The Clifford algebra $C\ell(V, Q)$ can then be constructed as the quotient algebra $T(V)/I$.

Remark 22.1.4. Looking at definition 20.8.15 we see that the exterior algebra $\Lambda^*(V)$ coincides with the Clifford algebra $C\ell(V, 0)$. If $Q \neq 0$ then the two algebras are still isomorphic (as vector spaces) when $\text{char}(V) \neq 2$.

Property 22.1.5 (Dimension). If V has dimension n then $C\ell(V, Q)$ has dimension 2^n.

22.2 Geometric algebra

Definition 22.2.1 (Geometric algebra). Let V be a vector space equipped with a symmetric bilinear form $g : V \times V \to K$. The geometric algebra (GA) over V is defined as the Clifford algebra $C\ell(V, g)$. If $\text{char}(V) \neq 2$ then the bilinear form uniquely determines a quadratic form $Q : v \mapsto g(v, v)$ as required in definition 22.1.1.

Definition 22.2.2 (Inner and exterior product). Analogous to the inner product in linear algebra and the wedge product in exterior algebras one can define an (a)symmetric product on the geometric algebra.

[1] See definition 3.2.14.

First of all we note that the product ab of two vectors a and b can be written as the sum of a symmetric and an antisymmetric part:

$$ab = \frac{1}{2}(ab + ba) + \frac{1}{2}(ab - ba) \tag{22.3}$$

We can then define the inner product as the symmetric part:

$$a \cdot b := \frac{1}{2}(ab + ba) = \frac{1}{2}\left((a+b)^2 - a^2 - b^2\right) = g(a,b) \tag{22.4}$$

Analogously we define the exterior (outer) product as the antisymmetric part:

$$a \wedge b := \frac{1}{2}(ab - ba) \tag{22.5}$$

These definitions allow us the rewrite formula 22.3 as:

$$\boxed{ab = a \cdot b + a \wedge b} \tag{22.6}$$

Remark. Looking at the last equality in the definition of the inner product 22.4 we see that condition 22.1 is satisfied when $a = b$.

Definition 22.2.3 (Multivector). Any element of the GA over V is called a multivector. The simple multivectors of grade k, i.e. elements of the form $v_1 v_2 ... v_k$ with $v_i \in V$ for all i, are called k-blades. This generalizes the remark underneath 20.8.3. Sums of multivectors of different grades are called mixed multivectors[2].

Let $n = \dim(V)$. Multivectors of grade n are also called **pseudoscalars** and multivectors of grade $n-1$ are also called **pseudovectors**.

Definition 22.2.4 (Grade projection operator). Let a be a general multivector. The grade (projection) operator $\langle\cdot\rangle_k : \mathcal{G} \to \mathcal{G}_k$ is defined as the projection of a on the k-vector part of a.

Using the grade operators we can extend the inner and exterior product to the complete GA as follows.

Formula 22.2.5. Let A, B be two multivectors of respectively grade m and n. Their inner product is defined as:

$$A \cdot B = \langle AB \rangle_{|m-n|} \tag{22.7}$$

Their exterior product is defined as:

$$A \wedge B = \langle AB \rangle_{m+n} \tag{22.8}$$

An explicit calculation gives us for $A \in \mathcal{G}_1, B \in \mathcal{G}_k$:

$$A \cdot B = \frac{1}{2}\left(AB - (-1)^k BA\right) \tag{22.9}$$

$$A \wedge B = \frac{1}{2}\left(AB + (-1)^k BA\right) \tag{22.10}$$

[2]These elements do not readily represent a geometric structure.

22.3 Classification

Formula 22.3.1 (Dimensional reduction).

$$\mathbb{R}_{p+1,q+1} \cong \mathbb{R}_{p,q} \otimes \mathbb{R}(2) \tag{22.11}$$

Formula 22.3.2.

$$\mathbb{R}_{p+1,q} \cong \mathbb{R}_{q+1,p} \tag{22.12}$$

Formula 22.3.3.

$$\mathbb{R}_{p,q+2} \cong \mathbb{R}_{q,p} \otimes \mathbb{H} \tag{22.13}$$

Theorem 22.3.4 (Bott periodicity). *The classification of (real) Clifford algebras is periodic modulo 8:*

$$\mathbb{R}_{p,q+8} \cong \mathbb{R}_{p+8,q} \cong \mathbb{R}_{p,q} \otimes \mathbb{R}(16) \tag{22.14}$$

This has deep implications in K-theory.

22.4 Pin group

22.4.1 Clifford group

Definition 22.4.1 (Transposition). Let $\{e_i\}_{i\leq n}$ be a basis for V. On the tensor algebra $T(V)$ there exists an anti-automorphism v^t that reverses the order of the basis vectors:

$$\cdot^t : e_i \otimes e_j \otimes \cdots \otimes e_k \mapsto e_k \otimes \cdots \otimes e_j \otimes e_i \tag{22.15}$$

Because the ideal in the definition of a Clifford algebra is invariant under this map, it induces an anti-automorphism, called the transposition or **reversal**, on $C\ell(V)$.

Definition 22.4.2 (Main involution). Let V_0, V_1 be respectively the grade 0 and 1 components of the Clifford algebra $C\ell(V,Q)$. Consider the following operator:

$$\hat{v} = \begin{cases} v & v \in V_0 \\ -v & v \in V_1 \end{cases} \tag{22.16}$$

This operation can be generalized to all of $C\ell(V,Q)$ using linearity. The resulting operator is called the main involution or **inversion** on $C\ell(V,Q)$. Furthermore it turns the Clifford algebra into a superalgebra[3].

Formula 22.4.3 (Twisted conjugation). Let $v \in V$ be a vector and let $s \in C\ell(V,Q)$ be a unit of the Clifford algebra over V, i.e. $Q(s) \neq 0$. The twisted conjugation of v by s is given by the map:

$$\chi : C\ell(V,Q) \times V \to \mathrm{Aut}(V) \qquad \text{with} \qquad \chi(s)v = sv\hat{s}^{-1} \tag{22.17}$$

This map clearly preserves the norm on V and hence $\chi(s)$ is an element of $O(V,Q)$ for all units s.

Definition 22.4.4 (Clifford group). The Clifford group $\Gamma(V,Q)$ is defined as follows:

$$\Gamma(V,Q) = \{s \in C\ell_{hom}(V,Q) : sv\hat{s}^{-1} \in V, v \in V\} \tag{22.18}$$

where the subscript *hom* indicates that we only look at the homogeneous Clifford vectors. Because the units of $C\ell(V)$ form a group, $\Gamma(V,Q)$ also forms a group.

[3] See definition 19.2.32.

Property 22.4.5. When looking at the units of $C\ell(V)$ that belong to V itself, the twisted conjugation is given by a Householder transformation[4].

Property 22.4.6. If V is finite-dimensional and Q non-degenerate, the map

$$\chi : \Gamma(V,Q) \to O(V,Q) : s \mapsto \chi(s) \tag{22.19}$$

defines a representation[5] called the **vectorial representation**. Furthermore, from the first isomorphism theorem 3.1.54 it follows that $O(V,Q)$ is isomorphic to $\Gamma(V,Q)/\ker\chi$ where $\ker\chi = \mathbb{R}_0$. This isomorphism also implies[6] that the Clifford group is given by the set of finite products of invertible elements $v \in V$:

$$\Gamma(V,Q) = \left\{ \prod_i^n s_i : s_i \text{ invertible in } V, n \in \mathbb{N} \right\} \tag{22.20}$$

Corollary 22.4.7. By nothing that pure rotations can be decomposed as an even number of reflections we find that:

$$\Gamma^+(V,Q)/\mathbb{R}_0 \cong \mathrm{SO}(V,Q) \tag{22.21}$$

where Γ^+ is the intersection of the even Clifford algebra and the Clifford group.

22.4.2 Pin and Spin

Formula 22.4.8 (Spinor norm). On $\Gamma(V,Q)$ one can define the spinor norm:

$$\mathcal{N}(x) : \Gamma(V,Q) \to K^\times : x \mapsto x^t x \tag{22.22}$$

where x^t is the transposition 22.15. On V, $\mathcal{N}$ coincides with the norm induced by Q.

Definition 22.4.9 (Pin and spin groups). Using the spinor norm $\mathcal{N}$ we can now define the pin and spins groups as follows:

$$\mathrm{Pin}(V) = \{ s \in \Gamma(V,Q) : \mathcal{N}(s) = 1 \} \tag{22.23}$$

and

$$\mathrm{Spin}(V) = \mathrm{Pin}(V) \cap \Gamma^+(V,Q) \tag{22.24}$$

Remark 22.4.10. In the literature one can sometimes find the following alternative definition of the spinor norm:

$$\mathcal{N}(x) = \hat{x}^t x \tag{22.25}$$

When using this convention the definition of the pinor group uses the condition $|\mathcal{N}(s)| = 1$.

Alternative Definition 22.4.11. The Pin group can also be defined as the set of elements in $\Gamma(V,Q)$ that can be written as a product of unit Clifford vectors. The Spin group is then defined as the elements that can be written as the product of an even number of unit Clifford vectors.

Property 22.4.12. The Pin group satisfies following isomorphism:

$$\mathrm{Pin}(V,Q)/\mathbb{Z}_2 \cong O(V,Q) \tag{22.26}$$

and analogously for the Spin group and $\mathrm{SO}(V,Q)$. These relations also imply that the Pin and Spin groups form a double cover[7] of respectively the orthogonal and special orthogonal groups.

[4] See definition 19.4.18.

[5] In $\mathrm{char}(K) \neq 2$, the surjectiveness of the map χ follows from the *Cartan-Dieudonné theorem*. But even in characteristic 2, the surjectiveness holds.

[6] Again using the *Cartan-Dieudonné theorem*, valid only when $\mathrm{char}(K) \neq 2$.

[7] A covering group is a topological group that is also a covering space. See definition 7.3.15 for more information about the latter.

Definition 22.4.13 (Spinor). Consider a vector space V equipped with a representation of the group $\mathrm{Spin}(n)$, called the **spin representation**. Elements of V are called spinors.

Example 22.4.14. The following table gives some group isomorphisms for the spin group in $\dim n$:

n	$\mathrm{Spin}(n)$
1	$\mathrm{O}(1)$
2	$\mathrm{U}(1)$
3	$\mathrm{SU}(2)$
4	$\mathrm{SU}(2) \times \mathrm{SU}(2)$

For quadratic forms of signature (p, q) we find the following table:

$(1, n)$	$\mathrm{Spin}(1, n)$
$(1, 1)$	$\mathrm{GL}(1, \mathbb{R})$
$(1, 2)$	$\mathrm{SL}(1, \mathbb{R})$
$(1, 3)$	$\mathrm{SL}(2, \mathbb{C})$

Formula 22.4.15. Consider the basis of $\mathfrak{su}(2)$ given by the Pauli matrices 55.12. An explicit (double) covering map $\rho : \mathrm{Spin}(3) \cong \mathrm{SU}(2) \to \mathrm{SO}(3)$ is given by:

$$\rho : U \mapsto \frac{1}{2}\mathrm{tr}(U\sigma_i U^\dagger \sigma^j) e_j \otimes \varepsilon^i \tag{22.27}$$

where ε^k is the dual[8] of the basis vector e_k.

[8] See equation 19.21.

Chapter 23

Operator Algebras

23.1 Operators on Banach & Hilbert spaces

23.1.1 Operator topologies

In this subsection all spaces will be assumed to be Banach unless stated otherwise.

Definition 23.1.1 (Weak topology). A sequence of operators $(T_n)_{n\in\mathbb{N}}$ on a space X converges to an operator T in the weak (Banach) topology if the sequence $(T_n x)_{n\in\mathbb{N}}$ converges to Tx weakly[1] for all x.

Definition 23.1.2 (Strong operator topology). A sequence of operators $(T_n)_{n\in\mathbb{N}}$ on a space X converges to an operator T in the strong (operator) topology if the sequence $(T_n x)_{n\in\mathbb{N}}$ converges to Tx for all x. Equivalently it is the topology generated by the seminorms $T \to ||Tx||, \forall x \in X$.

Definition 23.1.3 (Operator norm). The operator norm of L is defined as follows:

$$||L||_{op} = \inf\{M \in \mathbb{C} | \forall v \in V : ||Lv||_W \leq M||v||_V\} \tag{23.1}$$

As the name suggests it is a norm on $\mathcal{B}(V, W)$. Equivalent definitions of the operator norm are:

$$||L||_{op} = \sup_{||x||\leq 1} ||L(x)|| = \sup_{||x||=1} ||L(x)|| = \sup_{x\neq 0} \frac{||L(x)||}{||x||} \tag{23.2}$$

Definition 23.1.4 (Norm topology[2]). A sequence of operators $(T_n)_{n\in\mathbb{N}}$ on a space X converges to an operator T in the norm topology if the sequence $(||T_n - T||)_{n\in\mathbb{N}}$ converges to 0.

23.1.2 Bounded operators

Definition 23.1.5 (Bounded operator). Let $L : V \to W$ be a linear operator between two Banach spaces. The operator is said to be bounded if it satisfies the following condition:

$$||L||_{op} < \infty \tag{23.3}$$

Notation 23.1.6. The space of bounded linear operators from V to W is denoted by $\mathcal{B}(V, W)$.

[1] This is with respect to the weak topology 21.1.2.

[2] Also called the **uniform (operator) topology**.

Property 23.1.7. If V is a Banach space then $\mathcal{B}(V)$ is also a Banach space.

Following property reduces the problem of continuity to that of boundedness:

Property 23.1.8. Let $f \in \mathcal{L}(V, W)$. Following statements are equivalent:

- f is bounded.
- f is continuous at 0.
- f is continuous on V.
- f is uniformly continuous.
- f maps bounded sets to bounded sets.

Property 23.1.9. Let A be a bounded linear operator with eigenvalue λ. We then have:

$$|\lambda| \leq ||A||_{op} \tag{23.4}$$

Property 23.1.10. Let A be a bounded linear operator. Let $A^\dagger$ denote its adjoint[3]. Then $A^\dagger$ is bounded and $||A||_{op} = ||A^\dagger||_{op}$.

Definition 23.1.11 (Schatten class operator). The **Schatten p-norm** is defined as:

$$||T||_p = \text{tr}\left(\sqrt{T^\dagger T}^{\,p}\right)^{1/p} \tag{23.5}$$

Operators for which this norm is finite are elements of the p^{th} Schatten class $\mathcal{I}_p$.

Property 23.1.12. The Schatten classes are Banach spaces with respect to the associated Schatten norms.

We now consider the two most prominent Schatten classes:

Definition 23.1.13 (Trace class operator). The space of trace class operators on a Hilbert space $\mathcal{H}$ is defined as:

$$\mathcal{B}_1(\mathcal{H}) = \{S \in \mathcal{B}(\mathcal{H}) : \text{tr}(|S|) < \infty\} \tag{23.6}$$

The following theorem can be seen as the analogon of Riesz' theorem for trace class operators:

Theorem 23.1.14. *For every bounded linear functional ρ on the space of trace class operators $\mathcal{B}_1(\mathcal{H})$ there exists a unique bounded linear operator $T \in \mathcal{B}(\mathcal{H})$ such that:*

$$\rho(S) = tr(ST) \tag{23.7}$$

for all $S \in \mathcal{B}_1(\mathcal{H})$. Hence $\mathcal{B}_1(\mathcal{H})$ and $\mathcal{B}(\mathcal{H})$ are isometrically equivalent.

Definition 23.1.15 (Hilbert-Schmidt operator). Consider the Hilbert-Schmidt norm $||\cdot||_2$ from equation 19.43. An operator $T \in \mathcal{B}(\mathcal{H})$ is said to be a Hilbert-Schmidt operator if it satisfies:

$$||T||_2 < \infty \tag{23.8}$$

This space is closed under taking adjoints.

A more general, but still well-behaved class of operators, are the closed operators:

[3]See definition 19.4.6.

Definition 23.1.16 (Closed operator). A linear operator $f : X \to Y$ is said to be closed if for every sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathrm{dom}(f)$ converging to a point $x \in X$ such that $f(x_n)$ converges to a point $y \in Y$ one finds that $x \in \mathrm{dom}(f)$ and $f(x) = y$.

Equivalently one can define a closed linear operator as a linear operator for which its graph is a closed subset in the direct sum $X \oplus Y$.

Definition 23.1.17 (Closure). Let f be a linear operator. Its closure (if it exists) is the closed linear operator $\overline{f}$ such that the graph of $\overline{f}$ is the closure of the graph of f in $X \oplus Y$.

23.1.3 Compact operators

Definition 23.1.18 (Compact operator). Let V, W be Banach spaces. A linear operator $A : V \to W$ is compact if the image of any bounded set in V is relatively compact[4].

Alternative Definition 23.1.19 (Compact operator). Let V, W be Banach spaces. A linear operator $A : V \to W$ is compact if for every bounded sequence (x_n) in V the sequence $(A[x_n]) \subset W$ has a convergent subsequence.

Notation 23.1.20. The space of compact bounded linear operators between Banach spaces X, Y is denoted by $\mathcal{B}_0(X, Y)$. If $X = Y$ then this is abbreviated to $\mathcal{B}_0(X)$ as usual.

Property 23.1.21. $\mathcal{B}_0(X)$ is a two-sided ideal in the (Banach) algebra $\mathcal{B}(X)$.

23.1.4 Fredholm operators

Property 23.1.22. Every compact operator is bounded and hence continuous.

Corollary 23.1.23. Every linear map between finite-dimensional Banach spaces is bounded.

Definition 23.1.24 (Calkin algebra). Consider the algebra $B(V)$ of bounded linear operators on V together with its two-sided ideal $K(V)$ of compact operators. The quotient algebra $\mathcal{Q}(V) = B(V)/K(V)$ is called the Calkin algebra of V.

Definition 23.1.25 (Fredholm operator). Let V, W be Banach spaces. A Fredholm operator $F : V \to W$ is a bounded linear operator $F \in \mathcal{B}(V, W)$ for which the kernel and cokernel are finite-dimensional.

By a theorem of Atkinson one can characterize Fredholm operators using the Calkin algebra:

Property 23.1.26 (Atkinson). An operator $F : V \to W$ is a Fredholm operator if and only if it is invertible when projected onto the Calking algebra, i.e. there exists a bounded linear operator $G : W \to V$ and compact operators C_1, C_2 such that $\mathbb{1}_V - FG = C_1$ and $\mathbb{1}_W - GF = C_2$.

23.1.5 Spectrum

Definition 23.1.27 (Resolvent operator). Let A be a bounded linear operator on a normed space V. The resolvent operator of A is defined as the operator $(A - \lambda\mathbb{1}_V)^{-1}$, where $\lambda \in \mathbb{C}$.

Definition 23.1.28 (Resolvent set). The resolvent set $\rho(A)$ consists of all scalars $\lambda \in \mathbb{C}$ for which the resolvent operator of A is a bounded linear operator on a dense subset of V. These scalars λ are called **regular values** of A.

[4]See definition 7.5.11.

Definition 23.1.29 (Spectrum). The set of scalars $\mu \in \mathbb{C} \setminus \rho(A)$ is called the spectrum $\sigma(A)$.

Remark 23.1.30. It is obvious from the definition of an eigenvalue that every eigenvalue of A belongs to the spectrum of A. The converse however is not true.

Definition 23.1.31 (Point spectrum). The set of scalars $\mu \in \mathbb{C}$ for which $A - \mu \mathbb{1}_V$ fails to be injective is called the point spectrum $\sigma_p(A)$. This set coincides with the set of eigenvalues of A.

Definition 23.1.32 (Continuous spectrum). The set of scalars $\mu \in \mathbb{C}$ for which $A - \mu \mathbb{1}_V$ is injective, fails to be surjective and is dense in V is called the continuous spectrum of A.

Definition 23.1.33 (Residual spectrum). The set of scalars $\mu \in \mathbb{C}$ for which $A - \mu \mathbb{1}_V$ is injective but fails to have a dense range in V is called the residual spectrum $\sigma(A)$. It follows that $\sigma_r(A) \subseteq \sigma(A)$.

Definition 23.1.34 (Essential spectrum). The set of scalars $\mu \in \mathbb{C}$ for which $A - \mu \mathbb{1}_V$ is not a Fredholm operator is called the essential spectrum $\sigma_{\text{ess}}(A)$.

From Atkinson's theorem[5] one can derive the following result:

Property 23.1.35. Let A be a bounded linear operator and let T be a compact operator. The essential spectra of A and $A + T$ coincide.

23.2 Involutive algebras

Definition 23.2.1 (Involution). Let $*$ be an automorphism of an algebra A. If $*(*\ a) = a$ for all $a \in A$ then $*$ is called an involution of A.

Definition 23.2.2 (Involutive algebra[6]). An involutive algebra is an associative algebra A over a commutative involutive ring $(R, ^-)$ together with an operator $^* : A \to A$ such that:

- $(a + b)^* = a^* + b^*$
- $(ab)^* = b^* a^*$
- $(\lambda a)^* = \overline{\lambda} a^*$

where $\lambda \in R$.

23.3 C^*-algebras

Definition 23.3.1 (C*-algebra). A C*-algebra is an involutive Banach algebra[7] such that the **C*-identity**

$$||a^* a|| = ||a||\ ||a^*|| \tag{23.9}$$

is satisfied.

[5] In fact one could (equivalently) define the essential spectrum in terms of the Calkin algebra using Atkinson's theorem. Then this property would be an obvious consequence.

[6] Also called a ***-algebra**.

[7] See definition 21.1.6.

23.3.1 Self-adjoint and positive elements

Definition 23.3.2 (Positive element). An element of a C*-algebra is called positive if its spectrum is contained in $[0, +\infty[$.

Property 23.3.3. Every positive element a can be written as $a = b^*b$ for some element b. Hence every positive-element is self-adjoint.

Definition 23.3.4 (Cuntz algebra). The n^{th} Cuntz algebra $\mathcal{O}_n$ is defined as the (universal) unital C^*-algebra generated by n isometric elements s_i under the additional relation

$$\sum_{i=1}^{n} s_i^* s_i = 1 \tag{23.10}$$

where 1 is the unit element.

23.3.2 Positive maps

Definition 23.3.5 (Positive map). A morphism of C*-algebras is called positive if every positive element is mapped to a positive element.

Definition 23.3.6 (Completely positive map). A morphism of C*-algebras $T : A \to B$ is called completely positive if for all $k \in \mathbb{N}$ the following map is positive:

$$\mathbb{1}_k \otimes T : \mathbb{C}^{k\times k} \otimes A \to \mathbb{C}^{k\times k} \otimes B \tag{23.11}$$

If T satisfies this condition only up to an integer n then it is called n**-positive**.

Definition 23.3.7 (State). Let A be a C*-algebra. A state ψ on A is a positive linear functional of unit norm.

Definition 23.3.8 (Adjoint map). Consider a continuous linear map ϕ defined on the Schatten class $\mathcal{I}_p$. Given a trace functional tr on the C^*-algebra one can define the adjoint map ϕ^* defined on $\mathcal{I}_q$, where p, q are Hölder conjugate. This adjoint is given by the following equation:

$$\mathrm{tr}\Big((\phi^*(A))^* B\Big) = \mathrm{tr}\Big(A^* \phi(B)\Big) \tag{23.12}$$

where $A \in \mathcal{I}_q, B \in \mathcal{I}_p$.

Definition 23.3.9 (Trace preserving map). A map ϕ is said to be trace preserving if it satisfies:

$$\mathrm{tr}(\phi(A)) = \mathrm{tr}(A) \tag{23.13}$$

for all trace class elements A. Using the above definition it is easily seen that on a unital C^*-algebra this is equivalent to:

$$\phi^*(1) = 1 \tag{23.14}$$

Property 23.3.10. A completely positive, trace preserving map ϕ satisfies:

$$||\phi||_1 = 1 \tag{23.15}$$

where the subscript 1 denotes the fact that this operator is defined on trace class elements.

Definition 23.3.11 (Positivity improving). A positive map ϕ is said to be positivity improving if it satisfies:

$$A \geq 0, A \neq 0 \implies \phi(A) > 0 \tag{23.16}$$

Definition 23.3.12 (Ergodic map). A positive map ϕ is said to be ergodic if it satisfies:

$$\forall A \geq 0, A \neq 0 : \exists t_A > 0 : \exp(t_A \phi) A > 0 \tag{23.17}$$

23.3.3 Classification

Theorem 23.3.13. *Let $\mathcal{C}$ be a finite-dimensional C*-algebra. Then there exist unique integers $N, d_1, ..., d_N$ such that:*

$$\mathcal{C} \cong \bigoplus_{i=1}^{N} M_{d_i}(K) \tag{23.18}$$

This implies that every C-algebra can be represented using block matrices.*

23.4 von Neumann algebras

Definition 23.4.1 (Concrete von Neumann algebra). A weakly closed unital $*$-algebra of bounded operators on some Hilbert space.

Alternative Definition 23.4.2 (von Neumann algebra). A $*$-subalgebra of a C^*-algebra equal to its double commutant: $M'' = M$

Theorem 23.4.3 (Double Commutant theorem[8]). *The above definitions are equivalent.*

23.4.1 Factors

Definition 23.4.4 (Factor). Consider a von Neumann algebra M. A $*$-subalgebra A is called a factor of M if its center $Z(A)$ is given by the scalar multiples of the idenity.

Definition 23.4.5 (Powers index). Consider a Hilbert space $\mathcal{H}$ together with its von Neumann algebra of bounded operators $B(\mathcal{H})$. A unital $*$-endomorphism α has Powers index $n \in \mathbb{N}$ if the space $\alpha(B(\mathcal{H}))$ is isomorphic to a type I_n factor.

[8] Often called **von Neumann's double commutant theorem**.

Chapter 24

Representation Theory

References for this chapter are [8, 10].

24.1 Group representations

Definition 24.1.1 (Representation). A representation of a group G, acting on a vector space V, is a homomorphism $\rho : G \to GL(V)$ from G itself to the automorphism group[1] of V. This is a specific case of a group action[2].

Property 24.1.2. Because every linear map maps the zero vector to the zero vector, a group representation can never be free[3].

Definition 24.1.3 (Subrepresentation). A subrepresentation of a representation V is a subspace of V invariant under the action of the group G.

Example 24.1.4 (Permutation representation). Consider a vector space V equipped with a basis $\{e_i\}_{i\in I}$ with $|I| = n$. Let $G = S^n$ be the symmetric group of dimension n. Based on remark 3.1.56 we can consider the action of G on the index set I. This representation is given by

$$\rho(g) : \sum_{i\in I} v_i e_i \mapsto \sum_{i\in I} v_i e_{g\cdot i} \tag{24.1}$$

Example 24.1.5. Consider a representation ρ on V. There exists a natural representation on the dual space V^*. The homomorphism $\rho^* : G \to GL(V^*)$ is given by:

$$\rho^*(g) = \rho^T(g^{-1}) : V^* \to V^* \tag{24.2}$$

where ρ^T is the transpose as defined in 19.3.30. This map satifies the following defining property:

$$\Big\langle \rho^*(g)(v^*), \rho(g)(v) \Big\rangle = \langle v^*, v\rangle \tag{24.3}$$

where $\langle\cdot,\cdot\rangle$ is the natural pairing of V and its dual.

Example 24.1.6. A representation ρ which acts on spaces V, W can also be extended to the tensor product $V \otimes W$ in the following way:

$$g(v \otimes w) = g(v) \otimes g(w) \tag{24.4}$$

Definition 24.1.7 (Intertwiner). If we look at the representation of G as G-modules then the natural morphisms are the intertwining maps 3.1.68.

[1] See definition 19.3.9.
[2] See definition 3.1.55.
[3] See definition 3.1.61.

24.2 Irreducible representations

Definition 24.2.1 (Irreducibility). A representation is said to be irreducible if there exist no proper non-zero subrepresentation.

Example 24.2.2 (Standard representation). Consider the action of $\mathrm{Sym}(n)$ on a vector space V. The line generated by $v_1 + v_2 + ... + v_n$ is invariant under the permutation action of $\mathrm{Sym}(n)$. It follows that the permutation representation (on finite-dimensional spaces) is never irreducible.

The $(n-1)$-dimensional complementary subspace

$$W = \{a_1 v_1 + a_2 v_2 + ... + a_n v_n | a_1 + a_2 + ... + a_n = 0\} \tag{24.5}$$

does form an irreducible representation when we restrict ρ to W. It is called the standard representation of S^n.

Theorem 24.2.3 (Schur's lemma). *Let V, W be two finite-dimensional irreducible representations of a group G. Let $\varphi : V \to W$ be an intertwiner. We then have:*

- *φ is an isomorphism or $\varphi = 0$*
- *If $V = W$ then φ is constant, i.e. φ is a scalar multiple of the identity map $\mathbb{1}_V$.*

Property 24.2.4. If W is a subrepresentation of V then there exists an invariant complementary subspace W' such that $V = W \oplus W'$.

This space can be found as follows: Choose an arbitrary complement U such that $V = W \oplus U$. From this we construct a projection map $\pi_0 : V \to W$. Averaging over G gives

$$\pi(v) = \sum_{g \in G} g \circ \pi_0(g^{-1} v) \tag{24.6}$$

which is a G-linear map $V \to W$. On W it is given by the multiplication of W by $|G|$. Its kernel is then an invariant subspace of V complementary to W.

Theorem 24.2.5 (Maschke). *Let G be a finite group. A representation[4] V can be uniquely decomposed as*

$$V = V_1^{\oplus a_1} \oplus \cdots \oplus V_k^{\oplus a_k} \tag{24.7}$$

where all V_k's are distinct irreducible representations.

The following property is often used in quantum mechanics to quickly find forbidden transitions in atomic or molecular systems:

Property 24.2.6 (Selection rules). Let G be a semisimple group. Let W_1, W_2 be two (finite-dimensional) inequivalent irreducible unitary subrepresentations of $\mathcal{H}$. Let V be a representation equipped with an intertwiner $\rho : V \to \mathcal{L}(\mathcal{H})$. For all $v \in V, w_i \in W_i$ we have:

$$\langle w_1 | \rho(v) w_2 \rangle = 0 \tag{24.8}$$

unless $V \otimes W_2$ contains a subrepresentation equivalent to W_1.

[4] The characteristic of the base field should not divide the order of G.

Chapter 25

Non-Commutative Algebra

References for this chapter are [11].

25.1 Coalgebras

Dual (in the categorical sense) to the definition of a (unital associative) algebra we have:

Definition 25.1.1 (Coalgebra). A vector space C together with two linear maps $\Delta : C \to C \otimes C$ and $\varepsilon : C \to K$, called the **comultiplication** and **counit**, is called a coalgebra if it satisfies following two axioms:

1. $(\mathbb{1} \otimes \Delta) \circ \Delta = (\Delta \otimes \mathbb{1}) \circ \Delta$
2. $(\mathbb{1} \otimes \varepsilon) \circ \Delta = (\varepsilon \otimes \mathbb{1}) \circ \Delta = \mathbb{1}_C$

Example 25.1.2. The simplest example is given by the vector space V with basis $\{e_i\}_{i \in I}$ where the comultiplication and counit are defined as follows:

$$\Delta(e_i) = e_i \otimes e_i \tag{25.1}$$

and

$$\varepsilon(e_i) = 1 \tag{25.2}$$

By linearity these maps can be extended to all of V. Important cases are the tensor algebra and exterior algebra over a vector space. (See definitions 20.5.11 and 20.8.14.)

Definition 25.1.3 (Group-like element). An element c in a coalgebra (C, Δ, ε) that satisfies $\Delta(c) = c \otimes c$ and $\varepsilon(c) = 1$.

Remark. The name 'group-like' stems from the fact that the coalgebra structure on the group algebra $K[G]$ is given by defining $\Delta(g) = g \otimes g$ for all $g \in G$.

Definition 25.1.4 (Unital coalgebra). A coalgebra (C, Δ, ε) is said to be unital if it comes equipped with a coalgebra morphism $\eta : K \to C$. The element $\eta(1)$ is often also denoted by 1.

Definition 25.1.5 (Primitive element). An element c in a unital coalgebra (C, Δ, ε) that satisfies $\Delta(c) = c \otimes 1 + 1 \otimes c$.

Notation 25.1.6 (Sweedler notation). Let (C, Δ) be a coalgebra. For any element $c \in C$ the comultiplication $\Delta(c)$ is an element of $C \otimes C$ and can thus be written in the following form

$$\Delta(c) = \sum_i a_i \otimes b_i$$

However for long operations this gets tedious and hence we introduct the notation[1] $\Delta(c) = \sum_{(c)} c_{(1)} \otimes c_{(2)}$ or even $\Delta(c) = c_{(1)} \otimes c_{(2)}$. For example due to coassociativity $(\Delta \otimes \mathbb{1}) \circ \Delta = (\mathbb{1} \otimes \Delta) \circ \Delta$ we can write:

$$c_{(1)} \otimes c_{(2)} \otimes c_{(3)} = \sum_{(c)} c_{(1)(1)} \otimes c_{(1)(2)} \otimes c_{(2)} = \sum_{(c)} c_{(1)} \otimes c_{(2)(1)} \otimes c_{(2)(2)}$$

The counit law becomes:

$$c = c_{(1)}\varepsilon(c_{(2)}) = \varepsilon(c_{(1)})c_{(2)}$$

and hence one can freely move the counit ε around.

25.2 Hopf algebras

Definition 25.2.1 (Bialgebra). Let A be a vector space over a field K. Suppose that the triple (A, ∇, η) defines a unital associative algebra and that the triple (A, Δ, ε) defines a counital coassociative coalgebra. Then the quintuple $(A, \nabla, \eta, \Delta, \varepsilon)$ defines a bialgebra if ∇ and Δ satisfy the following commutative diagrams:

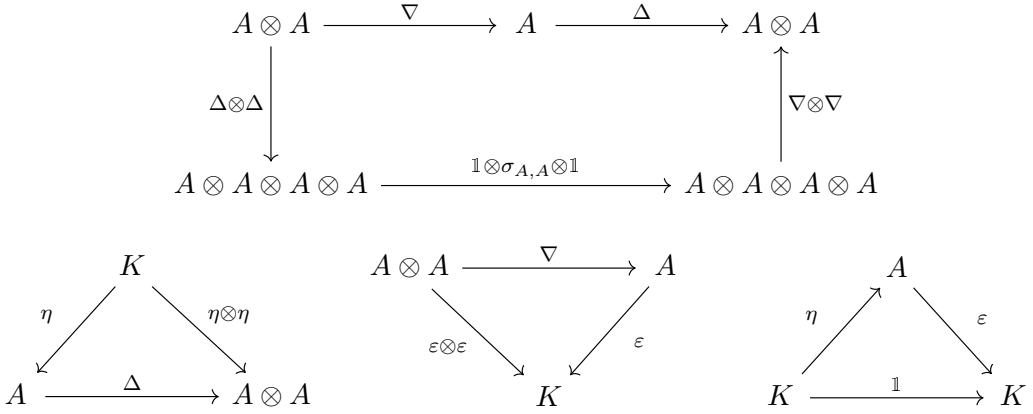


Figure 25.1: Bialgebra conditions.

where these diagrams state that ∇, η are coalgebra morphisms and Δ, ε are algebra morphisms.

Definition 25.2.2 (Hopf algebra). Let $(A, \nabla, \eta, \Delta, \varepsilon)$ be a bialgebra. A is called a Hopf algebra if it is equipped with a linear map $S : A \to A$ that satisfies:

$$\nabla \circ (\mathbb{1}_A \otimes S) \circ \Delta = \nabla \circ (S \otimes \mathbb{1}_A) \circ \Delta = \eta \circ \varepsilon \tag{25.3}$$

The map S is also called the **antipode** or **coinverse**.

Property 25.2.3. Given a Hopf algebra structure on a bialgebra, the antipode S is an antihomomorphism. Furthermore it is unique and hence being a Hopf algebra is a property, not a structure.

Remark 25.2.4. Sometimes one requires the antipode to be invertible.

Definition 25.2.5 (Quasi-triangular Hopf algebra). A Hopf algebra H for which there exists an invertible element $R \in H \otimes H$ that satisfies:

- $R\Delta(a) = \sigma\Big(\Delta(x)\Big)R$

[1] Sometimes one uss the notation $\Delta(c) = \Delta_1(c) \otimes \Delta_2(c)$.

- $(\Delta \otimes \mathbb{1})(R) = R_{13}R_{23}$
- $(\mathbb{1} \otimes \Delta)(R) = R_{13}R_{12}$

where $\sigma(x \otimes y) = y \otimes x$ is the braiding on H and where $R_{ij} \in H \otimes H \otimes H$ is defined using the components of R in the i^{th} and j^{th} position and the unit element $1 \in H$ in the other position, i.e. $(a \otimes b)_{13} = a \otimes 1 \otimes b$.

The element R is often called the **universal R-matrix**.

Remark 25.2.6 (Tensor product of modules). One could ask where bialgebras and especially Hopf algebras naturally arise. Consider an algebra A together with its category of modules **AMod**. Now one would like to define a monoidal structure on **AMod** induced by the tensor product on A. However this monoidal structure should be compatible with the action of A.

The intuitive (left) action

$$A \otimes (M \otimes_A N) \to M \otimes_A N : a \otimes m \otimes n \mapsto (am) \otimes n$$

does not admit a suitable tensor unit due to its asymmetric definition. To obtain the correct definition we are inspired by group representations: $g \cdot (m \otimes n) = gm \otimes gn$. In this case one has the diagonal map $\Delta : G \to G \times G$ which can be used to act on both sides of the tensor product. Now one could ask "Why not just define the action of an algebra in the same way?", namely:

$$a \otimes (m \otimes n) \mapsto (am) \otimes (an)$$

However because we require the action to be linear (after all it should be compatible with the algebra morphisms) this definition is not valid. Hence we require the existence of an algebra morphism $A \to A \otimes A$ with which we can construct a suitable action as follows:

$$A \otimes (M \otimes_A N) \to M \otimes_A N : a \otimes m \otimes n \mapsto (a_{(1)}m) \otimes (a_{(2)}n) \tag{25.4}$$

where we used Sweedler notation. Together with the usual conditions of an algebra action one obtains exactly the requirement that A should be a bialgebra. So if A is a bialgebra then **AMod** will be a monoidal category (this is in fact an equivalence known as *Tannaka duality*).

Now one could require some more structure on **AMod**, for example that it admits duals. Consider an A-module V together with its dual $V^* \cong \mathrm{Hom}(V, \mathbb{C})$. Given a linear map $S : A \to A$ one could define a general action as follows:

$$(af)(v) = f\big(S(a)v\big) \tag{25.5}$$

The requirement that this is indeed an action leads us to the requirement $S(ab) = S(b)S(a)$ on S, which is equivalent to requiring that S is an algebra antihomomorphism. Together with the other compatibility conditions, such as that the evaluation and coevaluation maps induced by the underlying vector spaces are also A-module morphisms, we are led to the requirement that A is a Hopf algebra. Hence if A is a Hopf algebra (with an invertible antipode) then **AMod** will be a rigid monoidal category.

25.3 Differential calculi

Definition 25.3.1 (First order differential calculus). Let A be an algebra and let Γ be an A-bimodule. An A-bimodule morphism $d : A \to \Gamma$ is called a first order differential calculus if it satisfies the following two conditions:

- Leibniz rule: $d(ab) = (da)b + a(db)$

- Standard form: Every element $g \in \Gamma$ can be written as

$$g = \sum_{i=1}^{n} a_i(db_i)$$

for some $n \in \mathbb{N}$ and (not necessarily unique) elements $\{a_i, b_i\}_{i \leq n}$.

Remark 25.3.2. The second condition can be rewritten in terms of a right action using the Leibniz rule.

25.4 Quantum groups

This section heavily builds upon the theory presented in chapter 28. The content is partially based on talks by André Henriques.

Construction 25.4.1 (Jimbo-Drinfeld). Consider a Lie algebra $\mathfrak{g}$ together with its universal enveloping algebra $U(\mathfrak{g})$ constructed as using the Serre relations 28.4.2:

1. $[H_i, H_j] = 0$
2. $[H_i, E_j] = a_{ij}E_j$
3. $[H_i, F_j] = -a_{ij}F_j$
4. $[E_i, F_j] = \delta_{ij}H_j$
5. $\mathrm{ad}_{E_i}^{|a_{ij}|-1}(E_j) = 0$
6. $\mathrm{ad}_{F_i}^{|a_{ij}|-1}(F_j) = 0$

To obtain the quantum group $U_q(\mathfrak{g})$, which is also called a **deformation** or **quantization** of $U(\mathfrak{g})$, one replaces the generators H_i by the following generators[2]:

$$K_i := q^{d_i H_i} \tag{25.6}$$

where $d_i = \frac{\langle \alpha_i, \alpha_i \rangle}{2}$ is related to the norm of the i^{th} simple root. Hence instead of the H_i being functionals on the root lattice, one gets gets linear functions from the root lattice to the Laurent polynomials in q, i.e. $\mathbb{C}[q, q^{-1}]$.

From this functional point of view one can rewrite the second and third relation as follows:

$$f \cdot E_i = E_i \tau_{\alpha_i}(f)$$
$$f \cdot F_i = F_i \tau_{-\alpha_i}(f)$$

where f is a polynomial in the H_i's and $\tau_{\alpha_i}(f)(\lambda) = f(\lambda + \alpha_i)$. Replacing H_i by K_i one obtains the following relations:

2*. $K_i E_j = q^{d_i a_{ij}} E_j K_i$

3*. $K_i F_j = q^{-d_i a_{ij}} F_j K_i$

The three relations relating E_i's and F_i's are deformed using q-analog numbers. First we define the q-numbers[3]:

$$[n]_q := \frac{q^n - q^{-n}}{q - q^{-1}} \tag{25.7}$$

Using this definition Serre relation 4 becomes:

[2] To be complete one should also add generators K_i^{-1} which act as inverses of the generators K_i.

[3] Note that q-numbers are often defined differently, this definition is equal to $\frac{1}{q^{n-1}}[n]_{q^2}$ when using the common definition.

4*. $[E_i, F_j] = \delta_{ij}[H_i]_{q^{d_i}} = \delta_{ij}\frac{K_i - K_i^{-1}}{q^{d_i} - q^{-d_i}}$

where the factor $[H_i]_{q^{d_i}}$ should be interpreted as first evaluating H_i on a root and then taking the q-analog. The adjoint action relations (5 and 6) on E_i and F_i can be rewritten by replacing binomial coefficients by their q-analog ($i \neq j$):

5*. $\sum_{k=1}^{1+|a_{ij}|}(-1)^k \begin{bmatrix} 1+|a_{ij}| \\ k \end{bmatrix}_{q^{d_i}} E_i^{1+|a_{ij}|-k} E_j E_i^k = 0$

6*. $\sum_{k=1}^{1+|a_{ij}|}(-1)^k \begin{bmatrix} 1+|a_{ij}| \\ k \end{bmatrix}_{q^{d_i}} F_i^{1+|a_{ij}|-k} F_j F_i^k = 0$

Part VI

Differential Geometry

Chapter 26

Curves and Surfaces

26.1 Curves

Property 26.1.1 (Regular curve). Let $\vec{\boldsymbol{c}}(t) : I \to \mathbb{R}^n$ be a curve defined on an interval I. $\vec{\boldsymbol{c}}(t)$ is said to be regular[1] if $\frac{d\vec{\boldsymbol{c}}}{dt}(t) \neq \vec{\mathbf{0}}$ for all $t \in I$.

Definition 26.1.2 (C^r-parameter transformation). A transformation $\varphi :]c, d[\to]a, b[: u \mapsto t$ such that $\varphi(u)$ is a C^r-diffeomorphism[2].

Definition 26.1.3 (Geometric property). A geometric property is a property that is invariant under:

1. parameter transformations
2. positive orthonormal changes of basis

Theorem 26.1.4. *Let $\vec{\boldsymbol{c}}(t), \vec{\boldsymbol{d}}(t)$ be two curves with the same image. The following relation holds for all t:*

$$\vec{\boldsymbol{c}}(t) \text{ regular} \iff \vec{\boldsymbol{d}}(t) \text{ regular} \tag{26.1}$$

26.1.1 Arc length

Definition 26.1.5 (Natural parameter). Let $\vec{\boldsymbol{c}}(t)$ be a curve. The parameter t is said to be a natural parameter if:

$$\left|\left|\frac{d\vec{\boldsymbol{c}}}{dt}\right|\right| \equiv 1 \tag{26.2}$$

Formula 26.1.6 (Arc length). The following function $\phi(t)$ is a bijective map and a natural parameter of $\vec{\boldsymbol{c}}(t)$:

$$\phi(t) = \int_{t_0}^{t} ||\dot{\vec{\boldsymbol{c}}}(t)||dt \tag{26.3}$$

Remark. The arc length as defined above is often denoted by 's'.

Theorem 26.1.7. *Let $\vec{\boldsymbol{c}}(t)$ be a curve. Let u be an alternative parameter of $\vec{\boldsymbol{c}}(t)$. It is a natural parameter if and only if there exists a constant α such that:*

$$u = \pm s + \alpha$$

where s is the integral as defined in equation 26.3.

Remark. As the last theorem implies, no unique natural parameter or arc length exists.

[1] See also property 27.3.8.
[2] See definition 7.3.11

26.1.2 Frenet-Serret frame

Definition 26.1.8 (Tangent vector). Let $\vec{\boldsymbol{c}}(s)$ be parametrized by arc length. The tangent vector $\vec{\boldsymbol{t}}(s)$ is defined as:

$$\vec{\boldsymbol{t}}(s) = \vec{\boldsymbol{c}}\,'(s) \tag{26.4}$$

Property 26.1.9. From the definition of the natural parametrization 26.2 and the previous definition it follows that the tangent vector is a unit vector.

Definition 26.1.10 (Principal normal vector). Let $\vec{\boldsymbol{c}}(s)$ be parametrized by arc length. The principal normal vector is defined as:

$$\vec{\boldsymbol{n}}(s) = \frac{\vec{\boldsymbol{t}}\,'(s)}{||\vec{\boldsymbol{t}}\,'(s)||} \tag{26.5}$$

Property 26.1.11. From property 26.1.9 and the definition of the principal normal vector it follows that the tangent vector and principal normal vector are orthogonal.

Definition 26.1.12 (Binormal vector). Let $\vec{\boldsymbol{c}}(s)$ be parametrized by arc length. The binormal vector is defined as:

$$\vec{\boldsymbol{b}}(s) = \vec{\boldsymbol{t}}(s) \times \vec{\boldsymbol{n}}(s) \tag{26.6}$$

Definition 26.1.13 (Frenet-Serret frame). As the vectors $\vec{\boldsymbol{t}}(s), \vec{\boldsymbol{n}}(s)$ and $\vec{\boldsymbol{b}}(s)$ are mutually orthonormal and linearly independent, we can use them to construct a positive orthonormal basis. The ordered basis $\left(\vec{\boldsymbol{t}}(s), \vec{\boldsymbol{n}}(s), \vec{\boldsymbol{b}}(s)\right)$ is called the **Frenet-Serret** frame.

Remark. This basis does not have to be the same in every point of $\vec{\boldsymbol{c}}(s)$.

Definition 26.1.14 (Curvature). Let $\vec{\boldsymbol{c}}(s)$ be parametrized by arc length. The curvature of $\vec{\boldsymbol{c}}(s)$ is defined as:

$$\frac{1}{\rho(s)} = ||\vec{\boldsymbol{t}}\,'(s)|| \tag{26.7}$$

Definition 26.1.15 (Torsion). Let $\vec{\boldsymbol{c}}(s)$ be a curve parametrized by arc length. The torsion of $\vec{\boldsymbol{c}}(s)$ is defined as:

$$\tau(s) = \rho(s)^2(\vec{\boldsymbol{t}} \;\; \vec{\boldsymbol{t}}\,' \;\; \vec{\boldsymbol{t}}\,'') \tag{26.8}$$

Formula 26.1.16 (Frenet formulas). The derivatives of the tangent, principal normal and binormal vectors can be written as a linear combination of the those vectors themself as:

$$\begin{cases} \vec{\boldsymbol{t}}\,'(s) &= & & \frac{1}{\rho(s)}\vec{\boldsymbol{n}}(s) & \\ \vec{\boldsymbol{n}}\,'(s) &= & -\frac{1}{\rho(s)}\vec{\boldsymbol{t}}(s) & + & \tau(s)\vec{\boldsymbol{b}}(s) \\ \vec{\boldsymbol{b}}\,'(s) &= & & -\tau(s)\vec{\boldsymbol{n}}(s) & \end{cases} \tag{26.9}$$

Theorem 26.1.17 (Fundamental theorem of curves). *Let $k(s), w(s) : U \to \mathbb{R}$ be two $\mathcal{C}^1$ functions with $k(s) \geq 0, \forall s$. There exists an interval $]-\varepsilon, \varepsilon[\subset U$ and a curve $\vec{\boldsymbol{c}}(s) :]-\varepsilon, \varepsilon[\to \mathbb{R}^3$ with natural parameter s such that $\vec{\boldsymbol{c}}(s)$ has $k(s)$ as its curvature and $w(s)$ as its torsion.*

26.2 Surfaces

Notation 26.2.1. Let $\vec{\boldsymbol{\sigma}}$ be a surface[3]. The derivative of $\vec{\boldsymbol{\sigma}}$ with respect to the coordinate q^i is written as follows:

$$\frac{\partial \vec{\boldsymbol{\sigma}}}{\partial q^i} = \vec{\boldsymbol{\sigma}}_i \tag{26.10}$$

[3] $\vec{\boldsymbol{\sigma}}$ denotes the surface as a vector field. Σ denotes the geometric image of $\vec{\boldsymbol{\sigma}}$.

26.2.1 Tangent vectors

Definition 26.2.2 (Tangent plane). Let $P(q_0^1, q_0^2)$ be a point on the surface Σ. The tangent space $T_P\Sigma$ to $\vec{\boldsymbol{\sigma}}$ in P is defined as follows:

$$\forall \vec{\boldsymbol{r}} \in T_P\Sigma : \left[\vec{\boldsymbol{r}} - \vec{\boldsymbol{\sigma}}(q_0^1, q_0^2)\right] \cdot \left[\vec{\boldsymbol{\sigma}}_1(q_0^1, q_0^2) \times \vec{\boldsymbol{\sigma}}_2(q_0^1, q_0^2)\right] = 0 \tag{26.11}$$

Definition 26.2.3 (Normal vector). The cross product in equation 26.11 is closely related to the normal vector to Σ in P. The normal vector in the point (q_0^1, q_0^2) is defined as:

$$\vec{\boldsymbol{N}}(q_0^1, q_0^2) = \frac{1}{||\vec{\boldsymbol{\sigma}}_1 \times \vec{\boldsymbol{\sigma}}_2||} (\vec{\boldsymbol{\sigma}}_1 \times \vec{\boldsymbol{\sigma}}_2) \tag{26.12}$$

26.2.2 First fundamental form

Definition 26.2.4 (Metric coefficients). Let $\vec{\boldsymbol{\sigma}}$ be a surface. The metric coefficients g_{ij} are defined as follows:

$$\boxed{g_{ij} = \vec{\boldsymbol{\sigma}}_i \cdot \vec{\boldsymbol{\sigma}}_j} \tag{26.13}$$

Definition 26.2.5 (Scale factor). The following factors are often used in vector calculus:

$$g_{ii} = h_i^2 \tag{26.14}$$

Definition 26.2.6 (First fundamental form). Let $\vec{\boldsymbol{\sigma}}$ be a surface. Define a bilinear form $I_P(\vec{\boldsymbol{v}}, \vec{\boldsymbol{w}}) : T_P\Sigma \times T_P\Sigma \to \mathbb{R}$ that restricts the inner product to $T_P\Sigma$:

$$I_P(\vec{\boldsymbol{v}}, \vec{\boldsymbol{w}}) = \vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{w}} \tag{26.15}$$

This bilinear form is called the first fundamental form or **metric**.

Corollary 26.2.7. All $\vec{\boldsymbol{v}}, \vec{\boldsymbol{w}} \in T_P\Sigma$ are linear combinations of the tangent vectors $\vec{\boldsymbol{\sigma}}_1, \vec{\boldsymbol{\sigma}}_2$. This leads to the following relation between the first fundamental form and the metric coefficients 26.13:

$$I_P(\vec{\boldsymbol{v}}, \vec{\boldsymbol{w}}) = v^i \vec{\boldsymbol{\sigma}}_i \cdot w^j \vec{\boldsymbol{\sigma}}_j = g_{ij} v^i w^j \tag{26.16}$$

Notation 26.2.8. The length of a cruve $\vec{\boldsymbol{c}}(t)$ can be written as follows:

$$s = \int \sqrt{||\dot{\vec{\boldsymbol{c}}}(t)||} dt = \int \sqrt{ds^2} \tag{26.17}$$

where the second equality is formally defined. The two equalities together can be combined into the following notation for the metric:

$$\boxed{ds^2 = g_{ij} dq^i dq^j} \tag{26.18}$$

Formula 26.2.9. Let (g_{ij}) be the metric tensor. We define the matrix (g^{ij}) as its inverse:

$$(g^{ij}) = \frac{1}{\det(g_{ij})} \begin{pmatrix} g_{22} & -g_{12} \\ -g_{12} & g_{11} \end{pmatrix} \tag{26.19}$$

26.2.3 Isometries

Definition 26.2.10 (Isometry). An isometry is a distance-preserving map, i.e. a diffeomorphism $\Phi : \Sigma \to \Sigma'$ that maps arc segments in Σ to arc segments with the same length in Σ'.

Property 26.2.11. A diffeomorphism Φ is an isometry if and only if the metric coefficients of σ and σ' are the same.

Definition 26.2.12 (Conformal map). A diffeomorphism $\Phi : \Sigma \to \Sigma'$ is said to be conformal or isogonal if it maps two intersecting curves in Σ to intersecting curves in Σ' with the same intersection angle.

Property 26.2.13. A diffeomorphism Φ is conformal if and only if the metric coefficients of σ and σ' are proportional.

Definition 26.2.14 (Surface preserving map). A diffeomorphism $\Phi : \Sigma \to \Sigma'$ is sadi to be surface-preserving if it maps a segment of Σ to a segment of Σ' with the same surface.

Property 26.2.15. A diffeomorphism Φ is surface-preserving if and only if the metric coefficients of σ and σ' satisfy:

$$g'_{11}g'_{22} - (g'_{12})^2 = g_{11}g_{22} - g_{12}^2 \tag{26.20}$$

for all points (q^1, q^2).

Corollary 26.2.16. A map that is surface-preserving and conformal is also isometric.

26.2.4 Second fundamental form

Definition 26.2.17 (Second fundamental form). Let $\vec{\boldsymbol{\sigma}}(q^1, q^2)$ be a surface. The second fundamental form is a bilinear form $II_P(\vec{\boldsymbol{v}}, \vec{\boldsymbol{w}}) : T_P\Sigma \times T_P\Sigma \to \mathbb{R}$ defined as follows:

$$\boxed{II_P(\vec{\boldsymbol{v}}, \vec{\boldsymbol{w}}) = L_{ij}(q^1, q^2)v^i w^j} \tag{26.21}$$

where $L_{ij} = \vec{\boldsymbol{N}} \cdot \vec{\boldsymbol{\sigma}}_{ij}$.

Definition 26.2.18 (Normal curvature). Let $\vec{\boldsymbol{c}}$ be a curve parametrized as

$$\vec{\boldsymbol{c}}(s) = \vec{\boldsymbol{\sigma}}\left(q^1(s), q^2(s)\right)$$

The normal curvature of $\vec{\boldsymbol{c}}(s)$ at a point $\left(q^1(s), q^2(s)\right)$ is defined as:

$$\boxed{\frac{1}{\rho_n(s)} = \vec{\boldsymbol{c}}\,''(s) \cdot \vec{\boldsymbol{N}}(s)} \tag{26.22}$$

From the definition of the second fundamental form it follows that the normal curvature can be written as:

$$\frac{1}{\rho_n(s)} = II(\vec{\boldsymbol{t}}, \vec{\boldsymbol{t}}) = \frac{II\left(\dot{\vec{\boldsymbol{c}}}(t), \dot{\vec{\boldsymbol{c}}}(t)\right)}{I\left(\dot{\vec{\boldsymbol{c}}}(t), \dot{\vec{\boldsymbol{c}}}(t)\right)} \tag{26.23}$$

where the last equality holds for any given parameter t.

Theorem 26.2.19 (Meusnier's theorem). *Let $\vec{\boldsymbol{c}}, \vec{\boldsymbol{d}}$ be two curves on a surface $\vec{\boldsymbol{\sigma}}$. The curves have the same normal curvature in a point $\left(q^1(t_0), q^2(t_0)\right)$ if the following two conditions are satisfied:*

- $\vec{\boldsymbol{c}}(t_0) = \vec{\boldsymbol{d}}(t_0)$
- $\dot{\vec{\boldsymbol{c}}}(t_0) \mid\mid \dot{\vec{\boldsymbol{d}}}(t_0)$

Furthermore, the osculating circles of all curves with the same normal curvature at a given point form a sphere.

Property 26.2.20. The normal curvature of at a given point is equal to the curvature of the normal section, i.e. the intersection of the surface with a normal plane at the point.

Definition 26.2.21 (Geodesic curvature). Let $\vec{\boldsymbol{c}}$ be a curve parametrized as:

$$\vec{\boldsymbol{c}}(s) = \vec{\boldsymbol{\sigma}}\left(q^1(s), q^2(s)\right)$$

The geodesic curvature of $\vec{\boldsymbol{c}}(s)$ at the point $\left(q^1(s), q^2(s)\right)$ is defined as:

$$\boxed{\frac{1}{\rho_g(s)} = \left(\vec{\boldsymbol{N}}(s)\ \vec{\boldsymbol{t}}(s)\ \vec{\boldsymbol{t}}\,'(s)\right)} \tag{26.24}$$

Formula 26.2.22. Let $\vec{\boldsymbol{c}}$ be a curve on a surface $\vec{\boldsymbol{\sigma}}$. From the definitions of the normal and geodesic curvature it follows that:

$$\boxed{\frac{1}{\rho^2} = \frac{1}{\rho_n^2} + \frac{1}{\rho_g^2}} \tag{26.25}$$

26.2.5 Curvature of a surface

Definition 26.2.23 (Weingarten map). Let P be a point of a surface Σ. The Weingarten map $L_P : T_P\Sigma \to T_P\Sigma$ is a linear map defined as:

$$L_P(\vec{\boldsymbol{\sigma}}_1) = -\vec{\boldsymbol{N}}_1 \qquad \text{and} \qquad L_P(\vec{\boldsymbol{\sigma}}_2) = -\vec{\boldsymbol{N}}_2 \tag{26.26}$$

Formula 26.2.24. Let $\vec{\boldsymbol{v}}, \vec{\boldsymbol{w}} \in T_P\Sigma$. The following equalities hold:

$$L_P(\vec{\boldsymbol{v}}) \cdot \vec{\boldsymbol{w}} = L_P(\vec{\boldsymbol{w}}) \cdot \vec{\boldsymbol{v}} = II_P(\vec{\boldsymbol{v}}, \vec{\boldsymbol{w}}) \tag{26.27}$$

Formula 26.2.25. Let $\left(g^{ij}\right)$ be the inverse of the metric tensor. The matrix elements of L_P are defined as:

$$L_j^k = g^{ki} L_{ij} \tag{26.28}$$

Formula 26.2.26 (Weingarten formulas).

$$\vec{\boldsymbol{N}}_j = -L_j^k \vec{\boldsymbol{\sigma}}_k \tag{26.29}$$

Property 26.2.27. For every point P on the surface Σ there exists a basis $\{\vec{\boldsymbol{h}}_1, \vec{\boldsymbol{h}}_2\} \subset T_P\Sigma$ of eigenvectors of L_P. Furthermore, the corresponding eigenvalues are given by $II_P(\vec{\boldsymbol{h}}_i, \vec{\boldsymbol{h}}_i)$ and these eigenvalues are the extreme values of the normal curvature at the point P.

Definition 26.2.28 (Principal curvatures). The eigenvalues of the Weingarten map are called the principal curvatures of the surface and they are denoted by $\frac{1}{R_1}$ and $\frac{1}{R_2}$. The tangent vectors corresponding to these curvatures are called the **principal directions**.

Property 26.2.29. If the principal curvatures are not equal, the principal directions are orthogonal. If they are equal, the point P is said to be an **umbilical point** or **umbilic**.

Property 26.2.30. [4] The principal directions satisfy[5]:

$$L_P(\vec{\boldsymbol{h}}_1) \cdot \vec{\boldsymbol{h}}_2 = 0 \tag{26.30}$$

If P is an umbilic then every tangent vector in P is a principal direction and the equality is satisfied for every two tangent vectors.

Definition 26.2.31 (Line of curvature). A curve is a line of curvature if the tangent vector in every point P is a principal direction of the surface in P.

Formula 26.2.32 (Rodrigues' formula). A curve is a line of curvature if and only if it satisfies the following formula:

$$\frac{d\vec{\boldsymbol{N}}}{dt}(t) = -\frac{1}{R(t)}\frac{d\vec{\boldsymbol{c}}}{dt}(t) \tag{26.31}$$

If the curve satisfies this formula, then the scalar function $1/R(t)$ coincides with the principal curvature along the curve.

Formula 26.2.33 (Differential equation for curvature lines).

$$\begin{vmatrix} (\dot{q}^2)^2 & -\dot{q}^1\dot{q}^2 & (\dot{q}^1)^2 \\ g_{11} & g_{12} & g_{22} \\ L_{11} & L_{12} & L_{22} \end{vmatrix} = 0 \tag{26.32}$$

Property 26.2.34. From theorem 26.2.27 we know that the principal directions are orthogonal vectors. It follows that on a surface containing no umbilics the curvature lines form an orthogonal web and in every point P we find 2 orthogonal curvature lines.

Definition 26.2.35 (Gaussian curvature). The Gaussian curvature K of a surface is defined as the determinant of the Weingarten map, i.e.:

$$K = \frac{1}{R_1 R_2} \tag{26.33}$$

Definition 26.2.36 (Mean curvature). The mean curvature H of a surface is defined as the trace of the Weingarten map, i.e.:

$$H = \frac{1}{2}\left(\frac{1}{R_1} + \frac{1}{R_2}\right) \tag{26.34}$$

Property 26.2.37. The principal curvatures are the solutions of the following equation:

$$x^2 - 2Hx + K = 0 \tag{26.35}$$

This is the characteristic equation (19.67) of the Weingarten map.

Definition 26.2.38. Let P be a point on the surface Σ.

- P is said to be **elliptic** if $K > 0$ in P.
- P is said to be **hyperbolic** if $K < 0$ in P.
- P is said to be **parabolic** if $K = 0$ and $\frac{1}{R_1}$ or $\frac{1}{R_2} \neq 0$ in P.
- P is said to be **flat** if $\frac{1}{R_1} = \frac{1}{R_2} = 0$ in P.

[4] adjoint!vectors

[5] Tangent vectors that satisfy this equation are called **adjoint** tangent vectors.

- P is said to be **umbilical** if $\frac{1}{R_1} = \frac{1}{R_2}$ in P.

Remark. From previous definition it follows that a flat point is a special type of umbilic.

Property 26.2.39. A surface Σ containing only umbilics is part of a sphere or part of a plane.

Theorem 26.2.40. *In the neighbourhood of a point P of a surface with principal curvatures $1/R_1$ and $1/R_2$, the surface is locally given by the following quadric:*

$$x_3 = \frac{1}{2}\left(\frac{x_1^2}{R_1} + \frac{x_2^2}{R_2}\right) \tag{26.36}$$

up to order $O(x^2)$.

Formula 26.2.41 (Euler's formula). The normal curvature of a couple $(P, \vec{e})$ where is $\vec{e} = \vec{h}_1 \cos\theta + \vec{h}_2 \sin\theta \in T_P\Sigma$ is given by:

$$\frac{1}{\rho_n} = \frac{\cos^2\theta}{R_1} + \frac{\sin^2\theta}{R_2} \tag{26.37}$$

Definition 26.2.42 (Asymptotic curve). An asymptotic curve is a curve which is in every point P tangent to a direction with zero normal curvature.

Formula 26.2.43 (Differential equation for asymptotic curves).

$$L_{11}\left(\dot{q}^1(t)\right)^2 + 2L_{12}\dot{q}^1(t)\dot{q}^2(t) + L_{22}\left(\dot{q}^2(t)\right)^2 = 0 \tag{26.38}$$

Property 26.2.44. A curve on a surface is an asymptotic curve if and only if the tangent plane and the osculation plane coincide in every point P of the surface.

26.2.6 Christoffel symbols and geodesics

Formula 26.2.45 (Gauss' formulas). The second derivatives of the surface $\vec{\sigma}$ are given by:

$$\vec{\sigma}_{ij} = L_{ij}\vec{N} + \Gamma^k_{\ ij}\vec{\sigma}_k \tag{26.39}$$

where the **Christoffel symbols** $\Gamma^k_{\ ij}$ are defined as:

$$\Gamma^k_{\ ij} = g^{kl}\vec{\sigma}_l \cdot \vec{\sigma}_{ij} \tag{26.40}$$

Corollary 26.2.46. From the expression of the Christoffel symbols we can derive an alternative expression using only the metric tensor g_{ij}:

$$\boxed{\Gamma^k_{\ ij} = \frac{1}{2}g^{kl}\left(\frac{\partial g_{il}}{\partial q^j} - \frac{\partial g_{ij}}{\partial q^l} + \frac{\partial g_{jl}}{\partial q^i}\right)} \tag{26.41}$$

Definition 26.2.47 (Geodesic). A geodesic is a curve with zero geodesic curvature[6].

Property 26.2.48. A curve on a surface is a geodesic if and only if the tangent plane and the osculation plane are orthogonal in every point P of the surface.

Formula 26.2.49 (Differential equation for geodesic). If the curve is parametrized by arc length, then it is a geodesic if the functions $q^1(s)$ and $q^2(s)$ satisfy the following differential equation:

$$\boxed{q''^k + \Gamma^k_{\ ij}q'^i q'^j = 0} \tag{26.42}$$

[6]See definition 26.24.

26.2.7 Theorema Egregium

See also section 33.2 for a generalization to Riemannian manifolds.

Formula 26.2.50 (Codazzi-Mainardi equations).

$$\frac{\partial L_{ij}}{\partial q^k} - \frac{\partial L_{ik}}{\partial q^j} = \Gamma^l{}_{ik} L_{lj} - \Gamma^l{}_{ij} L_{lk} \tag{26.43}$$

Definition 26.2.51 (Riemann curvature tensor).

$$\boxed{R^l{}_{ijk} = \frac{\partial \Gamma^l{}_{ik}}{\partial q^j} - \frac{\partial \Gamma^l{}_{ij}}{\partial q^k} + \Gamma^s{}_{ik}\Gamma^l{}_{sj} - \Gamma^s{}_{ij}\Gamma^l{}_{ks}} \tag{26.44}$$

Formula 26.2.52 (Gauss' equations).

$$R^l{}_{ijk} = L_{ik}L^l_j - L_{ij}L^l_k \tag{26.45}$$

Theorem 26.2.53 (Theorema Egregium). *The Gaussian curvature*[7] *K is completely determined by the metric tensor g_{ij} and its derivatives:*

$$\boxed{K = \frac{R^l{}_{121} g_{l2}}{g_{11}g_{22} - g_{12}^2}} \tag{26.46}$$

Remark. This theorem is remarkable due to the fact that the coefficients L_{ij}, which appear in the general formula of the Gaussian curvature, cannot be expressed in terms of the metric tensor.

Property 26.2.54. From the condition of isometries 26.2.11 and the previous theorem it follows that if two surfaces are connected by an isometric map, the corresponding points in Σ and Σ' have the same Gaussian curvature.

Corollary 26.2.55. There exists no isometric projection from the sphere to the plane. This also implies that a perfect (read: isometric) map of the Earth can not be created.

[7] See formula 26.33.

Chapter 27

Manifolds

References for this chapter (and Part VI in general) are [12--16].

27.1 Charts

Definition 27.1.1 (Chart). Let M be a set. Let U be an open subset of M and let O be an open subset of $\mathbb{R}^n$. Let $\varphi : U \to O$ be a homeomorphism. The pair (U, φ) is called a chart on M.

Definition 27.1.2 (Transition map). Let (U_1, φ_1) and (U_2, φ_2) be two charts in $\mathcal{A}$. The mapping $\varphi_1^{-1} \circ \varphi_2$ is called a transition map.

If $\varphi_1^{-1} \circ \varphi_2$ is continuous then the charts are said to be C^0-compatible. However the composition of any two continuous functions is also continuous so it follows that every two charts on a topological manifold are C^0-compatible.

Definition 27.1.3 (Atlas). Let M be a set. Let $\{(U_i, \varphi_i)\}_i$ be a set of (pairwise) $\diamond$-compatible charts (where $\diamond$ denotes any compatibility relation) such that $\bigcup_i U_i = M$. This set of charts is called a $\diamond$-atlas on M. From the remark on C^0-compatibility of charts in previous definition it is then obvious that every atlas is a C^0-atlas.

Definition 27.1.4 (Maximal Atlas). Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be two atlasses covering the same set M. If $\mathcal{A}_1 \cup \mathcal{A}_2 = \mathcal{A}$ is again an atlas then the atlasses are said to be equivalent or compatible. The largest such union is called a maximal atlas.

Definition 27.1.5 (Manifold). A set M equipped with a maximal C^0-atlas $\mathcal{A}$ is called a topological manifold. An alternative definition (often used in topology) is that of a locally Euclidean Hausdorff space. The topology on M is given by the collection of open sets contained in the charts.

Remark. In the literature second-countability is often added to the definition of a topological manifold. This ensures that the space has (among others) the property of paracompactness and hence lends itself to the construction of partitions of unity (necessary for the introduction of integration theory).

For an alternative definition of manifolds in the context of *diffeological spaces*, see definition 27.7.4.

Definition 27.1.6 (C^k-manifold). If all transition maps are C^k-diffeomorphisms then the manifold is called a C^k-manifold. A C^∞-manifold is also called a **smooth manifold**.

Definition 27.1.7 (Structure sheaf). Let M be a C^k-manifold. The structure sheaf $\mathcal{O}_M$ is defined as the sheaf that assigns to every open set $U \subseteq M$ the set of C^k-functions $f : U \to \mathbb{R}$.

Generally we define for all $j \leq k$ the sheaf $\mathcal{O}_M^j$ as the sheaf that assigns to every open set $U \subseteq M$ the set of C^j-functions $f : U \to \mathbb{R}$.

Theorem 27.1.8 (Whitney). *Every C^k-atlas contains a C^∞-atlas. Furthermore, if two C^k-atlasses contain the same C^∞-atlas then they are identical. It follows that every differentiable manifold is automatically smooth.*

Theorem 27.1.9 (Radó-Moise). *In the dimensions $1, 2$ and 3 there exists for every topological manifold a unique smooth structure.*

Theorem 27.1.10. *For dimensions higher than 4, there exist only finitely many distinct smooth structures.*

Remark. In $\dim M = 4$ there are only partial results. For non-compact manifolds there exist uncountably many distinct smooth structures. For compact manifolds there exists no complete characterization.

Formula 27.1.11 (Smooth[1]function). Let $f : M \to N$ be a function between two smooth manifolds. f is said to be smooth if there exist charts (U, φ) and (V, ψ) for M and N with $f(U) \subseteq V$ such that the function

$$f_{\varphi\psi} = \psi \circ f \circ \varphi^{-1} \tag{27.1}$$

is smooth on $\mathbb{R}^n$.

Remark. The function $f_{\varphi\psi}$ in equation 27.1 is called the **local representation** of f.

Notation 27.1.12. The set of all C^∞ functions on a manifold M defined on a neighbourhood of $m \in M$ is denoted by $C_m^\infty(M)$. This set forms a commutative unital ring when equipped with the usual sum and product (composition) of functions.

Definition 27.1.13 (Good cover[2]). Let M be an n-dimensional manifold with an open cover $\mathcal{U} = \{U_i\}_{i \in I}$. The cover $\mathcal{U}$ is called a good cover if every non-empty finite intersection $U_{i_1} \cap ... \cap U_{i_k}$ is diffeomorphic to $\mathbb{R}^n$. If M admits a finite good cover then M is said to be of **finite type**.

Property 27.1.14. Every smooth manifold admits a good cover. Furthermore if the manifold is compact, then it admits a finite good cover.

27.2 Tangent vectors

Definition 27.2.1 (Tangent vector). Let M be a smooth manifold and $p \in M$. Let $f, g : M \to \mathbb{R} \in C_p^\infty(M)$. A tangent vector on M is a differential operator v_p satisfying the following properties:

1. Linearity: $v_p(af + g) = av_p(f) + v_p(g)$
2. Leibniz property: $v_p(fg) = f(p)v_p(g) + g(p)v_p(f)$

Maps with these properties are also called **derivations**[3].

Property 27.2.2. For every constant function $c : p \mapsto c$ we have:

$$v_p(c) = 0 \tag{27.2}$$

[1] In this definition one can replace 'smooth' by 'C^k-differentiable'.

[2] Sometimes called a **nice cover**.

[3] Generally, every operation that satisfies the Leibniz property is called a derivation.

Definition 27.2.3 (Tangent space). Following from the previous definition, we can construct a tangent (vector) space T_pM in each point $p \in M$. The basis vectors are given by:

$$\boxed{\left.\frac{\partial}{\partial q^i}\right|_p : C_p^\infty(M,\mathbb{R}) \to \mathbb{R} : f \mapsto \frac{\partial}{\partial q^i}\left(f \circ \varphi^{-1}\right)(\varphi(p))} \tag{27.3}$$

where (U, φ) is a coordinate chart such that $p \in U$ and $(q^1, ..., q^n)$ are local coordinates.

Remark 27.2.4. Due to the explicit dependence of the tangent vectors on the point $p \in M$, it is clear that for curved manifolds the tangent spaces belonging to different points will not be the same.

Property 27.2.5. From the above tangent space construction it follows that:

$$\boxed{\dim(T_pM) = \dim(M)} \tag{27.4}$$

This also implies that the tangent spaces over two distinct points $p, q \in M$ are isomorphic.

Definition 27.2.6 (Curve). A smooth function $\gamma : \mathbb{R} \to M$ with $\gamma(0) = m$ is called a smooth curve through $m \in M$.

Alternative Definition 27.2.7 (Tangent space). The alternative construction goes as follows. Let (U, φ) be a chart for the point $p \in M$. Two smooth curves γ_1, γ_2 through $p \in M$ are said to be tangent at p if:

$$\frac{d(\varphi \circ \gamma_1)}{dt}(0) = \frac{d(\varphi \circ \gamma_2)}{dt}(0) \tag{27.5}$$

or equivalently, if their local representatives are tangent in 0. This relation imposes an equivalence
def relation[4] on the set of smooth curves through p. One then defines the tangent space at p as the set of equivalence classes of tangent curves through p. Explicitly these equivalence classes are constructed as follows:

We can define the following tangent vector to the curve $c(t)$ through p as:

$$\boxed{v_p(f) = \left.\frac{d(f \circ c)}{dt}\right|_{t=0}} \tag{27.6}$$

Applying the chain rule gives us

$$v_p(f) = \frac{\partial(f \circ \varphi^{-1})}{\partial q^i}(\varphi(p))\frac{dq^i}{dt}(0) \tag{27.7}$$

where $q^i = (\varphi \circ c)^i$. The first factor depends only on the point p and the second factor is equal for all tangent curves through p. We thus see that tangent curves define the same tangent vector.

The proof that both definitions of the tangent space are in fact equivalent is given in the appendices.

[4]The relation is well-defined (under a change of chart) because the transition maps (and their Jacobian matrices) are invertible and thus non-singular.

27.3 Submanifolds

27.3.1 Immersions and submersions

Definition 27.3.1 (Immersion). Let $f : M \to N$ be a differentiable function between smooth manifolds. f is called an immersion if its derivative[5] is everywhere injective, or equivalently if its derivative has maximal rank[6] everywhere:

$$\mathrm{rk}_p(f) = \dim(M), \forall p \in M \tag{27.8}$$

Definition 27.3.2 (Critical point). A point $m \in \mathrm{dom}(f)$ is said to be critical if $T_m f$ is not surjective. The image of a critical point is called a critical value.

Definition 27.3.3 (Non-degenerate critical point). A critical point $p \in M$ of a function f is said to be non-degenerate if the Hessian of f is non-singular at p.

Property 27.3.4. A point $m \in \mathrm{dom}(f)$ is critical if and only if there exists a chart $U \ni m$ for which $\frac{\partial f}{\partial x^i}(m) = 0$.

Theorem 27.3.5 (Sard). *Consider a map $\psi : M \to N$, where $\dim M = m$ and $\dim N = n$. Set $k_0 = \max\{1, m - n + 1\}$. If ψ is of class C^k with $k \geq k_0$ then the set of critical values of ψ has Lebesgue measure 0.*

Definition 27.3.6 (Regular point). A regular point of f is a point $x \in M$ such that $T_x f$ is surjective.

Definition 27.3.7 (Regular value). Let $f : M \to N$ be a differentiable map between smooth manifolds. A point $y \in N$ is called a **regular value** if every point in the preimage $f^{-1}(y)$ is a regular point or equivalently if it is not a critical value.

Corollary 27.3.8. It follows from property 27.3.4 that a point $m \in \mathrm{dom}(f)$ is regular if and only if $\frac{\partial f}{\partial x^i}(m) \neq 0$ in all charts $U \ni m$.

Definition 27.3.9 (Submersion). Let $f : M \to N$ be a differentiable map between smooth manifolds. A map f is called a submersion if all $x \in M$ are regular, or equivalently if

$$\mathrm{rk}_p(f) = \dim(N), \forall p \in M \tag{27.9}$$

Definition 27.3.10 (Embedding). A differentiable function between smooth manifolds is called a smooth embedding if its both an immersion and an embedding in the topological sense 7.3.12. This implies that the submanifold topology coincides with the subspace topology 7.1.

27.3.2 Submanifolds

Definition 27.3.11 (Submanifold). Let M be a manifold. A subset $N \subset M$ is called a submanifold of M if N, equipped with the subspace topology, is a topological manifold on its own.

Definition 27.3.12 (Immersive submanifold). Let M, N be smooth manifolds with $N \subset M$. N is said to be an immersive submanifold of M if the inclusion $i : N \hookrightarrow M$ is an immersion.

[5] This is formally defined in 30.1. For now it is the map represented by the Jacobian matrix.
[6] See definition 30.1.9.

Theorem 27.3.13 (Submersion theorem[7]). *Consider a smooth map $f : M_1 \to M_2$ between smooth manifolds. Let $y \in M_2$ be a regular value. Then $N = f^{-1}(y)$ is a submanifold of M_1 with codimension* $\dim(M_2)$.

Definition 27.3.14 (Embedded submanifold). Let M be a manifold. A subset N is an embedded[8] or **regular submanifold** if the inclusion map $f : M \hookrightarrow N$ is a smooth embedding.

Definition 27.3.15 (Slice). Let $m < n$ be two positive integers. The space $\mathbb{R}^m$ can be viewed as a subspace of $\mathbb{R}^n$ by identifying them in the following way:

$$\mathbb{R}^m \cong \mathbb{R}^m \times \{\underbrace{0, ..., 0}_{n-m}\} \overset{\iota}{\hookrightarrow} \mathbb{R}^m \times \mathbb{R}^{n-m} \cong \mathbb{R}^n \tag{27.10}$$

where $\iota : (x_1, ..., x_m) \mapsto (x_1, ..., x_m, \underbrace{0, ..., 0}_{n-m})$ is the canonical inclusion map.

Alternative Definition 27.3.16. A k-dimensional embedded manifold N of M can be defined equivalently as a subset of M such that there exists a positive integer k and such that for every point $p \in N$ there exists a chart (U, φ) that satisfies

$$\varphi(U \cap N) = \varphi(U) \cap (\mathbb{R}^k \times \{ \underbrace{0, ..., 0}_{\dim(M)-k} \}) \tag{27.11}$$

The set $U \cap N$ is called a slice of (U, φ) in analogy with the previous definition of a (standard) slice.

Definition 27.3.17 (Closed embedded manifold). Let N be an immersed submanifold of M. If the inclusion map $i : N \hookrightarrow M$ is closed, then N is a (closed) embedded manifold.

Definition 27.3.18 (Transversal intersection). Consider a smooth manifold M. Two submanifolds X, Y are said to be transversal (or intersect transversally) if at each intersection point p the following relation holds:

$$T_pX + T_pY = T_pM \tag{27.12}$$

If the dimensions of X and Y are complementary (in M) then the sum becomes a direct sum. If two submanifolds do not intersect at all then they are vacuously transversal (independent of their dimension).

Property 27.3.19. The codimension of the transversal intersection of two submanifolds is equal to the sum of the codimensions of the intersecting submanifolds. It follows that if the submanifolds have complentary dimensions the intersection consists of isolated points.

27.4 Manifolds with boundary

Definition 27.4.1 (Manifold with boundary). Let $\mathbb{H}^n$ denote the upper half space, i.e.:

$$\mathbb{H}^n \cong \mathbb{R}^{n-1} \times \mathbb{R}^+ = \{(x_1, ..., x_n) | x_n \geq 0\} \subset \mathbb{R}^n \tag{27.13}$$

An n-dimensional manifold with boundary is then given by a set M together with a maximal atlas consisting of (regular) charts (U, φ) such that U is diffeomorphic to $\mathbb{R}^n$, these points are called **interior points**, and (boundary) charts (V, ϕ) such that V is diffeomorphic to $\mathbb{H}^n$, these points are called **boundary points**.

[7] Also called the **regular value theorem**.

[8] An immersed submanifold is defined analogously. The requirement of the inclusion map being a smooth embedding is relaxed to it being an (injective) immersion. However the submanifold topology will no longer coincide with the subspace topology.

Remark 27.4.2 (Manifold boundary). The boundary ∂M, consisting of all boundary points of M as defined in the above definition, should not be confused with the topological boundary of M. In general these are different sets. Similarly, the interior $\text{Int}(M) = M \backslash \partial M$, in the sense of manifolds, should not be confused with the topological interior.

Property 27.4.3. Let M be an n-dimensional manifold with boundary. Let (U, φ) be a chart for $p \in \partial M$. Then

$$\varphi(p) \in \partial \mathbb{H}^n = \{(x_1, ..., x_n) | x_n = 0\} \tag{27.14}$$

Definition 27.4.4 (Manifold with corners). Analogous to the definition of a manifold with boundaries one can define a manifold wih corners using *corner charts* of the form

$$\varphi : U \to \mathbb{R}^k \times (\mathbb{R}^+)^l$$

In contrast to the case of manifolds with boundary one does need to add an extra requirement when working with higher order corners: For every two charts (U, φ) and (V, ψ) the transition function should preserve the corners:

$$\varphi \circ \psi^{-1}(V \cap \{0\} \times \mathbb{R}^k \subset \{0\} \times \mathbb{R}^k$$

Remark 27.4.5. In the topological setting every manifold with corners (ven higher order ones) is homeomorphic to a manifold with boundary. However when working with smooth structures this result fails, there exists no diffeomorphism.

27.4.1 Cobordisms

Definition 27.4.6 (Cobordism). Two manifolds X, Y are said to be **cobordant** if there exists a manifold with boundary M such that $\partial M = X \sqcup Y$. The manifold M is said to be a cobordism[9] between X and Y.

Remark. In the category of oriented manifolds one can also define a cobordism, but there the manifolds X, Y should respect the orientation of ∂M.

27.5 Morse theory

27.5.1 Morse functions

Definition 27.5.1 (Morse function). Let M be a smooth manifold. A smooth function $f \in C^\infty(M)$ is called a Morse function if it has no degenerate critical point.

Property 27.5.2. The set of Morse functions is open and dense in the C^2-topology[10].

27.6 Surgery theory

Definition 27.6.1 (Dehn twist). Consider an orientable surface M together with a simple closed curve c. A tubular neighbourhood T of c is homeomorphic to an annulus and hence allows a parametrization $(e^{i\alpha}, t)$ where $\alpha \in [0, 2\pi[, t \in [0, 1]$. A Dehn twist about c is given by an automorphism which is given by $(e^{i\alpha}, t) \mapsto (e^{i(\alpha + 2\pi t)}, t)$ on T and restricts to the identity outside of it.

[9] Some authors use the terms *bordism* and *bordant* in this context.

[10] See section 29.2 on jet spaces.

27.7 Diffeological spaces

Definition 27.7.1 (Diffeological space). Let X be a set. A diffeology $\mathcal{D}$ on X is defined as a collection of maps $f : U \subseteq \mathbb{R}^n \to X$, called **plots**, satisfying the following conditions:

- If f is constant then $f \in \mathcal{D}$.
- If every point in $\mathrm{dom}(f)$ has a neighbourhood V on which $f|_V \in \mathcal{D}$ then $f \in \mathcal{D}$.
- If $g \in \mathcal{D}$ and $h : W \subseteq \mathbb{R}^m \to \mathrm{dom}(g)$ is smooth then $g \circ h \in \mathcal{D}$.

where U, V and W are open. The set X can be turned into a topological space by equipping it with the **D-topology**, i.e. the final topology with respect to $\mathcal{D}$.

Remark 27.7.2. The domain of different plots can be subsets of different Euclidean spaces $\mathbb{R}^m$ and $\mathbb{R}^n$.

Definition 27.7.3 (Differentiable map). Let $(X, \mathcal{D})$ and $(Y, \mathcal{D}')$ be diffeological spaces. A map $g : X \to Y$ is said to be differentiable if for every $f \in \mathcal{D}$ the composition $g \circ f \in \mathcal{D}'$.

The diffeological spaces together with their differentiable morphisms form a category DiffeologicalSp.

Alternative Definition 27.7.4. Let M be a diffeological space[11]. M is an n-manifold if it is locally diffeomorphic to $\mathbb{R}^n$.

27.8 Path groupoid

Definition 27.8.1 (Thin homotopy). Let M be a smooth manifold. A homotopy $H : [0,1]^2 \to M$ is called thin if it is smooth and if it pulls back[12] every two-form to 0:

$$\forall \omega \in \Omega^2(M) : H^*\omega = 0 \tag{27.15}$$

Definition 27.8.2 (Lazy path[13]). Let M be a smooth manifold. A path $f : [0,1] \to M$ is said to be lazy if it is locally constant on a neighbourhood of 0 and 1.

Definition 27.8.3 (Path groupoid). Let M be a smooth manifold. The path groupoid $P_1(M)$ is the groupoid which has the points of M as objects and the thin homotopy classes of lazy paths on M with fixed endpoints as morphisms.

Remark. The restriction to lazy paths is required to ensure the smoothness of composite paths. The quotient by thin homotopies is required to ensure the validity of the associativity and invertibility properties.

[11] See definition 27.7.1.
[12] See section 30.4.
[13] Also called a path with **sitting instants**.

Chapter 28

Lie groups and Lie algebras

References for this chapter are [8, 10].

28.1 Lie groups

Definition 28.1.1 (Lie group). A Lie group is a group that is also a differentiable manifold such that both the multiplication and inversion are smooth functions.

Definition 28.1.2 (Lie subgroup). A subset of a Lie group is a Lie subgroup if it is both a subgroup and a closed submanifold.

Theorem 28.1.3 (Closed subgroup theorem[1]). *If H is a closed[2] subgroup of a Lie group G then H is a Lie subgroup of G.*

Property 28.1.4. Let G be a connected Lie group. Every neighbourhood U_e of the identity e generates G, i.e. every element $g \in G$ can be written as a word in U_e.

Definition 28.1.5 (Isogeny). Let G, H be two Lie groups. G and H are said to be isogenous if one is a covering space[3] of the other. The covering map is then called an isogeny between G and H.

28.1.1 Left invariant vector fields

Definition 28.1.6 (Left Invariant Vector Field (LIVF)). Let G be a Lie group. Let X be a vector field on G. X is left invariant if the following equivariance relation holds for all $g \in G$:

$$L_{g,*}X(h) = X(g \cdot h) \tag{28.1}$$

where L_g denotes the left action map associated with g.

Property 28.1.7. The set $\mathcal{L}(G)$ of LIVF's on a Lie group G is a vector space over $\mathbb{R}$.

Property 28.1.8. The map $L_{g,*}$ is an isomorphism for every $g \in G$. It follows that a LIVF is uniquely determined by its value at the identity of G. Furthermore, for every $v \in T_e(G)$, there exists a LIVF $X \in \mathcal{L}(G)$ such that $X(e) = v$ and this mapping is an isomorphism from $T_e(G)$ to $\mathcal{L}(G)$.

[1] Sometimes called **Cartan's theorem**.
[2] With respect to the group topology on G.
[3] See definition 7.3.15.

28.1.2 One-parameter subgroups

Definition 28.1.9 (One-parameter subgroup). A one-parameter (sub)group is a Lie group morphism $\Phi : \mathbb{R} \to G$ from the additive group of real numbers to a Lie group G.

Property 28.1.10. Let $\Phi : \mathbb{R} \to G$ be a one-parameter subgroup of G. Let $\Psi : G \to H$ be a continuous group morphism. Then $\Psi \circ \Phi : \mathbb{R} \to H$ is a one-parameter subgroup of H.

Property 28.1.11. All LIVF's X are complete[4]. Hence for every LIVF X we can find an integral curve γ_X with initial condition $\gamma_X(0) = e$ for which the maximal flow domain[5] $D(X)$ is $]-\infty, +\infty[$. This implies that the associated flow σ_t determines a one-parameter subgroup of G. Conversely, for every one-parameter subgroup $\phi(t)$ we can construct a LIVF $X = \phi'(0)$. This correspondence is a bijection.

28.1.3 Cocycles

Definition 28.1.12 (Cocycle). Let M be a smooth manifold and G a Lie group. A cocycle on M with values in G is a family of smooth functions $g_{ij} : U_i \cap U_j \to G$ that satisfy the following condition:

$$g_{ij} = g_{ik} \circ g_{kj} \tag{28.2}$$

Property 28.1.13. Let $\{g_{ij}\}_{i,j}$ be a cocycle on M. We have the following properties:

- $g_{ii}(x) = \mathbb{1}_M$
- $g_{ij}(x) = (g_{ji}(x))^{-1}$

for all $x \in M$.

28.2 Lie algebras

There are two ways to define a Lie algebra. The first one is a stand-alone definition using a vector space equipped with a multiplication operation. The second one establishes a direct relation between Lie groups (see 28.1.1) and real Lie algebras.

28.2.1 Definitions

Definition 28.2.1 (Lie algebra). Let V be a vector space equipped with a binary operation $[\cdot, \cdot] : V \times V \to V$ is a Lie algebra if the Lie bracket $[\cdot, \cdot]$ satisfies the following conditions:

1. Bilinearity: $[ax + y, z] = a[x, z] + [y, z]$
2. Alternativity: $[v, v] = 0$
3. Jacobi identity: $[a, [b, c]] + [b, [c, a]] + [c, [a, b]] = 0$

Remark 28.2.2. Note that often the alternativity condition is replaced by an antisymmetry condition. However this is only well-defined in a field of characteristic $\neq 2$, for char $= 2$ we only have that alternativity implies antisymmetry. As we mostly work over fields such as $\mathbb{R}$ or $\mathbb{C}$ this won't pose a problem and hence we will always use the antisymmetry condition as the defining property.

[4] See definition 30.3.9.
[5] See definition 30.3.7.

Definition 28.2.3 (Lie algebra of LIVF's). Consider the vector space $\mathcal{L}(G)$ of LIVF's on a Lie group G. Using property 30.22 we can show that the commutator (Lie bracket) also defines a LIVF on G. It follows that $\mathcal{L}(G)$ is closed under Lie brackets and hence is a Lie algebra.

Alternative Definition 28.2.4 (Lie algebra of Lie group). Let G be a Lie group. The tangent space $\mathfrak{g} := T_eG$ has the structure of a Lie algebra where the Lie bracket is induced by the commutator of vector fields 30.22 in the following way:

$$[\![A, B]\!] := [l_{g,*}A, l_{g,*}B]|_{g=e} \tag{28.3}$$

where $A, B \in T_eG$ and where $[\cdot, \cdot]$ is the Lie bracket on $\mathcal{L}(G)$. This induces an isomorphism of Lie algebras: $\mathfrak{g} \cong_{\text{Lie}} \mathcal{L}(G)$.

Notation 28.2.5. Lie algebras are generally denoted by fraktur symbols. For example, the Lie algebra associated with the Lie group G is often denoted by $\mathfrak{g}$.

Theorem 28.2.6 (Ado). *Every finite-dimensional Lie algebra can be embedded as a subalgebra of* $\mathfrak{gl}_n = \mathrm{M}_n$.

Theorem 28.2.7 (Lie's third theorem). *Every finite-dimensional Lie algebra* $\mathfrak{g}$ *is the Lie algebra of a unique simply-connected Lie group* G.

Definition 28.2.8 (Lie algebra morphism). A map $\Phi : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra morphism if it satisfies following condition

$$\Phi([X, Y]) = [\Phi(X), \Phi(Y)] \tag{28.4}$$

for all $X, Y \in \mathfrak{g}$.

Property 28.2.9 (Homomorphisms theorem[6]). Let G, H be Lie groups with G simply-connected. If a linear map $\Phi : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra morphism then there exists a unique Lie group morphism $\phi : G \to H$ such that $\Phi = \phi_*$.[7]

28.2.2 Exponential map

Formula 28.2.10 (Exponential map). Let $X \in \mathfrak{g}$ be a LIVF on G. We define the exponential map $\exp : \mathfrak{g} \to G$ as:

$$\boxed{\exp(X) := \gamma_X(1)} \tag{28.5}$$

where γ_X is the associated one-parameter subgroup defined in property 28.1.11.

Property 28.2.11. The exponential map is the unique map $\mathfrak{g} \to G$ such that $\exp(0) = e$ and for which the restrictions to the lines through the origin in $\mathfrak{g}$ are one-parameter subgroups of G.

Corollary 28.2.12. Because the identity element $\mathbb{1}_{\mathfrak{g}} = (\exp_*)_e$ is an isomorphism, the inverse function theorem 30.1.10 implies that the image of exp will contain a neighbourhood of the identity $e \in G$. If G is connected then property 28.1.4 implies that exp generates all of G.

Together with the property that $\psi \circ \exp = \exp \circ \psi_*$ for every Lie group morphism $\psi : G \to H$ it follows that *if* G *is connected, a Lie group morphism* $\psi : G \to H$ *is completely determined by its differential* ψ_* *at the identity* $e \in G$.

[6] See also formula 28.3.8.

[7] The converse is trivial: every Lie group morphism induces a Lie algebra morphism through its differential.

Example 28.2.13 (Matrix Lie groups). For matrix Lie groups we define the classic matrix exponential:

$$e^{tX} = \sum_{k=0}^{+\infty} \frac{(tX)^k}{k!} \tag{28.6}$$

This operation defines a curve $\gamma(t)$ which can be used as a one-parameter subgroup on G. It should be noted that this formula converges for every $X \in M_{m,n}$ and is invertible with the inverse given by $\exp(-X)$. Using Ado's theorem 28.2.6 one can then use this matrix exponential to represent the exponential map for any (finite-dimensional) Lie algebra.

Property 28.2.14. Let G be a compact Lie group. The exponential map is surjective. However, because the asocciated Lie algebra $\mathfrak{g}$ is non-compact, the exponential map cannot be homeomorphic and hence cannot be injective.

28.2.3 Structure

Definition 28.2.15 (Structure constants). As Lie algebras are closed under Lie brackets, every Lie bracket can be expanded in term of a basis $\{X_k\}_{k\in I}$ as follows:

$$[X_i, X_j] = \sum_{k\in I} {c_{ij}}^k X_k \tag{28.7}$$

where the factors ${c_{ij}}^k$ are called the structure constants[8] of the Lie algebra.

Property 28.2.16. Two Lie algebras $\mathfrak{g}, \mathfrak{h}$ are isomorphic if one can find bases $\mathcal{B}$ for $\mathfrak{g}$ and $\mathcal{C}$ for $\mathfrak{h}$ such that the associated structure constants are equal for all indices i, j and k.

Formula 28.2.17 (Baker-Campbell-Hausdorff formula). This formula is the solution of the equation

$$Z = \log(\exp(X)\exp(X)) \tag{28.8}$$

for $X, Y \in \mathfrak{g}$. The solution is given by following formula

$$e^X e^Y = \exp\left(X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] - \frac{1}{12}[Y,[X,Y]] + \cdot\right) \tag{28.9}$$

One should note that this formula will only converge if X, Y are sufficiently small (for matrix Lie algebras this means that $||X|| + ||Y|| < \frac{\ln(2)}{2}$ under the Hilbert-Schmidt norm 19.43). Due to the closure under commutators (see Lie algebra definition) the exponent in the BCH formula is also an element of the Lie algebra. So the formula gives an expression for Lie group multiplication in terms of Lie algebra elements (whenever the formula converges).

Corollary 28.2.18 (Lie product formula[9]). Let $\mathfrak{g}$ be a Lie algebra. The following formula applies to any $X, Y \in \mathfrak{g}$:

$$e^{X+Y} = \lim_{n\to+\infty} \left(e^{\frac{X}{n}} e^{\frac{Y}{n}}\right)^n \tag{28.10}$$

Definition 28.2.19 (Normalizer). The normalizer of a Lie subalgebra $\mathfrak{h}$ is the space of elements $g \in \mathfrak{g}$ that satisfy $[g, \mathfrak{h}] \subseteq \mathfrak{h}$.

[8] Note that these constants are basis-dependent.

[9] Also called the Lie-Trotter formula.

28.2.4 Examples

Example 28.2.20. The cross product $\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ turns $\mathbb{R}^3$ into a Lie algebra.

Example 28.2.21. An interesting example is the Lie algebra associated to the Lie group of invertible complex[10] matrices $GL(n, \mathbb{C})$. This Lie group is a subset of its own Lie algebra $\mathfrak{gl}(n, \mathbb{C}) = M_n(\mathbb{C})$. It follows that for every $A \in GL(n, \mathbb{C})$ and every $B \in \mathfrak{gl}(n, \mathbb{C})$ the following equality holds:

$$L_{A,*}(B) = L_A(B) \tag{28.11}$$

Corollary 28.2.22. By noting that the endomorphism ring $\mathrm{End}(V)$ of an n-dimensional vector space V is given by the matrix ring $M_n(K)$, we see that $\mathrm{End}(V)$ also forms a Lie algebra when equipped with the commutator of linear maps.

Following two examples of Lie algebras can be checked using condition 19.32:

Example 28.2.23 (Lie algebra of $\mathrm{O}(3)$). The set of 3×3 anti-symmetric matrices. It is also important to note that $\mathfrak{o}(3) = \mathfrak{so}(3)$. The structure constants of this Lie algebra are given by 20.42, i.e. $C_{ijk} = \varepsilon_{ijk}$.

Example 28.2.24 (Lie algebra of $\mathrm{SU}(2)$). The set of 2×2 traceless anti-Hermitian matrices. This result can be generalized to arbitrary $n \in \mathbb{N}$.

Example 28.2.25. To compute the Lie bracket in the Lie algebra $\mathfrak{sl}(2, \mathbb{C}) = T_e(\mathrm{SL}(2, \mathbb{C}))$ we need to find the action of $l_{g,*}$ on any vector $Y \in \mathfrak{sl}(2, \mathbb{C})$. This is given by:

$$l_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right),*} \left(\frac{\partial}{\partial x^i}\bigg|_e \right) = \begin{pmatrix} a & 0 & b \\ -b & a & 0 \\ c & 0 & \frac{1+bc}{a} \end{pmatrix}_i^m \frac{\partial}{\partial x^m}\bigg|_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)} \tag{28.12}$$

where we used the coordinate chart (U, ϕ) defined by:

$$U = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{C}) : a \neq 0 \right\}$$

and

$$\phi : U \to \mathbb{C}^3 : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto (a, b, c)$$

One can then use this formula to work out the Lie bracket of the basis vectors $X_i = \frac{\partial}{\partial x^i}\big|_e$ to obtain the structure constants:

$$\begin{cases} [\![X_1, X_2]\!] = 2X_2 \\ [\![X_1, X_3]\!] = -2X_3 \\ [\![X_2, X_3]\!] = X_1 \end{cases} \tag{28.13}$$

28.2.5 Solvable Lie algebras

Definition 28.2.26 (Derived algebra). Let $\mathfrak{g}$ be a Lie algebra with Lie bracket $[\cdot, \cdot]$. The derived Lie algebra is defined as follows:

$$[\mathfrak{g}, \mathfrak{g}] = \{[x, y] : x, y \in \mathfrak{g}\} \tag{28.14}$$

[10] As usual, this result is also valid for real matrices.

Definition 28.2.27 (Solvable Lie algebra). Consider the sequence of derived Lie algebras

$$g \geq [\mathfrak{g}, \mathfrak{g}] \geq [[\mathfrak{g}, \mathfrak{g}], [\mathfrak{g}, \mathfrak{g}]] \geq \cdots \tag{28.15}$$

If this sequence ends in the zero-space then the Lie algebra $\mathfrak{g}$ is said to be solvable.

Definition 28.2.28 (Radical). Let $\mathfrak{g}$ be a Lie algebra. The radical of $\mathfrak{g}$ is the largest solvable ideal in $\mathfrak{g}$.

28.2.6 Simple Lie algebras

Definition 28.2.29 (Direct sum). The direct sum of two Lie algebras $\mathfrak{g}, \mathfrak{h}$ is defined as the direct sum in the sense of vector spaces (see 19.2.20) together with the condition

$$[x, y] = 0 \tag{28.16}$$

for all $x \in \mathfrak{g}$ and $y \in \mathfrak{h}$.

Definition 28.2.30 (Semidirect product). The semidirect product $\mathfrak{g} \ltimes \mathfrak{h}$ of two Lie algebras $\mathfrak{g}, \mathfrak{h}$ is defined as the direct sum in the sense of vector spaces (see 19.2.20) together with the condition that $\mathfrak{g}$ is an ideal of $\mathfrak{h}$ under the Lie bracket.

Definition 28.2.31 (Simple Lie algebra). A Lie algebra is said to be simple if it is non-Abelian and if it has no non-trivial ideals.

Definition 28.2.32 (Semisimple Lie algebra). A Lie algebra is said to be semisimple if it is the direct sum of simple algebras.

Theorem 28.2.33 (Levi decomposition). *Let $\mathfrak{g}$ be a finite-dimensional Lie algebra. This algebra can be decomposed as follows:*

$$\mathfrak{g} = \mathfrak{R} \ltimes (\mathfrak{L}_1 \oplus \cdots \oplus \mathfrak{L}_n) \tag{28.17}$$

where $\mathfrak{R}$ is the radical of $\mathfrak{g}$ and the algebras $\mathfrak{L}_i$ are simple subalgebras.

Definition 28.2.34. The semisimple subalgebra $\mathfrak{L}_1 \oplus \cdots \oplus \mathfrak{L}_n$ in the Levi decomposition of $\mathfrak{g}$ is called the **Levi subalgebra** or **Levi factor** of $\mathfrak{g}$.

28.3 Representation theory

28.3.1 Lie groups

Definition 28.3.1 (Representation of Lie groups). Let G be a Lie group and let V be a vector space. A representation of G on V is a Lie group morphism $\rho : G \to \mathrm{GL}(V)$.

Definition 28.3.2 (Adjoint representation of Lie groups). Let G be a Lie group. Consider the conjugation map $\Psi_g : h \mapsto ghg^{-1}$. The adjoint representation of G is defined by the differential of the conjugation $T_e\Psi_g$. For matix Lie groups this becomes:

$$\mathrm{Ad}_g : T_eG \to T_eG : X \mapsto gXg^{-1} \tag{28.18}$$

The adjoint representation is a representation of G on its own tangent space $T_eG \equiv \mathfrak{g}$.

28.3.2 Lie algebras

Definition 28.3.3 (Representation of Lie algebras). Let $\mathfrak{g}$ be a Lie algebra and let V be a vector space. A representation of $\mathfrak{g}$ on V is a Lie algebra morphism $\rho : \mathfrak{g} \to \mathrm{End}(V)$.

Formula 28.3.4 (Adjoint representation of Lie algebras). Using the fact that the adjoint representation of Lie groups is smooth we can define the adjoint representation of Lie algebras as:

$$\mathrm{ad}_X := T_e(\mathrm{Ad}_g) \tag{28.19}$$

where $g = e^{tX}$. Explicitly, let $\mathfrak{g}$ be a Lie algebra. For every element $X \in \mathfrak{g}$ the adjoint map is given by:

$$\mathrm{ad}_X(Y) = [X, Y] \tag{28.20}$$

Property 28.3.5. The adjoint representation ad_X is faithful.

Property 28.3.6. Given the antisymmetry of the Lie bracket, the Jacobi identity is equivalent to ad: $\mathfrak{g} \to \mathrm{End}(\mathfrak{g})$ being a Lie algebra morphism, i.e. $\mathrm{ad}_{[X,Y]} = [\mathrm{ad}_X, \mathrm{ad}_Y]$.

Formula 28.3.7. Let $\{e_i\}_{i \leq n}$ be a basis of a Lie algebra $\mathfrak{g}$. The structure coefficients can be calculated using the adjoint map as follows:

$$(\mathrm{ad}_{e_i})^j_k = C_{ik}{}^j \tag{28.21}$$

Formula 28.3.8 (Induced morphism). Let $\phi : G \to H$ be a Lie group morphism[11] with G connected and simply-connected. This morphism induces a a Lie algebra morphism[12] $\Phi : \mathfrak{g} \to \mathfrak{h}$ given by:

$$\Phi(X) = \frac{d}{dt}\phi\left(e^{tX}\right)\Big|_{t=0} \tag{28.22}$$

or equivalently:

$$\phi\left(e^{tX}\right) = e^{t\Phi(X)} \tag{28.23}$$

Example 28.3.9. The morphism induced by $\mathrm{Ad} : G \to H$ is precisely $\mathrm{ad} : \mathfrak{g} \to \mathfrak{h}$. Informally we can thus say that the infinitesimal version of the similarity transformation is given by the commutator (when $G = \mathrm{GL}_n$):

Corollary 28.3.10 (Commutator). For the general linear group GL_n the Lie bracket is given by the commutator:

$$\boxed{[X, Y] = XY - YX} \tag{28.24}$$

This follows from definition 28.20: $[X, Y] = \frac{d}{dt}\mathrm{Ad}_{\gamma(t)}(Y)\big|_{t=0}$ with $\gamma(0) = e$ and $\gamma'(0) = X$.

28.3.3 Killing form

Definition 28.3.11 (Killing form[13]). Let $\mathfrak{g}$ be a finite-dimensional Lie algebra. The Killing form on $\mathfrak{g}$ is defined as the following symmetric bilinear form[14]:

$$\boxed{K(X, Y) = \mathrm{tr}(\mathrm{ad}_X \circ \mathrm{ad}_Y)} \tag{28.25}$$

[11] Continuity (inherent to the definition of a Lie group morphism) is needed to ensure that $\phi(e^{tX})$ is also a one-parameter subgroup (see 28.1.10).

[12] See also property 28.2.9.

[13] Also called the **Cartan-Killing form**.

The trace can be found by representing the Lie algebra elements as matrices using Ado's theorem 28.2.6. From equation 28.21 we can work out the action of the Killing form on the basis $\{e_i\}_{i\leq n}$:

$$K_{ij} = C_{ik}{}^{l} C_{jl}{}^{k} \tag{28.26}$$

where $C_{ij}{}^{k}$ are the structure constants of the Lie algebra.

Theorem 28.3.12 (Cartan's criterion). *A Lie algebra is semisimple if and only if its Killing form is non-degenerate.*

Property 28.3.13. If a Lie group G is compact then the Killing form of its associated Lie algebra $\mathfrak{g}$ is negative-definite.

Corollary 28.3.14. Let G be a compact Lie group. If its Lie algebra is semisimple the Killing form K induces a metric

$$g : (X, Y) \mapsto -\operatorname{tr}(\operatorname{ad}_X, \operatorname{ad}_Y) \tag{28.27}$$

which turns the corresponding Lie group G into a Riemannian manifold.

Property 28.3.15. The adjoint map ad_Z is antisymmetric with respect to the Killing form:

$$K(\operatorname{ad}_Z X, Y) = -K(X, \operatorname{ad}_Z Y) \tag{28.28}$$

or equivalently:

$$K([X, Z], Y) = K(X, [Z, Y]) \tag{28.29}$$

Property 28.3.16. The Killing-form is Ad-invariant, i.e.

$$K(\operatorname{Ad}_g(X), \operatorname{Ad}_g(Y)) = K(X, Y) \tag{28.30}$$

for all $g \in G$. From this it follows that Ad is a map from G to the isometry group $\operatorname{Isom}(\mathfrak{g})$.

Definition 28.3.17. Let $\mathfrak{g}$ be a Lie algebra and let V be a vector space equipped with a Lie algebra representation $\rho : \mathfrak{g} \to \operatorname{End}(V)$. One can then define a Killing form associated with ρ in the following way:

$$K_\rho(X, Y) = \operatorname{tr}\Big(\rho(X) \circ \rho(Y)\Big) \tag{28.31}$$

Remark 28.3.18. This definition is a generalization of 28.25 which reduces to the Killing form K when choosing V to be $\mathfrak{g}$ in the adjoint representation.

Property 28.3.19. On a simple Lie algebra, every invariant, i.e. satisfying equation 28.28, symmetric bilinear form is a scalar multiple of the Killing form.

Example 28.3.20. For $\mathfrak{su}(n)$ the relation is given by:

$$\operatorname{tr}(XY) = 2nK(X, Y) \tag{28.32}$$

Property 28.3.21. When the Lie algebra $\mathfrak{g}$ is compact and semisimple, hence when the Killing form induces a metric, one can find a basis of $\mathfrak{g}$, constructed by orthonormalizing a given basis with respect to the Killing metric[15], such that the structure constants are invariant under cyclic permutation of the indices:

$$c_{ijk} = c_{jki} \tag{28.33}$$

A corollary of this property is also that the structure constants become totally antisymmetric.

[14] i.e. a symmetric $(0, 2)$-tensor in $\mathfrak{g}^* \otimes \mathfrak{g}^*$ (See definition 20.5.9)

[15] The proof uses the ad-invariance of the Killing form.

28.3.4 Weights, roots and Dynkin diagrams

Definition 28.3.22 (Cartan subalgebra). Let $\mathfrak{g}$ be a Lie algebra. A subalgebra $\mathfrak{h}$ is called a Cartan subalgebra if satisfies following two conditions:

- Nilpotent: Its lower central series terminates:

$$\exists n \in \mathbb{N} : \underbrace{[\mathfrak{g}, [\mathfrak{g}, [\mathfrak{g}, ...]]]}_{n \text{ times}} = 0 \tag{28.34}$$

- Self-normalizing:

$$\forall X \in \mathfrak{h} : [X, Y] \in \mathfrak{h} \implies Y \in \mathfrak{h} \tag{28.35}$$

From here one we will assume that all Lie algebras are finite-dimensional. This assumption is motivated by the following property:

Property 28.3.23. Every finite-dimensional Lie algebra contains a Cartan subalgebra.

Property 28.3.24. If $\mathfrak{g}$ is semisimple then its Cartan subalgebra is Abelian.

Construction 28.3.25. Let $\mathfrak{g}$ be a semisimple Lie algebra. A Cartan subalgebra $\mathfrak{h}$ can be constructed as follows: Choose linearly independent vectors $\{h_i\}_{i\in I}$ such that for all[16] $i, j \in I$: $[h_i, h_j] = 0$. If this set can be extended to a basis $\{h_i\}_{i\in I} \cup \{g_j\}_{j\in J}$ of $\mathfrak{g}$ such that every g_j is a non-trivial eigenvector of the adjoint map ad_{h_i} for all $i \in I$ then the algebra $\mathfrak{h} = \mathrm{span}\{h_i\}_{i\in I}$ is a Cartan subalgebra.

Definition 28.3.26 (Weight space). Let V be a representation of a Lie algebra $\mathfrak{g}$ with Cartan subalgebra $\mathfrak{h}$. Let λ be a linear functional on $\mathfrak{h}$. The weight space V_λ of V with **weight** λ is defined as follows:

$$V_\lambda = \{v \in V : h \cdot v = \lambda(H)v, \forall h \in \mathfrak{h}\} \tag{28.36}$$

Non-zero elements of a weight space are called **weight vectors**. If the representation V can be decomposed as a direct sum of weigth spaces it is called a **weight module**:

$$V = \bigoplus_{\lambda\in\mathfrak{h}^*} V_\lambda \tag{28.37}$$

In the case that V is the adjoint representation, the (non-zero) weights are called **roots**:

Definition 28.3.27 (Root). Let $\mathfrak{g}$ be a Lie algebra with Cartan subalgebra $\mathfrak{h}$. From the definition of a Cartan subalgebra it follows that for all $h \in \mathfrak{h}$:

$$[h, g_j] = \alpha_j(h)g_j \tag{28.38}$$

where $\{g_j\}_{j\in J}$ is the basis extension of $\mathfrak{g}$ with respect to $\mathfrak{h}$. Because $\alpha_j(h)$ is an eigenvalue it is an element of the base field $\mathbb{C}$ and hence we can view α_j as a linear map $\mathfrak{h} \to \mathbb{C}$, or equivalently as a weight of the adjoint representation. The non-zero linear maps are called the roots of $\mathfrak{g}$ and form the so-called **root system** Φ.

It follows that there exists a weight space decomposition of $\mathfrak{g}$:

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\lambda\in\Phi} \mathfrak{g}_\lambda \tag{28.39}$$

where the one-dimensional $\mathfrak{g}_\lambda$ are the weight space associated to the roots λ ($\mathfrak{h}$ is equal to $\mathfrak{g}_0$).

[16] The existence of such a choice, equivalent to requiring simultaneous diagonalization, is only guaranteed for semisimple Lie algebras.

Property 28.3.28. If $\alpha \in \Phi$ then $-\alpha \in \Phi$. Furthermore if $\alpha \in \Phi$ and $c\alpha \in \Phi$ then $c = \pm 1$.

We conclude that the root system Φ is not linearly independent. To introduce some kind of basis we define the following:

Definition 28.3.29 (Simple root). The set of simple roots[17] Δ is a linearly independent subset of Φ such that every element $\lambda \in \Phi$ can be written as:

$$\lambda = \pm \sum_{i}^{n} a_i \lambda_i \tag{28.40}$$

where $a_i \in \mathbb{N}$ and $\lambda_i \in \Delta$. This definition requires the expansion coefficients a_i of a certain root λ to be either all positive or all negative.

More generally one can also define the following equivalence relation:

Definition 28.3.30 (Positive roots). Let Φ be the root system of a given Lie algebra $\mathfrak{g}$. Because the only scalar multiples of a root $\lambda \in \Phi$ in the root system are $\pm\lambda$ we can define the set of positive roots Φ^+ as follows:

1. $\lambda \in \Phi^+ \implies -\lambda \notin \Phi^+$
2. $\alpha, \beta \in \Phi^+ \wedge \alpha + \beta \in \Phi \implies \alpha + \beta \in \Phi^+$

The simple roots are then exactly the elements in Φ^+ that cannot be written as a sum of other elements in Φ^+.

Property 28.3.31. Let $\mathfrak{h}$ be a Cartan subalgebra. The set of simple roots Δ forms a basis for the dual space $\mathfrak{h}^*$ (over $\mathbb{C}$) and hence the cardinality of Δ is equal to the dimension of the Cartan subalgebra. This dimension is called the **rank** of the Lie algebra.

Definition 28.3.32 (Weyl group). For every simple root λ we construct the Householder transformation[18] σ_λ as follows:

$$\sigma_\lambda : \text{span}_{\mathbb{R}}(\Delta) \to \text{span}_{\mathbb{R}}(\Delta) : \mu \mapsto \mu - 2\frac{\langle \lambda, \mu \rangle}{\langle \lambda, \lambda \rangle}\lambda \tag{28.41}$$

The Weyl group W is then defined as the group generated by all the σ_λ's.

Property 28.3.33. Every root $\phi \in \Phi$ can be written as $\phi = \sigma(\mu)$ for some $\mu \in \Delta$ and $\sigma \in W$. Furthermore, the root system Φ is closed under the action of W. The Weyl group W is exactly the symmetry group of the root system Φ and the isometry group of the Killing form (and its dual).

Definition 28.3.34 (Coroot). Consider the real span $\mathfrak{h}_0^*$ of the roots of $\mathfrak{g}$. Using the restriction of the (dual) Killing form to this real subspace one can construct a dual space $\mathfrak{h}_0 \subset \mathfrak{h}$. The coroot $\alpha^\vee \in \mathfrak{h}_0$ associated to a root α is then defined by the following formula:

$$\alpha^\vee = 2\frac{\langle \alpha, \cdot \rangle}{\langle \alpha, \alpha \rangle} \tag{28.42}$$

With this definition the Weyl transformations can be rewritten as follows:

$$\sigma_\lambda : \text{span}_{\mathbb{R}}(\Delta) \to \text{span}_{\mathbb{R}}(\Delta) : \mu \mapsto \mu - \mu(\lambda^\vee)\lambda \tag{28.43}$$

[17] For every root set Φ one can find a set of simple roots.
[18] See definition 19.4.18.

Notation 28.3.35 (Coroot). Sometimes it is more favourable to denote the coroot associated to α by H^α.

Definition 28.3.36 (Algebraically integral). An element $H \in \mathfrak{h}_0$ is said to be algebraically integral if its value on every root is an integer. The set of all algebraically integral elements is called the **weight lattice**.

Definition 28.3.37 (Fundamental weight). Let $\Delta = \{\alpha_i\}$ be the set of simple roots. The fundamental weights $\{\omega_i\}_{i \leq |\Delta|}$ are defined as the elements of $\mathfrak{h}_0^*$ for which the following formula is satisfied for all $i, j \leq |\Delta|$:

$$\omega_i(H^{\alpha_j}) = \delta_{ij} \tag{28.44}$$

Hence an element $\lambda \in \mathfrak{h}_0$ is algebraically integral if it is an integral combination of fundamental weights.

Property 28.3.38. Let $\alpha \in \Phi$ be a root. Choose a generating element E^α of the weight space $\mathfrak{g}_\alpha$ associated to α. Let F^α be the generator of the weight space $\mathfrak{g}_{-\alpha}$ such that $\{E^\alpha, F^\alpha, [E^\alpha, F^\alpha]\}$ defines a one-dimensional simple Lie algebra. Then the following relations hold (for $\beta \neq \pm\alpha$):

- $\beta(H^\alpha) = 2\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} \in \mathbb{Z}$
- $[H^\alpha, E^\alpha] = \alpha(H^\alpha)E^\alpha = 2E^\alpha$
- $[H^\alpha, F^\alpha] = -\alpha(H^\alpha)F^\alpha = -2F^\alpha$

where the inner product $\langle \cdot, \cdot \rangle$ is given by the dual Killing form.[19]

Definition 28.3.39 (Cartan matrix). Let $\lambda_i, \lambda_j \in \Delta$ be simple roots. Because the Weyl group is the symmetry group of the root system

$$\sigma_{\lambda_i}(\lambda_j) = \lambda_j - 2\frac{\langle \lambda_i, \lambda_j \rangle}{\langle \lambda_i, \lambda_i \rangle}\lambda_i$$

is a root. From the properties above it then follows that the quantity

$$C_{ij} = 2\frac{\langle \lambda_i, \lambda_j \rangle}{\langle \lambda_i, \lambda_i \rangle} = \lambda_j(H^{\lambda_i}) \tag{28.45}$$

is an integer. The matrix formed by these numbers is called the Cartan matrix.

Property 28.3.40. The Cartan matrix C_{ij} has the following properties:

- $C_{ii} = 2$
- $C_{ij} \leq 0$ if $i \neq j$
- $C_{ij} = 0 \iff C_{ji} = 0$

This last property however does not imply that the Cartan matrix is symmetric. The fact that it is not symmetric can immediately be seen from its definition.

Definition 28.3.41 (Bond number). For all indices $i \neq j$ the bond number n_{ij} is defined as follows:

$$n_{ij} = C_{ij}C_{ji} \tag{28.46}$$

[19] Consider the *sharp* map 33.2 where one replaces the metric g by the Killing form K. The dual Killing form K^* is then a proper inner product (when restricted to the real span of Δ) defined as:

$$K^*(\cdot, \cdot) = K(\cdot^\sharp, \cdot^\sharp)$$

Using the definition of the coefficients C_{ij} we see that n_{ij} is an integer equal to $4\cos^2 \sphericalangle(\lambda_i, \lambda_j)$. This implies that n_{ij} can only take on the values $0, 1, 2, 3$.[20]

Remark 28.3.42. In the case of $n_{ij} = 2$ or $n_{ij} = 3$ there arise two possibilities. Namely that $C_{ij} > C_{ji}$ or $C_{ij} < C_{ji}$. From the definition of the Cartan integers and the symmetry of the dual Killing form these cases correspond to $\langle \lambda_i, \lambda_i \rangle < \langle \lambda_j, \lambda_j \rangle$ and $\langle \lambda_i, \lambda_i \rangle > \langle \lambda_j, \lambda_j \rangle$.

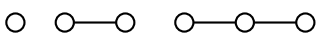
Construction 28.3.43 (Dynkin diagram). For a semisimple Lie algebra $\mathfrak{g}$ with simple root set Δ one can draw a so-called Dynkin diagram by using the following rules:

1. For every simple root $\lambda \in \Delta$ draw a circle: ◯
2. If ◯ and ◯ denote the simple roots λ_i and λ_j, draw n_{ij} lines between them.
3. When $n_{ij} = 2$ or $n_{ij} = 3$ add a $<$ or $>$ sign to relate the roots based on their lengths (see previous remark).

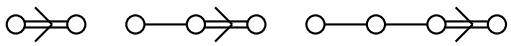
Theorem 28.3.44 (Cartan & Killing). *Every finite-dimensional simple $\mathbb{C}$-Lie algebra can be reconstructed from its set of simple roots Δ.*

Method 28.3.45. The Dynkin diagrams can be classified as follows (for every type the first three examples are given):

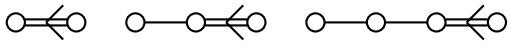
- A_n:



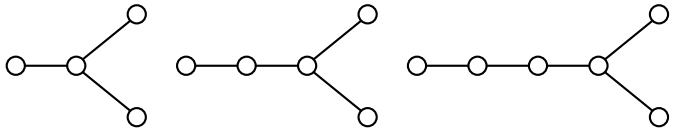
- $\mathrm{B}_n, n \geq 2$:



- $\mathrm{C}_n, n \geq 2$:



- $\mathrm{D}_n, n \geq 4$:



These are the only possible diagrams for simple Lie algebras.[21]

Example 28.3.46 (Special linear group). By looking at the Lie brackets in 28.13 we see that the one-element set $\{X_1\}$ forms a Cartan subalgebra of $\mathfrak{sl}(2, \mathbb{C})$. From 28.13 it is also immediately clear that the simple root set Δ is given by the one-element set $\{\lambda \in \mathfrak{sl}^*(2, \mathbb{C}) : \lambda(X_1) \mapsto 2\}$. Hence the Dynkin diagram for $\mathfrak{sl}(2, \mathbb{C})$ is A_1.

Construction 28.3.47 (Chevalley-Serre). Given a Dynkin diagram (of a simple Lie algebra) one can reconstruct the original Lie algebra $\mathfrak{g}$ (over $\mathbb{C}$) up to isomorphism. The number of nodes is equal to the number of simple roots and hence gives us the rank of $\mathfrak{g}$.

Let n denote the rank. We first construct the free Lie algebra on the $3n$ generators $\{E_i, F_i, H_i\}_{i \leq n}$. The Cartan subalgebra $\mathfrak{h} \leq \mathfrak{g}$ is constructed from the generators H_i by imposing the following relations:

- $[H_i, H_j] = 0$

[20] The value 4 would only be possible if the angle between λ_i and λ_j is 0 but this can only occur in the case that $i = j$, which was excluded from the definition.

[21] With exception of E_6, E_7, E_8, F_4 and G_2, the so-called *exceptional Lie algebras*.

- $[H_i, E_j] = a_{ij}E_j$
- $[H_i, F_j] = -a_{ij}F_j$

where the numbers a_{ij} form the Cartan matrix obtained by reversing construction 28.3.43. To complete the reconstruction one imposes following additional constraints:

- $[E_i, F_j] = \delta_{ij}H_j$
- $\mathrm{ad}_{E_i}^{|a_{ij}|-1}(E_j) = 0$
- $\mathrm{ad}_{F_i}^{|a_{ij}|-1}(F_j) = 0$

These relations are called the Chevalley-Serre or Serre relations.

Remark 28.3.48. For composite diagrams, and hence semisimple Lie algebras, one first constructs the Lie algebra corresponding to every simple diagram and then takes the direct sum.

Definition 28.3.49 (Kac-Moody algebra). Consider the Cartan matrix A associated to a finite-dimensional (semi)simple Lie algebra. This matrix has the properties listed in 28.3.40 together with the properties that $A_{ij} \geq -3$ and that it is irreducible and symmetrizable.

If one only retains the properties listed in 28.3.40 one obtains a so-called **generalized Cartan matrix**. Given such a generalized Cartan matrix A' one can construct a (possibly infinite-dimensional) Lie algebra in the following way:

Let $\tilde{L}$ be the Lie algebra obtained by applying the Chevalley-Serre construction to A'. Then there exists a maximal ideal I in $\tilde{L}$ such that $I \cap \mathfrak{h} = 0$. The quotient $\tilde{L}/I$ is called the Kac-Moody algebra associated to A'.

28.4 Universal enveloping algebra

Definition 28.4.1 (Universal enveloping algebra). Let $\mathfrak{g}$ be a Lie algebra with Lie bracket $[\cdot,\cdot]$. First construct the tensor algebra $T(\mathfrak{g})$. The universal enveloping algebra $U(\mathfrak{g})$ is defined as quotient of $T(\mathfrak{g})$ by the two-sided ideal generated by the elements $g \otimes h - h \otimes g - [g,h]$.

Construction 28.4.2. If one considers the generators and relations in the reconstruction theorem 28.3.47 as an (unital associative) algebra presentation instead of as a Lie algebra presentation[22] then one obtains the universal enveloping algebra $U(\mathfrak{g})$ of $\mathfrak{g}$.

Definition 28.4.3 (Casimir invariant[23]). Let $\mathfrak{g}$ be a Lie algebra. A Casimir invariant J is an element of the center of $U(\mathfrak{g})$.

Formula 28.4.4 (Quadratic Casimir invariant). Consider a Lie algebra representation $\rho : \mathfrak{g} \to \mathrm{End}(V)$ on an n-dimensional vector space V. Let $\{X_i\}_{i\leq n}$ be a basis for $\mathfrak{g}$. The (quadratic) Casimir invariant associated with ρ is given by:

$$\Omega_\rho = \sum_{i=0}^{n} \rho(X_i) \circ \rho(\xi_i) \tag{28.47}$$

where the set $\{\xi_i\}_{i\leq n}$ is defined by the relation $K_\rho(X_i, \xi_j) = \delta_{ij}$ using the Killing form 28.31.

Property 28.4.5. When the representation $\rho : \mathfrak{g} \to \mathrm{End}(V)$ is irreducible, Schur's lemma 24.2.3 tells us that:

$$\Omega_\rho = c_\rho \mathbb{1}_V \tag{28.48}$$

By taking the trace of this formula and using formula 28.31 we see that $c_\rho = \frac{\dim \mathfrak{g}}{\dim V}$.

[22] Where the Lie bracket is now the commutator constructed using the formal multiplication in $U(\mathfrak{g})$.

[23] Also known as a **Casimir operator** or **Casimir element**.

28.5 Poisson algebras and Lie superalgebras

Definition 28.5.1 (Internal Lie algebra). Let $(\mathbf{C}, \otimes, \mathbf{1})$ be a symmetric monoidal category with braiding σ. A Lie algebra internal to $\mathbf{C}$ is an object $A \in \mathrm{ob}(\mathbf{C})$ and a morphism $[\cdot,\cdot] : A \otimes A \to A$ satisyfing the following conditions:

- Antisymmetry: $[\cdot,\cdot] + [\cdot,\cdot] \circ \sigma_{A,A} = 0$
- Jacobi identity: $[\cdot,[\cdot,\cdot]] + [\cdot,[\cdot,\cdot]] \circ \tau + [\cdot,[\cdot,\cdot]] \circ \tau^2 = 0$

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

Example 28.5.2 (Lie superalgebra). If we use the braiding $\sigma(a \otimes b) = (-1)^{|a||b|} b \otimes a$ in the category **sVect** we obtain the notion of a Lie superalgebra.

Definition 28.5.3 (Poisson algebra). Let V be a vector space equipped with two bilinear operations $\star$ and $\{\cdot,\cdot\}$ that satisfy the following conditions:

- The couple $(V, \star)$ is an associative algebra.
- The couple $(V, \{\cdot,\cdot\})$ is a Lie algebra.
- the **Poisson bracket** $\{\cdot,\cdot\}$ acts as a derivation[24] with respect to the operation $\star$, i.e.

$$\{x, y \star z\} = \{x, y\} \star z + y \star \{x, z\}$$

[24] See definition 27.2.1.

Chapter 29

Fibre Bundles

29.1 General bundles

Definition 29.1.1 (Bundle). A bundle is a triple (E, B, π) where E and B are topological spaces and π is a continuous map.[1]

Definition 29.1.2 (Fibered manifold). A fibered manifold is a surjective submersion[2]

$$\pi : E \to B$$

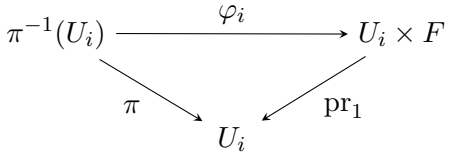
where E is called the **total space**, B the **base space** and π the **projection**. For every point $p \in B$, the set $\pi^{-1}(p)$ is called the **fibre** over p.

The most important example of a fibered manifold is a fibre bundle:

Definition 29.1.3 (Fibre bundle). A fibre bundle is a tuple (E, B, π, F, G) where E, B and F are topological spaces and G is a topological group (called the **structure group**), such that there exists a continuous surjective map

$$\pi : E \to B$$

and an open cover $\{U_i\}_{i \in I}$ of B for which there exists a family of homeomorphisms $\{\varphi_i : \pi^{-1}(U_i) \to U_i \times F\}_{i \in I}$ that make the following diagram commute:



As for topological bundles and fibered manifolds we call E and B the total space and base space respectively. The space F is called the **(typical) fibre**. We also call φ_i a **local trivialization**[3], (U_i, φ_i) a **bundle chart**[4] and the set $\{(U_i, \varphi_i)\}_{i \in I}$ a **trivializing cover**.

[1] Sometimes one requires that the map π is also surjective. However then we cannot make the association $\mathbf{Bundle}(X) \cong \mathbf{Top}/X$ of categories anymore.

[2] See definition 27.3.9.

[3] This name follows from the fact that the bundle is locally homeomorphic to a (trivial) product space: $E \cong U \times F$.

[4] This is due to the similarities with the charts defined for manifolds.

The transition maps $\varphi_j \circ \varphi_i^{-1} : (U_i \cap U_j) \times F \to (U_i \cap U_j) \times F$ can be identified with the cocycle[5] $g_{ji} : U_i \cap U_j \to G$, associated to the (left) action (which we require to be faithful[6]) of G on every fibre, by the following relation:

$$\varphi_j \circ \varphi_i^{-1}(b, x) = (b, g_{ji}(b) \cdot x) \tag{29.1}$$

Remark 29.1.4. One should pay attention that the bundle charts are not coordinate charts in the original sense 27.1.1 because the image of φ_i is not an open subset of $\mathbb{R}^n$. However they serve the same purpose and we can still use them to locally inspect the total space P.

Notation 29.1.5. A fibre bundle (E, B, π, F, G) is often indicated by the following diagram:

$$\begin{array}{ccc} F & \longrightarrow & E \\ & & \big\downarrow \pi \\ & & B \end{array}$$

or more compactly $F \hookrightarrow E \xrightarrow{\pi} B$. A drawback of these notations is that we do not immediately know what the structure group of the bundle is.

Definition 29.1.6 (Fibre). Let $F \hookrightarrow E \xrightarrow{\pi} B$ be a fibre bundle over a base space B. The fibre over $b \in B$ is defined as the set $\pi^{-1}(b)$. It is often denoted by F_b.

Definition 29.1.7 (Numerable fibre bundle). A fibre bundle which admits a local trivialization of a numerable open cover.

Definition 29.1.8 (Smooth fibre bundle). A smooth fibre bundle is a fibre bundle (E, B, π, F, G) with the following constraints:

- The base space B and typical fibre F are smooth manifolds.
- The structure goup G is a Lie group.
- The projection map, trivializing maps and transition functions are diffeomorphisms.

Remark 29.1.9. A smooth fibre bundle is also a smooth manifold.

Construction 29.1.10 (Fibre bundle construction theorem). Let M and F be topological spaces and let G be a topological group equipped with a left action on F. Suppose that we are given a cover $\{U_i\}_{i \in I}$ of M and a set of continuous functions $\{g_{ji} : U_i \cap U_j \to G\}$ that satisfy the cocycle condition 28.1.12. A fibre bundle over M can then be constructed as follows:

1. We first construct for every set U_i an associated set $U_i \times F$.
2. We then construct the disjoint union $T \equiv \bigsqcup_{i \in I} U_i \times F$ equipped with the disjoint union topology[7].
3. From this disjoint union we construct a quotient space[8] (equipped with the quotient space topology) by applying following equivalence relation for every i, j:

$$(p, f) \sim (p, g_{ji}(x) \cdot f) \tag{29.2}$$

for all $x \in U_i \cap U_j$ and $f \in F$. The fibre bundle is equal to this quotient space $T/\sim$ together with the projection π that maps the equivalence class of $(x, f) \in T$ to $x \in M$.

[5] See definition 28.1.12.
[6] See definition 3.1.62.
[7] See definition 7.1.4.
[8] See definition 7.4.

4. Local trivializations are given by the maps $\varphi_i : \pi^{-1}(U_i) \to U_i \times F$ that satisfy:

$$\varphi_i^{-1} : (x, f) \mapsto [(x, f)] \tag{29.3}$$

where $[A]$ means the equivalence class of A in $T/\sim$.

Definition 29.1.11 (Compatible[9] bundle charts). A bundle chart (V, ψ) is compatible with a trivializing cover $\{(U_i, \varphi_i)\}_{i\in I}$ if whenever $V \cap U_i \neq \emptyset$ there exists a map $h_i : V \cap U_i \to G$ such that:

$$\psi \circ \varphi_i^{-1}(b, x) = (b, h_i(b)x) \tag{29.4}$$

for all $b \in V \cap U_i$ and $x \in F$. Two trivializing covers are *equivalent* if all bundle charts are cross-compatible. As in the case of manifolds, this gives rise to the notion of a **G-atlas**. A **G-bundle** is then defined as a fibre bundle eqipped with an equivalence class of G-atlases.

29.1.1 Bundle maps

Definition 29.1.12 (Bundle map). A bundle map between two fibre bundles $\pi_1 : E_1 \to B_1$ and $\pi_2 : E_2 \to B_2$ is a pair (f_E, f_B) of continuous maps that make diagram 29.1 commute:

$$\begin{array}{ccc} E_1 & \xrightarrow{f_E} & E_2 \\ {\scriptstyle \pi_1}\downarrow & & \downarrow{\scriptstyle \pi_2} \\ B_1 & \xrightarrow[f_B]{} & B_2 \end{array}$$

Figure 29.1: Bundle map between fibre bundles.

Definition 29.1.13 (Isomorphic fibre bundles). Two fibre bundles F and G are isomorphic if there exist bundle maps $f : F \to G$ and $g : G \to F$ such that $f \circ g = \mathbb{1}_G$ and $g \circ f = \mathbb{1}_F$.

Definition 29.1.14 (Equivalent fibre bundles). Two fibre bundles $\pi_1 : E_1 \to B$ and $\pi_2 : E_2 \to B$ (with the same typical fibre and structure group) are equivalent if there exist trivializing covers[10] $\{(U_i, \varphi_i)\}_{i\in I}$ and $\{(U_i, \varphi_i')\}_{i\in I}$ and a family of smooth functions $\{\rho_i : U_i \to G\}_{i\in I}$ such that:

$$g'_{ji}(b) = \rho_j(b) \circ g_{ji}(b) \circ \rho_i^{-1}(b) \tag{29.5}$$

for every $b \in U_i \cap U_j$. An explicit form of these functions is given by:

$$\rho_i = \varphi_i' \circ \varphi_i^{-1} \tag{29.6}$$

Property 29.1.15. Two fibre bundles over the same base space are equivalent if and only if they are isomorphic. Furthermore, every bundle map between bundles over the same base space induces an equivalence (and thus also an isomorphism).

Definition 29.1.16 (Trivial bundle). A fibre bundle (E, B, π, F) is trivial if there exists an equivalence $E \cong B \times F$.

[9] Also called an **admissible chart**.

[10] Remark that the collection $\{U_i\}_{i\in I}$ is the same for both trivializing covers.

29.1.2 Fibre bundle operations

Definition 29.1.17 (Subbundle). A subbundle of a fibre bundle $\pi : E \to B$ is a triple (E', B', π') such that $E' \subset E$, $B' \subset B$ (where $\subset$ now means 'submanifold of') and $\pi' = \pi|_{E'}$.

Definition 29.1.18 (Pullback bundle). Let $\pi : E \to B$ be a fibre bundle. Let $f : B' \to B$ be a continuous map between topological spaces. The pullback bundle f^*E is defined as follows:

$$f^*E = \{(b', e) \in B' \times E : f(b') = \pi(e)\} \tag{29.7}$$

The topology on f^*E is induced by the subspace topology of the product $B' \times E$. The projection onto the second factor gives a map of total spaces $f^*E \to E$.

Definition 29.1.19 (Fibre product). Let (F_1, B, π_1) and (F_2, B, π_2) be two fibre bundles on a base space B. Their fibre product is defined as:

$$F_1 \diamond F_2 = \{(f, g) \in F_1 \times F_2 : \pi_1(f) = \pi_2(g)\} \tag{29.8}$$

29.1.3 Sections

Definition 29.1.20 (Section). A **global** section on a fibre bundle $\pi : E \to B$ is a smooth function $s : B \to E$ such that $\pi \circ s = \mathbb{1}_B$. For any open subset $U \subset B$ we define a local section as a smooth function $s_U : U \to E$ such that $\pi \circ s_U(b) = b$ for all $b \in U$.

Notation 29.1.21. The set of all global sections on a bundle E is denoted by $\Gamma(E)$. The set of local sections on $U \subset E$ is similarly denoted by $\Gamma(U)$.

Property 29.1.22. The sections on a fibre bundle E pullback to the pullback bundle f^*E by setting $f^*s = s \circ f$.

29.2 Jet bundles

Although the following constructions can be defined in the general context of smooth manifolds we will only consider them in the case of smooth fibre bundles. Only the notion of a jet will be defined in general for maps between smooth manifolds (to define a Whitney topology on the space of smooth functions).

Definition 29.2.1 (Jet). Consider two manifolds M, N. Two maps $\sigma, \xi \in C^\infty(M, N)$ with local coordinates (σ^i) and (ξ^i) define the same r-jet at a point $p \in M$ if and only if:

$$\left.\frac{\partial^\alpha \sigma^i}{\partial x^\alpha}\right|_p = \left.\frac{\partial^\alpha \xi^i}{\partial x^\alpha}\right|_p \tag{29.9}$$

for all $0 \leq i \leq \dim M$ and every multi-index α such that $0 \leq |\alpha| \leq r$. It is clear that this relation defines an equivalence relation. The r-jet at $p \in M$ with representative σ is denoted by $j^r_p\sigma$. The number r is called the **order** of the jet.

Definition 29.2.2 (Whitney C^k-topology). Let M, N be two smooth manifolds and consider the manifold[11] of k-jets $J^k(M, N)$. A basis for the Whitney C^k-topology on $C^\infty(M, N)$ is given by the sets

$$S^k(U) = \{f \in C^\infty(M, N) : J^k f \in U\} \tag{29.10}$$

where U is open in $J^k(M, N)$.

[11] This manifold will be constructed further down this section in the case of fibre bundles.

Remark 29.2.3. The k-jet construction is easily carried over to the theory of bundles. To construct the jet space of sections on a (smooth) fibre bundle (E, B, π) one just sets $M = B$ and $N = E$ in the previous definition. From this point on we will restrict to this case.

Definition 29.2.4 (Jet manifold). Consider a fibre bundle (E, B, π). The r-jet manifold $J^r(\pi)$ of the projection π is defined as:

$$J^r(\pi) = \{j^r_p\sigma : \sigma \in \Gamma(E), p \in B\} \tag{29.11}$$

The set $J^0(\pi)$ is identified with the total space E.

Definition 29.2.5 (Jet projections). Let (E, B, π) be a fibre bundle with r-jet manifolds $J^r(\pi)$. The **source projection** π_r and **target projection** $\pi_{r,0}$ are defined as the maps

$$\pi_r : J^r(\pi) \to B : j^r_p\sigma \mapsto p \tag{29.12}$$

$$\pi_{r,0} : J^r(\pi) \to E : j^r_p\sigma \mapsto \sigma(p) \tag{29.13}$$

These projections satisfy $\pi_r = \pi \circ \pi_{r,0}$. We can also define a **k-jet projection** $\pi_{r,k}$ as the map

$$\pi_{r,k} : J^r(\pi) \to J^k(\pi) : j^r_p\sigma \mapsto j^k_p\sigma \tag{29.14}$$

where $k \leq r$. The k-jet projections satisfy a transitivity property $j_{k,m} = j_{r,m} \circ j_{k,r}$.

Definition 29.2.6 (Jet prolongation). Let σ be a section on a fibre bundle (E, B, π). The r-jet prolongation $j^r\sigma$ corresponding to σ is defined as the following map:

$$j^r\sigma : B \to J^r(\pi) : p \mapsto j^r_p\sigma \tag{29.15}$$

Definition 29.2.7 (Jet bundle). The r-jet bundle corresponding to the projection π is then defined as the triple $(J^r(\pi), B, \pi_r)$. The bundle charts (U_i, φ_i)[12] on E define *induced* bundle charts on $J^r(\pi)$ in the following way:

$$U^r_i = \{j^r_p\sigma : \sigma(p) \in U_i\} \tag{29.16}$$

$$\varphi^r_i = \left(x^k, u^\alpha, \left.\frac{\partial^I u^\alpha}{\partial x^I}\right|_p\right) \tag{29.17}$$

where I is a multi-index such that $0 \leq |I| \leq r$. The partial derivatives $\left.\frac{\partial^I u^\alpha}{\partial x^I}\right|_p$ are called the **derivative coordinates** on $J^r(\pi)$.

Example 29.2.8 (Diffeomorphism jets). Let M be a smooth manifold. Consider the set $\mathcal{D}^{\mathrm{a}}_{\mathrm{local}}(M)$ of local analytic diffeomorphisms $\phi : M \to M : z \mapsto \phi(z) \equiv Z$. The locality property turns this set into a (smooth) pseudogroup[13].

By the inverse function theorem one can define the diffeomorphism jet bundle $D^r(M)$ as the subbundle of $J^r(M)$ for which

$$\det\left(\frac{\partial Z^\alpha}{\partial z^\beta}\right) \neq 0$$

It is also possible to endow this jet bundle with the structure of a groupoid[14]. Using the source and target projections one can check that two two elements $g^{(r)}, h^{(r)} \in D^r(M)$ can be multiplied if and only $\pi_r(h^{(r)}) = \pi_{r,0}(g^{(r)})$. The derivative coordinates can be found using the *Faá di Bruno formula*.

Furthermore every pseudogroup $\mathcal{G} \subset \mathcal{D}$ induces a jet bundle/Lie-subgroupoid $\mathcal{G}^{(r)} \subset \mathcal{D}^{(r)}$. This structure gives rise to the following notions:

[12] Where $\varphi_i = (x^k, u^\alpha)$ with x^k the base space coordinates and u^α the total space coordinates.
[13] See definition 7.3.22.
[14] See definition 4.12.2.

Definition 29.2.9 (Regular pseudogroup). Consider a smooth manifold M. Let $\mathcal{D}(M)$ be its diffeomorphism pseudogroup and let $\mathcal{G} \subset \mathcal{D}$ be another pseudogroup. If there exists an $N \in \mathbb{N}_0$, where N is called the **order**, such that for all $r \geq N$ the jets $\pi_r : \mathcal{G}^{(r)} \to M$ form an embedded submanifold of $\Pi_r : \mathcal{D}^{(r)} \to M$ and such that the jet projections $\pi_{r+1,r} : \mathcal{G}^{(r+1)} \to \mathcal{G}^{(r)}$ are fibrations then we say that $\mathcal{G}$ is a regular pseudogroup.

Definition 29.2.10 (Lie pseudogroup). Let $\mathcal{G} \subset \mathcal{D}$ be a regular analytic pseudogroup of order N. If every local diffeomorphism $\phi \in \mathcal{D}$ satisfying $j^N\phi \in \mathcal{G}^{(N)}$ is also an element of $\mathcal{G}$ then $\mathcal{G}$ is called a Lie pseudogroup.

Property 29.2.11. Let $\mathcal{G}$ be a Lie pseudogroup of order N. The regularity condition implies that the jet bundle $\mathcal{G}^{(r)}$ is described by a set of r^{th}-order partial differential equations

$$F\left(z, Z^{(r)}\right) = 0$$

The (local) solutions to these equations are exactly the analytic functions which have $(z_0, Z_0^{(r)})$ as local coordinates of their r-jet at $z_0 \in M$.

The Lie condition on $\mathcal{G}$ implies that every solution to the system is in fact an element of $\mathcal{G}$. This system of equations is called the **determining system** of the Lie pseudogroup.

Definition 29.2.12 (Lie completion). Let $\mathcal{G}$ be a regular pseudogroup. The Lie completion $\overline{\mathcal{G}}$ of $\mathcal{G}$ is defined as the set of all (local) analytic diffeomorphisms solving the determining system of $\mathcal{G}$. This completion is itself a Lie pseudogroup. If $\mathcal{H}$ is a Lie pseudogroup then $\overline{\mathcal{H}} = \mathcal{H}$.

Chapter 30

Vector Bundles

The tangent space, as introduced in subsection 27.2, can also be introduced in a more topological way:

30.1 Tangent bundle

Construction 30.1.1 (Tangent bundle). Let M be an n-dimensional manifold with atlas $\{(U_i, \varphi_i)\}_{i\leq n}$. Construct for every open set O an associated set $TO = O \times \mathbb{R}^n$ and construct for every smooth function f an associated smooth function on TO, called the **differential** or **derivative** of f, by:

$$Tf : O \times \mathbb{R}^n \to f(O) \times \mathbb{R}^n : (x, v) \mapsto (f(x), Df(x)v) \tag{30.1}$$

where $Df(x) : \mathbb{R}^n \to \mathbb{R}^n$ is the linear operator represented by the Jacobian matrix of f in x.

Applying this definition to the transition functions ψ_{ji} we obtain a new set of functions $\widetilde{\psi}_{ji} := T\psi_{ji} : U_i \times \mathbb{R}^n \to U_j \times \mathbb{R}^n$ given by:

$$\widetilde{\psi}_{ji}(\varphi_i(x), v) = \big(\varphi_j(x), D(\varphi_j \circ \varphi_i^{-1})(\varphi_i(x))v\big) \tag{30.2}$$

Because the transition functions are diffeomorphisms, the Jacobians are invertible. This implies that the maps $\widetilde{\psi}_{ji}$ are elements of $GL(\mathbb{R}^n)$. The tangent bundle is now obtained by applying the fibre bundle construction theorem 29.1.10 to the triple $(M, \mathbb{R}^n, GL(\mathbb{R}^n))$ together with the base cover $\{U_i\}_{i\leq n}$ and the cocycle $\{\widetilde{\psi}_{ji}\}_{i,j\leq n}$.

Definition 30.1.2 (Natural chart). The charts in the atlas of the constructed bundle are sometimes called **natural charts** or **adapted charts** because the first n coordinates are equal to the coordinates of the base space.

Alternative Definition 30.1.3. The above construction eventually comes down to the following, more intuitive, definition of the tangent bundle:

$$TM = \bigsqcup_{p\in M} T_pM \tag{30.3}$$

equipped with the disjoint union topology 7.1.4 and the projection map[1]

$$\pi : TM \to M : (p, X) \mapsto p \tag{30.4}$$

[1]The map π is single-valued because the tangent bundle is defined as the disjoint union of the tangent spaces.

where X is a tangent vector in T_pM. An atlas on TM is then given by the charts $(\pi^{-1}(U_i), \theta)$ with

$$\theta : TM \to \mathbb{R}^{2n} : (p, X) \mapsto (\varphi_i(p), X^1, ..., X^n) \tag{30.5}$$

where $X = X^i \left.\frac{\partial}{\partial x^i}\right|_p \in T_pM$ and where (U_i, φ_i) is a chart on M covering the point $p \in M$.

Property 30.1.4. Let M be an n-dimensional manifold. Using the natural charts on TM, which give a local homeomorphism

$$\psi_i : TM \to U_i \times \mathbb{R}^n \cong \mathbb{R}^n \times \mathbb{R}^n$$

we can see that TM is isomorphic to $\mathbb{R}^{2n}$. This implies that:

$$\boxed{\dim TM = 2 \dim M} \tag{30.6}$$

Definition 30.1.5 (Tangent space). Let $x \in M$. The topological definition of the tangent space is given by the fibre

$$T_xM := \tau_M^{-1}(x) \tag{30.7}$$

If we use the natural charts to map T_xM to the set $\varphi_i(x) \times \mathbb{R}^n$, we see that T_xM is isomorphic to $\mathbb{R}^n$ and thus also to M itself. Furthermore, we can equip every fibre with the following vector space structure:

$$(x, v_1) + (x, v_2) := (x, v_1 + v_2)$$
$$r(x, v) := (x, rv)$$

Remark 30.1.6. Now it is clear that the rule "*a vector is something that transforms like a vector*" stems from the fact that:

$$\text{a vector } v \in T_xM \text{ is tangent to } \varphi_i(x) \text{ in a chart } (U_i, \varphi_i)$$

if and only if

$$D(\varphi_j \circ \varphi_i^{-1})(\varphi_i(x))v \text{ is tangent to } \varphi_j(x) \text{ in a chart } (U_j, \varphi_j)$$

Definition 30.1.7 (Differential). The map T defined in 30.1 can be generalized to arbitrary smooth manifolds as the map $Tf : TM \to TN$. Furthermore, let $x \in U \subseteq M$ and let $V = f(U)$. By looking at the restriction of Tf to T_xM, denoted by T_xf, we see that it maps T_xU to $T_{f(x)}V$ linearly.

Property 30.1.8. The map $Tf : TM \to TN$ (see 30.1) has following properties[2]:

- $T(\mathbb{1}_M) = \mathbb{1}_{TM}$
- Let f, g be two smooth functions on smooth manifolds. Then $T(f \circ g) = Tf \circ Tg$.

Definition 30.1.9 (Rank). Let $f : M \to N$ be a differentiable map between smooth manifolds. Using the fact that Tf is a linear map of fibres[3], we define the rank of f at $p \in M$ as the rank (in the sense of 19.3.11) of the differential $Tf : T_pM \to T_{f(p)}N$.

Theorem 30.1.10 (Inverse function theorem). *A C^∞ map $f : M \to N$ between smooth manifolds is locally homeomorphic (resp. locally diffeomorphic) if and only if its differential $Tf : T_pM \to T_pN$ is an isomorphism (resp. diffeomorphism) at p.*

Definition 30.1.11 (Parallelizable manifold). A manifold is said to be parallelizable if its tangent bundle is trivial.

[2]This turns the map T into a functor on the category of smooth manifolds. Hence we can view T as a functorial derivative.

[3]See definition 30.1.7.

30.2 Vector bundles

Instead of restricting ourselves by letting the typical fibre be a Euclidean space with the same dimension as the base manifold, we can generalize the construction of the tangent bundle in the following way:

Construction 30.2.1 (Vector bundle). Consider a smooth n-dimensional manifold M with atlas $\{(U_i, \varphi_i)\}_{i\leq n}$, a cocycle $\{g_{ji} : U_i \cap U_j \to G\}_{i,j\leq n}$ with values in a Lie group G and a smooth representation $\rho : G \to GL(V)$, where V is a vector space. A bundle can then be constructed using 29.1.10

Remark 30.2.2. As is also the case for tangent bundles (which are specific cases of vector bundles where the typical fibre has the same dimension as the manifold) the choice of charts on E is not random. To preserve the structure of fibres, the use of the natural charts is imperative.

Example 30.2.3 (Line bundle). A line bundle is a vector bundle with a one-dimensional fibre V. A common example is the $\mathbb{C}$-line bundle over configuration space for which, in quantum mechanics, the sections correspond to the physical "wave functions".

30.2.1 Whitney sums

Definition 30.2.4 (Whitney sum). Consider two vector bundles E, E' with fibres W, W' respectively. Then we can construct a new vector bundle $E \oplus E'$ by defining the new typical fibre to be the direct sum $W \oplus W'$, i.e. the fibre above b is given by $W_b \oplus W'_b$. This operation is called the Whitney sum or direct sum of vector bundles.

Property 30.2.5. Let X be a paracompact Hausdorff space and let E be a vector bundle over X. Every vector subbundle F of E admits a orthogonal complement $F^\perp$ such that $F \oplus F^\perp \cong E$.

Property 30.2.6. Let X be a compact Hausdorff space. Every vector bundle E over X admits a complementary vector bundle E^c such that $E \oplus E^c \cong X \times \mathbb{R}^n$ for some $n \in \mathbb{N}$.

Definition 30.2.7 (Stable isomorphism). Two vector bundles E, E' over a base space B are said to be stably isomorphic if there exist integers $m, n \in \mathbb{N}$ such that $E \oplus (B \times \mathbb{R}^m) \cong E' \oplus (B \times \mathbb{R}^n)$.

30.2.2 Associated vector bundles

Definition 30.2.8 (Associated vector bundle). Consider a representation $\rho : GL(\mathbb{R}^n) \to GL(\mathbb{R}^l)$ together with the cocycle $\{t_{ji} := D(\psi_{ji}) \circ \varphi_i\}_{i,j\leq n}$ as defined for the tangent bundle. The composition $\rho \circ t_{ji} : U_i \cap U_j \xrightarrow{t_{ji}} GL(\mathbb{R}^n) \xrightarrow{\rho} GL(\mathbb{R}^l)$ is again a cocycle and can thus be used to define a new vector bundle on M through the fibre bundle construction theorem 29.1.10. The vector bundle $E = \rho(TM)$ so obtained is called the associated bundle of the tangent bundle induced by ρ.

Example 30.2.9 (Contravariant vectors). By noting that the k^{th} tensor power $\otimes^k$ induces a representation given by the tensor product of the representations, we can construct the bundle of k^{th} order contravariant vectors $\otimes^k(TM)$ with the cocycle given by $x \mapsto t_{ji}(x) \otimes \cdots \otimes t_{ji}(x)$.

Example 30.2.10 (Cotangent bundle). Another (smooth) representation is given by $A \mapsto (A^T)^{-1} = (A^{-1})^T$ for every linear map A. The vector bundle constructed this way, where the cocycle is given by $(t^T_{ji})^{-1}$, is called the cotangent bundle on M and is denoted by T^*M. Elements of the fibres are called **covariant vectors** or **covectors**.

Notation 30.2.11. A combination of the cocycle t_{ji} and its dual $(t^T_{ji})^{-1}$ can also be used to define the bundle of k^{th} order contravariant and l^{th} order covariant vectors on M. This bundle is denoted by $T^{(k,l)}M$.

Example 30.2.12 (Pseudovectors). If we consider the representation

$$\rho : A \mapsto \operatorname{sgn}\det(A)A \tag{30.8}$$

we can construct a bundle similar to the tangent bundle. The sign of the cocycle functions t_{ji} now has an influence on the fibres. Elements of these fibres are called **pseudovectors**.

Example 30.2.13 (Tensor density). Tensor densities are obtained by using the representation

$$\rho : A \mapsto \det(A)A \tag{30.9}$$

Pseudotensor densities are obtained by combining the above two representations.

30.2.3 Grassmann bundle

Looking at property 19.2.16 and noting that $\mathrm{GL}_n(\mathbb{R})$ is a Lie group, we can endow the Grassmannian $\mathrm{Gr}(k, \mathbb{R}^n)$ 19.2.15 with a differentiable structure, turning it into a smooth manifold. This allows us to construct a new bundle[4] by applying the usual bundle construction theorem 29.1.10:

Construction 30.2.14 (Grassmann bundle). We first define a new set of transition functions:

$$\begin{aligned}\psi_{ji} : \varphi_i(U_i \cap U_j) \times \mathrm{Gr}(k, \mathbb{R}^n) &\to \varphi_j(U_i \cap U_j) \times \mathrm{Gr}(k, \mathbb{R}^n) \\ : (\varphi_i(x), V) &\mapsto (\varphi_j(x), t_{ji}(x) \cdot V)\end{aligned} \tag{30.10}$$

where $\{t_{ji}\}_{i,j\leq n}$ is the tangent bundle cocycle, but now with an action on the compact manifold $\mathrm{Gr}(k, \mathbb{R}^n)$ instead of the vector space $\mathbb{R}^n$. These transition functions can then be used to create a new fibre bundle with typical fibre $\mathrm{Gr}(k, \mathbb{R}^n)$: The fibre over a point $p \in M$ is the Grassmannian $\mathrm{Gr}(k, T_pM)$ associated to the tangent space over p.

Notation 30.2.15. The Grassmann k-plane bundle is denoted by $\mathrm{Gr}(k, TM)$.

Definition 30.2.16 (Tautological bundle). Consider the Grassmannian $\mathrm{Gr}(k, V)$ of an $(n + k)$-dimensional vector space V. The total space of the tautological k-bundle $\gamma_{n,k}$ is defined as the set of points (W, w) where $W \in \mathrm{Gr}(k, V)$ and $w \in W$. Local trivializations are constructed as follows:

$$\varphi : \pi^{-1}(U) \to \mathrm{Gr}(k, V) \times Z : (W, w) \mapsto (W, \mathrm{proj}_Z(w)) \tag{30.11}$$

where proj_Z is the orthogonal projection[5] onto Z.

30.2.4 Sections

Definition 30.2.17 (Frame). A frame of a vector bundle E is a tuple $(s_1, ..., s_n)$ of smooth sections such that $(s_1(b), ..., s_n(b))$ is a basis of the fibre $\pi^{-1}(b)$ for all $b \in B$.

Property 30.2.18. A vector bundle is trivial if and only if there exists a frame of global sections.

Theorem 30.2.19 (Serre & Swan). *The set of all smooth sections $\Gamma(E)$ over a vector bundle E with base space M is a finitely generated projective $C^\infty(M)$-module.*

[4] Due to the fact that the Grassmannian is not a vector space, we construct a general fibre bundle and not a vector bundle.

[5] See definition 19.4.15.

30.3 Vector fields

Definition 30.3.1 (Vector field). A smooth section $s \in \Gamma(TM)$ of the tangent bundle is called a vector field. The set of vector fields forms a $C^\infty(M)$-module.

Notation 30.3.2. The set of all vector fields on a manifold M is often denoted by $\mathfrak{X}(M)$.

Theorem 30.3.3 (Hairy ball theorem). *There exists no nowhere vanishing vector field on an even-dimensional sphere S^{2n}.*

Definition 30.3.4 (Pullback). Let X be vector field on M and let $\varphi : M \to N$ be a diffeomorphism between smooth manifolds. The pullback of X along φ is defined as:

$$(\varphi^* X)_p = T\varphi^{-1}(X_{\varphi(p)}) \tag{30.12}$$

Definition 30.3.5 (Pushforward). Let $X \in \mathfrak{X}(M)$ and let $\varphi : M \to N$ be a diffeomorphism between smooth manifolds. Using the differential $T\varphi$ we can define the pushforward of X along φ as:

$$(\varphi_* X)_{\varphi(p)} = T\varphi(X_p) \tag{30.13}$$

which we can rewrite using the pullback as:

$$\varphi_* X = \varphi^{-1*} X \tag{30.14}$$

Or equivalently we can define a vector field on N by:

$$(\varphi_* X)_q(f) = X_{\varphi^{-1}(q)}(f \circ \varphi) \tag{30.15}$$

for all smooth functions $f : N \to \mathbb{R}$ and points $q \in N$.

30.3.1 Integral curves

Definition 30.3.6 (Integral curve). Let $X \in \mathfrak{X}(M)$ and let $\gamma :]a, b[\to M$ be a smooth curve on M. γ is said to be an integral curve of X if:

$$\boxed{\gamma'(t) = X(\gamma(t))} \tag{30.16}$$

for all $t \in]a, b[$ where we defined $\gamma'(t) := T\gamma(t, 1)$ using the functorial derivative 30.1.

This equation can be seen as a system of ordinary differential equations in the second argument. Using Picard's existence theorem[6] together with the initial value condition $\gamma(0) = p$ we can find a unique curve on $]a, b[$ satisfying the defining equation 30.16. Furthermore we can extend the interval $]a, b[$ to a maximal interval such that the solution is still unique. This solution, denoted by γ_p, is called the **integral curve of X through p.**

Definition 30.3.7 (Flow). Let $X \in \mathfrak{X}(M)$. The function σ_t:

$$\sigma_t(p) = \gamma_p(t) \tag{30.17}$$

is called the flow of X at time t. The **flow domain** is defined as the set $D(X) = \{(t, p) \in \mathbb{R} \times M \mid t \in]a_p, b_p[\}$ where $]a_p, b_p[$ is the maximal interval on which $\gamma_p(t)$ is defined.

Property 30.3.8. Suppose that $D(X) = \mathbb{R} \times M$. The flow σ_t has following properties for all $s, t \in \mathbb{R}$:

[6] Also Picard-Lindelöf theorem.

- $\sigma_0 = \mathbb{1}_M$
- $\sigma_{s+t} = \sigma_s \circ \sigma_t$
- $\sigma_{-t} = (\sigma_t)^{-1}$

These three properties[7] say that σ_t is a bijective group action from M to the additive group of real numbers. This implies that σ_t is indeed a **flow** in the general mathematical sense.

Definition 30.3.9 (Complete vector field). A vector field X is called complete if the flow domain for every flow is all of $\mathbb{R}$.

Property 30.3.10. The flow σ_t of a vector field is of class C^∞. If X is complete it follows from previous definition that the flow is a diffeomorphism from M onto itself.

Property 30.3.11. If the manifold M is compact then every vector field $X \in \mathfrak{X}(M)$ is complete.

30.3.2 Lie derivative

Formula 30.3.12 (Lie derivative for smooth functions). Let $X \in \mathfrak{X}(M)$ and let $f \in \mathcal{F}(M)$. The Lie derivative of f with respect to X at $p \in M$ is defined as:

$$\boxed{(\mathcal{L}_X f)(p) = \lim_{t\to 0} \frac{f(\gamma_p(t)) - f(p)}{t}} \tag{30.18}$$

which closely resembles the standard derivative in Euclidean space.

Formula 30.3.13 (†). Working out previous formula and rewriting it as an operator equality gives:

$$\boxed{\mathcal{L}_X = \sum_k X_k \frac{\partial}{\partial x^k}} \tag{30.19}$$

It is clear that this is just the vector field X expanded in the basis 27.2.3. We also recover the behaviour of a tangent vector as a derivation. So for smooth functions $f : M \to \mathbb{R}$ we obtain:

$$\mathcal{L}_X f(p) = X_p(f) \tag{30.20}$$

Formula 30.3.14 (Lie derivative for vector fields†). Let $X, Y \in \mathfrak{X}(M)$

$$\boxed{\mathcal{L}_X Y = \frac{d}{dt}(\sigma_t^* X)(\gamma_p(t))\Big|_{t=0}} \tag{30.21}$$

Property 30.3.15. Let $X, Y \in \mathfrak{X}(M)$ be vector fields of class C^k. The Lie derivative has following properties:

- $\mathcal{L}_X Y$ is a vector field.
- **Lie bracket**:

  $$\mathcal{L}_X Y = [X, Y] \tag{30.22}$$

  which is also a derivation on $C^{k-1}(M, \mathbb{R})$ due to the cancellation of second-order derivatives in the local representation. It follows that the Lie derivative on vector fields turns the space $\mathfrak{X}(M)$ into a real Lie algebra.

[7] The third property follows from the other two.

- The Lie derivative is antisymmetric:

$$\mathcal{L}_X Y = -\mathcal{L}_Y X \tag{30.23}$$

This follows from the previous property.

Definition 30.3.16 (Holonomic basis). A basis $\{e_i\}$ of T_pM is said to be holonomic (on a neighbourhood $U \subseteq M$) if all the Lie derivatives vanish on U:

$$\mathcal{L}_{e_i} e_j = 0 \tag{30.24}$$

Equivalently, if the structure coefficients of the Lie algebra $\mathfrak{X}(M)$ vanish on U.

Property 30.3.17. For every holonomic basis there exists a coordinate system for M such that the basis coincides with coordinate-induced basis.

30.3.3 Frobenius' theorem

Definition 30.3.18 (Distribution). A smooth section of the Grassman k-plane bundle[8] is called a distribution of k-planes.

Definition 30.3.19 (Integrable). Let M be a smooth manifold and let $W \in \Gamma(\mathrm{Gr}(k, TM))$ be a distribution of k-planes. A submanifold $N \subseteq M$ is said to integrate W with initial condition $p_0 \in M$ if for every $p \in N$ we find that $W(p) = T_pN$ and $p_0 \in N$. W is said to be integrable if there exists such a submanifold N.

Definition 30.3.20 (Frobenius' integrability condition). A distribution W over a smooth manifold M is said to satisfy the Frobenius integrability condition in an open set $U \subseteq M$ if for every two vector fields X, Y defined on U, such that $X(p) \in W(p)$ and $Y(p) \in W(p)$ for all $p \in U$, there Lie bracket $[X, Y](p)$ is also an element of $W(p)$ for all $p \in U$.

Theorem 30.3.21 (Frobenius' integrability theorem). *Let W be a distribution over a smooth manifold M. Then W is integrable if and only if W satisfies the Frobenius integrability condition.*

30.4 Differential k-forms

Definition 30.4.1 (Differential form). A differential k-form is a map

$$\boxed{\omega : T^{\diamond k} M \to \mathbb{R}} \tag{30.25}$$

such that the restriction of ω to each fibre of the fibre product[9] $T^{\diamond k}M$ is multilinear and antisymmetric.

The space of all differential k-forms on a manifold M is denoted by $\Omega^k(M)$. Just like $\mathfrak{X}(M)$ it forms a $C^\infty(M)$-module. The set $\Omega^0(M)$ is defined as the space of smooth functions $C^\infty(M)$.

Alternative Definition 30.4.2. An alternative definition goes as follows. Consider the representation

$$\rho_k : GL(R^{m*}) \to GL(\Lambda^k(\mathbb{R}^{m*})) : T \mapsto T \wedge ... \wedge T$$

[8] See definition 30.2.14.
[9] See definition 29.8.

where T is a linear map. This representation induces an associated vector bundle[10] $\rho_k(\tau_M^*)$ of the cotangent bundle on M. A differential k-form is then given by a section of $\rho_k(\tau_M^*)$. $\Omega^k(M)$ can then be defined as follows:

$$\Omega^k(M) = \Gamma(\rho_k(\tau_M^*))$$

Construction 30.4.3. We can construct a Grassmann algebra[11] by equipping the graded vector space

$$\Omega(M) = \bigoplus_{k\geq 0} \Omega^k(M) \tag{30.26}$$

with the wedge product of differential forms (which is induced by the wedge product on $\Lambda^k(\mathbb{R}^m)$ through the alternative definition). This graded algebra is associative, graded-commutative and unital with the constant function $1 \in C^\infty(M)$ as identity element.

Definition 30.4.4 (Pullback). Let $f : M \to N$ be a smooth function between smooth manifolds and let ω be a differential k-form on N. The pullback of ω by f is defined as:

$$\boxed{f^*(\omega) = \omega \circ f_* : TM \to \mathbb{R}} \tag{30.27}$$

So f^* can be seen as a map pulling elements from T^*N back to T^*M.

Definition 30.4.5 (Pushforward). Let $f : M \to N$ be a diffeomorphism between smooth manifolds and let ω be a differential k-form on M. The pushforward ω by f is defined as:

$$f_*(\omega) : \omega \circ (f^{-1})_* : TN \to \mathbb{R} \tag{30.28}$$

Remark. Note that the pushforward of differential k-forms is only defined for diffeomorphisms, in constrast to pullbacks which only require smooth functions. This also explains why differential forms are the most valuable elements in differential geomeotry. Vector fields can't even be pulled back in general by smooth maps.

Formula 30.4.6 (Dual basis). Consider the coordinate basis from definition 27.2.3 for the tangent space T_pM. From this set we can construct[12] a natural dual basis for the cotangent space T_p^*M using the natural pairing of these spaces:

$$\left\langle \frac{\partial}{\partial x^i}, dx^j \right\rangle = \delta_i^j \tag{30.29}$$

30.4.1 Vector-valued differential forms

Definition 30.4.7 (Vector-valued differential form). Let V be a vector space and E a vector bundle with V as typical fibre. A vector-valued differential form can be defined in two ways. Firstly we can define a vector-valued k-form as a map $\omega : \bigotimes^k TM \to V$. A more general definition is based on sections of a corresponding vector bundle:

$$\Omega^k(M, E) = \Gamma(E \otimes \Lambda^k T^*M) \tag{30.30}$$

[10] See definition 30.2.8.

[11] As in definition 20.8.14.

[12] It should however be noted that dx^i is not just a notation. These basis vectors are in fact constructed by applying the exterior derivative 30.4.13 to the coordinate maps x^i.

Formula 30.4.8 (Wedge product). Let $\omega \in \Omega^k(M, E_1)$ and $\nu \in \Omega^p(M, E_2)$. The wedge product of these differential forms is defined as:

$$\omega \wedge \nu(v_1, ..., v_{k+p}) = \frac{1}{k!p!} \sum_{\sigma \in S_{k+p}} \text{sgn}(\sigma) \omega(v_{\sigma(1)}, ..., v_{\sigma(k)}) \otimes \nu(v_{\sigma(k+1)}, ..., v_{\sigma(p)}) \tag{30.31}$$

This is a direct generalization of the formula for the wedge product of ordinary differential forms where we replaced the (scalar) product (product in the algebra $\mathbb{R}$) by the tensor product (product in the tensor algebra). It should be noted that the result of this operation is not an element of any of the original bundles E_1 or E_2 but of the product bundle $E_1 \otimes E_2$.

Definition 30.4.9 (Lie algebra-valued). A vector-valued differential form where the vector space V is equipped with a Lie algebra structure.

Formula 30.4.10 (Wedge product). Let $\omega \in \Omega^k(M, \mathfrak{g})$ and $\nu \in \Omega^p(M, \mathfrak{g})$. The wedge product of these differential forms is defined as:

$$[\omega \wedge \nu](v_1, ..., v_{k+p}) = \frac{1}{k!p!} \sum_{\sigma \in S_{k+p}} \text{sgn}(\sigma) [\omega(v_{\sigma(1)}, ..., v_{\sigma(k)}), \nu(v_{\sigma(k+1)}, ..., v_{\sigma(p)})] \tag{30.32}$$

where $[\cdot, \cdot]$ is the Lie bracket in $\mathfrak{g}$.

Formula 30.4.11. Let $\{e_a\}$ be a basis for the Lie algebra $\mathfrak{g}$. We can then write the Lie algebra-valued differential forms as follows: $\phi = \phi^\mu \otimes e_\mu$ and $\psi = \psi^\nu \otimes e_\nu$ where ϕ^μ, ψ^ν are ordinary differential forms. The above formula for their wedge product can now be rewritten more elegantly as:

$$\boxed{[\phi \wedge \psi] = (\phi^\mu \wedge \psi^\nu) \otimes [e_\mu, e_\nu]} \tag{30.33}$$

where $\wedge$ is the wedge product in $\Omega(M)$ and $[\cdot, \cdot]$ is the Lie bracket in $\mathfrak{g}$.

Corollary 30.4.12. Using this formula it is easy to verify a number of properties similar to the ones of ordinary differential forms. As an example we give the analogon to the graded commutativity on $\Omega(M)$:

$$[\phi \wedge \psi] = (-1)^{pq+1} [\psi \wedge \phi] \tag{30.34}$$

where $\phi \in \Omega^p(M, \mathfrak{g}), \psi \in \Omega^q(M, \mathfrak{g})$. Here the extra factor -1 arises due to the antisymmetry of the Lie bracket.

Analogously one can prove that the Lie algebra-valued wedge product satisfies a graded Jacobi-type identity:

$$(-1)^{pr}[\phi \wedge [\psi \wedge \theta]] + (-1)^{pq}[\psi \wedge [\theta \wedge \phi]] + (-1)^{qr}[\theta \wedge [\phi \wedge \psi]] = 0 \tag{30.35}$$

where $\theta \in \Omega^r(M, \mathfrak{g})$.

30.4.2 Exterior derivative

Definition 30.4.13 (Exterior derivative). The exterior derivative d_k is a map defined on the graded algebra of differential k-forms:

$$d_k : \Omega^k(M) \to \Omega^{k+1}(M) \tag{30.36}$$

For $k = 0$ it is given by[13]:

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i \tag{30.37}$$

where we remark that the 'infinitesimals' are in fact unit vectors with norm 1. This formula can be generalized to higher dimensions as follows:

$$\boxed{d(f dx_{i_1} \wedge ... \wedge dx_{i_k}) = df \wedge dx_{i_1} \wedge ... \wedge dx_{i_k}} \tag{30.38}$$

Corollary 30.4.14. It follows immediately from 30.38 that

$$d(dx_i) = 0 \tag{30.39}$$

for all $i \leq n$.

Property 30.4.15. The exterior derivatives have following properties:

- For all $k \geq 0$, for all $\omega \in \Omega^k(M)$: $d_k \circ d_{k+1} = 0$, so $\text{im}(d_k) \subseteq \ker(d_{k+1})$.
- The exterior derivative is an $\mathbb{R}$-linear map.
- Graded **Leibniz rule**:

$$d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^j \omega_1 \wedge d\omega_2 \tag{30.40}$$

  where $\omega_1 \in \Omega^j(M), \omega_2 \in \Omega^k(M)$.
- Let $f \in C^\infty(M)$: $f^*(d\omega) = d(f^*\omega)$ where f^* denotes the pullback 30.27.

Remark 30.4.16 (†). The gradient, rotor (curl) and divergence from standard vector calculus[14] can be rewritten using exterior derivatives as follows: Let $\vec{\boldsymbol{f}} = (f_1, f_2, f_3)$ with f_i smooth for every i and let f be a smooth function. Denote the canonical isomorphism between $\mathbb{R}^3$ and $\mathbb{R}^{3*}$ by $\sim$.

$$\boxed{\begin{aligned} \sim df &= \nabla f \qquad (30.41)\\ \sim (*d\alpha) &= \nabla \times \vec{\boldsymbol{f}} \qquad (30.42)\\ *d\omega &= \nabla \cdot \vec{\boldsymbol{f}} \qquad (30.43) \end{aligned}}$$

The properties in section 20.1.2 then follow from the identity $d^2 = 0$.

Example 30.4.17. Let $f \in C^\infty(M, \mathbb{R})$. Let γ be a curve on M. From the definition 30.29 of the basis $\{dx_k\}_{k \leq n}$ we obtain following result:

$$\langle df(x), \gamma'(t) \rangle = \sum_k \frac{\partial f}{\partial x_k}(x) \gamma_k'(t) = (f \circ \gamma')(t) \tag{30.44}$$

Example 30.4.18. An explicit formula for the exterior derivative of a k-form Φ is:

$$\begin{aligned} d\Phi(X_1, ..., X_{k+1}) = &\sum_{i=0}^{k+1} (-1)^{i+1} X_i(\Phi(X_1, ..., \hat{X}_i, ..., X_{k+1})) \\ &+ \sum_{i<j} (-1)^{i+j} \Phi([X_i, X_j], X_1, ..., \hat{X}_i, ..., \hat{X}_j, ..., X_{k+1}) \end{aligned} \tag{30.45}$$

where $\hat{X}$ means that this argument is omitted.

Definition 30.4.19 (Codifferential). Using the de Rham differential d and the Hodge star operator 20.56 we can define a boundary operator[15] $\delta : \Omega^k(M) \to \Omega^{k-1}(M)$:

$$\delta = (-1)^k *^{-1} d* = (-1)^{n(k+1)+1} * d* \tag{30.46}$$

Definition 30.4.20 (Hodge Laplacian[16]). Using the de Rham differential and codifferential one can define a derivation $\Delta : \Omega^k(M) \to \Omega^k(M)$:

$$\boxed{\Delta = d\delta + \delta d} \tag{30.47}$$

[13] For $f \in \Omega^0(M)$, we call df the **differential** of f.
[14] See section 20.1.
[15] This means that $\delta^2 = 0$ (see 9.11).
[16] Sometimes called the Hodge-de Rham or Laplace-de Rham operator.

30.4.3 Lie derivative

Formula 30.4.21 (Lie derivative of differential forms).

$$\boxed{\mathcal{L}_X\omega(p) = \lim_{t\to 0}\frac{\sigma_t^*\omega - \omega}{t}(p)} \tag{30.48}$$

Formula 30.4.22 (Lie derivative of smooth functions). Using the definition of the exterior derivative of smooth functions 30.37 and the definition of the dual (cotangent) basis 30.29 we can rewrite the Lie derivative 30.19 as:

$$Xf(p) = df_p(X(p)) \tag{30.49}$$

Property 30.4.23. The Lie derivative also has following Leibniz-type property with respect to differential forms (this also follows from equation 30.45 and 30.22):

$$\mathcal{L}_X(\omega(Y)) = (\mathcal{L}_X\omega)(Y) + \omega(\mathcal{L}_X Y) \tag{30.50}$$

where X, Y are two vector fields and ω is a 1-form.

30.4.4 Interior product

Definition 30.4.24 (Interior product). Aside from the differential (exterior derivative) we can also define another operation on the algebra of differential forms:

$$\iota_X : (\iota_X\omega)(v_1, ..., v_{k-1}) \mapsto \omega(X, v_1, ..., v_{k-1}) \tag{30.51}$$

This antiderivation (of degree -1) from $\Omega^k(M)$ to $\Omega^{k-1}(M)$ is called the **interior product** or **interior derivative**. This can be seen as a generalization of the contraction map 20.36.

Formula 30.4.25 (Cartan's magic formula[17]). Let X be a vector field and let ω be a differential k-form. The Lie derivative of ω along X is given by the following formula:

$$\mathcal{L}_X\omega = \iota_X(d\omega) + d(\iota_X\omega) \tag{30.52}$$

30.4.5 Lie derivative in general

Formula 30.4.26 (Lie derivative of tensor fields). By comparing the definitions of the Lie derivatives of vector fields 30.21 and differential forms 30.48 we can see that both definitions are identical upon replacing X by ω. This leads to the definition of a Lie derivative of a general tensor field $\mathcal{T} \in \Gamma(T^{(k,l)}M)$:

$$\boxed{\mathcal{L}_X\mathcal{T}(p) = \left.\frac{d}{dt}\sigma_t^*\mathcal{T}(\gamma_p(t))\right|_{t=0}} \tag{30.53}$$

Alternative Definition 30.4.27 (Lie derivative of tensor fields). The Lie derivative of tensor fields can also be defined as the unique differential operator satisfying following axioms:

- $\mathcal{L}_X$ coincides with X on $\mathcal{F}(M)$.
- $\mathcal{L}_X$ satisfies the product rule with respect to the tensor product.
- $\mathcal{L}_X$ satisfies the Leibniz rule with respect to contraction.

[17] Sometimes called **Cartan's (infinitesimal) homotopy formula**.

- $\mathcal{L}_X$ commutes with the exterior derivative.

Property 30.4.28. Every derivation D of the tensor algebra can be decomposed as

$$D = \mathcal{L}_X + S \tag{30.54}$$

for some vector field $X \in TM$ and endomorphisms $S \in T^{1,1}M$.

30.5 de Rham cohomology

Definition 30.5.1 (Exact form). Let $\omega \in \Omega^k(M)$. If ω can be written as $\omega = d\chi$ for some $\chi \in \Omega^{k-1}(M)$ then ω is said to be exact. It follows that:

$$\mathrm{im}(d_k) = \{\omega \in \Omega^{k+1}(M) : \omega \text{ is exact}\} \tag{30.55}$$

Definition 30.5.2 (Closed form). Let $\omega \in \Omega^k(M)$. If $d\omega = 0$ then ω i said to be closed. It follows that:

$$\{\omega \in \Omega^k(M) : \omega \text{ is closed}\} \subseteq \ker(d_k) \tag{30.56}$$

Remark 30.5.3. From the first item of property 30.4.15 it follows that every exact form is closed. The converse however is not true[18].

Definition 30.5.4 (de Rham complex). The structure given by the sequence

$$0 \to \Omega^0(M) \to \Omega^1(M) \to ... \to \Omega^{\dim(M)}(M) \to 0 \tag{30.57}$$

together with the sequence of exterior derivatives d_k forms a cochain complex. This complex is called the de Rham complex.

The relation between closed and exact forms can be used to define the de Rham cohomology groups.

Definition 30.5.5 (de Rham cohomology). The k^{th} de Rham cohomology group on M is defined as the following quotient space:

$$\boxed{H^k_{\mathrm{dr}}(M) = \frac{\ker(d_k)}{\mathrm{im}(d_{k-1})}} \tag{30.58}$$

This quotient space is a vector space. Two elements of the same equivalence class in $H^k_{dr}(M)$ are said to be **cohomologous**.

Formula 30.5.6 (Cup product). Let $[\nu] \in H^k_{\mathrm{dr}}$ and $[\omega] \in H^l_{\mathrm{dr}}$. The cup product in de Rham cohomology is given by $[\nu] \smile [\omega] = [\nu \wedge \omega]$.

Theorem 30.5.7 (de Rham). *The de Rham cohomology over a smooth manifold M is isomorphic to the singular cohomology[19] over M.*

Theorem 30.5.8 (Poincaré lemma[20]). *For every point $p \in M$ there exists a neighbourhood on which the de Rham cohomology is trivial:*

$$\forall p \in M : \exists U \subseteq M : H^k_{dr}(U) = 0 \tag{30.59}$$

This implies that every closed form is locally exact.

More generally this lemma says that the following isomorphism exists for every smooth manifold M:

$$H^*_{dr}(M \times \mathbb{R}^n) \cong H^*_{dr}(M) \tag{30.60}$$

[18] See result 30.5.8 for more information.

[19] See section 9.3.

[20] The original theorem states that on a contractible space (see definition 9.1.7) every closed form is exact.

30.5.1 Compact support

Theorem 30.5.9 (Poincaré duality). *Let M be a smooth orientable n-dimensional manifold. The pairing $\int : H^k(M) \otimes H_c^{n-k}(M) \to \mathbb{R}$ induces an isomorphism on cohomology:*

$$H^k(M) \cong \left(H_c^{n-k}(M)\right)^* \tag{30.61}$$

If M is of finite type the converse holds:

$$H_c^k(M) \cong \left(H^{n-k}(M)\right)^* \tag{30.62}$$

Definition 30.5.10 (Poincaré dual). Let M be a smooth oriented n-dimensional manifold and let $i : S \to M$ be a closed[21] oriented k-dimensional submanifold. The Poincaré dual of S in M is the unique cohomology class $[\eta_S] \in H^{n-k}(M)$ such that:

$$\int_S i^*\omega = \int_M \omega \wedge \eta_S \tag{30.63}$$

for all compactly supported $\omega \in H_c^k(M)$.

If S is compact in M two Poincaré duals exist:

- Closed dual: The Poincaré dual obtained by using the fact that S is compact and hence closed in M.
- Compact dual: Because S is compact all (hence not only compactly supported) forms $\omega \in H^k(M)$ can be integrated over S and, assuming M is of finite type, Poincaré duality implies that there exists a unique cohomology class with compact support η'_S such that:

  $$\int_S i^*\omega = \int_M \omega \wedge \eta'_S \tag{30.64}$$

  for all $\omega \in H^k(M)$.

Remark 30.5.11. Because the compact Poincaré dual induces a pairing on all closed forms ω, which include the compactly supported ones, the compact dual is, as a form, equal to the closed Poincaré dual. However as classes in cohomology these can be quite different.

Property 30.5.12 (Localization principle). The support of the compact Poincaré dual of a submanifold S may be shrunk to any neighbourhood of S.

30.5.2 Compact support in the vertical direction

Definition 30.5.13. Let $\pi : E \to M$ be a smooth vector bundle over M. A differential form $\omega \in \Omega^*(E)$ is an element of $\Omega^*_{cv}(E)$ if $\mathrm{supp}(\omega) \cap \pi^{-1}(K)$ is compact for every compact $K \subset M$. The cohomology of this complex is called the **de Rham cohomology with compact support in the vertical direction**.

Corollary 30.5.14. The definition implies that $\omega \in \Omega^*_{cv}(E)$ is compactly supported on each fibre $\pi^{-1}(x), x \in M$. This observation explains the name of the cohomology theory.

[21] in the sense of topology

30.6 Linear connections

Definition 30.6.1 (Koszul connection). Let $\pi : E \to M$ be a vector bundle over a smooth manifold M. A Koszul connection on E is a (smooth) linear map $\nabla : \Gamma(E) \to \Gamma(T^*M \otimes E)$ satisfying the Leibniz property:

$$\nabla(f\sigma) = f\nabla\sigma + df \otimes \sigma \tag{30.65}$$

for all $f \in C^\infty(M)$.

Property 30.6.2. Because $\nabla\sigma$ *eats* a vector field, which is a $C^\infty(M)$-linear operation, we obtain:

$$\nabla_{fX+Y}\sigma = f\nabla_X\sigma + \nabla_Y\sigma \tag{30.66}$$

Formula 30.6.3. Let E, E' be two vector bundles over a base manifold M. Koszul connections on E, E' induce a connection on the tensor product bundle $E \otimes E'$ as follows:

$$\nabla(X \otimes Y) = \nabla X \otimes Y + X \otimes \nabla Y \tag{30.67}$$

for $X \in \Gamma(E), Y \in \Gamma(E')$.

Definition 30.6.4 (Affine connection). Let M be a smooth manifold. An affine connection $\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ is a Koszul connection on the tangent bundle.

Property 30.6.5 (Local behaviour). Let $v \in T_pM$. If two vector fields $X, Y \in \Gamma(TM)$ are equal on some neighbourhood U of p then $\nabla_v X = \nabla_v Y$ on U. Furthermore, given a curve $c : [0,1] \to M$ and two vector fields $X, Y \in \Gamma(TM)$ such that $X \circ c = Y \circ c$ we find that $\nabla_{\dot{c}} X = \nabla_{\dot{c}} Y$. Hence we see that an affine connection only depends on the local behaviour of the given section.

Remark 30.6.6. The above property shows the major difference between the Lie derivative and the covariant derivative when acting on sections of the tangent bundle σ. Lie derivatives depend on the local behaviour of both X and σ. The covariant derivative on the other hand only depends on the value of X at $p \in M$ and on the local behaviour of σ.

Property 30.6.7 (Affinity). Consider two affine connections $\nabla, \overline{\nabla}$ on a smooth manifold M. The operator $\nabla - \overline{\nabla}$ is an endormorphism of E, i.e. $\nabla - \overline{\nabla} \in \Omega^1(M; \mathrm{End}(E))$. It follows that the set of affine connections forms an affine space (hence the name).

Definition 30.6.8 (Parallel tensor fields). A tensor field T is said to be parallel with respect to a connection ∇ if it satisfies $\nabla T = 0$.

Example 30.6.9. Important examples are the volume form Vol and the metric g with respect to the Levi-Civita connection on a Riemannian manifold (see definition 33.1.12).

Definition 30.6.10 (Connection coefficients). Let E be smooth vector bundle. Consider a Koszul connection ∇ and a (local) frame $\{e_i\}$ and coframe $\{f^i\}$ of E. For every vector field e_i one can write the following:

$$\nabla e_i = \Gamma^k{}_{ji} e_k \otimes f^j \tag{30.68}$$

The quantities $\Gamma^k{}_{ji}$ are also often called the **Christoffel symbols** of ∇. For a general vector field $\sigma = \sigma^i e_i$ one then obtains (if $\{e_i\}$ and $\{f_i\}$ are coordinate-induced):

$$\begin{aligned} \nabla\sigma &= (\nabla\sigma^i) \otimes e_i + \sigma^i(\nabla e_i) \\ &= (\partial_j \sigma^k) e_k \otimes f^j + \sigma^i(\Gamma^k{}_{ji} e_k \otimes f^j) \\ &= (\partial_j \sigma^k + \Gamma^k{}_{ji}\sigma^i) e_k \otimes f^j \end{aligned} \tag{30.69}$$

Definition 30.6.11 (Hessian). Consider a smooth manifold with connection ∇. The Hessian of a function $f \in C^\infty(M)$ is defined as the iterated covariant derivative:

$$\text{Hess}(f) = \nabla\nabla f \tag{30.70}$$

For a scalar function we know that $\nabla f = df$ and for covector fields we know that (in local coordinates):

$$\nabla_i \sigma_j = \partial_i \sigma_j - \Gamma^k_{ij}\sigma_k$$

where Γ^k_{ij} are the connection symbols. Combining these facts we obtain the following local formula for the Hessian of f:

$$\text{Hess}(f) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} - \Gamma^k_{ij}\frac{\partial f}{\partial x_k}\right) dx^i \otimes dx^j \tag{30.71}$$

30.6.1 Induced connections

Formula 30.6.12 (Connection on differential forms). Applying the Leibniz property of a Koszul connection to tensor contractions gives us the following form of the induced connection on the (s,k)-tensor bundle:

$$\begin{aligned}\nabla_Y T(\omega^1,...,\omega^s,X_1,...,X_k) = Y\Big(T(\omega^1,...,\omega^s,X_1,...,X_k)\Big) \\ - \sum_{i=1}^{s} T(\omega^1,...,\nabla_Y\omega^i,...,\omega^s,X_1,...,X_k) \\ - \sum_{i=1}^{k} T(\omega^1,...,\omega^s,X_1,...,\nabla_Y X_i,...,X_k)\end{aligned} \tag{30.72}$$

where $Y, X_1, ..., X_k \in \mathfrak{X}(M)$ and $\omega^1, ..., \omega^s \in \Omega^1(M)$.

Corollary 30.6.13. By noting that the covariant derivative of a vector field is a vector-valued differential form we can use the previous formula to compute the covariant derivative of the covariant derivative:

$$(\nabla_X\nabla)_Y Z = \nabla_X\nabla_Y Z - \nabla_{\nabla_X Y} Z - \nabla_Y\nabla_X Z \tag{30.73}$$

Definition 30.6.14 (Pullback connection). Let $E \to M$ be a vector bundle with Koszul connection ∇. Let $f : M' \to M$ be a smooth map. On the pullback bundle[22] f^*E there exists a unique Koszul connection ∇' satisfying:

$$\nabla'(f^*\chi) = f^*(\nabla\chi) \tag{30.74}$$

for any section of E.

Definition 30.6.15 (Invariant connection). Let G be a Lie group acting on a vector bundle $E \to M$. A Koszul connection ∇ on E is said to be invariant with respect to the G-action if it satisfies:

$$g^*\nabla = \nabla \tag{30.75}$$

for all $g \in G$.

[22] See definition 29.1.18.

30.7 Characteristic classes

Definition 30.7.1 (Characteristic class). Let M be a smooth manifold. A characteristic class is a map $c : \mathrm{Vect}(M) \to H^*(M)$ such that if $E, E' \in \mathrm{Vect}(M)$ are equivalent, then $c(E) = c(E')$.

30.7.1 Chern-Weil theory

The characteristic classes of a vector bundle can be constructed from the connection and curvature forms on the vector bundle. This is done using a class of polynomials in the Lie algebra $\mathfrak{g}$ of the structure group.

Definition 30.7.2 (Invariant polynomial). Let G be a Lie group with Lie algebra $\mathfrak{g}$. A polynomial $P \in K[\mathfrak{g}]$, where $K = \mathbb{R}$ or $K = \mathbb{C}$, is said to be invariant (or sometimes Ad-invariant) if:

$$P(X) = P(gXg^{-1}) \tag{30.76}$$

for all $X \in \mathfrak{g}$ and $g \in G$. This subalgebra of $K[\mathfrak{g}]$ is denoted by $K[\mathfrak{g}]^G$.

Definition 30.7.3 (Chern-Weil morphism). Let $E \to M$ be a vector bundle with connection A and curvature F. There exists a morphism of algebras $K[\mathfrak{g}]^G \to H^*_{dR}(M) : P \mapsto P(F)$ satisfying:

- $P(F)$ is closed.
- $P(F)$ pulls back uniquely to a (closed) form $\overline{P}(F) := \pi^* P(F)$ on M.
- $\overline{P}(F)$ does not depend on the connection A, i.e. for connections A, A' the difference $\overline{P}(F_A) - \overline{P}(F_{A'})$ is exact.

Chapter 31

Principal Bundles

References for this chapter are [9].

31.1 Principal bundles

Definition 31.1.1 (Principal bundle). A principal bundle is a fibre bundle P together with a right action $\rho : P \times G \to P$ that satisfies two properties:

- Free action[1]: This implies that the orbits are isomorphic to the structure group.
- Fibrewise transitivity[2]: This action preserves fibres, i.e. $y \cdot g \in F_b$ for all $y \in F_b, g \in G$ which implies that the fibres over B are exactly the orbits of ρ.

Together these properties imply that the typical fibre F and structure group G can be identified.

Property 31.1.2. A corrolary of this definition is that the bundle $E \xrightarrow{\pi} M$ is isomorphic to the bundle $E \xrightarrow{\rho} E/G$ where E/G denotes the orbit space of E with respect to the G-action and ρ is the projection onto an equivalence class in the orbit space. (This property could have been used as part of the definition instead of the fibrewise transitivity.)

Remark 31.1.3. We remark that although the fibres are homeomorphic to G, they do not carry a group structure due to the lack of a distinct identity element. This turns them into G-torsors[3]. However it is possible to locally (i.e. in a neighbourhood of a point $p \in M$), but not globally, endow the fibres with a group structure by choosing an element of every fibre to be identity element.

Property 31.1.4 (Dimension). The dimension of P is given by:

$$\dim P = \dim M + \dim G \tag{31.1}$$

Property 31.1.5. Every local trivialization φ_i is G-equivariant:

$$\varphi_i(z \cdot g) = \varphi_i(z) \cdot g \tag{31.2}$$

Definition 31.1.6 (Principal bundle map). A bundle map $F : P_1 \to P_2$ between principal G-bundles is a pair of smooth maps (f_B, f_P) such that:

1. (f_B, f_P) is a bundle map in the sense of fibre bundles.

[1] See definition 3.1.61.
[2] See definition 3.1.63.
[3] See definition 3.1.66.

2. f_P is G-equivariant[4].

The map f_P is said to **cover** f_B.

Definition 31.1.7 (Vertical automorphism). Consider a principal G-bundle $\pi : P \to M$. An automorphism f of this bundle is said to be vertical if it covers the identity, i.e. $\pi \circ f = \pi$. It is exactly this subgroup $\mathrm{Aut}_V(P) \subset \mathrm{Aut}(P)$ which is known as the **group of gauge transformations** or **gauge group**[5] in physics.

31.1.1 Associated bundles

Construction 31.1.8 (Associated principal bundle). For every fibre bundle we can construct an associated principal G-bundle by replacing the fibre F by G itself using the fibre bundle construction theorem 29.1.10 where the left action of G is given by left multiplication in G.

Property 31.1.9. A fibre bundle ξ is trivial if and only if the associated principal bundle is trivial. More generally, two fibre bundles are isomorphic if and only if their associated principal bundles are isomorphic.

Example 31.1.10 (Frame bundle). Let V be an n-dimensional vector space. Denote the set of ordered bases (or **frames**) of V by $F(V)$. It follows from the fact that every basis transformation is given by the action of an element of the general linear group that $F(V)$ is isomorphic to $\mathrm{GL}(V) \cong \mathrm{GL}(\mathbb{R}^n)$.

Given an n-dimensional vector bundle E we can thus construct an associated principal bundle by replacing every fibre $\pi^{-1}(b)$ by $F(\pi^{-1}(b)) \cong \mathrm{GL}(\mathbb{R}^n)$. The right action on this bundle by $g \in \mathrm{GL}(\mathbb{R}^n)$ is given by the basis transformation $\widetilde{e}_j = g^i_j e_i$.

Property 31.1.11 (Trivial vector bundles). Property 30.2.18 can now be reformulated as follows: A vector bundle is trivial if and only if its associated frame bundle admits a global section.

Construction 31.1.12 (Associated bundle to a principal bundle). Consider a principal bundle $G \hookrightarrow P \to M$ and let F be a smooth manifold equipped with a left G-action $\triangleright$. One can then construct an associated bundle $P_F \equiv P \times_\triangleright F$ in the following way:

1. Define an equivalence relation $\sim_G$ on the product manifold $P \times F$ by:

$$(p, f) \sim_G (p', f') \iff \exists g \in G : (p', f') = (p \cdot g, g^{-1} \triangleright f) \tag{31.3}$$

2. The total space of the associated bundle is then given by the following quotient manifold:

$$P_F := (P \times F)/\sim_G \tag{31.4}$$

3. The projection map $\pi_F : P_F \to M$ is defined as

$$\pi_F : [p, f] \mapsto \pi(p) \tag{31.5}$$

where $[p, f]$ is the equivalence class of $(p, f) \in P \times F$ in the quotient manifold P_F.

[4] See definition 3.1.68.

[5] Do not confuse with the structure group G which is also sometimes called the gauge group in physics.

Example 31.1.13 (Tangent bundle). Starting from the frame bundle $F(M)$ over a manifold M one can reconstruct (up to a bundle isomorphism) the tangent bundle TM in the following way:

Consider the left G-action $\triangleright$ given by:

$$\triangleright\colon G \times \mathbb{R}^n \to \mathbb{R}^n : (g \triangleright f)^i = g^i{}_j f^j \tag{31.6}$$

The tangent bundle is bundle isomorphic to the associated bundle $LM \times_{\triangleright} \mathbb{R}^n$ where the bundle map is defined as $[e, v] \mapsto v^i e_i \in TM$.

Construction 31.1.14 (Associated bundle map). Given principal bundle map, denoted by (f_P, f_B), between two principal bundles one can construct an associated bundle map between any two of their associated bundles with the same typical fibre in the following way:

1. The total space map $\widetilde{f}_P : P \times_G F \to P \times_{G'} F$ is given by:
$$\widetilde{f}_P([p, f]) = [f_P(p), f] \tag{31.7}$$
2. the base space map is simply given by f_B itself:
$$\widetilde{f}_B(b) = f_B(b) \tag{31.8}$$

31.1.2 Sections

Although every vector bundle has at least one global section, the **zero section**[6], a general principal bundle does not necessarily have a global section. This is made clear by the following property:

Property 31.1.15. A principal G-bundle P is trivial if and only if there exists a global section on P. Furthermore, there exists a bijection between the set of all global sections $\Gamma(P)$ and the set of trivializations $\mathrm{Triv}(P)$.

Corollary 31.1.16. Every local section $\sigma : U \to P$ induces a local trivialization φ by:

$$\varphi^{-1} : (m, g) \mapsto \sigma(m) \cdot g \tag{31.9}$$

Property 31.1.11 can also be stated as follows:

Theorem 31.1.17. *A vector bundle is trivial if and only if its associated principal bundle is trivial.*

Property 31.1.18. Let (P, M, π, G) be a principal bundle and let P_F be an associated bundle. There exists a bijection between the sections of P_F and the G-equivariant maps $\phi : P \to F$, i.e. maps satisfying $\phi(p \cdot g) = g^{-1} \cdot \phi(p)$.

An explicit correspondence is given by:

$$\sigma_\phi : M \to P_F : m \mapsto [p, \phi(p)] \tag{31.10}$$

where p is any point[7] in $\pi^{-1}(\{m\})$. In the other direction we find:

$$\phi_\sigma : P \to F : p \mapsto j_p^{-1} \circ \sigma(\pi(p)) \tag{31.11}$$

where $j_p : F \to P_F : f \mapsto [p, f]$ is a bijection. Either of these maps (ϕ_σ or σ_ϕ) is called a **Higgs field** in the physics literature.

[6] This is the map $s : b \to \vec{\mathbf{0}}$ for all $b \in B$.

[7] This is well defined due to equation 31.3.

31.2 Reduction of the structure group

Construction 31.2.1. Consider a fibre bundle $\mathcal{F} = (E, B, \pi, F, G)$. Let H be a subgroup of G. If the transition functions of $\mathcal{F}$ can be chosen as elements of H then we say that the structure group G can be reduced to H.

Definition 31.2.2 (G-structure). Consider a manifold M. A G-structure on M is the reduction of the structure group $\mathrm{GL}(n)$ of the frame bundle $F(M)$ to a subgroup $G \subset \mathrm{GL}(n)$.

Example 31.2.3 (Orientable manifold). An n-dimensional manifold is orientable if and only if the structure group $\mathrm{GL}(n)$ of its frame bundle $F(M)$ is reducible to $\mathrm{GL}^+(n)$, i.e. the group of invertible matrices with positive determinant.

Example 31.2.4 (Riemannian manifold). An $\mathrm{O}(n)$-structure on M turns the manifold into a Riemannian manifold[8]. Because the cotangent bundle T^*M transforms[9] using the transpose inverse, which leaves $\mathrm{O}(n)$ invariant, of the transition maps of the tangent bundle TM these two bundles are equivalent. The isomorphism is given by the musical isomorphism(s)[10].

31.2.1 Spinor bundles

In this subsection we only work with Riemannian manifolds[11] (M, g) because this ensures the existence of an $\mathrm{O}(n)$ reduction of the tangent bundle TM (see the example above).

Definition 31.2.5 (Spin structure). Consider the (oriented) orthonormal frame bundle $\pi_{SO} : F_{SO}(M) \to M$ which is obtained by reducing the structure group of the frame bundle $F(M)$ from $\mathrm{GL}(n)$ to $\mathrm{SO}(n)$. Furthermore, let $\pi_{spin} : P_{spin} \to M$ be a principal $\mathrm{Spin}(n)$-bundle over M.

The smooth manifold M is said to have a spin structure if there exists an equivariant 2-fold lifting of F_{SO} to P_{spin}, i.e. a morphism $\xi : P_{spin} \to F_{SO}(M)$ together with the 2-fold covering map $\rho : \mathrm{Spin}(n) \to \mathrm{SO}(n)$ that satisfy:

- $\pi_{SO} \circ \xi = \pi_{spin}$
- $\xi(p \lhd g) = \xi(p) \cdot \rho(g)$

for all $g \in \mathrm{Spin}(n)$, where $\lhd$ and $\cdot$ denote the right actions of the respective structure groups. If M admits a spin structure it is often called a **spin manifold** and the principal $\mathrm{Spin}(n)$-bundle P is called the **spin frame bundle**.

Property 31.2.6. A smooth orientable manifold M is spin if and only if its second Stiefel-Whitney class vanishes.

Property 31.2.7. A special case occurs when $\dim M = 3$. We then have that M is spin if it is compact and orientable.

Definition 31.2.8 (Spin bundle). A spin bundle is a vector bundle associated to a spin frame bundle.

Definition 31.2.9 (Spinor field). A spinor field is a (smooth) section of a spin bundle.

[8] See definition 33.1.3.
[9] See example 30.2.10.
[10] See definition 33.1.2.
[11] This also works for Lorentzian manifolds and $\mathrm{Spin}(1, n-1)$ groups.

31.3 Universal bundle

Definition 31.3.1 (Universal bundle). Consider a (compact) Lie group G. The universal bundle of G is the principal bundle

$$G \hookrightarrow EG \to BG$$

where EG is weakly contractible and G acts freely on EG. The space BG is called the **classifying space** of G.

Definition 31.3.2 (n-universal bundle). A principal bundle with $(n-1)$-connected total space.

Property 31.3.3 (Classification). The collection of principal G-bundles over a CW-complex X is in bijection with $[X, BG]$, i.e. the homotopy classes of continuous functions $f : X \to BG$. This bijection is given by $f \mapsto f^*EG$.

Property 31.3.4. A principal G-bundle $EG \to BG$ is universal if and only if EG is weakly contractible.

Remark 31.3.5. There also exists a subtly different notion of universal bundles and their associated classifying property. When we require the total space of the universal bundle to be contractible instead of weakly contractible, the mapping space $[X, BG]$ only classifies numerable principal bundles[12], but now over arbitrary base spaces X.

An explicit construction of the numerable universal bundle for any topological group G was provided by Milnor:

Construction 31.3.6 (Milnor). First consider the infinite join E_∞ constructed as the direct limit of finite joins $E_n = \underbrace{G \circ \cdots \circ G}_{n \text{ times}}$ where E_n is embedded in E_{n+1} using the identity element , equipped with the strong topology. Then construct the quotient of E_n (resp. E_∞) by the canonical right action of G on E_n. The bundle $p_n : E_n \to B_n$ (resp. $p : E_\infty \to B_\infty$) is an n-universal bundle (resp. ∞-universal bundle). It follows from the above property that $p : E_\infty \to B_\infty$ is a universal bundle for G.

31.4 Connections

31.4.1 Vertical vectors

Because smooth fibre bundles (which include smooth principal G-bundles) are also smooth manifolds we can define the traditional notions for them, such as the tangent bundle. We use these to construct the horizontal and vertical (sub)bundles:

Definition 31.4.1 (Vertical vector). Let $\pi : E \to B$ be a smooth fibre bundle. The subbundle $\ker(\pi_*)$ of TE is called the vertical bundle of E. Fibrewise this gives us $V_x = T_x(E_{\pi(x)})$.

For principal G-bundles we can use an equivalent definition:

Alternative Definition 31.4.2. Consider a smooth principal G-bundle $G \hookrightarrow P \xrightarrow{\pi} M$. We first construct a map ι_p for every element $p \in P$:

$$\iota_p : G \to P : g \mapsto p \cdot g \tag{31.12}$$

[12] See definition 29.1.7.

We then define a tangent vector $v \in T_pP$ to be vertical if it lies in the image of $\iota_{p,*}$, i.e. $\mathrm{Vert}(T_pP) = \mathrm{im}(\iota_{p,*})$. This construction is supported by the exactness of following short sequence:

$$0 \to \mathfrak{g} \xrightarrow{\iota_{p,*}} T_pP \xrightarrow{\pi_*} T_xM \to 0 \tag{31.13}$$

Property 31.4.3 (Dimension). It follows from the second definition that the vertical vectors of a principal G-bundle are nothing but the pushforward of the Lie algebra $\mathfrak{g}$ under the right action of G on P. Furthermore, the exactness of the sequence implies that $\iota_{p,*} : \mathfrak{g} \to \mathrm{Vert}(T_pP)$ is an isomorphism of vector spaces. In particular, it implies that

$$\dim \mathrm{Vert}(T_pP) = \dim \mathfrak{g} = \dim G \tag{31.14}$$

Definition 31.4.4 (Fundamental vector field). Consider a principal G-bundle. Let $A \in \mathfrak{g}$, where $\mathfrak{g}$ is the Lie algebra corresponding to G. The vertical vector field $A^\# : P \to TP$ given by

$$A^\#(p) = \iota_{p,*}(A) \in \mathrm{Vert}(T_pP) \tag{31.15}$$

is called the fundamental vector field associated to A.

Alternative Definition 31.4.5. An equivalent definition of the fundamental vector field $A^\#(p)$ is given by:

$$A^\#_p(f) = \frac{d}{dt} f(p \cdot \exp(tA))\bigg|_{t=0} \tag{31.16}$$

where $f \in C^\infty(P)$.

Property 31.4.6. The map $(\cdot)^\# : \mathfrak{g} \to \Gamma(TP)$ is a Lie algebra morphism:

$$[A, B]^\# = [A^\#, B^\#] \tag{31.17}$$

where the Lie bracket on the left is that in $\mathfrak{g}$ and the Lie bracket on the right is that in $\mathfrak{X}(M)$ given by 30.22.

Property 31.4.7. The vertical bundle satisfies the following G-equivariance condition:

$$R_{g,*}(\mathrm{Vert}(T_pP)) = \mathrm{Vert}(T_{pg}P) \tag{31.18}$$

By differentiating the equality

$$R_g \circ \iota_p = \iota_{pg} \circ \mathrm{ad}_{g^{-1}}$$

and using 28.3.2, 31.15 we obtain the following algebraic formulation of the G-equivariance condition:

$$R_{g,*}\left(A^\#(p)\right) = \left(\mathrm{Ad}_{g^{-1}} A\right)^\#(pg) \tag{31.19}$$

31.4.2 Ehresmann connections

Definition 31.4.8 (Ehresmann connection). Consider a smooth fibre bundle P. An (Ehresmann) connection on P is the selection of a subspace $\mathrm{Hor}(T_pP) \leq T_pP$ for every $p \in P$ such that:

- $\mathrm{Vert}(T_pP) \oplus \mathrm{Hor}(T_pP) = T_pP$
- The selection depends smoothly on p.[13]

The elements of $\mathrm{Hor}(T_pP)$ are said to be **horizontal vectors** with respect to the connection.

[13] See the definiton of a (smooth) distribution 30.3.18.

Definition 31.4.9 (Principal connection). A principal connection on a smooth principal G-bundle P is a G-equivariant Ehresmann connection, i.e. an Ehresmann connection for which the horizontal subspaces satisfy following G-equivariance condition:

$$R_{g,*}(\mathrm{Hor}(T_pP)) = \mathrm{Hor}(T_{pg}P) \tag{31.20}$$

where R_g denotes the right action of G on P

Remark 31.4.10. Remark that this condition is automatically satisfied for vertical bundles (see equation 31.18).

Definition 31.4.11 (Horizontal bundle). The horizontal (sub)bundle $\mathrm{Hor}(TP)$ is defined as $\bigsqcup_{p\in P}\mathrm{Hor}(T_pP)$. The G-equivariance condition then implies that this subbundle is invariant under (the pushforward of) the right action of G.

Property 31.4.12 (Dimension). Properties 31.1, 31.14 and the direct sum decomposition of T_pP imply the following relation:

$$\dim \mathrm{Hor}(T_pP) = \dim M \tag{31.21}$$

Here we briefly summarize all dimensional relations between the components of a principal G-bundle over a base manifold M:

$$\boxed{\begin{aligned} \dim P &= \dim M + \dim G && (31.22)\\ \dim M &= \dim \mathrm{Hor}(T_pP) && (31.23)\\ \dim G &= \dim \mathrm{Vert}(T_pP) && (31.24)\end{aligned}}$$

for all $p \in P$.

Definition 31.4.13 (Horizontal and vertical forms). Let $\theta \in \Omega^k(P)$ be a differential k-form. We define following notions:

- θ is said to be horizontal if
$$\theta(v_1, ..., v_k) = 0 \tag{31.25}$$
whenever at least 1 of the v_i lies in $\mathrm{Vert}(T_pP)$.
- θ is said to be vertical if
$$\theta(v_1, ..., v_k) = 0 \tag{31.26}$$
whenever at least 1 of the v_i lies in $\mathrm{Hor}(T_pP)$.

For functions $f \in \Omega^0(P)$ it is vacuously true that they are both vertical and horizontal.

Definition 31.4.14 (Dual connection). First we define the dual of the horizontal bundle:

$$\mathrm{Hor}(T_p^*P) = \{h^* \in T_p^*P | h^*(v) = 0, v \in \mathrm{Vert}(T_pP)\} \tag{31.27}$$

It is the set horizontal 1-forms. A dual connection can then be defined as the selection of a vertical covector bundle $\mathrm{Vert}(T_p^*P)$ satisfying the conditions of definition 31.4.8 and 31.4.9 (where Vert and Hor should be interchanged).

31.4.3 Connection form

Definition 31.4.15 (Connection one-form). Let (P, M, π, G) be a principal bundle. A connection one-form, related to a given principal connection, is a $\mathfrak{g}$-valued 1-form $\omega : \Gamma(TP) \to \mathfrak{g}$ that satisfies the following 2 conditions:

1. Cancellation of fundamental vector fields:
$$\omega(A^{\#}) = A \tag{31.28}$$
2. G-equivariance:
$$\omega \circ R_{g,*} = \mathrm{Ad}_{g^{-1}} \circ \omega \tag{31.29}$$

The horizontal subspaces are then defined as $\mathrm{Hor}(T_pP) = \ker \omega|_p$.

Formula 31.4.16. Consider a principal G-bundle P. Given a principal connection on P, the associated connection one-form is given by the following map:
$$\omega = (\iota_{p,*})^{-1} \circ \mathrm{pr}_V \tag{31.30}$$
where pr_V is the projection $TP \to \mathrm{Vert}(TP)$ associated to the decomposition from definition 31.4.8.

Formula 31.4.17. Consider a principal bundle (P, M, π, G) and an associated vector bundle $P \times_G V$. For every G-equivariant map $\phi : P \to V$ and any $X \in \mathfrak{g}$ we find that
$$d\phi(X^{\#}) + [\omega \wedge \phi](X^{\#}) = 0 \tag{31.31}$$
where the left action of $\mathfrak{g}$ is induced by the representation of G on V.

Property 31.4.18. Consider two principal G-bundles ξ_1 and ξ_2. Let ω be a connection one-form on ξ_1 and let $F : \xi_1 \to \xi_2$ be a bundle map. The map $F^*\omega$ defines an Ehresmann connection on ξ_2.

31.4.4 Maurer-Cartan form

Definition 31.4.19 (Maurer-Cartan form). For every $g \in G$ we have that the tangent space T_gG is isomorphic to $T_eG = \mathfrak{g}$. The isomorphism $T_gG \to \mathfrak{g}$ is given by the Maurer-Cartan form:
$$\boxed{\Omega := L_{g^{-1},*}} \tag{31.32}$$

Construction 31.4.20. Consider a manifold $M = \{x\}$. When constructing a principal G-bundle over M we see that the total space $P = \{x\} \times G$ can be identified with the structure group G. From the relations in property 31.4.12 it follows that the horizontal spaces are null-spaces (which indeed defines a smooth distribution and thus a connection according to 31.4.8) and that the vertical spaces are equal to the tangent spaces, i.e. $\mathrm{Vert}(T_gG) = T_gG$ (where we used the association $P \cong G$).

The simplest way to define a connection form ω on this bundle would be the trivial projection $TP \to \mathrm{Vert}(TP) = \mathbb{1}_{TP}$. The image of this map would however be T_gG and not $\mathfrak{g}$ as required. This can be solved by using the Maurer-Cartan form $\Omega : T_gG \to \mathfrak{g}$, i.e. we define $\omega(v) = \Omega(v)$.

Property 31.4.21. The Maurer-Cartan form is the unique Principal connection on the bundle $G \hookrightarrow G \to \{x\}$.

31.4.5 Local representation

Definition 31.4.22 (Yang-Mills field). Consider a principal bundle (P, M, π, G) and an open subset $U \subseteq M$. Given an Ehresmann connection ω on P and a local section $\sigma : U \to P$, we define the Yang-Mills field $\omega^U : \Gamma(TU) \to \mathfrak{g}$ as follows:

$$\omega^U = \sigma^* \omega \tag{31.33}$$

Definition 31.4.23 (Local representation). Consider a principal bundle (P, M, π, G). Let (U, φ) be a bundle chart on P. The local representation of an Ehresmann connection ω on P with respect to the chart (U, φ) is given by $(\varphi^{-1})^* \omega$.

Formula 31.4.24. Consider an Ehresmann connection ω on a principal bundle (P, M, π, G). According to property 31.1.16 every local section $\sigma : U \to P$ induces both a Yang-Mills field ω^U and a local representation of ω. These two forms are related by the following equation:

$$h^* \omega_{(m,g)}(v, X) = \mathrm{Ad}_{g^{-1}}(\omega^U_m(v)) + \Omega_g(X) \tag{31.34}$$

where $v \in T_m U, X \in \mathfrak{g}$, Ω is the Maurer-Cartan form on G and h is the local trivialization induced by σ.

Formula 31.4.25 (Compatibility condition). Consider a principal bundle (P, M, π, G) and 2 open subsets U, V of M. Given 2 local sections $\sigma_U : U \to P$ and $\sigma_V : V \to P$ and an Ehresmann connection ω on P, we can define two Yang-Mills field ω^U and ω^V on M.

On the intersection $U \cap V$ we can find a (unique) gauge transformation $\xi : U \cap V \to G$ such that $\sigma_V(m) = \sigma_U(m) \cdot \xi(m)$. Using this gauge transformation we can relate ω^U and ω^V as follows:

$$\omega^V_m = \mathrm{Ad}_{\xi(m)^{-1}} \omega^U_m + (\xi^* \Omega)_m \tag{31.35}$$

where Ω is the Maurer-Cartan form on G.

Example 31.4.26 (General linear group[14]). Let $G = \mathrm{GL}(\mathbb{R}^n)$. The second term in equation 31.35 can be written as follows:

$$(\xi^* \Omega)^i{}_j = (\xi(m)^{-1})^i{}_k \frac{\partial}{\partial x^\mu} \xi(p)^k{}_j dx^\mu \tag{31.36}$$

at every point $m \in M$. Formally this can be written coordinate-independently as:

$$\xi^* \Omega = \xi^{-1} d\xi \tag{31.37}$$

Example 31.4.27 (Christoffel symbols). Let $\Gamma^i{}_{j\mu}, \overline{\Gamma}^k{}_{l\nu}$ be the Yang-Mills fields corresponding to a connection of a frame bundle, where the sections are induced by a choice of coordinates (x^i and y^i respectively). In this case, the expansion coefficients of the Yang-Mills field are called the **Christoffel symbols**[15]. Using equations 31.35 and 31.37 this becomes:

$$\overline{\Gamma}^i{}_{j\mu} = \frac{\partial y^\nu}{\partial x^\mu} \left(\frac{\partial x^i}{\partial y^k} \Gamma^k{}_{l\nu} \frac{\partial y^l}{\partial x^j} + \frac{\partial x^i}{\partial y^k} \frac{\partial^2 y^k}{\partial x^j \partial x^\nu} \right) \tag{31.38}$$

[14] A derivation can be found in lecture 22 of [16].

[15] See also equation 26.40.

31.4.6 Horizontal lifts and parallel transport

Definition 31.4.28 (Horizontal lift). Consider a principal G-bundle $G \hookrightarrow P \to M$ and a curve $\gamma : [0,1] \to M$. Let $p_0 \in \pi^{-1}(\gamma(0))$. There exists a unique curve $\widetilde{\gamma}_{p_0} : [0,1] \to P$ satisfying the following conditions:

- $\widetilde{\gamma}_{p_0}(0) = p_0$
- $\pi \circ \widetilde{\gamma}_{p_0} = \gamma$
- $\widetilde{\gamma}'_{p_0}(t) \in \mathrm{Hor}(TP)$ for all $t \in [0,1]$

The curve $\widetilde{\gamma}_{p_0}$ is said to be the **horizontal lift** of γ starting at p_0. When it is clear from the context what the basepoint p_0 is, the subscript is often ommited and we write $\widetilde{\gamma}$ instead of $\widetilde{\gamma}_{p_0}$.

Remark 31.4.29 (Horizontal curve). Curves satisfying the last condition in the above property are said to be horizontal.

Method 31.4.30. Consider a principal bundle $G \hookrightarrow P \to M$. Let $\gamma(t)$ be a curve in M and let ω be an Ehresmann connection on P. For general structure groups G, the horizontal lift can be found as follows: Let $\delta(t)$ be a curve in P that projects onto $\gamma(t)$, i.e. $\pi \circ \delta = \gamma$, such that $\widetilde{\gamma}_{p_0}(t) = \delta(t) \cdot g(t)$ for some curve $g(t)$ in G. The curve $g(t)$ can then be found as the unique solution of the following first order ODE:

$$\mathrm{Ad}_{g(t)^{-1}} \omega_{\delta(t)}(X_{\delta,\delta(t)}) + \Omega_{g(t)}(Y_{g,g(t)}) = 0 \tag{31.39}$$

where X_δ, Y_g are tangent vectors to respectively the curves $\delta(t)$ and $g(t)$ and where Ω is the Maurer-Cartan form on G. As initial value condition we use $\delta(0) \cdot g(0) = p_0$.

Remark 31.4.31. When given a local section $\sigma : U \to P$ we can rewrite the ODE in a more explicit form. First we remark that the section induces a curve $\delta = \sigma \circ \gamma$. Taking the derivative yields $X_\delta = \sigma_*(X_\gamma)$. Using this we can rewrite the ODE as

$$\mathrm{Ad}_{g(t)^{-1}} \omega_{\delta(t)}(\sigma_* X_{\gamma,\gamma(t)}) + \Omega_{g(t)}(Y_{g,g(t)}) = 0 \tag{31.40}$$

By using the equality $f^*\omega = \omega \circ f_*$ and introducing the Yang-Mills field $A = \sigma^*\omega$ this becomes:

$$\mathrm{Ad}_{g(t)^{-1}} A(X_{\gamma,\gamma(t)}) + \Omega_{g(t)}(Y_{g,g(t)}) = 0 \tag{31.41}$$

Example 31.4.32. For matrix Lie groups the above ODE can be reformulated as follows: Given the trivial section $s : U \to U \times G : x \mapsto (x, e)$, where U is an open subset of M, the horizontal lift of $\gamma(t)$ can locally be parametrized as $\widetilde{\gamma}(t) = \underbrace{(s \circ \gamma)(t)}_{\delta(t)} \cdot g(t) = (\gamma(t), g(t))$ where $g(t)$ is a curve in G. To determine $\widetilde{\gamma}(t)$ it is thus sufficient to find $g(t)$. The ODE 31.39 then becomes:

$$g'(t) = -\omega(\gamma(t), e, \gamma'(t), 0) g(t) \tag{31.42}$$

Using the trivial section we can rewrite this formula. First we consider the action of the Yang-Mills field $s^*\omega$ on the derivative $\gamma_* = (\gamma(t), \gamma'(t))$. Using the fact that it is linear in the second argument we can write:

$$s^*\omega(\gamma(t), \gamma'(t)) = A(\gamma(t))\gamma'(t)$$

where $A : M \to \mathrm{Hom}(\mathbb{R}^{\dim M}, \mathfrak{g})$ gives a linear map for each point $\gamma(t) \in M$. The action can also be rewritten using the relation $f^*\omega = \omega \circ f_*$ as

$$s^*\omega(\gamma(t), \gamma'(t)) = \omega\Big(s_*(\gamma(t), \gamma'(t))\Big) = \omega(\gamma(t), e, \gamma'(t), 0)$$

Combining these relations with the ODE 31.42 gives:

$$\left(\frac{d}{dt} + A(\gamma(t))\gamma'(t)\right) g(t) = 0 \tag{31.43}$$

where $\frac{d}{dt}$ is the matrix given by the scalar multiplication of the derivative $\frac{d}{dt}$ and the identity matrix I.

Method 31.4.33. The ODE 31.39 can now be solved. We explicitly assume that G is a matrix Lie group such that we can start from equation 31.43. Direct intergation and iteration gives us:

$$g(t) = \left[I - \int_0^t dt_1 A(\gamma'(t_1)) + \int_0^t dt_1 \int_0^{t_1} dt_2 A(\gamma'(t_1))A(\gamma'(t_2)) - ...\right] g(0) \tag{31.44}$$

where A is the Yang-Mills field corresponding to the local section σ. This can be rewritten using the standard "square integration" trick[16] as:

$$g(t) = \left[I - \int_0^t dt_1 A(\gamma'(t_1)) + \frac{1}{2!}\int_0^t dt_1 \int_0^t dt_2 \mathcal{T}\Big(A(\gamma'(t_1))A(\gamma'(t_2))\Big) - ...\right] g(0) \tag{31.45}$$

By noting that this formula is equal to the path-ordered exponential series we find:

$$\boxed{g(t) = \mathcal{T}\exp\left(-\int_0^t dt' A(\gamma'(t'))\right) g(0)} \tag{31.46}$$

Definition 31.4.34 (Parallel transport). The parallel transport map with respect to the curve γ is defined as follows:

$$\mathrm{Par}_t^\gamma : \pi^{-1}(\gamma(0)) \to \pi^{-1}(\gamma(t)) : p_0 \mapsto \widetilde{\gamma}_{p_0}(t) \tag{31.47}$$

This map is G-equivariant and it is an isomorphism of fibres.

31.4.7 Holonomy

Definition 31.4.35 (Holonomy group). Consider a principal bundle $G \hookrightarrow P \to M$. Let $\Omega_m^{ps}M \subset \Omega_m M$ be the subset of the loop space with basepoint $m \in M$ consisting of piecewise smooth loops. The holonomy group $\mathrm{Hol}_p(\omega)$ based at $p \in \pi^{-1}(m) \subset P$ with respect to the connection form ω is given by:

$$\mathrm{Hol}_p(\omega) = \{g \in G : p \sim p \cdot g\} \tag{31.48}$$

where two point $p, q \in P$ are equivalent if there exists a loop $\gamma \in \Omega_m^{ps}M$ such that the horizontal lift $\widetilde{\gamma}$ connects p and q.

Definition 31.4.36 (Reduced holonomy group). The reduced holonomy group $\mathrm{Hol}_p^0(\omega)$ is defined as the subset of $\mathrm{Hol}_p(\omega)$ using only contractible loops.

31.4.8 Koszul connections and covariant derivatives

Definition 31.4.37 (Horizontal lifts on associated bundles). Let $P_F = P \times_G F$ be an associated bundle of a principal bundle (P, M, π, G). Let γ be a curve in M with horizontal lift $\widetilde{\gamma}_p$ in P. The horizontal lift of γ in P_F through a point $[p, f] \in P_F$ is defined as follows:

$$\widetilde{\gamma}_{[p,f]}^{P_F}(t) = [\widetilde{\gamma}_p(t), f] \tag{31.49}$$

Although the fibre element f seems to stay constant along the horizontal lift, it in fact changes according to formula 31.3.

[16] Well known from the Dyson series 57.9.

Definition 31.4.38 (Parallel transport on associated bundles). Similar to the case of principal bundles P, the parallel transport map on an associated bundle P_F is defined as:

$$\mathrm{Par}_t^\gamma : \pi_F^{-1}(\gamma(0)) \to \pi_F^{-1}(\gamma(t)) : [p, f] \mapsto \widetilde{\gamma}_{[p,f]}^{P_F}(t) \tag{31.50}$$

Example 31.4.39 (Parallel transport on vector bundles). Consider a principal bundle $G \hookrightarrow P \to M$. Suppose that the Lie group G acts on a vector space V by a representation $\rho : G \to \mathrm{GL}_m$. We can then construct an associated vector bundle $\pi_1 : P \times_{GL(V)} V \to M$. Assume further that we work on a chart (U, φ) such that we can locally write P, P_V as product bundles.

Parallel transport on this vector bundle is then defined as follows. Let $\gamma(t)$ be a curve in M such that $\gamma(0) = x_0$ and $\gamma(1) = x_1$. Furthermore, let the horizontal lift $\widetilde{\gamma}(t) = (\gamma(t), g(t))$ satisfy $\widetilde{\gamma}(0) = (x_0, h)$ as initial condition. The parallel transport of the point $(x_0, v_0) \in U \times V$ along γ is given by the following map:

$$\mathrm{Par}_t^\gamma : \pi_1^{-1}(x_0) \to \pi_1^{-1}(\gamma(t)) : (x_0, v_0) \mapsto \big(\gamma(t), \rho\big(g(t)h^{-1}\big)v_0\big) \tag{31.51}$$

It should be noted that this map is independent of the initial element $h \in G$ (despite the presence of the factor h^{-1}). Furthermore, Par_t^γ is an isomorphism of vector spaces and can thus be used to identify distant fibers (as long as they lie in the same path-component).

Remark 31.4.40. For every vector bundle one can construct the frame bundle and use the parallel transport map on this bundle to define parallel transport of vectors. Hence the previous construction is valid for any vector bundle.

Definition 31.4.41 (Covariant derivative). Consider a vector bundle with model fibre space V and its associated principal $\mathrm{GL}(V)$-bundle with Ehresmann connection ω, both over a base manifold M. Let $\sigma : M \to E$ be a section of the vector bundle and let X be a vector field on M. The covariant derivative of σ with respect to X is defined as:

$$\boxed{\nabla_X \sigma|_{x_0} = \lim_{t \to +\infty} \frac{(\mathrm{Par}_t^\gamma)^{-1}\sigma(\gamma(t)) - \sigma(x_0)}{t}} \tag{31.52}$$

where $\gamma(t)$ is any curve such that $\gamma(0) = x_0$ and $\gamma'(0) = X(x_0)$.

Property 31.4.42. The map

$$\Gamma(TM) \times \Gamma(E) \to \Gamma(E) : (X, \sigma) \mapsto \nabla_X \sigma \tag{31.53}$$

gives a Koszul connection 30.6.1. It follows that every Ehresmann connection on a principal bundle induces a Koszul connection on all of its associated vector bundles.

31.4.9 Exterior covariant derivative

Definition 31.4.43 (Exterior covariant derivative). Consider a principal bundle $G \hookrightarrow P \to M$ equipped with an Ehresmann connection ω. Let $\theta \in \Omega^k(P)$ be a differential k-form. The exterior covariant derivative $D\theta$ is defined as follows:

$$D\theta(v_0, ..., v_k) = d\theta(v_0^H, ..., v_k^H) \tag{31.54}$$

where d is the exterior derivative 30.4.13 and v_i^H is the projection of v_i on the horizontal subspace $\mathrm{Hor}(T_pP)$ associated to the Ehresmann connection ω. From the definition it follows that the exterior covariant derivative $D\theta$ is a horizontal form[17].

[17] See definition 31.4.13.

Remark 31.4.44. The exterior covariant derivative can also be defined for general W-valued k-forms where W is a vector space. This can be done by defining it component-wise with respect to a given basis on W. Afterwards one can prove that the choice of basis plays no role.

Property 31.4.45. If ϕ is equivariant than $D\phi$ is a tensorial form.

Formula 31.4.46. Using the Koszul connection on the tangent bundle TP we can rewrite the action of the exterior covariant derivative as follows:

$$\boxed{D\theta(v_0,...,v_k) = \sum_i^k (-1)^i \nabla_{v_i} \theta(v_0,...,\hat{v}_i,...,v_k) + \sum_{i<j}^k (-1)^{i+j} \theta([v_i,v_j],v_0,...,\hat{v}_i,...,\hat{v}_j,...,v_k)} \tag{31.55}$$

where $\hat{v}_i$ means that this vector is omitted. As an example we explicitly give the formula for a 1-form Φ:

$$D\Phi(X,Y) = \nabla_X(\Phi(Y)) - \nabla_Y(\Phi(X)) - \Phi([X,Y]) \tag{31.56}$$

which should remind the reader of the analogous formula for the ordinary exterior derivative 30.45.

By property 31.1.18 we can use the following construction to find an explicit expression for the covariant derivative on an associated vector bundle:

Construction 31.4.47. Let (P,M,π,G) be a principal bundle and let $P_V := P \times_G V$ be an associated vector bundle. Given a section $\sigma : M \to P_V$ we can construct a G-equivariant map $\phi : P \to V$ using formula 31.11.

First we construct the exterior covariant derivative of ϕ:

$$D\phi(X) = d\phi(X) + [\omega \wedge \phi](X) \tag{31.57}$$

where $X \in T_pP$. Now given an additional (local) section $\varphi : U \subseteq M \to P$ we can pull back the previous structure. This gives us:

$$(\varphi^* D\phi)(Y) = d(\varphi^*\phi)(Y) + [\varphi^*\omega \wedge \varphi^*\phi](Y) \tag{31.58}$$

where $Y \in T_mM$. After introducing the notations $S := \varphi^*\phi$ and $\nabla_Y S := (\varphi^* D\phi)(Y)$ and remembering the definition of the Yang-Mills field 31.33 this becomes:

$$\boxed{\nabla_Y S = dS(Y) + \omega^U(Y) \cdot S} \tag{31.59}$$

Example 31.4.48. Let $G = \mathrm{GL}(\mathbb{R}^n)$. In local coordinates equation 31.59 becomes:

$$(\nabla_Y S)^i = \frac{\partial S^i}{\partial x^k} Y^k + \Gamma^i{}_{jk} S^j Y^k \tag{31.60}$$

which is exactly the formula known from classic differential geometry and relativity.

31.4.10 Curvature

Definition 31.4.49 (Curvature). Let ω be an Ehresmann connection on a principal bundle $G \hookrightarrow P \to M$. The curvature Ω of ω is defined as the exterior covariant derivative $D\omega$.

Definition 31.4.50 (Flat connection). An Ehresmann connection ω is said to be flat if its curvature Ω vanishes everywhere.

Example 31.4.51. Let ω_G be the Maurer-Cartan form on a Lie group G. It follows from the fact that the only horizontal vector on the bundle $G \hookrightarrow G \to \{x\}$ is the zero vector, that the curvature of ω_G is 0. Hence the Maurer-Cartan form is a flat connection.

Property 31.4.52 (Second Bianchi identity). Let ω be an Ehresmann connection with curvature Ω.

$$\boxed{D\Omega = 0} \tag{31.61}$$

Remark 31.4.53. One should however pay attention not to generalize this result to general differential forms. Only the exterior derivative satisfies the coboundary condition $d^2 \equiv 0$, the exterior covariant derivative does not.

Formula 31.4.54 (Cartan structure equation). Let ω be an Ehresmann connection and let Ω be its curvature form.

$$\boxed{\Omega = d\omega + \frac{1}{2}[\omega \wedge \omega]} \tag{31.62}$$

The following property is an immediate consequence of the Frobenius integrability theorem 30.3.21 and the fact that an Ehresmann connection vanishes on the horizontal subbundle.

Property 31.4.55. Let ω be an Ehresmann connection. The associated horizontal distribution[18]

$$p \mapsto \mathrm{Hor}(T_pP)$$

is integrable if and only if the connection ω is flat. Furthermore, the vertical distribution is always integrable.

Definition 31.4.56 (Yang-Mills field strength). Let (P, M, π, G) be a principal bundle equipped with an Ehresmann connection ω. Given a local section $\sigma : U \subseteq M \to P$ we define the Yang-Mills field strength F as the pullback $\sigma^*\Omega$, where $\Omega = D\omega$ is the curvature of ω.

31.4.11 Torsion

Definition 31.4.57 (Solder form). Let (P, M, π, G) be a principal bundle. Let V be a $\dim M$-dimensional vector space equipped with a representation $\rho : G \to \mathrm{GL}(V)$ such that $TM \cong P \times_G V$ in the sense of associated bundles. A solder(ing) form θ on P is a V-valued one-form that satisfies the following conditions:

- Horizontal: $\forall p \in P$: $\ker \theta_p \leq \mathrm{Vert}(T_pP)$
- G-equivariant: $R_g^*\theta = \rho(g)^{-1}\theta$

where R_g is the right action of G on P.

Definition 31.4.58 (Torsion). Let (P, M, π, G) be a principal bundle equipped with an Ehresmann connection ω and a solder form θ. The torsion Θ is defined as the exterior covariant derivative $D\theta$.

Formula 31.4.59 (Cartan structure equation). Let ω be an Ehresmann connection, θ a solder form and Θ its torsion form.

$$\Theta = d\theta + \omega \barwedge \theta \tag{31.63}$$

[18] See 30.3.18 for the definition of a distribution of vector spaces.

where the wedge product is defined analogously[19] to 30.31 and 30.32 using the representation of $\mathfrak{g}$ on V induced by the representation $\rho : G \to \mathrm{GL}(V)$:

$$\omega \barwedge \theta(v, w) = \Big[\omega(v) \cdot \theta(w) - \omega(w) \cdot \theta(v)\Big] \tag{31.64}$$

where $\cdot$ denotes the representation of $\mathfrak{g}$ on V.

Property 31.4.60 (First Bianchi identity). Let ω be an Ehresmann connection, Ω its curvature, θ the solder form and Θ its torsion.

$$\boxed{D\Theta = \Omega \barwedge \theta} \tag{31.65}$$

[19]For forms with deg ≥ 1 we sum over all permutations of the arguments.

Chapter 32

Integration Theory

For the theory on measure spaces and Lebesgue integration see chapter 14.

32.1 Orientation

Definition 32.1.1 (Orientation). Similar to definition 20.8.19 we can define an orientation on a differentiable manifold M. First we modify the definition of the volume element a little bit. A **volume form** on M is a nowhere-vanishing top-dimensional differential form $\mathrm{Vol} \in \Omega^n(M)$ where $n = \dim(M)$. The definition of an orientation is then equivalent to that in 20.8.19.

An **oriented atlas** is given by all charts of M for which the pullback of the Euclidean volume form is a positive multiple of Vol. This also means that the transition functions have a positive Jacobian determinant[1]. The existence of a volume form turns a differentiable manifold into an **orientable manifold**.

Alternatively an orientable manifold with volume form ω is said to be **positively oriented** if $\omega(v_1, ..., v_n) > 0$ where $(v_1, ..., v_n)$ is a basis for T_pM.

Example 32.1.2. Let $M = \mathbb{R}^n$. The canonical Euclidean volume form is given by the determinant map

$$\det : (u_1, ..., u_n) \mapsto \det(u_1, ..., u_n) \tag{32.1}$$

where the u_n's are expressed in the canonical basis $(e_1, ..., e_n)$. The name 'volume form' is justified by noting that the determinant map gives the signed volume of the n-dimensional parallelotope spanned by the vectors $\{u_1, ..., u_n\}$.

Property 32.1.3. Let ω_1, ω_2 be two volume forms on M. Then there exists a smooth function f such that

$$\omega_1 = f\omega_2$$

Furthermore, the sign of this function is constant on every connected component of M.

32.2 Integration of top-dimensional forms

Definition 32.2.1 (Measure zero). A subset $U \subset M$ of an orientable manifold is said to be of measure zero if it is the countable union of inverse images of null sets in $\mathbb{R}^n$.

[1] This is in fact an equivalent definition.

Definition 32.2.2 (Integrable form). A differential form is said to be integrable if its components with respect to any basis of $\Omega^k(M)$ are Lebesgue integrable on $\mathbb{R}^n$.

Formula 32.2.3 (Compact support). Let ω be a top-dimensional form on M with compact support on $U \subset M$.

$$\int_M \omega = \int_U \omega = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \omega_{12\ldots n}(x) dx^1 dx^2 ... dx^n \tag{32.2}$$

This integral is well defined because under a (orientation preserving) change of coordinates the component $\omega_{1\ldots n}$ transforms as $\omega'_{1\ldots n} = J\omega_{1\ldots n}$ where J is the Jacobian determinant. Inserting this in the integral and replacing dx_i by dx'_i then gives us the well-known change-of-variables formula from Lebesgue integration theory.

Formula 32.2.4 (General). Let ω be a top-dimensional form on M and let $\{\varphi_i\}_i$ be a partition of unity[2] subordinate to an atlas on M.

$$\int_M \omega = \sum_i \int_M \varphi_i \omega \tag{32.3}$$

where the integrals on the right-hand side are defined as in equation 32.2.

Property 32.2.5 (Compact manifolds). Let M be a smooth compact manifold. Because every continuous form on M is obviously compactly supported all continuous forms are integrable on M.

Property 32.2.6 (Pullback). Let $f : M \to N$ be an orientation-preserving diffeomorphism.

$$\int_M f^* \omega = \int_N \omega \tag{32.4}$$

Notation 32.2.7. Because the integral of differential forms satisfies properties similar to the ones listed in 14.2.12 we introduce the following notation:

$$\int_M \omega = \langle M, \omega \rangle \tag{32.5}$$

32.3 General integration

32.4 Stokes' theorem

Theorem 32.4.1 (Stokes' theorem). *Let Σ be an orientable smooth manifold. Denote the boundary of Σ by $\partial\Sigma$. Let ω be a differential k-form on Σ. We have the following equality:*

$$\boxed{\int_{\partial\Sigma} \omega = \int_\Sigma d\omega} \tag{32.6}$$

Corollary 32.4.2. The Kelvin-Stokes theorem 20.22, the divergence theorem 20.23 and Green's identity 20.24 are immediate results of this (generalized) Stokes' theorem.

[2]This always exists if we require M to be paracompact. See definition 7.5.17 and property 7.5.19.

32.5 de Rham Cohomology

Now we can also give a little side note about why the de Rham cohomology groups 30.58 really form a cohomology theory. For this we need some concepts from homology which can be found in section 9.2. Let M be a compact differentiable manifold and let $\{\lambda_i : \Delta^k \to M\}$ be the set of singular k-simplexes on M.

Now suppose that we want to integrate over a singular k-chain C on M, i.e. $C = \sum_{i=0}^{k} a_i \lambda_i$. Formula 32.3 says that we can pair the k-form ω and the chain C such that they act as duals to each other (hence p-forms are also called p-**cochains**), producing a real number[3]:

$$\langle \cdot, C \rangle : \Omega^n(M) \to \mathbb{R} : \omega \mapsto \int_C \omega = \sum_{i=0}^{k} a_i \int_{\Delta_k} \lambda_i^* \omega \tag{32.7}$$

where λ_i^* pulls back ω to Δ^k which is a subset of $\mathbb{R}^k$ as required. Now Stokes' theorem 32.6 tells us that

$$\int_C d\omega = \int_{\partial C} \omega \tag{32.8}$$

Using the pairing $\langle \cdot, \cdot \rangle$ this becomes

$$\langle d\omega, C \rangle = \langle \omega, \partial C \rangle \tag{32.9}$$

The operators d and ∂ can thus be interpreted as formal adjoints. After checking (again using Stokes' theorem) that all chains C and cochains ω belonging to the same equivalence classes $[C] \in H_k(M, \mathbb{R})$ and $[\omega] \in H^k(M, \mathbb{R})$ give rise to the same number[4] $\langle \omega, C \rangle$ we see that the singular homology groups and the de Rham cohomology groups on M are well-defined dual groups. The name cohomology is thus wel chosen for 30.58.

[3] This requires the chain group to have real coefficients instead of integer coefficients as is mostly used in homology.

[4] Suppose that $A, B \in [C]$ and $\phi, \chi \in [\omega]$ then $\langle \phi, A \rangle = \langle \chi, B \rangle$.

Chapter 33

Riemannian Geometry

33.1 Riemannian manifolds

33.1.1 Metric

Definition 33.1.1 (Bundle metric). Consider the bundle of second order covariant vectors. Following from 20.5.1 every section g of this bundle gives a bilinear map

$$g_x : T_xM \times T_xM \to \mathbb{R}$$

for all $x \in M$. If this map is symmetric and non-degenerate and if it depends smoothly on p it is called a **(Lorentzian)** metric.[1]

The maps $\{g_x\}_{x\in M}$ can be 'glued' together to form a global metric g, defined on the fibre product[2] $TM \diamond TM$. Defining this map on $TM \times TM$ is not possible as tangent vectors belonging to different points in M cannot be 'compared'. The collection $\{\langle\cdot|\cdot\rangle_x | x \in M\}$ is called a **bundle metric**.

A Riemannian metric also induces a duality between TM and T^*M. This is given by the *flat* and *sharp* isomorphisms:

Definition 33.1.2 (Musical isomorphisms). Let $g : TM \times TM \to \mathbb{R}$ be the Riemannian metric on M. The **flat** isomorphism is defined as:

$$\flat : v \mapsto g(v, \cdot) \tag{33.1}$$

The **sharp** isomorphism is defined as the inverse map:

$$\sharp : g(v, \cdot) \mapsto v \tag{33.2}$$

These 'musical' isomorphisms can be used to lower and raise tensor indices.

33.1.2 Riemannian manifold

Definition 33.1.3 (Pseudo-Riemannian manifold). Let M be a smooth manifold. This manifold is called pseudo-Riemannian if it is equipped with a pseudo-Riemannian metric. A **Riemannian manifold** is similarly defined.

[1] See also the section about Hermitian forms and metric forms 19.4.
[2] See definition 29.8.

Definition 33.1.4 (Riemannian isometry). Let (M, g_M) and (N, g_N) be two Riemannian manifolds. An isometry 26.2.10 $f : M \to N$ is said to be Riemannian if $F^* g_N = g_M$.

Property 33.1.5. Let M be a pseudo-Riemannian manifold. For every $p \in M$ there exists a splitting $T_pM = P \oplus N$ where P is a subspace on which the pseudometric is positive-definite and N is a subspace on which the pseudometric is negative-definite. This splitting is however not unique, only the dimensions of the two subspaces are well-defined.

Due to the continuity of the pseudometric, the dimensions of this splitting wil be the same for points in the same neighbourhood. For connected manifolds this amounts to a global invariant:

Definition 33.1.6 (Index). Let M be a connected pseudo-Riemannian manifold. The dimension of the *negative* subspace N in the above splitting $T_pP = P \oplus N$ is called the index of the pseudo-Riemannian manifold.

Definition 33.1.7 (Null coordinate). Consider a coordinate x on a pseudo-Riemannian manifold (M, g). This coordinate is said to be a null coordinate if it satisfies the following condition at every point in M:

$$g\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial x}\right) = 0 \tag{33.3}$$

Theorem 33.1.8 (Whitney's embedding theorem). *Every smooth paracompact*[3] *manifold M can be embedded in $\mathbb{R}^{2\dim M}$.*

Theorem 33.1.9 (Whitney's immersion theorem). *Every smooth paracompact manifold M can be immersed in $\mathbb{R}^{2\dim M - 1}$.*

Theorem 33.1.10 (Immersion conjecture). *Every smooth paracompact manifold M can be immersed in $\mathbb{R}^{2\dim M - a(\dim M)}$ where $a(n)$ is the number of 1's in the binary expansion of n.*

Definition 33.1.11 (Riemannian cone). Let (M, g) be a Riemannian manifold. Consider the product manifold $M \times]0, +\infty[$. This manifold can also be turned into a Riemannian manifold by equipping it with the metric $t^2 g + dt^2$. This manifold is called the Riemannian cone or **metric cone** of (M, g).

33.1.3 Levi-Civita connection

Definition 33.1.12 (Riemannian connection). An affine connection ∇ on a Riemannian manifold (M, g) is said to be Riemannian if it satisfies following two conditions:

1. ∇ is metric:
$$X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \tag{33.4}$$

2. ∇ is torsion-free:
$$\nabla_X Y - \nabla_Y X = [X, Y] \tag{33.5}$$

A Riemannian connection is often called a **Levi-Civita connection**.

Theorem 33.1.13 (Fundamental theorem of Riemannian geometry). *The Levi-Civita connection on a Riemannian manifold (M, g) is unique.*

[3] See definition 7.5.14.

Formula 33.1.14 (Koszul formula). The Levi-Civita connection ∇ on a Riemannian manifold (M, g) is implicitly (and uniquely[4]) given by the following formula:

$$\begin{aligned} 2g(\nabla_X Y, Z) &= \mathcal{L}_X g(Y, Z) + d(\iota_X g)(Y, Z) && (33.6)\\ &= X(g(Y, Z)) + Y(g(Z, X)) - Z(g(X, Y)) \\ &\qquad + g([X, Y], Z) - g([Z, X], Y) - g([Y, Z], X) && (33.7)\end{aligned}$$

33.1.4 Killing vectors

Definition 33.1.15 (Killing vector). Let (M, g) be a Riemannian manifold. A vector field X satisfying

$$\boxed{\mathcal{L}_X g = 0} \tag{33.8}$$

is called a Killing vector field.

A simple calculation gives us the following coordinate expression:

$$(\mathcal{L}_X g)_{\mu\nu} = X^\lambda \partial_\lambda g_{\mu\nu} + g_{\lambda\mu}\partial_\nu X^\lambda + g_{\lambda\nu}\partial_\mu X^\lambda \tag{33.9}$$

Formula 33.1.16. Given a Levi-Civita connection ∇ on (M, g) we can rewrite the Killing condition as follows:

$$\nabla_{(m} X_{n)} = 0 \tag{33.10}$$

Definition 33.1.17 (Killing tensor). Let ∇ be the Levi-Civita connection on (M, g). A tensor T satisyfing

$$\nabla_{(m_N} T_{m_1 \ldots m_{N-1})} = 0 \tag{33.11}$$

is called a Killing tensor. It is obvious that this **generalized Killing condition** is a direct generalization of the Killing condition as given above.

33.2 Curvature

Formula 33.2.1 (Riemann curvature tensor). The Riemann (curvature) tensor R is defined as following $(1, 3)$-tensor:

$$R(v, w)z = [\nabla_v, \nabla_w]z - \nabla_{[v,w]}z \tag{33.12}$$

where ∇ is the Levi-Civita connection. In index notation it is given by:

$$R^i_{jkl} = dx^i\big(R(e_k, e_l)e_j\big) \tag{33.13}$$

Formula 33.2.2 (Directional curvature operator[5]).

$$R_v(w) = R(w, v)v \tag{33.14}$$

Formula 33.2.3 (Sectional curvature).

$$\sec(v, w) = \frac{g(R(w, v)v, w)}{g(v, v)g(w, w) - g(v, w)^2} = \frac{g(R_v(w), w)}{g(v \wedge w, v \wedge w)} \tag{33.15}$$

An important result states that the sectional curvature only depends on the span of v, w.

[4] Any connection satisfying this formula necessarily coincides with the Levi-Civita connection.

[5] Also called the **tidal force operator** (mostly in physics).

Remark 33.2.4. For surfaces the sectional curvature coincides with the Gaussian curvature K (see Theorema Egregium 26.46). Generally the sectional curvature gives the Gaussian curvature of the plane spanned by the vector v, w.

Formula 33.2.5 (Ricci tensor).

$$R_{\mu\nu} = R^{\lambda}{}_{\mu\lambda\nu} \tag{33.16}$$

Formula 33.2.6 (Ricci scalar).

$$R = R^{\mu}{}_{\mu} \tag{33.17}$$

This scalar quantity is also called the **scalar curvature**.

Formula 33.2.7 (Einstein tensor).

$$\boxed{G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R} \tag{33.18}$$

Theorem 33.2.8. *For 4-dimensional manifolds the Einstein tensor $G_{\mu\nu}$ is the only tensor containing at most second derivatives of the metric $g_{\mu\nu}$ and satisfying:*

$$\nabla_{\mu} G^{\mu\nu} = 0 \tag{33.19}$$

33.3 Sphere bundle

Definition 33.3.1 (Unit sphere bundle). Let V be a normed vector space. Consider a vector bundle $V \hookrightarrow E \to B$. From this bundle we can derive a new bundle where we replace the typical fibre V by the unit sphere $\{v \in V : ||v|| = 1\}$. It should be noted that this new bundle is not a vector bundle as the unit sphere is not a vector space.

Remark 33.3.2 (Unit disk bundle). A similar construction can be made by replacing the unit sphere by the unit disk $\{v \in V : ||v|| \leq 1\}$.

33.4 Hilbert bundles

Definition 33.4.1 (Hilbert bundle). A vector bundle for which the typical fibre is a Hilbert space is called a Hilbert bundle.

Definition 33.4.2 (Compatible Hilbert bundle). Consider the isomorphisms

$$l_x : F_x \to \mathcal{H} : h \mapsto \varphi_i(x, h) \in \pi(x) \tag{33.20}$$

where $\mathcal{H}$ is the typical fibre and where $\{(U_i, \varphi_i)\}_{i \in I}$ is a trivializing cover. These maps l_x are called **point-trivializing maps**.

Using these maps we can extend the metric structure of the typical fibre $\mathcal{H}$ to the fibres F_x for all x by:

$$\langle v|w\rangle_x = \langle l_x(v)|l_x(w)\rangle_{\mathcal{H}} \tag{33.21}$$

The Hilbert bundle is said to be compatible (with the metric structure on $\mathcal{H}$) if the above extension is valid for all $v, w \in F_x$.

Remark. For compatible Hilbert bundles, the transition maps $l_{x \to y} = l_y^{-1} \circ l_x : \pi^{-1}(x) \to \pi^{-1}(y)$ are easily seen to be isometries.

33.5 Conformal transformations

33.5.1 Conformal Killing vectors

Definition 33.5.1 (Conformal Killing vector). Consider a pseudo-Riemannian manifold (M, g). A vector field X is said to be conformal with conformal factor $\Omega : M \to \mathbb{R}$ if it satisfies:

$$\mathcal{L}_X g = \Omega g \tag{33.22}$$

In local coordinates this amounts to:

$$\nabla_\mu X_\nu + \nabla_\nu X_\mu = \Omega g_{\mu\nu} \tag{33.23}$$

where ∇ is the Levi-Civita connection associated to (M, g).

Chapter 34

Symplectic Geometry

34.1 Symplectic manifolds

Definition 34.1.1 (Symplectic form). Let $\omega \in \Omega^2(M)$ be a differential 2-form. ω is said to be symplectic if it satisfies following properties:

- Closedness: $d\omega = 0$
- Non-degeneracy: if $\omega(u, v) = 0, \forall u \in TM$ then $v = 0$

Definition 34.1.2 (Symplectic manifold). A manifold M equipped with a symplectic 2-form ω is called a symplectic manifold. This structure is often denoted as the pair (M, ω).

Property 34.1.3 (Dimension). From the antisymmetry and the non-degeneracy of the symplectic form it follows that M is even-dimensional.

Theorem 34.1.4 (Darboux). *Let (M, ω) be a symplectic manifold. For every neighbourhood Ω in T^*M there exists an adapted coordinate system (q^i, p^i) such that:*

$$\omega|_\Omega = \sum_i dp^i \wedge dq^i \tag{34.1}$$

The charts in Darboux's theorem are called **Darboux charts** and they cover M.

Remark 34.1.5. This theorem shows that all symplectic manifolds of the same dimension are locally isomorphic, i.e. there exist no local invariants, in contrast to for example Riemannian manifolds.

Formula 34.1.6. In Darboux coordinates the components of the symplectic form ω read:

$$\omega_{ij} = \left(\begin{array}{c|c} 0 & -\mathbb{1} \\ \hline \mathbb{1} & 0 \end{array}\right) \tag{34.2}$$

By the non-degeneracy we can define the 'dual' $\omega^\sharp$ as:

$$(\omega^\sharp)^{ij} = \left(\begin{array}{c|c} 0 & \mathbb{1} \\ \hline -\mathbb{1} & 0 \end{array}\right) \tag{34.3}$$

34.1.1 Symplectomorphisms

Definition 34.1.7 (Symplectomorphism). A symplectomorphism is a morphism of symplectic manifolds, i.e. a diffeomorphism $f : (M, \omega_M) \to (N, \omega_N)$ satisfying:

$$f^*\omega_N = \omega_M \tag{34.4}$$

These maps form a *semigroup* called the **symplectomorphism group**[1]

Definition 34.1.8 (Symplectic vector field). A vector field is said to be symplectic if its flow preserves the symplectic form ω:

$$\mathcal{L}_X\omega = 0 \tag{34.5}$$

Equivalently a vector field is symplectic if its flow is a symplectomorphism. These vector fields form a Lie subalgebra of $\mathfrak{X}(M)$.

Definition 34.1.9 (Hamiltonian vector field). Let (M, ω) be a symplectic manifold. For every function $f \in C^\infty(M)$ we define the associated Hamiltonian vector field X_f by the following relation[2]:

$$\omega(X_f, \cdot) = -df(\cdot) \tag{34.6}$$

or by using $\omega^\sharp$:

$$X_f(\cdot) = \omega^\sharp(-df, \cdot) \tag{34.7}$$

These vector fields form a Lie subalgebra of the Lie algebra of symplectic vector fields. The flow associated to a Hamiltonian vector field is sometimes called a **Hamiltonian symplectomorphism**.[3]

Definition 34.1.10 (Poisson bracket). Let (M, ω) be a symplectic manifold. The Poisson bracket of two functions $f, g \in C^\infty(M)$ is defined as:

$$\{f, g\} = X_f(g) \tag{34.8}$$

or equivalently:

$$X_{\{f,g\}} = [X_f, X_g] \tag{34.9}$$

Property 34.1.11. The Poisson brackets induced by the symplectic form turns the structure $(C^\infty(M), \{\cdot, \cdot\})$ into a Lie algebra[4] and the second equation in fact gives a (surjective) Lie algebra morphism[5] $(C^\infty(M), \{\cdot, \cdot\}) \to (\{\text{X : X is a HVF on M}\}, [\cdot, \cdot])$. Furthermore, together with the pointwise multiplication the structure becomes a Poisson algebra[6].

Definition 34.1.12 (Poisson manifold). A smooth manifold on which the algebra of smooth functions can be equipped with a Poisson algebra structure.

Property 34.1.13. From the property above every symplectic manifold is a Poisson manifold. The converse however is not true.

[1] Not to be confused with the symplectic group $\mathrm{Sp}(n)$.

[2] A lot of different conventions exist in the literature. We use the one compatible with the Hamiltonian equations 43.13 which are universally accepted.

[3] The fact that the Hamiltonian flow indeed preserves the symplectic form follows from the closedness of ω.

[4] The antisymmetry follows from equation 30.23 and the Jacobi-identity follows from the closedness of ω.

[5] The kernel is given by the constant functions.

[6] See definition 28.5.3.

34.2 Lagrangian submanifolds

Definition 34.2.1 (Symplectic complement). Let (M, ω) be a symplectic manifold and let $S \subset M$ be an embedded submanifold $\iota : S \hookrightarrow M$. The symplectic orthogonal complement $T_p^{\perp} S$ at the point $p \in S$ is defined as[7]:

$$T_p^{\perp} S = \{v \in T_p M : \omega(v, \iota_* w) = 0, \forall w \in T_p S\} \tag{34.10}$$

Definition 34.2.2 (Isotropic submanifold). Let (M, ω) be a symplectic manifold. An embedded submanifold $\iota : S \hookrightarrow M$ is called isotropic if $T_p S \subset T_p^{\perp} S$.

Definition 34.2.3 (Isotropic submanifold). Let (M, ω) be a symplectic manifold. An embedded submanifold $\iota : S \hookrightarrow M$ is called co-isotropic if $T_p^{\perp} S \subset T_p S$.

Definition 34.2.4 (Larangian submanifold). Let (M, ω) be a symplectic manifold. An embedded submanifold $\iota : S \hookrightarrow M$ is called Lagrangian if $T_p S = T_p^{\perp} S$. Therefore they are sometimes called maximal isotropic submanifolds.

34.3 Cotangent bundle

Construction 34.3.1 (Liouville one-form). Let M be a smooth manifold. The cotangent bundle T^*M comes equipped with a canonical symplectic form:

Let (q, p) denote the local coordinates on T^*M. Define a 1-form α by

$$\alpha = p_i dq^i$$

In coordinate-free notation this can be written as

$$\alpha(z) = \pi_2(z)\Big(\pi_*(z)\Big) \tag{34.11}$$

where $z \in TT^*M$, $\pi : T^*M \to M$ and $\pi_2 : TT^*M \to T^*M$. This 1-form, called the **canonical 1-form**[8], induces a symplectic form in the following way:

$$\omega = d\alpha$$

34.4 Hamiltonian dynamics

Definition 34.4.1 (Completely integrable system). Consider a multimap

$$F \equiv (F_1, ..., F_n) : M \to \mathbb{R}^n$$

on a symplectic manifold (M, ω). This map defines a completely integrable system (CIS) if it satisfies the following conditions:

- $\dim M = 2n$
- The Hamiltonian vector fields $\{X_{F_i}\}_{i \leq n}$ are almost everywhere linearly independent.[9]
- For every $i, j \leq n$ we have that $\{F_i, F_j\} = 0$.

[7] The notation $T_p^{\omega} S$ is also sometimes used.
[8] Also known as the **Liouville 1-form**.
[9] Equivalently we require that the Jacobian DF has full rank almost everywhere.

34.4.1 Hamiltonian actions

Definition 34.4.2 (Hamiltonian torus action). Let (M, ω) be a symplectic manifold. Let T^n be a torus group acting on M. The action of T^n is said to be Hamiltonian if there exists a CIS on M such that the action arises as the Hamiltonian flow of F.

Chapter 35

Contact Geometry

35.1 Contact structure

35.1.1 Contact form

Definition 35.1.1 (Contact element). Let M be a smooth n-dimensional manifold. A contact element at the point $p \in M$, called the **contact point**, is a $(n-1)$-dimensional subspace of the tangent space T_pM.

Property 35.1.2. As every $(n-1)$-dimensional subspace of the tangent space can be constructed as the kernel of a linear functional (living in T_p^*M) one can construct the space of contact elements as a quotient of the cotangent bundle:

$$PT^*M = (T^*M \backslash \{0_M\}) / \sim \tag{35.1}$$

where the equivalence relation $\sim$ is defined by $\omega \sim \rho \iff \exists \lambda \in \mathbb{R}_0 : \omega = \lambda\rho$.

Definition 35.1.3 (Contact structure). Let M be a $(2n+1)$-dimensional smooth manifold. A distribution ξ of contact elements on M is called a contact structure on M if the (locally) defining one-form α satisfies the following non-integrability condition[1]:

$$\alpha \wedge (d\alpha)^n \neq 0 \tag{35.2}$$

If the one-form α is defined globally on M then it is called a **contact form** and the pair (M, α) is called a **contact manifold**.

Property 35.1.4 (Coorientable distribution). A global contact form α such that $\xi = \ker(\alpha)$ can be defined globally if and only if the distribution ξ is coorientable, i.e. the line bundle TM/ξ is trivial (orientable).

35.1.2 Reeb vector fields

Definition 35.1.5 (Reeb vector field). Let (M, α) be a contact manifold. A Reeb vector field on M is a vector field X such that $\alpha(X) = 1$ and $\iota_X d\alpha = 0$.

Property 35.1.6. Given a contact manifold, there exists a unique Reeb vector field associated to it.

[1] In fact it is maximally non-integrable. (Compare with Frobenius' theorem **??**TODO.)

Chapter 36

Complex Geometry

36.1 Complex structures

Definition 36.1.1 (Almost complex structure). Let M be a smooth manifold. An almost complex structure on M is a smooth $(1,1)$-tensor field $J : TM^{\mathbb{C}} \to TM^{\mathbb{C}}$ such that $J|_p : T_pM^{\mathbb{C}} \to T_pM^{\mathbb{C}}$ satisfies $J|_p^2 = -1$ for all $p \in M$.

This structure induces a decomposition in so-called holomorphic and anti-holomorphic components:

$$TM^{\mathbb{C}} = TM^{+} \oplus TM^{-}$$

where both bundles have the same dimension. When the coordinates on M are denoted by $\{x^k\}_{k\leq 2n}$ the basis for these two subbundles is given by $\{\frac{\partial}{\partial z^k} = \frac{1}{2}\left(\frac{\partial}{\partial x^{2k-1}} - i\frac{\partial}{\partial x^{2k}}\right)\}_{k\leq n}$ and $\{\frac{\partial}{\partial \bar{z}^k} = \frac{1}{2}\left(\frac{\partial}{\partial x^{2k-1}} + i\frac{\partial}{\partial x^{2k}}\right)\}_{k\leq n}$ respectively.

Remark 36.1.2. The reason that we define the almost complex structure on the complexified tangent bundle has to do with the fact that J is only diagonalizable on a complex vector space (because they square to a negative value).

Property 36.1.3. An almost complex manifold is even-dimensional and orientable.

Property 36.1.4. A manifold M admits an almost complex structure if and only if the structure group of the tangent bundle TM can be reduced from $GL(\mathbb{R}^{2n})$ to $GL(\mathbb{C}^n)$.

Definition 36.1.5 (Complex manifold). A topological space M for which there exists an open cover $\{U_i\}_i$ such that for every U_i there exists a homeomorphism $\varphi_i : U_i \to \mathbb{C}^n$ onto some open subset of $\mathbb{C}^n$. Furthermore the transition functions $\varphi_{ji} : \varphi_i(U_i \cap U_j) \to \varphi_j(U_i \cap U_j)$ are also required to be holomorphic.

Definition 36.1.6 (Complex dimension). The integer n in previous definition is called the complex dimension of M, denoted by $\dim_{\mathbb{C}}(M)$.

Property 36.1.7. An almost complex manifold is complex if and only if the structure group can be further reduced to $\mathrm{U}(n)$.

Theorem 36.1.8 (Newlander-Nirenberg). *An almost complex manifold is complex if and only if the* ***Nijenhuis tensor*** N_J *vanishes for all vector fields:*

$$N_J(X,Y) = [JX, JY] - J[JX, Y] - J[X, JY] - [X, Y] = 0 \tag{36.1}$$

When working in a local coordinate-induced basis we obtain the following condition:

$$J^\nu_\rho \partial_\nu J^\mu_\sigma - J^\nu_\sigma \partial_\nu J^\mu_\rho - J^\mu_\nu \partial_\rho J^\nu_\sigma + J^\mu_\nu \partial_\sigma J^\mu_\rho = 0 \tag{36.2}$$

An almost complex structure satisfying this condition is also said to be integrable. (See also next property.)

36.2 Complex differential forms

Property 36.2.1. On a complex manifold there exist coordinates $\{z^i\}_{i\leq n}$ such that the almost complex structure J can be written as:

$$J = i\partial_k \otimes dz^k - i\partial_{\overline{k}} \otimes d\overline{z}^k \tag{36.3}$$

This coordinate expansion can be used to, given an almost complex manifold satisfying 36.1, find a coordinate transformation from the real coordinates $\{x^i\}_{i\leq 2n}$ to the complex coordinates $\{z^i\}_{i\leq n}$.

Using the basis forms $dz^i, d\overline{z}^i$ one can also define complex Grassmann spaces $\Omega^{p,q}_m(M)$, analogous to $\Omega^k(X)$ for smooth manifolds, for any point $m \in M$:

$$\Omega^{1,0}_m(M) = \mathrm{span}_{\mathbb{C}}\{dz^i_m\} \tag{36.4}$$

$$\Omega^{0,1}_m(M) = \mathrm{span}_{\mathbb{C}}\{d\overline{z}^i_m\} \tag{36.5}$$

$$\Omega^{p,q}_m(M) = \left(\bigwedge_{i=1}^{p} \Omega^{1,0}_m\right) \wedge \left(\bigwedge_{j=1}^{q} \Omega^{0,1}_m\right) \tag{36.6}$$

Property 36.2.2. The spaces $\Omega^{1,0}(M)$ and $\Omega^{0,1}(M)$ are stable under holomorphic coordinate transformations, i.e. they transform tensorially and hence they are complex vector bundles.

On the bundle $\Omega^k(M) = \bigoplus_{p+q=k} \Omega^{p,q}(M)$ of forms of total degree k one can then define the canonical projection maps $\pi^{p,q} : \Omega^k \to \Omega^{p,q}$.

Definition 36.2.3 (Dolbeault operator). Consider a general $(p+q)$-form $\omega \in \Omega^{p,q}(M)$. The de Rham differential maps this form to a $(p+q+1)$-form. This form is in general an element of $\sum_{r+s=p+q+1} \Omega^{r,s}(M)$. Using the projection maps $\pi^{p,q}$ one obtains the Dolbeault operators:

$$\partial = \pi^{p+1,q} \circ d \tag{36.7}$$

$$\overline{\partial} = \pi^{p,q+1} \circ d \tag{36.8}$$

Property 36.2.4. By explicitly writing out the action of the de Rham differential d on a general (p,q)-form one obtains the following decomposition:

$$d = \partial + \overline{\partial} \tag{36.9}$$

and by using the coboundary property of d:

$$\partial^2 = \overline{\partial}^2 = 0 \tag{36.10}$$

$$\partial\overline{\partial} + \overline{\partial}\partial = 0 \tag{36.11}$$

Formula 36.2.5. Analogous to the definition of the de Rham codifferential 30.46 one can define the adjoint of the ∂ and $\overline{\partial}$ operators:

$$\partial^\dagger = - * \partial *$$ (36.12)

$$\overline{\partial}^\dagger = - * \overline{\partial} *$$ (36.13)

where we used the fact that the real dimension of a complex manifold is even: $(-1)^{n(k+1)+1} = -1$.

Corollary 36.2.6. Using these definitions one can rewrite the Hodge Laplacian as:

$$\Delta = 2(\partial\partial^\dagger + \partial^\dagger\partial) = 2(\overline{\partial}\overline{\partial}^\dagger + \overline{\partial}^\dagger\overline{\partial})$$ (36.14)

36.3 Kähler manifolds

Definition 36.3.1 (Kähler manifold). Consider a smooth manifold equipped with a Riemannian structure (M, g), a symplectic structure (M, ω) and an almost complex structure J. This manifold is a Kähler manifold if the structures satisfy following equivalent sets of compatibility conditions:

- The almost complex structure J is integrable[1].
- The symplectic form is compatible with the almost complex structure:

$$\omega(v, w) = \omega(Jv, Jw)$$ (36.15)

and

$$\omega(v, Jv) > 0$$ (36.16)

or

- M is Hermitian[2]:

$$g_{\mathbb{C}}(v, w) = g_{\mathbb{C}}(Jv, Jw)$$ (36.17)

- The fundamental two-form $\omega(v, w) = g_{\mathbb{C}}(v, Jw)$ is closed[3].

or

- M is Hermitian.
- J is parallel with respect to the Levi-Civita connection on (M, g):

$$\nabla_X J = 0$$ (36.18)

Remark 36.3.2. The property, which says that J acts isometrically, can be interchanged for the statement that J acts as a symplectomorphism. These two statements are equivalent on a Kähler manifold.

Definition 36.3.3 (Kähler form). The central object in all these definitions is the Kähler form[4]:

$$\boxed{\omega(v, w) = g_{\mathbb{C}}(v, Jw)}$$ (36.19)

[1] If not then the manifold is said to be almost Kähler.
[2] $g_{\mathbb{C}}$ is the extension (by linearity) of g from TM to $T^{\mathbb{C}}M$. This map has $\mathbb{C}$ as codomain.
[3] The non-degeneracy follows from the non-degeneracy of the metric.
[4] Sometimes called the **Hermitian form** or **fundamental form**.

Formula 36.3.4. The metric $g = g_{ij}dx^i \otimes dx^j$ can be rewritten as:

$$g = g_{i\bar{j}}\big(dz^i \otimes d\bar{z}^j + d\bar{z}^j \otimes dz^i\big) \tag{36.20}$$

The Kähler form can then be written as:

$$\omega = ig_{i\bar{j}}dz^i \wedge d\bar{z}^j \tag{36.21}$$

Definition 36.3.5 (Kähler potential). Using the $\partial\overline{\partial}$-lemma 36.4.4 we can locally write the Kähler form as:

$$\omega = i\partial\overline{\partial}K(z,\overline{z}) \tag{36.22}$$

where the real function $K \in \Omega^0(M)$ is called the **Kähler potential**.

Corollary 36.3.6. It follows from formula 36.21 that we can (locally) rewrite the Hermitian metric as:

$$g_{i\bar{j}} = \partial_i\partial_{\bar{j}}K(z,\overline{z}) \tag{36.23}$$

Property 36.3.7. The Christoffel symbols associated to the Levi-Civita connection on (M,g) are also quite easy when M is Kähler. Only the ${\Gamma_{ij}}^k$ and ${\Gamma_{\bar{i}\bar{j}}}^{\bar{k}}$ components remain and they are given by:

$${\Gamma_{ij}}^k = g^{k\bar{m}}\partial_i g_{j\bar{m}} \tag{36.24}$$

$${\Gamma_{\bar{i}\bar{j}}}^{\bar{k}} = g^{\bar{k}m}\partial_{\bar{i}} g_{\bar{j}m} \tag{36.25}$$

Accordingly the only non-vanishing component of the Riemann curvature tensor is:

$$R_{\bar{i}j\bar{k}l} = g_{\bar{k}m}\partial_{\bar{i}}{\Gamma_{jl}}^m \tag{36.26}$$

Definition 36.3.8 (Kähler transformation). From definition 36.22 one can see that the Kähler potential is not defined unambiguously. The following transformation leaves the Kähler form invariant:

$$K'(z,\overline{z}) = K(z,\overline{z}) + f(z) + \overline{f}(\overline{z}) \tag{36.27}$$

On overlapping coordinate charts the transformation between Kähler potentials is exactly of this form.

36.3.1 Killing vectors

Definition 36.3.9 (Holomorphic Killing vector). Consider the set of Killing vectors X_A associated to the metric g, i.e. $\mathcal{L}_{X_A}g = 0$. Within this set of vectors one can look at the set of vectors k_A satisfying

$$\mathcal{L}_{k_A}J = 0 \tag{36.28}$$

These are called holomorphic Killing vectors because there components are holomorphic in the sense of complex analysis. This can easily be shown by writing the Killing condition in terms of covariant derivatives and by using expression 36.3.

Remark 36.3.10. Equivalently one could require the following condition:

$$\mathcal{L}_{k_A}\omega = 0 \tag{36.29}$$

Definition 36.3.11 (Moment map). Let k be a holomorphic Killing vector. From $d\omega = 0$ one can, using Cartan's magic formula 30.52 and the above condition, find that $\iota_k\omega$ is closed. Poincaré's lemma then implies that here exists a real function $\mathcal{P}(z,\overline{z})$ such that:

$$\iota_k\omega = d\mathcal{P} \tag{36.30}$$

Using the expression 36.21 one can then find the following expression for the Killing vector:

$$k^i = -ig^{i\bar{j}}\partial_{\bar{j}}\mathcal{P} \tag{36.31}$$

36.4 Cohomology

36.4.1 Hodge-de Rham cohomology

36.4.2 Dolbeault cohomology

Theorem 36.4.1 (Hodge decomposition). *Let M be a Kähler manifold. For all $k \in \mathbb{N}$ we have:*

$$H^k(M) = \bigoplus_{p+q=k} H^{p,q}(M) \tag{36.32}$$

By analogy with the Poincaré lemma for smooth manifolds we have following theorems:

Theorem 36.4.2 (∂-lemma). *Let $\alpha \in \Omega^{p,q}(M)$. If $\partial\alpha = 0$ then locally there exists a complex form $\beta \in \Omega^{p-1,q}$ such that $\alpha = \partial\beta$.*

Theorem 36.4.3 ($\overline{\partial}$-lemma). *Let $\alpha \in \Omega^{p,q}(M)$. If $\overline{\partial}\alpha = 0$ then locally there exists a complex form $\beta \in \Omega^{p,q-1}$ such that $\alpha = \overline{\partial}\beta$.*

Theorem 36.4.4 ($\partial\overline{\partial}$-lemma). *Let $\alpha \in \Omega^{p,q}(M)$. If $d\alpha = 0$ then locally there exists a complex form $\beta \in \Omega^{p-1,q-1}$ such that $\alpha = \partial\overline{\partial}\beta$.*

Chapter 37

Calculus of Variations

37.1 Constrained systems

37.1.1 Holonomic constraints

Definition 37.1.1 (Holonomic constraint). A constraint $f(q,t)=0$ is called holonomic if it only depends on the coordinates q^i and t.

Method 37.1.2 (Holonomic constraints). The Euler-Lagrange equations of a system with k holonomic constraints $f_k(q,t)=0$ can be obtained from the generalized action functional

$$\int_a^b \Big[L(q(t),\dot{q}(t),t)+\sum_{j=1}^{k}\lambda_j(t)f_j(q(t),t)\Big]dt \tag{37.1}$$

where $\lambda_j(t)$ are undetermined (Lagrange) multipliers.

37.2 Noether symmetries

37.2.1 Classical systems

Definition 37.2.1 (Noether symmetry). Let $L(q^i,\dot{q}^i,t)$ be the Lagrangian function describing some system $\mathcal{S}$. An infinitesimal transformation[1]

$$\begin{aligned} q^i &\longrightarrow q^i+\varepsilon\xi^i(q^k,t)\\ t &\longrightarrow t+\varepsilon\tau(q^k,t)\end{aligned}$$

$$\dot{q}^i \longrightarrow \dot{q}^i+\varepsilon(\dot{\xi}^i-\dot{q}^i\dot{\tau})$$

is called a Noether symmetry for L if it satisfies:

$$\int_{\tilde{t}_0}^{\tilde{t}_1}\tilde{L}(\tilde{q}^i,\dot{\tilde{q}}^i,\tilde{t})d\tilde{t}=\int_{t_0}^{t_1}L(q^i,\dot{q}^i,t)dt+\varepsilon\int_{t_0}^{t_1}\frac{df}{dt}dt \tag{37.2}$$

for every curve $q:[a,b]\to\mathbb{R}$, for every subinterval $[t_0,t_1]\subseteq[a,b]$ and for some function $f(q^i,t)$.

[1] The transformation for $\dot{q}^i$ is in fact induced by the ones for q^i and t.

Chapter 38

K-theory

Important: In this chapter all topological (base) spaces are supposed to be both compact and Hausdorff. This ensures that the complex of K-theories satisfies the Eilenberg-Steenrod axioms 9.4.1.

38.1 Basic definitions

Definition 38.1.1 (K-theory). Let $\mathrm{Vect}(X)/\sim$ be the set of isomorphism classes of finite-dimensional vector bundles over a topological space X. Because this set is well-behaved with respect to Whitney sums, the structure $(\mathrm{Vect}(X)/\sim, \oplus)$ forms an Abelian monoid. The Grothendieck completion[1] of $(\mathrm{Vect}(X)/\sim, \oplus)$ is called the K-theory of X.

Notation 38.1.2. The K-theory of a space X is denoted by $K^0(X)$.

Example 38.1.3. Let $\{x_0\}$ be a one-point space. The K-theory $K^0(\{x_0\})$ is isomorphic to the additive group of integers $(\mathbb{Z}, +)$.

Definition 38.1.4 (Virtual vector bundle). The elements of $K^0(X)$ are pairs $([E], [E'])$ that can be formally written as a difference $[E] - [E']$. These elements are called virtual (vector) bundles.

Definition 38.1.5 (Virtual rank). The virtual rank of the virtual bundle $([E], [E'])$ is defined as follows:

$$\mathrm{rk}([E] - [E']) = \mathrm{rk}(E) - \mathrm{rk}(E') \tag{38.1}$$

Property 38.1.6. Property 30.2.6 implies that every virtual bundle is of the form $[E] - [X \times \mathbb{R}^n]$ for some vector bundle E and integer $n \in \mathbb{N}$.

Definition 38.1.7 (Reduced K-theory). Let (X, x_0) be a pointed space. The inclusion $\{x_0\} \hookrightarrow X$ induces a group morphism $M : K^0(X) \to K^0(x_0)$ given by restriction of virtual bundles to the basepoint x_0. The reduced K-theory $\widetilde{K^0}(X)$ is given by $\ker(M)$.

Alternative Definition 38.1.8. One can define the reduced K-theory $\widetilde{K}(X)$ equivalently as follows: Consider the stable isomorphism classes[2] of vector bundles over X. Under Whitney sums these define a commutative group $(\mathrm{Vect}(X)/\sim_{stable}, \oplus)$ which is (naturally) isomorphic to $\widetilde{K^0}(X)$.

[1] See definition 3.1.6.
[2] See definition 30.2.7.

Chapter 39

Synthetic Differential Geometry

39.1 Neighbourhoods

Definition 39.1.1 (Neighbourhood relation). A reflexive and symmetric relation $\sim$ with the additional property that the morphisms in the category under consideration preserve this relationship.

Example 39.1.2 (Monad). Let M be a set. Given a neighbour relation $\sim$ on M, the (first order) monad around $x \in M$ is defined as:

$$\underline{\mathfrak{M}}(x) = \{y \in M : y \sim x\} \tag{39.1}$$

Definition 39.1.3 (Infinitesimal simplex). An infinitesimal k-simplex with respect to a neighbour relation $\sim$ is a collection of $k+1$ points $\{x_i\}_{i \leq k}$ such that $x_i \sim x_j$ for every $i, j \leq k$.

Definition 39.1.4 (Geometric distribution). Let M be a set equipped with a neighbour relation $\sim$. A (geometric) distribution on M is a reflexive symmetric refinement $\approx$ of $\sim$. A distribution is said to be **involutive** if

$$(x \approx y) \wedge (y \approx z) \wedge (x \sim z) \implies x \approx z \tag{39.2}$$

for all $x, y, z \in M$.

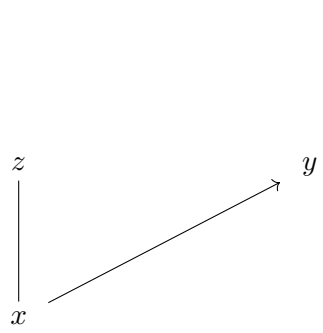
Definition 39.1.5 (Integral subset). Let M be a set equipped with a neighbour relation $\sim$ and a distribution $\approx$. A subset $N \subseteq M$ is said to be integral with respect to $\approx$ if $\approx$ and $\sim$ coincide on N.

Theorem 39.1.6 (Frobenius' theorem). *An involutive distribution admits maximal connected integral subsets, these are called* ***leaves****.*

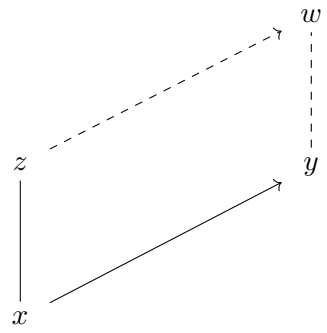
39.2 Affine connections

Definition 39.2.1 (Affine connection). An affine connection is a map $\lambda(x, y, z)$ which for every three points $x, y, z \in M$ such that $y \sim x$ and $z \sim x$ gives a point $w \in M$ such that $w \sim z$ and $w \sim y$. Graphically this is given by a completion of diagram 39.1a to diagram 39.1b.

Remark. By looking at these diagrams the concept of parallel transport can be made a lot more intuitive than in classic differential geometry, e.g. diagram 39.1b shows the parallel transport of the point z along xy.



(a) Neighbouring points.



(b) Connection in synthetic theories.

Definition 39.2.2 (Symmetric connection). An affine connection λ is said to be symmetric or **torsion-free** if $\lambda(x, y, z) = \lambda(x, z, y)$.

Definition 39.2.3 (Flat connection). An affine connection λ is said to be flat or **curvature-free** if parallel transporting a point around an infinitesimal 2-simplex gives that same point again.

Definition 39.2.4 (Curvature). Let λ be an affine connection on M. The curvature of λ is the map $\mathcal{R}$ which assigns to every infinitesimal 2-simplex $\{x_0, x_1, x_2\}$ the following automorphism:

$$\underline{\mathfrak{M}}(x_0) \to \underline{\mathfrak{M}}(x_0) : z \mapsto \text{result of parallel transporting z around } \{x_0, x_1, x_2\}$$

Definition 39.2.5 (Geodesic). A subset $S \subseteq M$ stable under the affine connection λ.

39.3 Euclidean geometry

39.3.1 Infinitesimal elements

Definition 39.3.1 (Infinitesimal line). Let R be the line. By picking two distinct points, labelled 0 and 1, one can turn the line into a commutative ring[1] $(R, +, \cdot)$. The infinitesimal line is then defined as the following set:

$$\Delta = \{x \in R : x^2 = 0\} \tag{39.3}$$

A neighbourhood relation on R is then induced by setting $\underline{\mathfrak{M}}(0) \equiv \Delta$.

Remark 39.3.2. If one would follow the Euclidean point of view, this set would be $\{0\}$. However by not requiring R to be a field we obtain a larger set.

Axiom 39.1. For every map $f : \Delta \to R$ there exists a unique element $b \in R$, called the **slope** of f, such that:

$$f(d) = f(0) + d \cdot b \tag{39.4}$$

for all $d \in \Delta$.

Corollary 39.3.3. The map $\alpha : R \times R \to R^{\Delta} : (a, b) \mapsto (f : d \mapsto a + d \cdot b)$ is invertible and hence an isomorphism[2].

Corollary 39.3.4. Let $a, b \in R$. If $d \cdot a = d \cdot b$ for all $d \in \Delta$ then $a = b$.

Notation 39.3.5. Analogous to the infinitesimal line one defines the subsets D_k in the following way:

$$D_k = \{x \in R : x^{k+1} = 0\} \tag{39.5}$$

[1] Or more explicitly an algebra over the rationals $\mathbb{Q}$.

[2] If one equips the set $R \times R$ with the multiplication rule $(a, b) \cdot (a', b') = (a \cdot a', a \cdot b' + a' \cdot b)$ then this becomes an R-algebra isomorphism.

39.3.2 Calculus

Formula 39.3.6 (Taylor expansion). From axiom 39.1 we derive the following exact Taylor expansion:

$$f(x+d) = f(x) + d \cdot f'(x) \tag{39.6}$$

where $f'(x)$ can be interpreted as the derivative of f at the point $x \in R$.

Remark. If f also depends on additional parameters in R, then we can define the partial derivatives in a similar fashion.

Property 39.3.7. Using axiom 39.1 it can be easily proven that the derivative is linear and satisfies Leibniz's rule.

Using the sets D_k one can prove higher order expansions. First we generalize axiom 39.1:

Axiom 39.2. For every $g : D_k \to R$ there exist unique elements $\{b_1, \ldots, b_k\}$ in R such that:

$$f(d) = f(0) + \sum_j d^j \cdot b_j \tag{39.7}$$

for all $d \in D_k$.

Corollary 39.3.8. Let $f : R \to R$ and $d \in D_k$.

$$\boxed{f(x+d) = f(x) + d \cdot f'(x) + \cdots + \frac{d^k}{k!} \cdot f^{(k)}(x)} \tag{39.8}$$

Part VII

Probability Theory & Statistics

Chapter 40

Probability

40.1 Probability

Definition 40.1.1 (Axioms of probability).

- $P(E) \geq 0$
- $P(E_1 \text{ or } E_2) = P(E_1) + P(E_2)$ if E_1 and E_2 are exclusive.
- $\sum_S P(E_i) = 1$ where the summation runs over all exclusive events.

Remark 40.1.2. The second axiom can be defined more generally by saying that the probability P should be σ-additive. Together with the first axiom and the consequence that $P(\emptyset) = 0$ means that the probability is a measure 14.1.1.

Definition 40.1.3 (Sample space). Let X be a random variable. The set of all possible outcomes of X is called the sample space. The sample space is often denoted by Ω.

Definition 40.1.4 (Probability space). Let (Ω, Σ, P) be a measure space[1]. This measure space is called a probability space if $P(X) = 1$. Furthermore, the measure P is called a probability measure or simply probability.

Definition 40.1.5 (Event). Let (Ω, Σ, P) be a probability space. An element S of the σ-algebra Σ is called an event.

Remark. From the definition of an event 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 event can occur simultaneously.

Remark. When working with measure-theoretic probability spaces it is more convenient to use the σ-algebra (see 2.4.2) of events instead of the power set (see 2.1.2) of all events. Intuitively this seems to mean that some possible outcomes are not treated as events. However we can make sure that the σ-algebra still contains all 'useful' events by using a 'nice' definition of the used probability space. Further information concerning probability spaces can be found in chapter 14.

Formula 40.1.6. Let A, B be two events.

$$P(A \cup B) = P(A) + P(B) + P(A \cap B) \tag{40.1}$$

[1] See definition 14.1.2.

Definition 40.1.7 (Disjoint events). Two events A and B are said to be disjoint if they cannot happen at the same time:

$$P(A \cap B) = 0 \tag{40.2}$$

Corollary 40.1.8. If A and B are disjoint, the probability that both A and B are true is just the sum of their individual propabilities.

Formula 40.1.9 (Complement). Let A be an event. The probability of A being false is denoted as $P\left(\overline{A}\right)$ and is given by:

$$P\left(\overline{A}\right) = 1 - P(A) \tag{40.3}$$

Corollary 40.1.10. From the previous equation and de Morgan's laws (equations 2.6 and 2.7) we derive the following formula[2]:

$$P\left(\overline{A} \cap \overline{B}\right) = 1 - P(A \cup B) \tag{40.4}$$

40.2 Conditional probability

Definition 40.2.1 (Conditional probability). Let A, B be two events. The probability of A given that B is true is denoted as $P(A|B)$.

$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)}} \tag{40.5}$$

Corollary 40.2.2. By interchanging A and B in previous equation and by remarking that this has no effect on the quantity $P(A \cap B)$ the following rsult can be deducued:

$$P(A|B)P(B) = P(B|A)P(A) \tag{40.6}$$

Theorem 40.2.3 (Bayes' theorem). *Let A, B be two events. From the conditional probability 40.5 it is possible to derive following important theorem:*

$$\boxed{P(A|B) = \frac{P(B|A)P(A)}{P(B)}} \tag{40.7}$$

Definition 40.2.4 (Independent events). Let A, B be two events. A and B are said to be independent if they satisfy the following relation:

$$P(A \cap B) = P(A)P(B) \tag{40.8}$$

Corollary 40.2.5. If A and B are two independent events, then equation 40.7 simplifies to:

$$P(A|B) = P(A) \tag{40.9}$$

Property 40.2.6. The events $A_1, ..., A_n$ are independent if for all $k \leq n$ for each choice of k events the probability of their intersection is equal to the product of their indivudal propabilities.

Property 40.2.7. The σ-algebras $\mathcal{F}_1, ..., \mathcal{F}_n$ defined on probability space $(\Omega, \mathcal{F}, P)$ are independent if for all choices of distinct indices $i_1, ..., i_k$ from $\{1, ..., n\}$ and for all choices of sets $F_{i_n} \in \mathcal{F}_{i_n}$ the following equation holds:

$$P(F_{i_1} \cap ... \cap F_{i_k}) = P(F_{i_1})...P(F_{i_k}) \tag{40.10}$$

Formula 40.2.8. Let $(B_i)_{i \in \mathbb{N}}$ be a sequence of pairwise disjoint events. If $\bigcup_{i=1}^{+\infty} B_i = \Omega$ then the total probability of a given event A can be calculated as follows:

$$P(A) = \sum_{i=1}^{+\infty} P(A|B_i)P(B_i) \tag{40.11}$$

[2]Switching the union and intersection has no impact on the validity of the formula.

40.3 Random variables

Definition 40.3.1 (Random variable). Let (Ω, Σ, P) be a probability space. A function $X : \Omega \to \mathbb{R}$ is called a random variable if $\forall a \in \mathbb{R} : X^{-1}\big([a, +\infty[\big) \in \Sigma$.[3]

Definition 40.3.2 (σ-algebra of a random variable). Let X be a random variable defined on a probability space (Ω, Σ, P). The following family of sets is a σ-algebra:

$$X^{-1}(\mathcal{B}) = \{S \in \Sigma : S = X^{-1}(B \in \mathcal{B})\} \tag{40.12}$$

Notation 40.3.3. The σ-algebra generated by the random variable X is often denoted by $\mathcal{F}_X$, analogous to notation 2.4.7.

Theorem 40.3.4. *Let X, Y be two random variables. X and Y are independent if the σ-algebras generated by them are independent*[4]*.*

40.4 Probability distribution

Definition 40.4.1 (Probability distribution). Let X be a random variable defined on a probability space (Ω, Σ, P). The following function is a measure on the σ-algebra of Borel sets:

$$\boxed{P_X(B) = P(X^{-1}(B))} \tag{40.13}$$

This measure is called the probability distribution of X.

Formula 40.4.2 (Change of variable). Let X be a random variable defined on a probability space (Ω, Σ, P).

$$\int_\Omega g(X(\omega))dP(\omega) = \int_{\mathbb{R}} g(x)dP_X(x) \tag{40.14}$$

Definition 40.4.3 (Density). Let f be a non-negative integrable function and recall theorem 14.2.21. The function f is called the **density** of P with respect to the Lebesgue measure m.

For P to be a probability, f should satisfy the following condition:

$$\int f dm = 1 \tag{40.15}$$

Definition 40.4.4 (Cumulative distribution function). Let f be a density. The c.d.f. corresponding to f is given by:

$$F(y) = \int_{-\infty}^{y} f(x)dx \tag{40.16}$$

Theorem 40.4.5 (Skorokhod's representation theorem). *Let $F : \mathbb{R} \to [0, 1]$ be a function that satisfies following 3 properties:*

- *$F(x)$ is non-decreasing.*
- *$\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to+\infty} F(x) = 1$*
- *$F(x)$ is right-continuous: $y \geq y_0, y \to y_0 \implies F(y) \to F(y_0)$.*

There exists a random variable $X : [0, 1] \to \mathbb{R}$ defined on the probability space $([0, 1], \mathcal{B}, m_{[0,1]})$ such that $F = F_X$.

[3] $X^{-1}\big([a, +\infty[\big) = \{\omega \in \Omega : X(\omega) \geq a\}$.
[4] See equation 40.10.

Formula 40.4.6. Let the absolutely continuous probability P_X be defined on the product space $\mathbb{R}^n$. Let f_X be the density associated with P_X. Let $g : \mathbb{R}^n \to \mathbb{R}$ be integrable with respect to P_X.

$$\int_{\mathbb{R}^\kappa} g(x)dP_X(x) = \int_{\mathbb{R}^n} f_X(x)g(x)dx \tag{40.17}$$

Corollary 40.4.7. Previous formula together with formula 40.14 gives rise to:

$$\int_{\Omega} g(X)dP = \int_{\mathbb{R}^\kappa} f_X(x)g(x)dx \tag{40.18}$$

40.5 Moments

40.5.1 Expectation value

Definition 40.5.1 (Expectation value). Let X be random variable defined on a probability space (Ω, Σ, P).

$$\boxed{E(X) = \int_{\Omega} XdP} \tag{40.19}$$

Notation 40.5.2. Other often used notations are $\langle X \rangle$ and μ.

Definition 40.5.3 (Moment of order r). The moment of order r is defined as the expectation value of the r^{th} power of X and by equation 40.18 this becomes:

$$E(X^r) = \int x^r f_X(x)dx \tag{40.20}$$

Definition 40.5.4 (Central moment of order r).

$$E((X-\mu)^r) = \int (x-\mu)^r f_X(x)dx \tag{40.21}$$

Definition 40.5.5 (Variance). The central moment of order 2 is called the variance: $V(X) = E((X-\mu)^2)$.

Property 40.5.6. If $E(X^n)$ are finite for $n > 0$ then for all $k \leq n$, $E(X^k)$ are also finite. If $E(X^n)$ is infinite then for all $k \geq n$, $E(X^k)$ are also infinite.

Property 40.5.7. Moments of order n are determined by central moments of order $k \leq n$ and central moments of order n are determined by moments of order $k \leq n$.

Definition 40.5.8 (Moment generating function).

$$M_X(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tX} P(X)dX \tag{40.22}$$

Theorem 40.5.9. *If the above function exists we can derive the following useful result*[5] *by using the series expansion 12.15:*

$$E[X^n] = \left.\frac{d^n M_X(t)}{dt^n}\right|_{t=0} \tag{40.23}$$

[5]This property is the reason why 40.22 is called the moment generating function.

Definition 40.5.10 (Characteristic function). Let X be a random variable. The characteristic function of X is defined as follows:

$$\varphi_X(t) = E(e^{itX}) \tag{40.24}$$

Property 40.5.11. The characteristic function has the following properties:

- $\varphi_X(0) = 1$
- $|\varphi_X(t)| \leq 1$
- $\varphi_{aX+b}(t) = e^{itb}\varphi_X(at)$

Formula 40.5.12. If $\varphi_X(t)$ is k times continuously differentiable then X has finite k^{th} moment and

$$E(X^k) = \frac{1}{i^k}\frac{d^k}{dt^k}\varphi_X(0) \tag{40.25}$$

Conversely, if X has finite k^{th} moment then $\varphi_X(t)$ is k times continuously differentiable and the above formula holds.

Formula 40.5.13 (Inversion formula). Let X be a random varibale. If the c.d.f. of X is continuous at $a, b \in \mathbb{R}$ then

$$F_X(b) - F_X(a) = \lim_{c\to+\infty}\frac{1}{2\pi}\int_{-c}^{c}\frac{e^{-ita} - e^{-itb}}{it}\varphi_X(t)dt \tag{40.26}$$

Formula 40.5.14. If $\varphi_X(t)$ is integrable then the c.d.f. is given by:

$$f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-itx}\varphi_X(t)dt \tag{40.27}$$

Remark 40.5.15. From previous formula it is clear that the density function and the characteristic function are Fourier transformed quantities.

40.5.2 Correlation

Theorem 40.5.16. *Let X, Y be two random variables. They are independent if and only if $E(f(X)g(Y)) = E(f(X))E(g(Y))$ holds for all Borel measurable[6] bounded functions f, g.*

The value $E(XY)$ is equal to the inner product $\langle X|Y\rangle$ as defined in 14.40. It follows that independence of random variables implies orthogonality. To generalize this concept, we introduce following notions.

Definition 40.5.17 (Centred random variable). Let X be a random variable with finite expectation value $E(X)$. The centred random variable X_c is defined as $X_c = X - E(X)$.

Definition 40.5.18 (Covariance). Let X, Y be two random variables. The covariance of X and Y is defined as follows:

$$\mathrm{cov}(X, Y) = \langle X_c|Y_c\rangle = E\big((X - E(X))(Y - E(Y))\big) \tag{40.28}$$

Some basic math gives:

$$\mathrm{cov}(X, Y) = E(XY) - E(X)E(Y) \tag{40.29}$$

[6] See definition 14.1.35.

Definition 40.5.19 (Correlation). Let X, Y be two random variables. The correlation is defined as the cosine of the angle between X_c and Y_c:

$$\rho_{XY} = \frac{\text{cov}(X,Y)}{||X||_2||Y||_2} \tag{40.30}$$

Corollary 40.5.20. From theorem 40.5.16 it follows that independent random variables are also uncorrelated.

Corollary 40.5.21. Uncorrelated X and Y satisfy the following equality: $E(XY) = E(X)E(Y)$.

Property 40.5.22. Let $(X_i)_{i\in\mathbb{N}}$ be a sequence of independent random variables. Their variances satisfy the following equation:

$$V\left(\sum_{i=1}^{+\infty} X_i\right) = \sum_{i=1}^{+\infty} V(X_i) \tag{40.31}$$

40.5.3 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 $\mathcal{G} \subset \Sigma$. We know that the spaces $L^2(\Sigma)$ and $L^2(\mathcal{G})$ are complete[7] and hence the projection theorem 21.2.22 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:

$$\forall Z \in L^2(\mathcal{G}) : \langle X - Y | Z \rangle = \int_\Omega (X - Y) Z dP = 0 \tag{40.32}$$

Since $\mathbb{1}_G \in L^2(\mathcal{G})$ for every $G \in \mathcal{G}$ we find by applying 14.29:

$$\int_G X dP = \int_G Y dP \tag{40.33}$$

This leads us to introducing the following notion of conditional expectations:

Definition 40.5.23 (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 40.33 for every $G \in \mathcal{G}$. This Y is called the conditional expectation of X given $\mathcal{G}$ and it is denoted by $Y := E(X|\mathcal{G})$:

$$\boxed{\int_G E(X|\mathcal{G}) dP = \int_G X dP} \tag{40.34}$$

Remark 40.5.24. Although our derivation was based on random variables from the L^2 class, it is also possible to construct (unique) conditional expectations for random variables from the L^1 class by using method 14.2.24.

40.6 Joint distributions

Definition 40.6.1 (Joint distribution). Let X, Y be two random variables defined on the same probability space (Ω, Σ, P). Consider the vector $(X, Y) : \Omega \to \mathbb{R}^2$. The distribution of (X, Y) is defined on the Borel sets of the plane $\mathbb{R}^2$ and it is given by the following measure:

$$P_{(X,Y)}(B) = P((X, Y) \in B) \tag{40.35}$$

[7] See property 14.4.3.

Definition 40.6.2 (Joint density). If the probability measure from previous definition can be written as

$$P_{(X,Y)}(B) = \int_B f_{(X,Y)}(x,y)dm_2(x,y) \tag{40.36}$$

for some integrable $f_{(X,Y)}$ it is said that X and Y have a joint density.

Definition 40.6.3 (Marginal distribution). The distributions of one-dimensional random variables is determined by the joint distribution:

$$P_X(A) = P_{(X,Y)}(A \times \mathbb{R}) \tag{40.37}$$

$$P_Y(A) = P_{(X,Y)}(\mathbb{R} \times A) \tag{40.38}$$

Corollary 40.6.4. If the joint density exists then the marginal distributions are absolutely continuous and given by

$$f_X(x) = \int_{\mathbb{R}} f_{(X,Y)}(x,y)dy \tag{40.39}$$

$$f_Y(y) = \int_{\mathbb{R}} f_{(X,Y)}(x,y)dx \tag{40.40}$$

The converse however is not always true. The one-dimensional densities can be absolutely continuous without the existence of the joint density.

40.6.1 Independence

Theorem 40.6.5. *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:*

$$P_{(X,Y)} = P_X \times P_Y \tag{40.41}$$

Remark 40.6.6. If X and Y are absolutely continuous then the previous theorem also applies with the densities instead of the distributions.

40.6.2 Conditional probability

Formula 40.6.7 (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:

$$h(y|X \in A) = \frac{\int_A f_{(X,Y)}(x,y)dx}{\int_A f_X(x)dx} \tag{40.42}$$

For $X = \{a\}$ this equation fails as the denominator would become 0. However it is possible to avoid this problem by formally putting

$$h(y|A = a) = \frac{f_{(X,Y)}(a,y)}{f_X(a)} \tag{40.43}$$

with $f_X(a) \neq 0$ which is non-restrictive because the probability of having a measurement $(X,Y) \in \{(x,y) : f_X(x) = 0\}$ is 0. We can thus define the conditional probability of Y given $X = a$:

$$P(Y \in B|X = a) = \int_B h(y|X = a)dy \tag{40.44}$$

Formula 40.6.8 (Conditional expectation).

$$E(Y|X)(\omega) = \int_{\mathbb{R}} yh(y|X(\omega))dy \tag{40.45}$$

Furthermore, let $\mathcal{F}_X$ denote the σ-algebra generated by the random variable X. Using Fubini's theorem we can prove that for all sets $A \in \mathcal{F}_X$ the following equality, which should be compared with equation 40.34, holds:

$$\int_A E(Y|X)dP = \int_A YdP \tag{40.46}$$

Remark 40.6.9. Following from previous two equations we can say that the conditional expectation $E(Y|X)$ is the best representation of the random variable Y as a function of X (i.e. measurable with respect to $\mathcal{F}_X$).

Property 40.6.10. As mentioned above, applying Fubini's theorem gives:

$$\boxed{E(E(Y|X)) = E(Y)} \tag{40.47}$$

Chapter 41

Statistics

In this chapter, most definitions and formulas will be based on either a standard calculus approach or a data-driven approach. For a measure-theory based approach see chapter 40.

41.1 Data samples

41.1.1 Moment

Formula 41.1.1 (r^{th} sample moment).

$$\overline{x^r} = \frac{1}{N}\sum_{i=1}^{N} x_i^r \tag{41.1}$$

Formula 41.1.2 (r^{th} central sample moment).

$$m_r = \frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})^r \tag{41.2}$$

41.1.2 Mean

Definition 41.1.3 (Arithmetic mean). The arithmetic mean is used to average out differences between measurements. It is equal to the 1^{st} sample moment:

$$\boxed{\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i} \tag{41.3}$$

Formula 41.1.4 (Weighted mean). Let $f : \mathbb{R} \to \mathbb{R}^+$ be a weight function. The weighted mean is given by:

$$\overline{x} = \frac{\sum_i f(x_i)x_i}{\sum_i f(x_i)} \tag{41.4}$$

Corollary 41.1.5. If the data has been grouped in bins, the weight function is given by the number of elements in each bin. Knowing this the (binned) mean becomes:

$$\overline{x} = \frac{1}{N}\sum_{i=1} n_i x_i \tag{41.5}$$

Remark 41.1.6. In the above defintions the measurements x_i can be replaced by function values $f(x_i)$ to calculate the mean of the function $f(x)$.

Remark 41.1.7. It is also important to notice that $\overline{f}(x) \neq f(\overline{x})$. The equality only holds for linear functions.

Definition 41.1.8 (Geometric mean). Let $\{x_i\}$ be a positive data set[1]. The geometric mean is used to average out *normalized* measurements, i.e. ratios with respect to a reference value.

$$g = \left(\prod_{i=1}^{N} x_i \right)^{1/N} \tag{41.6}$$

The following relation exists between the arithmetic and geometic mean:

$$\ln g = \overline{\ln x} \tag{41.7}$$

Definition 41.1.9 (Harmonic mean).

$$h = \left(\frac{1}{N} \sum_{i=1}^{N} x_i^{-1} \right)^{-1} \tag{41.8}$$

The following relation exists between the arithmetic and harmonic mean:

$$\frac{1}{h} = \overline{x^{-1}} \tag{41.9}$$

Property 41.1.10. Let $\{x_i\}$ be a positive data set.

$$h \leq g \leq \overline{x} \tag{41.10}$$

where the equalities only hold when all x_i are equal.

Definition 41.1.11 (Mode). The most occurring value in a dataset.

Definition 41.1.12 (Median). The median of dataset is the value x_i such that half of the values is greater than x_i and the other half is smaller than x_i.

41.1.3 Dispersion

Definition 41.1.13 (Range). The simplest indicator for statistical dispersion. It is however very sensitive for extreme values.

$$R = x_{max} - x_{min} \tag{41.11}$$

Definition 41.1.14 (Mean absolute difference).

$$\mathrm{MD} = \frac{1}{N} \sum_{i=1}^{N} |x_i - \overline{x}| \tag{41.12}$$

Definition 41.1.15 (Sample variance).

$$V(X) = \frac{1}{N} \sum_{i=1}^{N} (x_i - \overline{x})^2 \tag{41.13}$$

[1] A negative data set is also allowed. The real condition is that all values should have the same sign.

Formula 41.1.16. The variance can also be written in the following way:

$$\boxed{V(X) = \overline{x^2} - \overline{x}^2} \tag{41.14}$$

Remark 41.1.17. A better estimator for the variance of a sample is the following formula:

$$\hat{s} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2 \tag{41.15}$$

Equation 41.13 gives a good estimation when the sample mean $\overline{x}$ is replaced by the "true" mean μ. Otherwise one should use the estimator 41.15.

Definition 41.1.18 (Standard deviation).

$$\sigma(X) = \sqrt{V(x)} \tag{41.16}$$

Definition 41.1.19 (Skewness). The skewness γ describes the asymmetry of a distribution. It is defined in relation to the third central moment m_3;

$$m_3 = \gamma\sigma^3 \tag{41.17}$$

where σ is the standard deviation. A positive skewness indicates a tail to the right or alternatively a median smaller than $\overline{x}$. A negative skewness indicates a median larger than $\overline{x}$.

Definition 41.1.20 (Pearson's mode skewness).

$$\gamma_P = \frac{\overline{x} - \text{mode}}{\sigma} \tag{41.18}$$

Definition 41.1.21 (Kurtosis). The kurtosis c is an indicator for the "tailedness". It is defined in relation to the fourth central moment m_4:

$$m_4 = c\sigma^4 \tag{41.19}$$

Definition 41.1.22 (Excess kurtosis). The excess kurtosis is defined as $c - 3$. This fixes the excess kurtosis of all univariate normal distributions at 0. A positive excess is an indicator for long "fat" tails, a negative excess indicates short "thin" tails.

Definition 41.1.23 (Percentile). The p^{th} percentile c_p is defined as the value that is larger than $p\%$ of the measurements. The median is the 50^{th} percentile.

Definition 41.1.24 (Interquartile range). The interquartile range is the difference between the upper and lower quartile (75^{th} and 25^{th} percentile respectively).

Definition 41.1.25 (FWHM). The **Full Width at Half Maximum** is the difference between the two values of the independent variable where the dependent variable is half of its maximum.

Property 41.1.26. For Gaussian distributions the following relation exists between the FWHM and the standard deviation σ:

$$\text{FWHM} = 2.35\sigma \tag{41.20}$$

41.1.4 Multivariate datasets

When working with bivariate (or even multivariate) distributions it is useful to describe the relationship between the different random variables. The following two definitions are often used.

Definition 41.1.27 (Covariance). Let X, Y be two random variables. The covariance of X and Y is defined as follows:

$$\operatorname{cov}(X,Y) = \frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})(y_i - \overline{y}) = \overline{xy} - \overline{x}\ \overline{y} \tag{41.21}$$

The covariance of X and Y is often denoted by σ_{XY}.

Formula 41.1.28. The covariance and standard deviation are related by the following equality:

$$\sigma_X^2 = \sigma_{XX} \tag{41.22}$$

Definition 41.1.29 (Correlation coefficent).

$$\rho_{XY} = \frac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_Y} \tag{41.23}$$

The correlation coefficient is bounded to the interval $[-1, 1]$. It should be noted that its magnitude is only an indicator for the linear dependence.

Remark 41.1.30. For multivariate distributions the above definitions can be generalized using matrices:

$$V_{ij} = \operatorname{cov}(x_{(i)}, x_{(j)}) \tag{41.24}$$

$$\rho_{ij} = \rho_{(i)(j)} \tag{41.25}$$

where $\operatorname{cov}(x_{(i)}, x_{(j)})$ and $\rho_{(i)(j)}$ are defined using equations 41.21 and 41.23. The following general equality exists:

$$V_{ij} = \rho_{ij}\sigma_i\sigma_j \tag{41.26}$$

41.2 Law of large numbers

Theorem 41.2.1 (Law of large numbers). *If the size N of a sample tends towards infinity, then the observed frequencies tend towards the theoretical propabilities.*

Corollary 41.2.2 (Frequentist probability[2]).

$$P(X) = \lim_{n\to\infty} \frac{f_n(X)}{n} \tag{41.27}$$

41.3 Probability densities

Remark. In the following sections and subsections, all distributions will be taken to be continuous. The formulas can be modified for use with discrete distributions by replacing the integral with a summation.

[2] Also called the **empirical probability**.

Definition 41.3.1 (Probability density function). Let X be a random variable and $P(X)$ the associated probability distribution. The p.d.f. $f(X)$ is defined as follows:

$$\boxed{P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f(X)dX} \tag{41.28}$$

An alternative definition[3] is the following:

$$f(X) = \lim_{\delta x \to 0} \frac{P(x \leq X \leq x + \delta x)}{\delta x} \tag{41.29}$$

Definition 41.3.2 (Cumulative distribution function). Let X be a random variable and $f(X)$ the associated p.d.f. The cumulative distribution function $F(X)$ is defined as follows:

$$F(x) = \int_{-\infty}^{x} f(X)dX \tag{41.30}$$

Theorem 41.3.3. *Let X be a random variable. Let $P(X)$ and $F(X)$ be the associated p.d.f. and c.d.f. Using standard calculus the following equality can be proven:*

$$P(x_1 \leq X \leq x_2) = F(x_2) - F(x_1) \tag{41.31}$$

Theorem 41.3.4. *$F(X)$ is continuous if and only if $P_X(\{y\}) = 0$ for every $y \in \mathbb{R}$.*

Remark 41.3.5 (Normalization).

$$F(\infty) = 1 \tag{41.32}$$

Formula 41.3.6. The p^{th} percentile c_p can be computed as follows[4]:

$$c_p = F^{-1}(p) \tag{41.33}$$

Definition 41.3.7 (Parametric family). A parametric family of probability densities $f(X; \vec{\theta})$ is a set of densities described by one or more parameters $\vec{\theta}$.

41.3.1 Function of a random variable

Formula 41.3.8. Let X be random variable and $f(X)$ the associated p.d.f. Let $a(X)$ be a function of X. The random variable $A = a(X)$ has an associated p.d.f. $g(A)$. If the function $a(x)$ can be inverted, then $g(A)$ can be computed as follows:

$$\boxed{g(a) = f(x(a)) \left| \frac{dx}{da} \right|} \tag{41.34}$$

41.3.2 Multivariate distributions

Remark. In this section all defintions and thereoms will be given for bivariate distributions, but can be easily generalized to more random variables.

[3] A more formal definition uses *measure theory* and the *Radon-Nikodym derivative*.

[4] This is clear from the definition of a percentile, as this implies that $F(c_p) = p$.

Definition 41.3.9 (Joint density). Let X, Y be two random variables. The joint p.d.f. $f_{XY}(x, y)$ is defined as follows:

$$f_{XY}(x,y)dxdy = \begin{cases} f_x(x \in [x, x+dx]) \\ f_y(y \in [y, y+dy]) \end{cases} \tag{41.35}$$

Remark 41.3.10. As f_{XY} is a probability density, the normalization condition 41.32 should be fulfilled.

Definition 41.3.11 (Conditional density). The conditional p.d.f. of X when Y has the value y is given by the following formula:

$$g(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} \tag{41.36}$$

where we should pay attention to the remark made when we defined 40.43.

Corollary 41.3.12. If X and Y are independent, then by remark 40.6.6 the marginal p.d.f is equal to the conditional p.d.f.

Theorem 41.3.13 (Bayes' theorem). *The conditional p.d.f. can be computed without prior knowledge of the joint p.d.f:*

$$\boxed{g(x|y) = \frac{h(y|x)f_X(x)}{f_Y(y)}} \tag{41.37}$$

Remark. This theorem is the statistical (random variable) analogon of theorem 40.7.

Formula 41.3.14. Let $Z = XY$ with X, Y two independent random variables. The distribution $f(z)$ is given by

$$f(z) = \int_{-\infty}^{+\infty} g(x)h(z/x)\frac{dx}{|x|} = \int_{-\infty}^{+\infty} g(z/y)h(y)\frac{dy}{|y|} \tag{41.38}$$

Corollary 41.3.15. Taking the Mellin transform 15.27 of both the positive and negative part of the above integrand (to be able to handle the absolute value) gives following relation

$$\mathcal{M}\{f\} = \mathcal{M}\{g\}\mathcal{M}\{h\} \tag{41.39}$$

Formula 41.3.16. Let $Z = X + Y$ with X, Y two independent random variables. The distribution $f(z)$ is given by the convolution of $g(x)$ and $h(y)$:

$$f(z) = \int_{-\infty}^{+\infty} g(x)h(z-x)dx = \int_{-\infty}^{+\infty} g(z-y)h(y)dy \tag{41.40}$$

41.3.3 Important distributions

Formula 41.3.17 (Uniform distribution).

$$\boxed{P(x;a,b) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{elsewhere} \end{cases}} \tag{41.41}$$

$$E(x) = \frac{a+b}{2} \tag{41.42}$$

$$V(x) = \frac{(b-a)^2}{12} \tag{41.43}$$

Formula 41.3.18 (Normal distribution). Also called the Gaussian distribution.

$$\boxed{\mathcal{G}(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}} \tag{41.44}$$

Formula 41.3.19 (Standard normal distribution).

$$\boxed{\mathcal{N}(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}} \tag{41.45}$$

Remark 41.3.20. Every Gaussian distribution can be rewritten as a standard normal distribution by setting $Z = \frac{X-\mu}{\sigma}$.

Remark 41.3.21. The c.d.f. of the standard normal distribution is given by the error function: $F(z) = \text{Erf}(z)$.

Formula 41.3.22 (Exponential distribution).

$$\boxed{P(x;\tau) = \frac{1}{\tau} e^{-\frac{x}{\tau}}} \tag{41.46}$$

$$E(x) = \tau \tag{41.47}$$

$$V(x) = \tau^2 \tag{41.48}$$

Theorem 41.3.23. *The exponential distribution is memoryless:*

$$P(X > x_1 + x_2 | X > x_2) = P(X > x_1) \tag{41.49}$$

Formula 41.3.24 (Bernoulli distribution). A radnom variable that can only take 2 possible values is described by a Bernoulli distribution. When the possible values are 0 and 1, with respective chances p and $1-p$, the distribution is given by:

$$\boxed{P(x;p) = p^x (1-p)^{1-x}} \tag{41.50}$$

$$E(x) = p \tag{41.51}$$

$$V(x) = p(1-p) \tag{41.52}$$

Formula 41.3.25 (Binomial distribution). A process with n identical independent trials, all Bernoulli processes $P(x;p)$, is described by a binomial distribution:

$$\boxed{P(r;p,n) = p^r (1-p)^{n-r} \frac{n!}{r!(n-r)!}} \tag{41.53}$$

$$E(r) = np \tag{41.54}$$

$$V(r) = np(1-p) \tag{41.55}$$

Formula 41.3.26 (Poisson distribution). A process with known possible outcomes but an unknown number of events is described by a Poisson distribution $P(r;\lambda)$ with λ the average expected number of events.

$$\boxed{P(r;\lambda) = \frac{e^{-\lambda}\lambda^r}{r!}} \tag{41.56}$$

$$E(r) = \lambda \tag{41.57}$$

$$V(r) = \lambda \tag{41.58}$$

Theorem 41.3.27. *If two Poisson processes $P(r;\lambda_a)$ and $P(r;\lambda_b)$ occur simultaneously and if there is no distinction between the two, then the probability of r events is also described by a Poisson distribution with average $\lambda_a + \lambda_b$.*

Corollary 41.3.28. The number of events coming from A is given by a binomial distribution $P(r_a;\Lambda_a,r)$ where $\Lambda_a = \frac{\lambda_a}{\lambda_a+\lambda_b}$.

Remark 41.3.29. For large values of λ ($\lambda \to \infty$), the Poisson distribution $P(r;\lambda)$ can be approximated by a Gaussian distribution $\mathcal{G}(x;\lambda,\sqrt{\lambda})$.

Formula 41.3.30 (χ^2 distribution). The sum of k squared independent (standard) normally distributed random variables Y_i defines the random variable:

$$\chi^2_k = \sum_{i=1}^{k} Y_i^2 \tag{41.59}$$

where k is said to be the number of **degrees of freedom**.

$$\boxed{P(\chi^2;n) = \frac{\chi^{n-2}e^{-\frac{\chi^2}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}} \tag{41.60}$$

Remark 41.3.31. Due to the CLT the χ^2 distribution approximates a Guassian distribution for large k: $P(\chi^2;k) \xrightarrow{k>30} \mathcal{G}(\sqrt{2\chi^2};\sqrt{2k-1},1)$

Formula 41.3.32 (Student-t distribution). The Student-t distribution describes a sample with estimated standard deviation $\hat{\sigma}$.

$$\boxed{P(t;n) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)\left(1+\frac{t^2}{n}\right)^{\frac{n+1}{2}}}} \tag{41.61}$$

where

$$t = \frac{(x-\mu)/\sigma}{\hat{\sigma}/\sigma} = \frac{z}{\sqrt{\chi^2/n}} \tag{41.62}$$

Remark. The significance of a difference between the sample mean $\overline{x}$ and the true mean μ is smaller due to the (extra) uncertainty of the estimated standard deviation.

Formula 41.3.33 (Cauchy distribution[5]). The general form $f(x; x_0, \gamma)$ is given by:

$$f(x; x_0, \gamma) = \frac{1}{\pi} \frac{\gamma}{(x - x_0)^2 + \gamma^2} \tag{41.63}$$

The characteristic function 40.24 is given by:

$$E\left(e^{itx}\right) = e^{ix_0 t - \gamma|t|} \tag{41.64}$$

Property 41.3.34. Both the mean and variance of the Cauchy distribution are undefined.

41.4 Central limit theorem (CLT)

Theorem 41.4.1 (Central limit theorem). *A sum of n independent random variables X_i has the following properties:*

1. $\mu = \sum_i \mu_i$
2. $V(X) = \sum_i V_i$
3. *The sum will be approximately (!!) normally distributed.*

Remark 41.4.2. If the random variables are not independent, property 2 will not be fulfilled.

Remark 41.4.3. The sum of Gaussians will be exactly Gaussian.

41.4.1 Distribution of sample mean

The difference between a sample mean $\overline{x}$ and the true mean μ is described by a distribution with following mean and variance:

Property 41.4.4.

$$< \overline{x} >= \mu \tag{41.65}$$

Property 41.4.5.

$$V(\overline{x}) = \frac{\sigma^2}{N} \tag{41.66}$$

41.5 Errors

41.5.1 Different measurement types

When performing a sequence of measurements x_i with different variances σ_i^2, it is impossible to use the arithmetic mean 41.3 in a meaningful way because the measurements are not of the same type. Therefore it is also impossible to apply the CLT 41.4.1.

Definition 41.5.1 (Weighted mean). The appropriate alternative is the *weighted mean*:

$$\boxed{\overline{x} = \frac{\sum_i \frac{x_i}{\sigma_i^2}}{\sum_i \frac{1}{\sigma_i^2}}} \tag{41.67}$$

[5] Also known (especially in particle physics) as the **Breit-Wigner** distribution.

The resolution of the weighted mean is given by:

$$V(\overline{x}) = \frac{1}{\sum_i \sigma_i^{-2}} \tag{41.68}$$

41.5.2 Propagation of errors

Formula 41.5.2. Let X be random variable with variance $V(x)$. The variance of a linear function $f(X) = aX + b$ is given by:

$$V(f) = a^2 V(x) \tag{41.69}$$

Formula 41.5.3. Let X be random variable with **small** (!!) variance $V(x)$. The variance of a general function $f(X)$ is given by:

$$V(f) \approx \left(\frac{df}{dx}\right)^2 V(x) \tag{41.70}$$

Corollary 41.5.4. The correlation coefficient ρ (41.23) of a random variable X and a **linear** function of X is independent of σ_x and is always equal to ± 1.

Formula 41.5.5 (Law of error propagation). Let $\vec{X}$ be a set of random variables with **small** variances. The variance of a general function $f(\vec{X})$ is given by:

$$\boxed{V(f) = \sum_p \left(\frac{\partial f}{\partial X_{(p)}}\right)^2 V(X_{(p)}) + \sum_p \sum_{q \neq p} \left(\frac{\partial f}{\partial X_{(p)}}\right)\left(\frac{\partial f}{\partial X_{(q)}}\right) \operatorname{cov}(X_{(p)}, X_{(q)})} \tag{41.71}$$

Definition 41.5.6 (Fractional error). Let X, Y be two **independent** random variables. The standard deviation of $f(x, y) = xy$ is given by the fractional error:

$$\left(\frac{\sigma_f}{f}\right)^2 = \left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2 \tag{41.72}$$

Remark 41.5.7. The fractional error of quantity is equal to the fractional error of the reciprocal of that quantity.

Property 41.5.8. Let X be a random variable. The error of the logarithm of X is equal to the fractional error of X.

Definition 41.5.9 (Covariance of functions).

$$\operatorname{cov}(f_1, f_2) = \sum_p \sum_q \left(\frac{\partial f_1}{\partial X_{(p)}}\right)\left(\frac{\partial f_2}{\partial X_{(q)}}\right) \operatorname{cov}(X_{(p)}, X_{(q)}) \tag{41.73}$$

Corollary 41.5.10. Let $\vec{f} = \{f_1, ..., f_k\}$. The covariance matrix V_f of the k functions is given by:

$$V_f = G\ V_X\ G^T \tag{41.74}$$

where G is the Jacobian matrix of $\vec{f}$.

41.5.3 Systematic errors

Systematic errors are errors that always have the same influence (they shift all values in the same way), that are not independent of eachother and that cannot be directly inferred from the measurements.

41.6 Estimators

Definition 41.6.1 (Estimator). An estimator is a procedure that, given a sample, produces a numerical value for a property of the parent population.

41.6.1 General properties

Definition 41.6.2 (Consistency). An estimator $\hat{a}$ is said to be consistent if

$$\boxed{\lim_{N\to\infty} \hat{a} = a} \tag{41.75}$$

Definition 41.6.3 (Unbiased estimator). An estimator $\hat{a}$ is said to be unbiased if

$$\boxed{<\hat{a}>= a} \tag{41.76}$$

Definition 41.6.4 (Bias).

$$B(\hat{a}) = |<\hat{a}> - a| \tag{41.77}$$

Definition 41.6.5 (Efficiency). An estimator $\hat{a}$ is said to be efficient if its variance $V(\hat{a})$ is equal to the minimum variance bound 41.86.

Definition 41.6.6 (Mean squared error).

$$\Upsilon(\hat{a}) = B(\hat{a})^2 + V(\hat{a}) \tag{41.78}$$

Remark 41.6.7. If an estimator is unbiased, the MSE is equal to the variance of the estimator.

41.6.2 Fundamental estimators

Property 41.6.8 (Mean estimator). The sample mean $\overline{x}$ is a consistent and unbiased estimator for the true mean μ due to the CLT. The variance $V(\overline{x})$ of the estimator is given by equation 41.66.

Formula 41.6.9 (Variance estimator for known mean). If the true mean μ is known then a consistent and unbiased estimator for the variance is given by:

$$\widehat{V(x)} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2 \tag{41.79}$$

Formula 41.6.10 (Variance estimator for unknown mean). If the true mean μ is unknown and the sample mean has been used to estimate μ, then a consistent and unbiased estimator is given by[6]:

$$s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2 \tag{41.80}$$

[6]The modified factor $\frac{1}{N-1}$ is called the Bessel correction. It corrects the bias of the estimator given by the sample variance 41.13. The consistency however is guaranteed by the CLT.

41.6.3 Estimation error

Formula 41.6.11 (Variance of estimator of variance).

$$V(\widehat{V(x)}) = \frac{(N-1)^2}{N^3} < (x- < x >)^4 > - \frac{(N-1)(N-3)}{N^3} < (x- < x >)^2 >^2 \tag{41.81}$$

Formula 41.6.12 (Variance of estimator of standard deviation).

$$V(\hat{\sigma}) = \frac{1}{4\sigma^2} V(\widehat{V(x)}) \tag{41.82}$$

Remark 41.6.13. The previous result is a little odd, as one has to know the true standard deviation to compute the variance of the estimator. This problem can be solved in two ways. Or a value (hopefully close to the real one) inferred from the sample is used as an estimator or a theoretical one is used in the design phase of an experiment to see what the possible outcomes are.

41.6.4 Likelihood function

Definition 41.6.14 (Likelihood). The likelihood $\mathcal{L}(a;\vec{x})$ is the probability to find a set of measurements $\vec{x} = \{x_1, ..., x_N\}$ given a distribution $P(X;a)$:

$$\mathcal{L}(a;\vec{x}) = \prod_{i=1}^{N} P(x_i;a) \tag{41.83}$$

Definition 41.6.15 (Log-likelihood).

$$\log \mathcal{L}(a;\vec{x}) = \sum_i \ln P(x_i;a) \tag{41.84}$$

Property 41.6.16. The expectation value of an estimator $\hat{a}$ is given by:

$$< \hat{a} >= \int \hat{a}\mathcal{L}(\hat{a};X)dX \tag{41.85}$$

Theorem 41.6.17 (Minimum variance bound). *The variance of an **unbiased** estimator has a lower bound: the minimum variance bound*[7] *(MVB).*

$$\boxed{V(\hat{a}) \geq \frac{1}{\left\langle \left(\frac{d \ln \mathcal{L}}{da} \right)^2 \right\rangle}} \tag{41.86}$$

For a biased estimator with bias b the MVB takes on the following form:

$$V(\hat{a}) \geq \frac{\left(1 + \frac{db}{da}\right)^2}{\left\langle \left(\frac{d \ln \mathcal{L}}{da} \right)^2 \right\rangle} \tag{41.87}$$

Remark 41.6.18.

$$\left\langle \left(\frac{d \ln \mathcal{L}}{da} \right)^2 \right\rangle = - \left\langle \frac{d^2 \ln \mathcal{L}}{da^2} \right\rangle \tag{41.88}$$

Definition 41.6.19 (Fisher information).

$$I_X(a) = \left\langle \left(\frac{d \ln \mathcal{L}}{da} \right)^2 \right\rangle = N \int \left(\frac{d \ln P}{da} \right)^2 P dX \tag{41.89}$$

[7] It is also known as the **Cramer-Rao bound**.

41.6.5 Maximum likelihood estimator

From definition 41.6.14 it follows that the estimator $\hat{a}_{MLH}$ that makes the given measurements most probable is the value of a for which the likelihood function is maximal. It is therefore not the most probable estimator.

Using Bayes' theorem we find $P(a;x) = P(x;a)\frac{P(a)}{P(x)}$. The prior probability $P(x)$ is fixed since the values x_i are given by our measurement and hence does not vary and the probability $P(a)$ is generally assumed to be a uniform distribution if there is no prior knowledge about a. It follows that $P(a;x)$ and $P(x;a)$ are proportional and hence the logarithms of these functions differ only by an additive constant. This leads us to following method for finding an estimator $\hat{a}$:

Method 41.6.20 (Maximum likelihood estimator). The maximum likelihood estimator $\hat{a}$ is obtained by solving following equation:

$$\left.\frac{d\ln\mathcal{L}}{da}\right|_{a=\hat{a}} = 0 \tag{41.90}$$

Remark 41.6.21. MLH estimators are mostly consistent but often biased.

Property 41.6.22. MLH estimators are invariant under parameter transformations.

Corollary 41.6.23. The invariance implies that the two estimators $\hat{a}$ and $\widehat{f(a)}$ cannot both be unbiased at the same time.

Property 41.6.24. Asymptotically ($N \to \infty$) every consistent estimator becomes unbiased and efficient.

41.6.6 Least squares

Method 41.6.25 (Least squares).

1. Fitting a function $y = f(x;a)$ to a set of 2 variables (x,y) where the x values are exact and the y valiues have an uncertainty σ_i to estimate the value a.
2. For every event (x_i, y_i) define the residual $d_i = y_i - f(x_i;a)$.
3. Determine (analytically) the χ^2 value:

$$\chi^2 = \sum_i \left(\frac{d_i}{\sigma_i}\right)^2 \tag{41.91}$$

4. Find the most propably value of $\hat{a}$ by solving the equation $\frac{d\chi^2}{da} = 0$.

Property 41.6.26. The optimal (minimal) χ^2 is distributed according to a χ^2 dsitribution 41.60 $P(\chi^2;n)$. The number of degrees of freedom n is equal to the number of events N minus the number of fitted parameters k.

Formula 41.6.27 (Linear fit). When all uncertainties σ_i are equal, the slope $\hat{m}$ and intercept $\hat{c}$ are given by following formulas:

$$\hat{m} = \frac{\overline{xy} - \overline{x}\ \overline{y}}{\overline{x^2} - \overline{x}^2} = \frac{\mathrm{cov}(x,y)}{V(x)} \tag{41.92}$$

$$\hat{c} = \overline{y} - \hat{m}\overline{x} = \frac{\overline{x^2} - \overline{x}\ \overline{y}}{\overline{x^2} - \overline{x}^2} \tag{41.93}$$

Remark 41.6.28. The equation $\overline{y} = \hat{c} + \hat{m}\overline{x}$ means that the linear fit passes through the center of mass $(\overline{x}, \overline{y})$.

Formula 41.6.29 (Errors of linear fit).

$$V(\hat{m}) = \frac{1}{N(\overline{x^2} - \overline{x}^2)}\sigma^2 \tag{41.94}$$

$$V(\hat{c}) = \frac{\overline{x^2}}{N(\overline{x^2} - \overline{x}^2)}\sigma^2 \tag{41.95}$$

$$\mathrm{cov}(\hat{m}, \hat{c}) = \frac{-\overline{x}}{N(\overline{x^2} - \overline{x}^2)}\sigma^2 \tag{41.96}$$

Remark 41.6.30. When the uncertainties σ_i are different, the arithmetic means have to be replaced by weighted means and the quantity σ^2 has to be replaced by its weighted variant:

$$\overline{\sigma^2} = \frac{\sum \sigma_i^2/\sigma_i^2}{\sum \sigma_i^{-2}} = \frac{N}{\sum \sigma_i^{-2}} \tag{41.97}$$

41.6.7 Binned least squares

The least squares method is very useful to fit data which has been grouped in bins (histograms):

Method 41.6.31 (Binned least squares).

1. N events with distributions $P(X; a)$ divided in N_B intervals. Interval j is centered on the value x_j, has a width W_j and contains n_j events.
2. The ideally expected number of events in the j^{th} interval: $f_j = NW_jP(x_j; a)$
3. The real number of events has a Poisson distribution: $\overline{n}_j = \sigma_j^2 = f_j$
4. Define the binned χ^2 as: $\chi^2 = \sum_i^{N_B} \frac{(n_i - f_i)^2}{f_i^2}$

41.7 Confidence

The real value of a parameter ε can never be known exactly. But it is possible to construct an interval I in which the real value should lie with a certain confidence C.

Example 41.7.1. Let X be a random variable with distribution $\mathcal{G}(x; \mu, \sigma)$. The measurement x lies in the interval $[\mu - 2\sigma; \mu + 2\sigma]$ with 95% **chance**. The real value μ lies in the interval $[x - 2\sigma; x + 2\sigma]$ with 95% **confidence**.

Remark. In the previous example there are some Bayesian assumptions: all possible values (left or right side of peak) are given the same possiblity due to the Gaussian distribution, but if one removes the symmetry it is mandatory to use a more careful approach. The symmetry between uncertainties σ and confidence levels is only valid for Gaussian distributions.

41.7.1 Interval types

Definition 41.7.2 (Two-sided confidence interval).

$$\boxed{P(x_- \leq X \leq x_+) = \int_{x_-}^{x_+} P(x)dx = C} \tag{41.98}$$

There are three possible (often used) two-sided intervals:

- Symmetric interval: $x_+ - \mu = \mu - x_-$
- Shortest interval: $|x_+ - x_-|$ is minimal
- Central interval: $\int_{-\infty}^{x_-} P(x)dx = \int_{x_+}^{\infty} P(x)dx = \frac{1-C}{2}$

The central interval is the (best and) most widely used confidence interval.

Remark 41.7.3. For Gaussian distributions these three definitions are equivalent.

Definition 41.7.4 (One-sided confidence interval).

$$\boxed{P(x \geq x_-) = \int_{x_-}^{+\infty} P(x)dx = C} \tag{41.99}$$

$$\boxed{P(x \leq x_+) = \int_{-\infty}^{x_+} P(x)dx = C} \tag{41.100}$$

Remark 41.7.5. For a discrete distribution it is often impossible to find integers $x_\pm$ such that the real value lies with exact confidence C in the interval $[x_-; x_+]$.

Definition 41.7.6 (Discrete central confidence interval).

$$x_- = \max_\theta \left[\sum_{x=0}^{\theta-1} P(x;X) \right] \leq \frac{1-C}{2} \tag{41.101}$$

$$x_+ = \min_\theta \left[\sum_{x=\theta+1}^{+\infty} P(x;X) \right] \leq \frac{1-C}{2} \tag{41.102}$$

41.7.2 General construction

For every value of the true parameter X it is possible to construct a confidence interval. This leads to the construction of two functions $x_-(X)$ and $x_+(X)$. The 2D diagram obtained by plotting $x_-(X)$ and $x_+(X)$ with the x-axis horizontally and X-axis vertically is called the **confidence region**.

Method 41.7.7. Let x_0 be a measurement of the parameter X. From the confidence region it is possible to infere a confidence interval $[X_-(x); X_+(x)]$, where the upper limit X_+ is not the limit such that there is only a $\frac{1-C}{2}$ chance of having a true parameter $X \geq X_+$, but the limit such that if the true parameter $X \geq X_+$ then there is a chance of $\frac{1-C}{2}$ to have a measurement x_0 or smaller.

41.7.3 Extra conditions

Method 41.7.8 (Bayesian statistics).

$$p(\text{theory}|\text{result}) = p(\text{result}|\text{theory})\frac{p(\text{theory})}{p(\text{result})} \tag{41.103}$$

or more mathematically:

$$\boxed{p(X|x) = p(x|X)\frac{p(X)}{p(x)}} \tag{41.104}$$

where we remark that:

- The denominator $p(\text{result})$ is only a normalization constant.
- The probability $p(x|X)$ to obtain a measurement x when the true parameter is X is a Gaussian distribution $\mathcal{G}(x; X, \sigma)$

Remark. If nothing is known about the theory, we often assume that $p(X)$ is a uniform probability 41.41.

41.7.4 Interval for a sample mean

Formula 41.7.9 (Interval with known variance). If the sample size is large enough, the real distribution is unimportant, because the CLT ensures a Gaussian distribution of the sample mean $\overline{X}$. The α-level confidence interval such that $P(-z_{\alpha/2} < Z < z_{\alpha/2})$ with $Z = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{N}}$ is given by:

$$\left[\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{N}}; \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{N}}\right] \tag{41.105}$$

Remark 41.7.10. If the sample size is not sufficiently large, the measured quantity must follow a normal distribution.

Formula 41.7.11 (Interval with unknown variance). To account for the uncertainty of the estimated standard deviation $\hat{\sigma}$, the student-t distribution 41.61 is used instead of a Gaussian distribution to describe the sample mean $\overline{X}$. The α-level confidence interval is given by:

$$\left[\overline{X} - t_{\alpha/2;(n-1)}\frac{s}{\sqrt{N}}; \overline{X} + t_{\alpha/2;(n-1)}\frac{s}{\sqrt{N}}\right] \tag{41.106}$$

where s is the estimated standard deviation 41.80.

Formula 41.7.12 (Wilson score interval). For a sufficiently large sample, a sample proportion $\hat{P}$ is approximately Gaussian distributed with expectation value π and variance $\frac{\pi(\pi-1)}{N}$. The α-level confidence interval is given by:

$$\left[\frac{(2N\hat{P} + z_{\alpha/2}^2) - z_{\alpha/2}\sqrt{z_{\alpha/2}^2 + 4N\hat{P}(1-\hat{P})}}{2(N + z_{\alpha/2}^2)}; \frac{(2N\hat{P} + z_{\alpha/2}^2) + z_{\alpha/2}\sqrt{z_{\alpha/2}^2 + 4N\hat{P}(1-\hat{P})}}{2(N + z_{\alpha/2}^2)}\right] \tag{41.107}$$

Remark. The expectation value and variance are these of a binomial distribution 41.53 with $r = X/N$.

41.8 Hypotheses and testing

41.8.1 Hypothesis

Definition 41.8.1 (Simple hypothesis). A hypothesis is called simple if the distribution is fully specified.

Definition 41.8.2 (Composite hypothesis). A hypothesis is called composite if the distribution is given relative to some parameter values.

41.8.2 Testing

Definition 41.8.3 (Type I error). Rejecting a true null hypothesis.

Definition 41.8.4 (Type II error). Accepting/retaining a false null hypothesis.

Definition 41.8.5 (Significance). The probability of making a type I error:

$$\boxed{\alpha = \int P_{H_0}(x)dx} \tag{41.108}$$

Property 41.8.6. Let $a_1 > a_2$. An a_2-level test is also significant at the a_1-level.

Remark 41.8.7. For discrete distributions it is not always possible to achieve an exact level of significance.

Remark. Type I errors occur occasionally. They cannot be prevented, they should however be controlled.

Definition 41.8.8 (Power). The probability of not making a type II error:

$$\boxed{\beta = \int P_{H_1}(x)dx \qquad \rightarrow \qquad \text{power: } 1-\beta} \tag{41.109}$$

Remark 41.8.9. A good test is a test with a small significance and a large power. The propabilities $P_{H_0}(x)$ and $P_{H_1}(x)$ should be as different as possible.

Theorem 41.8.10 (Neyman-Pearson test). *The following test is the most powerful test at significance level α for a threshold η:*

The null hypothesis H_0 is rejected in favour of the alternative hypothesis H_1 if the likelihood ratio Λ satisfies the following condition:

$$\Lambda(x) = \frac{L(x|H_0)}{L(x|H_1)} \leq \eta \tag{41.110}$$

where $P(\Lambda(x) \leq \eta|H_0) = \alpha$

Remark. In some references the reciprocal of $\Lambda(x)$ is used as the definition of the likelihood ratio.

41.9 Goodness of fit

Let $f(x|\vec{\theta})$ be the fitted function with N measurements.

41.9.1 χ^2-test

Formula 41.9.1.

$$\chi^2 = \sum_{i=1}^{N} \frac{[y_i - f(x_i)]^2}{\sigma_i^2} \tag{41.111}$$

Property 41.9.2. If there are $N - n$ fitted parameters we have:

$$\int_{\chi^2}^{\infty} f(\chi^2|n)d\chi^2 \approx 1 \implies \begin{cases} \circ \text{ good fit} \\ \circ \text{ errors were overestimated} \\ \circ \text{ selected measurements} \\ \circ \text{ lucky shot} \end{cases} \tag{41.112}$$

Property 41.9.3 (Reduced chi-squared χ^2_{red}). Define the reduced chi-squared value as follows: $\chi^2_{\text{red}} = \chi^2/n$ where n is the number of degrees of freedom.

- $\chi^2_{\text{red}} >> 1$: Poor modelling.
- $\chi^2_{\text{red}} > 1$: Bad modelling or underestimation of the uncertainties.
- $\chi^2_{\text{red}} = 1$: Good fit.
- $\chi^2_{\text{red}} < 1$: Impropable, overestimation of the uncertainties.

41.9.2 Runs test

A good χ^2-test does not mean that the fit is good. As mentioned in property 41.112 it is possible that the errors were overestimated. Another condition for a good fit is that the data points vary around the fit, i.e. there are no long sequences of points that lie above/underneath the fit. (It is a result of the 'randomness' of a data sample') This condition is tested with a runs test 41.113/41.114.

Remark 41.9.4. The χ^2-test and runs test are complementary. The χ^2-test only takes the absolute value of the differences between the fit and data points into account, the runs test only takes the signs of the differences into account.

Formula 41.9.5 (Runs distribution).

$$\boxed{P(r_{even}) = 2\frac{C^{N_B-1}_{\frac{r}{2}-1} C^{N_O-1}_{\frac{r}{2}-1}}{C^N_{N_B}}} \tag{41.113}$$

$$\boxed{P(r_{odd}) = \frac{C^{N_B-1}_{\frac{r-3}{2}} C^{N_O-1}_{\frac{r-1}{2}} + C^{N_O-1}_{\frac{r-3}{2}} C^{N_B-1}_{\frac{r-1}{2}}}{C^N_{N_B}}} \tag{41.114}$$

$$E(r) = 1 + 2\frac{N_B N_O}{N} \tag{41.115}$$

$$V(r) = 2\frac{N_B N_O}{N}\frac{2N_B N_O - 1}{N(N-1)} \tag{41.116}$$

Remark 41.9.6. For $r > 10 - 15$ the runs distribution approximates a Gaussian distribution.

41.9.3 Kolmogorov test

Definition 41.9.7 (Empirical distribution function).

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{]-\infty,x]}(x_i) \tag{41.117}$$

where $\mathbb{1}_A(x)$ is the indicator function 14.23.

Definition 41.9.8 (Kolmogorov-Smirnov statistic). Let $F(x)$ be a given cumulative distribution function. The n^{th} Kolmogorov-Smirnov statistic is defined as:

$$D_n = \sup_x |F_n(x) - F(x)| \tag{41.118}$$

Definition 41.9.9 (Kolmogorov distribution).

$$P(K \leq x) = 1 - 2\sum_{i=1}^{+\infty}(-1)^{i-1}e^{-2i^2x^2} = \frac{\sqrt{2\pi}}{x}\sum_{i=1}^{+\infty}e^{-(2i-1)^2\pi^2/(8x^2)} \tag{41.119}$$

Property 41.9.10 (Kolmogorov-Smirnov test). Let the null hypothesis H_0 state that a given data sample is described by a continuous distribution $P(x)$ with cumulative distribution function $F(x)$. The null hypothesis is rejected at significance level α if:

$$D_n\sqrt{n} > K_\alpha \tag{41.120}$$

where K_α is defined by using the Kolmogorov distribution: $P(K \leq K_\alpha) = 1 - \alpha$

Part VIII

Classical Mechanics

Chapter 42

Equations of Motion

42.1 General quantities

42.1.1 Linear quantities

Formula 42.1.1 (Force).

$$\boxed{\vec{\boldsymbol{F}} = \frac{d\vec{\boldsymbol{p}}}{dt}} \tag{42.1}$$

Remark. In classical mechanics, this formula is given by Newton's second law.

Formula 42.1.2 (Work).

$$W = \int \vec{\boldsymbol{F}} \cdot d\vec{\boldsymbol{l}} \tag{42.2}$$

Definition 42.1.3 (Conservative force). If the work done by a force is independent of the path taken, the force is said to be **conservative**.

$$\oint_C \vec{\boldsymbol{F}} \cdot d\vec{\boldsymbol{l}} = 0 \tag{42.3}$$

Stokes' theorem 20.22 together with relation 20.15 lets us rewrite the conservative force as the gradient of a scalar field:

$$\boxed{\vec{\boldsymbol{F}} = -\nabla V} \tag{42.4}$$

Formula 42.1.4 (Kinetic energy).

$$E_{kin} = \frac{p^2}{2m} \tag{42.5}$$

42.1.2 Angular quantities

Formula 42.1.5 (Angular velocity).

$$\omega = \frac{v}{r} \tag{42.6}$$

Formula 42.1.6 (Angular frequency).

$$\nu = \frac{\omega}{2\pi} \tag{42.7}$$

Formula 42.1.7 (Moment of inertia). For a symmetric object the moment of inertia is given by:

$$I = \int_V r^2 \rho(r) dV \tag{42.8}$$

For a general body we can define the moment of inertia tensor:

$$\boxed{\mathcal{I} = \int_V \rho(\vec{\boldsymbol{r}}) \left(r^2 \mathbb{1} - \vec{\boldsymbol{r}} \otimes \vec{\boldsymbol{r}}\right) dV} \tag{42.9}$$

Definition 42.1.8 (Principal axes of inertia). Let $[I]$ be the matrix of inertia[1]. This is a real symmetric matrix, which means that it admits an eigendecomposition[2] of the form:

$$[I] = [Q][\Lambda][Q]^T \tag{42.10}$$

The columns of $[Q]$ are called the principal axes of inertia. The eigenvalues are called the **principal moments of inertia**.

Example 42.1.9 (Objects with azimuthal symmetry†). Let r denote the radius of the object.

- Solid disk: $I = \frac{1}{2}mr^2$
- Cylindrical shell: $I = mr^2$
- Hollow sphere: $I = \frac{2}{3}mr^2$
- Solid sphere: $I = \frac{2}{5}mr^2$

Theorem 42.1.10 (Parallel axis theorem). *Consider a rotation about an axis ω through a point A. Let ω_{CM} be a parallel axis through the center of mass. The moment of inertia about ω is related to the moment of inertia about ω_{CM} in the following way:*

$$\boxed{I_A = I_{CM} + M||\vec{\boldsymbol{r}}_A - \vec{\boldsymbol{r}}_{CM}||^2} \tag{42.11}$$

where M is the mass of the rotating body.

Formula 42.1.11 (Angular momentum).

$$\boxed{\vec{\boldsymbol{L}} = \vec{\boldsymbol{r}} \times \vec{\boldsymbol{p}}} \tag{42.12}$$

Given the angular velocity vector we can compute the angular momentum as follows:

$$\vec{\boldsymbol{L}} = \mathcal{I}(\vec{\boldsymbol{\omega}}) \tag{42.13}$$

where $\mathcal{I}$ is the moment of inertia tensor. If $\vec{\boldsymbol{\omega}}$ is parallel to a principal axis, then the formula reduces to:

$$\vec{\boldsymbol{L}} = I\vec{\boldsymbol{\omega}} \tag{42.14}$$

Formula 42.1.12 (Torque).

$$\vec{\boldsymbol{\tau}} = \frac{d\vec{\boldsymbol{L}}}{dt} \tag{42.15}$$

For constant bodies, this formula can be rewritten as follows:

$$\vec{\boldsymbol{\tau}} = I\vec{\boldsymbol{\alpha}} = \vec{\boldsymbol{r}} \times \vec{\boldsymbol{F}} \tag{42.16}$$

[1] The matrix associated with the inertia tensor 42.9.
[2] See 19.6.16.

Remark 42.1.13. From the previous definitions it follows that both the angular momentum and torque vectors are in fact pseudo-vectors and thus change sign under coordinate transforms with $\det = -1$.

Formula 42.1.14 (Rotational energy).

$$E_{\text{rot}} = \frac{1}{2} I \omega^2 \tag{42.17}$$

42.2 Central force

Definition 42.2.1 (Central force). A central force is a force that only depends on the relative position of two objects:

$$\vec{\boldsymbol{F}}_c \equiv F\left(||\vec{\boldsymbol{r}}_2 - \vec{\boldsymbol{r}}_1|\,|\right)\hat{\boldsymbol{e}}_{\boldsymbol{r}} \tag{42.18}$$

42.3 Kepler problem

Formula 42.3.1 (Potential for a point mass).

$$\boxed{V = -G\frac{M}{r}} \tag{42.19}$$

where $G = 6.67 \times 10^{-11} \frac{Nm^2}{\text{kg}^2}$ is the **gravitational constant**.

42.4 Harmonic oscillator

Formula 42.4.1 (Harmonic potential).

$$\boxed{V = \frac{1}{2} k x^2} \tag{42.20}$$

or

$$V = \frac{1}{2} m \omega^2 x^2 \tag{42.21}$$

where we have set $\omega = \sqrt{\frac{k}{m}}$.

Formula 42.4.2 (Solution).

$$x(t) = A \sin \omega t + B \cos \omega t \tag{42.22}$$

$$= Ce^{i\omega t} + De^{-i\omega t} \tag{42.23}$$

Chapter 43

Lagrangian and Hamiltonian Mechanics

Definition 43.0.1 (Generalized coordinates). The generalized coordinates q_k are independent coordinates that completely describe the current configuration of a system relative to a reference configuration.

When a system has N degrees of freedom and n_c constraints, there are $(N - n_c)$ generalized coordinates. Furthermore, every set of generalized coordinates, describing the same system, should contain exactly $(N - n_c)$ coordinates.

Definition 43.0.2 (Generalized velocities). The generalized velocities $\dot{q}_k$ are the derivatives of the generalized coordinates with respect to time.

Notation 43.0.3.

$$L\left(\vec{\boldsymbol{q}}(t), \dot{\vec{\boldsymbol{q}}}(t), t\right) \equiv L\left(q_1(t), ..., q_n(t), \dot{q}_1(t), ..., \dot{q}_n(t), t\right) \tag{43.1}$$

Definition 43.0.4 (Action).

$$\boxed{S = \int_{t_1}^{t_2} L\left(\vec{\boldsymbol{q}}(t), \dot{\vec{\boldsymbol{q}}}(t), t\right) dt} \tag{43.2}$$

43.1 Euler-Lagrange equations†

Formula 43.1.1 (Euler-Lagrange equation of the first kind).

$$\boxed{\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}^k}\right) - \frac{\partial T}{\partial q^k} = Q_k} \tag{43.3}$$

where T is the total kinetic energy.

Formula 43.1.2 (Euler-Lagrange equation of the second kind).

$$\boxed{\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\right) - \frac{\partial L}{\partial q^k} = 0} \tag{43.4}$$

43.2 Conservation laws and symmetry properties

Definition 43.2.1 (Conjugate momentum). Also called the **canonically conjugate momentum**.

$$p_k = \frac{\partial L}{\partial \dot{q}^k} \tag{43.5}$$

Definition 43.2.2 (Cyclic coordinate). If the lagrangian L does not explicitly depend on a coordinate q_k, the coordinate is called a cyclic coordinate.

Property 43.2.3. The conjugate momentum of a cyclic coordinate is a conserved quantity.

$$\dot{p}_k \overset{43.5}{=} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\right) \overset{43.4}{=} \frac{\partial L}{\partial q^k} \overset{\text{cyclic coord.}}{=} 0 \tag{43.6}$$

43.3 Noether's theorem

Theorem 43.3.1 (Noether's theorem†). *Consider a field transformation*

$$\phi(x) \to \phi(x) + \alpha\delta\phi(x) \tag{43.7}$$

where α is an infinitesimal quantity and $\delta\phi$ is a small deformation. In case of a symmetry we obtain the following conservation law:

$$\partial_\mu\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi - \mathcal{J}^\mu\right) = 0 \tag{43.8}$$

The factor between parentheses can be interpreted as a conserved current $j^\mu(x)$. Noether's theorem states that every symmetry of the form 43.7 leads to such a current.

The conservation can also be expressed in terms of a charge[1]*:*

$$\frac{dQ}{dt} = \frac{d}{dt}\int j^0 d^3x = 0 \tag{43.9}$$

Definition 43.3.2 (Stress-energy tensor). Consider a field transformation

$$\phi(x) \to \phi(x+a) = \phi(x) + a^\mu\partial_\mu\phi(x)$$

Because the Lagrangian is a scalar it transforms similarly:

$$\mathcal{L} \to \mathcal{L} + a^\mu\partial_\mu\mathcal{L} = \mathcal{L} + a^\nu\partial_\mu(\delta^\mu{}_\nu\mathcal{L}) \tag{43.10}$$

This leads to the existence of 4 conserved currents. These can be used to define the stress-energy tensor:

$$\boxed{T^\mu{}_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\nu\phi - \mathcal{L}\delta^\mu{}_\nu} \tag{43.11}$$

[1]The conserved current and its associated charge are called the **Noether current** and **Noether charge**.

43.4 Hamilton's equations

Definition 43.4.1 (Canonical coordinates). Consider the generalized coordinates $(q, \dot{p}, t)$ from the Lagrangian formalism. Using these we can define a new set of coordinates, called canonical coordinates, by exchanging the time-derivatives $\dot{q}^i$ in favour of the conjugate momenta p_i (see definition 43.5) and leaving the coordinates q^i and t invariant.

Definition 43.4.2 (Hamiltonian function). The (classical) Hamiltonian function is defined as follows:

$$H(q, p, t) = \sum_i p_i \dot{q}^i - L(q, p, t) \tag{43.12}$$

Formula 43.4.3 (Hamilton's equations[2]).

$$\boxed{\begin{aligned} \dot{q}^i &= \frac{\partial H}{\partial p_i} \qquad (43.13)\\ \dot{p}_i &= -\frac{\partial H}{\partial q^i} \qquad (43.14)\\ \frac{\partial L}{\partial t} &= -\frac{\partial H}{\partial t} \qquad (43.15) \end{aligned}}$$

The formula to obtain the Hamiltonian from the Lagrangian is an application of the following more general **Legendre transformation**:

Definition 43.4.4 (Legendre transformation). Consider an equation of the following form:

$$df = udx + vdy \tag{43.16}$$

where $u = \frac{\partial f}{\partial x}$ and $v = \frac{\partial f}{\partial y}$.

Suppose we want to perform a coordinate transformation $(x, y) \to (u, y)$ while preserving the general form of 43.16 for differential quantities. To do this we consider the function

$$g = f - ux \tag{43.17}$$

Differentiating gives

$$\begin{aligned} dg &= df - udx - xdu \\ &= (udx + vdy) - udx - xdu \\ &= vdy - xdu \end{aligned}$$

which has the form of 43.16 as desired. The quantities v and x are now given by

$$x = -\frac{\partial g}{\partial u} \qquad \text{and} \qquad v = \frac{\partial g}{\partial y} \tag{43.18}$$

The transition $f \longrightarrow g$ defined by equations 43.16 and 43.17 is called a Legendre transformation.

Remark 43.4.5. Although the previous derivation used only 2 coordinates, the definition of Legendre transformations can easily be generalized to more coordinates.

Definition 43.4.6 (Poisson bracket).

$$\boxed{\{A, B\} = \frac{\partial A}{\partial p}\frac{\partial B}{\partial q} - \frac{\partial B}{\partial p}\frac{\partial A}{\partial q}} \tag{43.19}$$

where q, p are the generalized coordinates in the Hamiltonian formalism.

[2] Also known as the *canonical equations of Hamilton*.

Formula 43.4.7 (Total time derivative).

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + \{H, F\} \tag{43.20}$$

where $\{\cdot, \cdot\}$ is the Poisson bracket as defined above and H is the Hamiltonian 43.12.

43.5 Hamilton-Jacobi equation

43.5.1 Canonical transformations

Definition 43.5.1 (Canonical transformations). A canonical transformation is a transformation that leaves the Hamiltonian equations of motion unchanged. Mathematically this means that the transformations leave the action invariant up to a constant, or equivalently, they leave the Lagrangian invariant up to a complete time-derivative:

$$\sum_i \dot{q}^i p_i - H(q, p, t) = \sum_i \dot{Q}^i P_i - K(Q, P, t) - \frac{dG}{dt}(Q, P, t) \tag{43.21}$$

The function G is called the generating function of the canonical transformation. The choice of G uniquely determines the transformation.

Formula 43.5.2 (Hamilton-Jacobi equation). Sufficient conditions for the generating function S are given by:

$$P_i = \frac{\partial S}{\partial Q^i}$$
$$Q^i = \frac{\partial S}{\partial P_i}$$

and

$$K = H + \frac{\partial S}{\partial t}$$

Choosing the new Hamiltonian function K to be 0 gives the Hamilton-Jacobi equation:

$$\boxed{H\left(q, \frac{\partial S}{\partial q}\right) + \frac{\partial S}{\partial t} = 0} \tag{43.22}$$

The function S is called **Hamilton's principal function**.

Property 43.5.3. The new coordinates P_i and Q^i are all constants of motion. This follows immediately from the choice $K = 0$.

Definition 43.5.4 (Hamilton's characteristic function). If the system is time-independent it follows from the HJE that the principal function is of the form

$$S(q, p, t) = W(q, p) - Et \tag{43.23}$$

where E is a constant. The time-independent function W is called Hamilton's characteristic function.

Substituting this result in the HJE results in

$$H\left(q, \frac{\partial S}{\partial q}\right) = E \tag{43.24}$$

In time-independent systems the Hamiltonian function is thus a constant of motion and we call it the **energy** of the system.

43.5.2 Stäckel potential

Remark 43.5.5. If the principal function can be separated into n equations, the HJE splits up into n equations of the form

$$h_i\left(q^i, \frac{dS}{dq^i}, \alpha_i\right) = 0 \tag{43.25}$$

The partial differential equation for S can thus be rewritten as a system of n ordinary differential equations.

Theorem 43.5.6 (Stäckel condition). *Using an orthogonal coordinate system, the Hamilton-Jacobi equation is separable if and only if the potential is of the following form:*

$$V(q) = \sum_{i=1}^{n} \frac{1}{G_i^2(q)} W_i(q^i) \tag{43.26}$$

whenever the Hamiltonian function can be written as

$$H(q,p) = \frac{1}{2}\sum_i \frac{p_i^2}{G_i^2(q)} + V(q) \tag{43.27}$$

These potentials are called ***Stäckel potentials****.*

Chapter 44

Phase Space

44.1 Phase space

Definition 44.1.1 (Phase space). The set of all possible n-tuples[1] (q^i, p_i) of generalized coordinates and associated momenta is called the phase space of the system.

Definition 44.1.2 (Rotation). A rotation is the change of a coordinate for which every possible value is allowed.

Definition 44.1.3 (Libration). A libration is the change of coordinate for which only a subset of the total range is allowed. It is the generalization of an oscillation.

44.2 Material derivative

Definition 44.2.1 (Lagrangian derivative[2]). Let $a(\vec{\boldsymbol{r}}, \vec{\boldsymbol{v}}, t)$ be a property of a system defined at every point of the system. The Lagrangian derivative along a path $(\vec{\boldsymbol{r}}(t), \vec{\boldsymbol{v}}(t))$ in phase space is given by:

$$\begin{aligned}\frac{Da}{Dt} &= \lim_{\Delta t \to 0} \frac{a(\vec{\boldsymbol{r}} + \Delta\vec{\boldsymbol{r}}, \vec{\boldsymbol{v}} + \Delta\vec{\boldsymbol{v}}, t + \Delta t) - a(\vec{\boldsymbol{r}}, \vec{\boldsymbol{v}}, t)}{\Delta t} \\ &= \frac{\partial a}{\partial t} + \frac{d\vec{\boldsymbol{r}}}{dt} \cdot \frac{\partial a}{\partial \vec{\boldsymbol{r}}} + \frac{d\vec{\boldsymbol{v}}}{dt} \cdot \frac{\partial a}{\partial \vec{\boldsymbol{v}}} \\ &= \frac{\partial a}{\partial t} + \vec{\boldsymbol{v}} \cdot \nabla a + \frac{d\vec{\boldsymbol{v}}}{dt} \cdot \frac{\partial a}{\partial \vec{\boldsymbol{v}}} \end{aligned} \tag{44.1}$$

The second term $\vec{\boldsymbol{v}} \cdot \nabla a$ in this equation is called the **advective** term.

Remark 44.2.2. In the case that $a(\vec{\boldsymbol{r}}, \vec{\boldsymbol{v}}, t)$ is a tensor field the gradient ∇ has to be replaced by the covariant derivative. The advective term is then called the **convective** term.

Corollary 44.2.3. If we take $a(\vec{\boldsymbol{r}}, \vec{\boldsymbol{v}}, t) = \vec{\boldsymbol{r}}$ we obtain:

$$\frac{D\vec{\boldsymbol{r}}}{Dt} = \vec{\boldsymbol{v}} \tag{44.2}$$

[1] Not only those as given by the equations of motion.

[2] Also known as the **material derivative**, especially when applied to fluidum mechanics.

44.3 Liouville's theorem

Formula 44.3.1 (Liouville's lemma). Consider a phase space volume element dV_0 moving along a path $(\vec{\boldsymbol{r}}(t), \vec{\boldsymbol{v}}(t)) \equiv (\vec{\boldsymbol{x}}(t))$. The Jacobian $J(\vec{\boldsymbol{x}}, t)$ associated with this motion is given by:

$$J(\vec{\boldsymbol{x}}, t) = \frac{dV}{dV_0} = \det\left(\frac{\partial \vec{\boldsymbol{x}}}{\partial \vec{\boldsymbol{x}}_0}\right) = \sum_{ijklmn} \varepsilon_{ijklmn} \frac{\partial x^1}{\partial x_0^i} \frac{\partial x^2}{\partial x_0^j} \frac{\partial x^3}{\partial x_0^k} \frac{\partial x^4}{\partial x_0^l} \frac{\partial x^5}{\partial x_0^m} \frac{\partial x^6}{\partial x_0^n} \tag{44.3}$$

The Lagrangian derivative of this Jacobian then becomes:

$$\frac{DJ}{Dt} = (\nabla \cdot \vec{\boldsymbol{x}})J \tag{44.4}$$

Furthermore using the Hamiltonian equations 43.13 it is easy to prove that

$$\nabla \cdot \vec{\boldsymbol{x}} = 0 \tag{44.5}$$

Theorem 44.3.2 (Liouville's theorem). *Let $V(t)$ be a phase space volume containing a fixed set of particles. Application of Liouville's lemma gives:*

$$\frac{DV}{Dt} = \frac{D}{Dt}\int_{\Omega(t)} d^6x = \frac{D}{Dt}\int_{\Omega_0} J(\vec{\boldsymbol{x}}, t) d^6x_0 = 0 \tag{44.6}$$

It follows that the phase space volume of a Hamiltonian system[3] is invariant with respect to time-evolution.

Formula 44.3.3 (Boltzmann's transport equation). Let $F(\vec{\boldsymbol{r}}, \vec{\boldsymbol{v}}, t)$ be the mass distribution function:

$$M_{tot} = \int_{\Omega(t)} F(\vec{\boldsymbol{x}}, t) d^6x \tag{44.7}$$

From the conservation of mass we can derive the following formula:

$$\frac{DF}{Dt} = \frac{\partial F}{\partial t} + \frac{d\vec{\boldsymbol{r}}}{dt} \cdot \frac{\partial F}{\partial \vec{\boldsymbol{r}}} - \nabla V \cdot \frac{\partial F}{\partial \vec{\boldsymbol{v}}} = \left[\frac{\partial F}{\partial t}\right]_{col} \tag{44.8}$$

where the right hand side gives the change of F due to collisions.[4] This formula is a partial differential equation in 7 variables which can be solved to obtain $F(\vec{\boldsymbol{x}}, t)$.

Consider a Hamiltonian system with a finite phase space $\mathcal{V}$. By Liouville's theorem, the phase flow generated by the equations of motion is a measure (volume) preserving map $g : \mathcal{V} \to \mathcal{V}$. This leads us to the following theorem:

Theorem 44.3.4 (Poincaré recurrence theorem). *Let $\mathcal{V}_0$ be the phase space volume of the system. For every point $x_0 \in \mathcal{V}_0$ and for every neighbourhood U of x_0 there exists a point $y \in U$ such that $g^n y \in U$ for every $n \in \mathbb{N}$.*

Theorem 44.3.5 (Strong Jeans theorem[5]). *The distribution function $F(\vec{\boldsymbol{r}}, \vec{\boldsymbol{v}})$ of a time-independent system for which almost all orbits are regular can be expressed in terms of 3 integrals of motion, called* **isolating integrals**.

[3] A system that satisfies Hamilton's equations of motion.

[4] The collisionless form of this equation is sometimes called the *Vlasov equation*.

[5] Actually due to Donald Lynden-Bell.

44.4 Continuity equation

Formula 44.4.1 (Reynolds transport theorem[6]). Consider a quantity

$$F = \int_{V(t)} f(\vec{\boldsymbol{r}}, \vec{\boldsymbol{v}}, t) dV$$

Using equation 44.4 and the divergence theorem 20.23 we can obtain:

$$\boxed{\frac{DF}{Dt} = \int_V \frac{\partial f}{\partial t} dV + \oint_S f\vec{\boldsymbol{v}} \cdot d\vec{\boldsymbol{S}}} \tag{44.9}$$

Formula 44.4.2 (Continuity equations). For a conserved quantity the equation above becomes:

$$\frac{Df}{Dt} + (\nabla \cdot \vec{\boldsymbol{v}}) f = 0 \tag{44.10}$$

$$\frac{\partial f}{\partial t} + \nabla \cdot (f\vec{\boldsymbol{v}}) = 0 \tag{44.11}$$

If we set $f = \rho$ (mass density) then the first equation is called the **Lagrangian continuity equation** and the second equation is called the **Eulerian continuity equation**. Both equations can be found by pulling the Lagrangian derivative inside the integral on the left-hand side of 44.9.

Corollary 44.4.3. Combining the Reynolds transport theorem with the Lagrangian continuity equation gives the following identity for an arbitrary function f:

$$\frac{D}{Dt} \int_V \rho f dV = \int_V \rho \frac{Df}{Dt} dV \tag{44.12}$$

44.5 Dynamical systems

This section gives a formal treatment of the concepts presented in chapter 43 and mainly uses notions and formulas from chapter 34.

Definition 44.5.1 (Dynamical system). Let (M, ω) be a symplectic manifold and let $H \in C^\infty(M)$ be a distinguished observable. The triple (M, ω, H) is called a dynamical system with Hamiltonian H. The time derivative of any observable $F \in C^\infty(M)$ is defined by:

$$\dot{F} = \{H, F\} \tag{44.13}$$

where $\{\cdot, \cdot\}$ is the Poisson bracket on M. The time evolution is then governed by the Hamiltonian flow $\exp(tX_H)$.

Definition 44.5.2 (Conserved quantity). Let (M, ω, H) be a dynamical system. An observable $F \in C^\infty(M)$ is said to be conserved if it satisfies $\{H, F\} = 0$.

Corollary 44.5.3 (Noether's theorem). Noether's theorem is now quite trivial.

Proof. Let Q be a conserved quantity, then:

$$\begin{aligned} &\{H, Q\} = 0 \\ \iff &\{Q, H\} = 0 \end{aligned} \tag{44.14}$$

[6] This is a 3D extension of the *Leibniz integral rule*.

□

From now on we consider a specific type of Hamiltonian function H. Let (Q, g) be a Riemannian manifold and let the cotangent bundle $M := T^*Q \xrightarrow{\pi} Q$ be a symplectic manifold. The Hamiltonian under consideration is of the form (in local Darboux coordinates):

$$H(q,p) = \frac{1}{2}g^{ij}(q)p_i p_j + V(q) \tag{44.15}$$

These Hamiltonians have two types of symmetries (conserved quantities):

Definition 44.5.4 (Kinematical). Consider a conserved quantity C. If $\pi_*(X_C) \in \Gamma(TQ)$ exists and $\mathcal{L}_{\pi_*(X_C)}g = 0$ then the symmetry is said to be kinematical.

Remark. The second condition says that $\pi_*(X_C)$ is a *Killing vector*.!!COMPLETE!!

Definition 44.5.5 (Dynamical). Any symmetry that is not a kinematical symmetry is said to be dynamical.

The following algorithm gives us a way to find conditions to check whether a given observable is conserved:

Method 44.5.6 (Van Holten's algorithm). Let the conserved quantity be analytic, i.e. $C(q,p) = \sum_{k=0}^{N} \frac{1}{k!}a^{(n_1\dots n_k)}(q)p_{n_1}\dots p_{n_k}$ for some $N \in \mathbb{N}$, where the brackets $(\cdots)$ around the indices denote symmetrization. For a flat manifold, i.e. g does not depend on q, we can rewrite $\{C, T+V\} = 0$ to obtain

$$\sum_{n=1}^{N}\left[\frac{1}{(k-1)!}a^{n_1\dots n_{k-1}i}p_{n_1}\dots p_{n_{k-1}}\frac{\partial V}{\partial q^i} - \frac{2}{k!}\frac{\partial}{\partial q^i}a^{n_1\dots n_k}p_{n_1}\dots p_{n_k}g^{im}p_m\right] = 0$$

Because two polynomials are equal if and only if their corresponding coefficients are equal we obtain following equations:

1. 0^{th} order: $a^k\dfrac{\partial V}{\partial q^k} = 0$
2. 1^{st} order: $a^{(n_1 i)}\dfrac{\partial V}{\partial q^i} - 2\dfrac{\partial a}{\partial q^i}g^{in_1} = 0$
3. N^{th} order: $\dfrac{1}{N!}a^{(n_1\dots n_N i)}\dfrac{\partial V}{\partial q^i} - \dfrac{2}{(N-1)!}\dfrac{\partial}{\partial q^i}a^{(n_1\dots n_{N-1}}g^{i)n_N} = 0$

where one should pay attention to the symmetrization brackets in the second term of the last equation. Pulling down the indices by multiplying with the metric $g_{n_i m_i}$ gives:

$$a_{(m_1\dots m_N)}{}^i\partial_i V - 2N\partial_{(m_N}a_{m_1\dots m_{N-1})} = 0 \tag{44.16}$$

The upper bound N in the series expansion is given by the generalized Killing condition[7]:

$$\partial_{(m_{N+1}}a_{m_1\dots m_N)} = 0 \implies a_{(m_1\dots m_{(}N+1))} = 0 \tag{44.17}$$

Remark 44.5.7. The above algorithm still holds for curved manifolds when replacing all partial derivatives ∂_i by (Levi-Civita) covariant derivatives ∇_i.

[7] See equation 33.11.

Chapter 45

Fluid Mechanics

45.1 Cauchy stress tensor

Theorem 45.1.1 (Cauchy's stress theorem[1]). *Knowing the stress vectors acting on the coordinate planes through a point A is sufficient to calculate the stress vector acting on an arbitrary plane passing through A.*

The *Cauchy stress theorem* is equivalent to the existence of the following tensor:

Definition 45.1.2 (Cauchy stress tensor). The Cauchy stress tensor is a $(0,2)$-tensor $\mathbf{T}$ that gives the relation between a stress vector associated to a plane and the normal vector $\vec{\boldsymbol{n}}$ to that plane:

$$\vec{\boldsymbol{t}}_{(\vec{\boldsymbol{n}})} = \mathbf{T}(\vec{\boldsymbol{n}}) \tag{45.1}$$

Example 45.1.3. For identical particles, the stress tensor is given by:

$$\mathbf{T} = -\rho\langle \vec{\boldsymbol{w}} \otimes \vec{\boldsymbol{w}} \rangle \tag{45.2}$$

where $\vec{\boldsymbol{w}}$ is the random component of the velocity vector and $\langle\cdot\rangle$ denotes the expectation value (see 40.19).

Theorem 45.1.4 (Cauchy's lemma). *The stress vectors acting on opposite planes are equal in magnitude but opposite in direction:*

$$\vec{\boldsymbol{t}}_{(-\vec{\boldsymbol{n}})} = -\vec{\boldsymbol{t}}_{(\vec{\boldsymbol{n}})} \tag{45.3}$$

Formula 45.1.5 (Cauchy momentum equation). From Newton's second law 42.1 it follows that:

$$\frac{D\vec{\boldsymbol{P}}}{Dt} = \int_V \vec{\boldsymbol{f}}(\vec{\boldsymbol{x}},t)dV + \oint_S \vec{\boldsymbol{t}}(\vec{\boldsymbol{x}},t)dS \tag{45.4}$$

where $\vec{\boldsymbol{P}}$ is the momentum density, $\vec{\boldsymbol{f}}$ are body forces and $\vec{\boldsymbol{t}}$ are surface forces (such as shear stress). Using Cauchy's stress theorem and the divergence theorem 20.23 we get

$$\frac{D\vec{\boldsymbol{P}}}{Dt} = \int_V \left[\vec{\boldsymbol{f}}(\vec{\boldsymbol{x}},t) + \nabla\cdot\mathbf{T}(\vec{\boldsymbol{x}},t)\right] dV \tag{45.5}$$

The left-hand side can be rewritten using 44.12 as

$$\int_V \rho\frac{D\vec{\boldsymbol{v}}}{Dt}dV = \int_V \left[\vec{\boldsymbol{f}}(\vec{\boldsymbol{x}},t) + \nabla\cdot\mathbf{T}(\vec{\boldsymbol{x}},t)\right] dV \tag{45.6}$$

[1] Also known as **Cauchy's fundamental theorem**.

Chapter 46

Optics

46.1 General

46.1.1 Conservation of energy

From the law of conservation of energy we can derive the following formula:

$$\boxed{T + R + A = 1} \tag{46.1}$$

where

T : Transmission coefficient
R : Reflection coefficient
A : Absorption coefficient

46.1.2 Photon

Formula 46.1.1 (Energy).

$$E = h\nu = \hbar\omega = \frac{hc}{\lambda} \tag{46.2}$$

Formula 46.1.2 (Momentum).

$$p = \frac{h}{\lambda} = \hbar k \tag{46.3}$$

where formula 46.4 was used in the last step.

Remark. These formulas can also be (approximately) used for particles for which the rest mass (energy) is negligible.

46.2 Plane wave

Formula 46.2.1 (Wave number).

$$k = \frac{2\pi}{\lambda} \tag{46.4}$$

Formula 46.2.2 (Plane wave). Following equations represent a plane wave moving in the x-direction and polarized in the xy-plane:

$$\vec{\boldsymbol{E}}(x,t) = \mathrm{Re}\left\{A\,\exp\left[i\left(kx-\omega t+\phi\right)\right]\right\}\vec{\boldsymbol{e}}_y \tag{46.5}$$

$$\vec{\boldsymbol{E}}(x,t) = \mathrm{Re}\left\{A\,\exp\left[2\pi i\left(\frac{x}{\lambda}-\frac{t}{T}+\frac{\phi}{2\pi}\right)\right]\right\}\vec{\boldsymbol{e}}_y \tag{46.6}$$

46.3 Refraction

Formula 46.3.1 (Refraction).

$$v_2 = \frac{v_1}{n} \tag{46.7}$$

Formula 46.3.2 (Diëlectric function). In the case of non-magnetic materials ($\mu_r \approx 1$), we can write the diëlectric function as following:

$$\epsilon = \epsilon_r + i\epsilon_i = \widetilde{n}^2 = (n+ik)^2 \tag{46.8}$$

Where $\widetilde{n}$ is the **complex refractive index** and k is the **extinction coefficient**.

46.4 Absorption

Theorem 46.4.1 (Law of Lambert-Beer[†]).

$$\frac{I(x)}{I(0)} = exp\left(-\frac{4\pi\nu k}{c}x\right) \tag{46.9}$$

Definition 46.4.2 (Absorption coefficient). The constant factor in the Lambert-Beer law is called the absorption coefficient.

$$\alpha = \frac{4\pi\nu k}{c} \tag{46.10}$$

46.5 Diffraction

Chapter 47

Astronomy

47.1 Ellipsoidal coordinates

We start from folowing parametrized equation:

$$f(\tau) = \frac{x^2}{\tau+\alpha} + \frac{y^2}{\tau+\beta} + \frac{z^2}{\tau+\gamma} - 1 \tag{47.1}$$

where $\alpha < \beta < \gamma < 0$. By multiplying away the denominators and setting $f(\tau) = 0$ we obtain a polynomial equation of degree 3 in τ. This polynomial can be formally factorised as

$$-(\tau-\lambda)(\tau-\mu)(\tau-\nu) \tag{47.2}$$

such that the solutions (λ, μ, ν) obey following rules:

$$\begin{cases} \nu & \in &]-\gamma, -\beta[\\ \mu & \in &]-\beta, -\alpha[\\ \lambda & \in &]-\alpha, +\infty[\end{cases}$$

From previous two equations we can find a solution for x^2 by multiplying by $(\tau+\alpha)$ and letting $\tau \to -\alpha$. Solutions for y^2 and z^2 can be found in a similar way:

$$\begin{cases} x^2 &= \dfrac{(\lambda+\alpha)(\mu+\alpha)(\nu+\alpha)}{(\beta-\alpha)(\gamma-\alpha)} \\ y^2 &= \dfrac{(\lambda+\beta)(\mu+\beta)(\nu+\beta)}{(\beta-\alpha)(\beta-\gamma)} \\ z^2 &= \dfrac{(\lambda+\gamma)(\mu+\gamma)(\nu+\gamma)}{(\alpha-\gamma)(\beta-\gamma)} \end{cases} \tag{47.3}$$

For these solutions multiple cases can be considered. We can define different surfaces by fixing τ at different values.

47.1.1 Ellipsoid: $\tau = \lambda$

First we look at the surfaces defined by fixing $\tau = \lambda$ in equation 47.1. By noting that all denominators are positive in this case, we see that the obtained surface is an ellipsoid with the x-axis as the shortest axis. By letting $\lambda \to +\infty$ we obtain the equation of a sphere with radius $\sqrt{\lambda}$. If $\lambda \to -\alpha$ we get an ellipse in the yz-plane. This ellipse is called the **focal ellipse**.

47.1.2 One-sheet hyperboloid: $\tau = \mu$

By fixing $\tau = \mu$ in 47.1 we obtain the equation of one-sheet hyperboloid (also called a **hyperbolic hyperboloid**) around the x-axis. By letting $\mu \to -\alpha$ the hyperboloid collapses in the yz-plane and we obtain the surface outside the focal ellipse. If $\mu \to -\beta$ the hyperboloid becomes degenerate and we get the surface inside the **focal hyperbola** defined by

$$\frac{x^2}{\alpha - \beta} + \frac{z^2}{\gamma - \beta} = 1 \tag{47.4}$$

This hyperbola intersects the z-plane in the foci of the focal ellipse.

47.1.3 Two-sheet hyperboloid: $\tau = \nu$

By fixing $\tau = \nu$ in 47.1 we obtain the equation of two-sheet hyperboloid (also called an **elliptic hyperboloid**) around the z-axis. By letting $\nu \to -\beta$ the hyperboloid becomes degenerate and we obtain the surface outside the focal hyperbola 47.4. If $\nu \to -\gamma$ the two sheets coincide in the xy-plane.

47.1.4 Hamiltonian function

When wrting out the kinetic energy in ellipsoidal coordinates by applying the chain rule for differentiation to the Cartesian kinetic energy while noting that mixed terms of the form $\frac{\partial x^a}{\partial \lambda^i}\frac{\partial x^a}{\partial \lambda^j}$ cancel out when writing them out using 47.3 it is clear that the Hamiltonian function can be spearated as follows:

$$H = \frac{1}{2}\left(\frac{p_\lambda^2}{Q_\lambda^2} + \frac{p_\mu^2}{Q_\mu^2} + \frac{p_\nu^2}{Q_\nu^2}\right) + V \tag{47.5}$$

where $Q_j^2 = \sum_i \left(\frac{\partial x^i}{\partial \lambda^j}\right)^2$ are the metric coefficients in ellipsoidal coordinates.

These coefficients can be calculated by noting that $\frac{\partial x^i}{\partial \lambda} = \frac{1}{x^i}\frac{\partial (x^i)^2}{\partial \lambda}$ and putting $\frac{1}{(\lambda+\alpha)(\lambda+\beta)(\lambda+\gamma)}$ in the front. Furthermore the coefficient belonging to λ^2, μ^2, ν^2, mixed terms and others can be calculated easily. By doing so we obtain following result

$$Q_\lambda^2 = \frac{1}{4}\frac{(\lambda-\mu)(\lambda-\nu)}{(\lambda+\alpha)(\lambda+\beta)(\lambda+\gamma)} \tag{47.6}$$

which is also valid for μ and ν by applying cyclic permutation to the coordinates.

Following from the Stäckel conditions 43.26 the potential must be of the form

$$V = \sum_i \frac{W_i(\lambda^i)}{Q_i^2} \tag{47.7}$$

if we want to obtain a seperable Hamilton-Jacobi equation. Due to the disjunct nature of λ, μ and ν we can consider W_λ, W_μ and W_ν as three parts of a single function $G(\tau)$ given by:

$$G(\tau) = -4(\tau + \beta)W_\tau(\tau) \tag{47.8}$$

The 3D potential is thus completely determined by a 1D function $G(\tau)$.

47.1.5 Hamilton-Jacobi equation

If we consider a time-independent system we can use 43.24 as our starting point. If we multiply this equation by $(\lambda-\mu)(\lambda-\nu)(\mu-\nu)$ we obtain

$$(\mu-\nu)\left[2(\lambda+\alpha)(\lambda+\beta)(\lambda+\gamma)\left(\frac{dS^\lambda(\lambda)^2}{d\lambda}\right)\right.$$
$$\left.-(\lambda+\alpha)(\lambda+\gamma)G(\lambda)-\lambda^2E\right]+\text{cyclic permutations}=0 \quad (47.9)$$

where we rewrote the multiplication factor in the form $a\lambda^2+b\mu^2+c\nu^2$ before multiplying the RHS of 43.24. This equation can be elegantly rewritten as

$$(\mu-\nu)U(\lambda)+(\lambda-\mu)U(\nu)+(\nu-\lambda)U(\mu)=0 \quad (47.10)$$

Differentiating twice with respect to any λ^i gives $U''(\tau)=0$ or equivalently

$$U(\tau)=I_3-I_2\tau \quad (47.11)$$

where I_2 and I_3 are two new first integrals of motion.

From the Hamiltonian-Jacobi equations of motion one can calculate the conjugate momenta $p_\tau=\frac{dS^\tau}{d\tau}$. After a lengthy calculation we obtain

$$p_\tau^2=\frac{1}{2(\tau+\beta)}\left[E-V_{\text{eff}}(\tau)\right] \quad (47.12)$$

where the effective potential is given by

$$\boxed{V_{\text{eff}}=\frac{J}{\tau+\alpha}+\frac{K}{\tau+\gamma}-G(\tau)} \quad (47.13)$$

where J and K are two conserved quantities given by

$$J=\frac{\alpha^2E+\alpha I_2+I_3}{\alpha-\gamma} \quad \text{and} \quad K=\frac{\gamma^2E+\gamma I_2+I_3}{\gamma-\alpha}$$

To be physically acceptable, p_τ^2 should be positive. This leads to following conditions on the energy:

$$\begin{cases}E\geq V_{\text{eff}}(\lambda)\\ E\geq V_{\text{eff}}(\mu)\\ E\leq V_{\text{eff}}(\nu)\end{cases} \quad (47.14)$$

The generating $G(\tau)$ function should also satisfy some conditions. First we note that we can rewrite our Stäckel potential $V(\lambda,\mu,\nu)$ as

$$V=-\frac{1}{\lambda-\nu}\left(\frac{F(\lambda)-F(\mu)}{\lambda-\mu}-\frac{F(\mu)-F(\nu)}{\mu-\nu}\right)\leq 0 \quad (47.15)$$

where $F(\tau)=(\tau+\alpha)(\tau+\gamma)G(\tau)$.

For $\lambda\to+\infty$ (or $r^2\to+\infty$) we get $V\approx-\frac{F(\lambda)}{\lambda^2}\approx-G(\lambda)$. Because $V\sim\lambda^{-1}$ it is clear that $G(\tau)$ cannot decay faster than $\lambda^{-1/2}$ at infity. Furthermore we can interpret 47.15 as an approximation of $-F''(\tau)$. So it follows that $F(\tau)$ should be convex. For $\tau\to-\gamma$ we get

$$\begin{cases}\alpha+\tau<0\\ \tau+\gamma\to 0\end{cases}$$

So if $G(\tau)$ decays faster than $\dfrac{1}{\tau+\gamma}$ then $F(\tau) \to -\infty$ which is not possible for a convex function.

To fullfil these conditions we assume that the generating function can be written as

$$\boxed{G(\tau) = \frac{GM}{\sqrt{\gamma_0 + \tau}}} \tag{47.16}$$

where G is the gravitational constant and M is the galactic mass.

Theorem 47.1.1 (Kuzmin's theorem). *The spatial mass density function generated by a Stäckel potential is completely determined by a function of the form $\rho(z)$.*

Corollary 47.1.2. For triaxial mass models in ellipsoidal coordinates the axial ratios are inversely proportional to the axial ratios of the coordinate system.

Part IX

Electromagnetism

Chapter 48

Electricity

48.1 Resistance R

48.1.1 Conductivity

Definition 48.1.1 (Drift velocity). The average speed of the independent charge carriers is the drift velocity $\vec{v_d}$. It is important to remark that v_d is not equal to the propagation speed of the electric signal[1].

Formula 48.1.2 (Mobility).

$$\mu = \frac{v_d}{E} \tag{48.1}$$

Formula 48.1.3 (Conductivity).

$$\sigma = nq\mu \tag{48.2}$$

Formula 48.1.4 (Resistivity).

$$\rho = \frac{1}{\sigma} \tag{48.3}$$

48.1.2 Current density

Formula 48.1.5. Let A be the cross-section of a conductor. Let $\vec{\boldsymbol{J}}$ be the current density through A. The current through A is then given by:

$$I = \iint_A \vec{\boldsymbol{J}} \cdot \hat{\mathbf{n}} dS \tag{48.4}$$

Formula 48.1.6 (Free current). The current density generated by free charges is given by:

$$\vec{\boldsymbol{J}} = nq\vec{\boldsymbol{v}}_d \tag{48.5}$$

48.1.3 Pouillet's law

$$R = \frac{l}{A}\rho \tag{48.6}$$

where:

[1] It is several orders of magnitude smaller.

- ρ: resistivity of the material
- l: length of the resistor
- A: cross-sectional area

48.2 Ohm's law

Formula 48.2.1 (Ohm's law).

$$\boxed{\vec{\boldsymbol{J}} = \sigma \cdot \vec{\boldsymbol{E}}} \tag{48.7}$$

where σ is the conductivity tensor.

Formula 48.2.2 (Ohm's law in wires). The following formula can be found by combining equations 48.2, 48.3,48.4 and 48.7 and by assuming that the conductivity tensor is a scalar (follows from the isotropic behaviour of common resistors):

$$U = RI \tag{48.8}$$

48.3 Capacitance C

Definition 48.3.1 (Capacitance). The capacitance is a (geometrical) value that reflects the amount of charge an object can store:

$$C = \frac{q}{V} \tag{48.9}$$

48.4 Electric dipole

Formula 48.4.1 (Electric dipole).

$$\vec{\boldsymbol{p}} = q\vec{\boldsymbol{a}} \tag{48.10}$$

where:

- q: charge of the positive particle
- $\vec{\boldsymbol{a}}$: vector pointing from the negative to the positive particle

Formula 48.4.2 (Energy). If an electric dipole is placed in an electric field, its potential energy is:

$$U = -\vec{\boldsymbol{p}} \cdot \vec{\boldsymbol{E}} \tag{48.11}$$

Formula 48.4.3 (Torque). If an electric dipole is placed in an electric field, a torque is generated:

$$\vec{\boldsymbol{\tau}} = \vec{\boldsymbol{p}} \times \vec{\boldsymbol{E}} \tag{48.12}$$

Chapter 49

Magnetism

49.1 Magnetic field

49.1.1 Fields

The **magnetizing field** $\vec{\boldsymbol{H}}$ is the field resulting from all external sources. When applying an external (magnetic) field, some materials will try to oppose this external influence. Similar to polarization in electricity one can measure the **magnetization**:

$$\vec{\boldsymbol{M}} = \chi \vec{\boldsymbol{H}} \tag{49.1}$$

where χ is the magnetic susceptibility.

The **magnetic induction** $\vec{\boldsymbol{B}}$ is the field resulting from external sources and internal magnetization. (It is the 'real', detectable field.) In vacuum we have the following relation between the magnetic induction, the magnetizing field and the magnetization:

$$\vec{\boldsymbol{B}} = \mu_0 \left(\vec{\boldsymbol{H}} + \vec{\boldsymbol{M}} \right) \tag{49.2}$$

By combining this formula with formula 49.1 we get[1]:

$$\vec{\boldsymbol{B}} = \mu_0 \left(1 + \chi\right) \vec{\boldsymbol{H}} \tag{49.3}$$

The proportionality constant in formula 49.3 is called the **magnetic permeability**:

$$\mu = \mu_0 (1 + \chi) \tag{49.4}$$

where μ_0 is the magnetic permeability of the vacuum. The factor $1 + \chi$ is called the **relative permeability** and it is often denoted by μ_r.

49.1.2 Tensorial formulation

In anistropic materials we have to use the tensorial formulation.

$$B_i = \sum_j \mu_{ij} H_j \tag{49.5}$$

$$M_i = \sum_j \chi_{ij} H_j \tag{49.6}$$

Both μ and χ are tensors of rank 2.

[1] This equation is only valid in linear media.

49.2 Magnetic multipoles

49.2.1 Dipole

$$\vec{\boldsymbol{m}} = IS\vec{\boldsymbol{u}}_n \tag{49.7}$$

49.3 Electric charges in a magnetic field

49.3.1 Cyclotron

Formula 49.3.1 (Gyroradius).

$$r = \frac{mv_\perp}{|q|B} \tag{49.8}$$

Formula 49.3.2 (Gyrofrequency[2]).

$$\omega = \frac{|q|B}{m} \tag{49.9}$$

[2]Also called the Larmor frequency.

Chapter 50

Maxwell Equations

50.1 Lorentz force

Formula 50.1.1 (Lorentz force).

$$\vec{\boldsymbol{F}} = q\left(\vec{\boldsymbol{E}} + \vec{\boldsymbol{v}} \times \vec{\boldsymbol{B}}\right) \tag{50.1}$$

Formula 50.1.2 (Lorentz force density).

$$\vec{\boldsymbol{f}} = \rho\vec{\boldsymbol{E}} + \vec{\boldsymbol{J}} \times \vec{\boldsymbol{B}} \tag{50.2}$$

50.2 Differential Maxwell equations

Formula 50.2.1 (Gauss' law for electricity).

$$\nabla \cdot \vec{\boldsymbol{E}} = \frac{\rho}{\varepsilon} \tag{50.3}$$

Formula 50.2.2 (Gauss' law for magnetism).

$$\nabla \cdot \vec{\boldsymbol{B}} = 0 \tag{50.4}$$

Formula 50.2.3 (Faraday's law).

$$\nabla \times \vec{\boldsymbol{E}} = -\frac{\partial \vec{\boldsymbol{B}}}{\partial t} \tag{50.5}$$

Formula 50.2.4 (Maxwell's law[1]).

$$\nabla \times \vec{\boldsymbol{B}} = \varepsilon\mu\frac{\partial \vec{\boldsymbol{E}}}{\partial t} + \mu\vec{\boldsymbol{J}} \tag{50.6}$$

50.3 Potentials

50.3.1 Decomposition in potentials

Following the Helmholtz decomposition 20.18 we can derive the following general form for $\vec{\boldsymbol{B}}$ starting from Gauss' law 50.4:

$$\boxed{\vec{\boldsymbol{B}} = \nabla \times \vec{\boldsymbol{A}}} \tag{50.7}$$

[1] Also called the law of Maxwell-Ampère.

where $\vec{\boldsymbol{A}}$ is the magnetic potential.

Combining equation 50.7 with Faraday's law 50.5 and rewriting it a bit gives the following general form for $\vec{\boldsymbol{E}}$:

$$\boxed{\vec{\boldsymbol{E}} = -\nabla V - \frac{\partial \vec{\boldsymbol{A}}}{\partial t}} \tag{50.8}$$

where V is the electrostatic potential.

Property 50.3.1. Substituting the expressions 50.7 and 50.8 into Gauss' law 50.3 and Maxwell's law 50.6 gives the following two (coupled) conditions for the electromagnetic potentials:

$$\triangle \vec{\boldsymbol{A}} - \varepsilon\mu \frac{\partial^2 \vec{\boldsymbol{A}}}{\partial t^2} = \nabla \left(\nabla \cdot \vec{\boldsymbol{A}} + \varepsilon\mu \frac{\partial V}{\partial t} \right) - \mu \vec{\boldsymbol{J}} \tag{50.9}$$

$$\triangle V - \varepsilon\mu \frac{\partial^2 V}{\partial t^2} = -\frac{\partial}{\partial t} \left(\nabla \cdot \vec{\boldsymbol{A}} + \varepsilon\mu \frac{\partial V}{\partial t} \right) - \frac{\rho}{\varepsilon} \tag{50.10}$$

50.3.2 Gauge transformations

Looking at equation 50.7, it is clear that a transformation $\vec{\boldsymbol{A}} \longrightarrow \vec{\boldsymbol{A}} + \nabla\psi$ has no effect on $\vec{\boldsymbol{B}}$ due to property 20.15. To compensate this in equation 50.8, we also have to perform a transformation $V \longrightarrow V - \frac{\partial \psi}{\partial t}$.

The (scalar) function $\psi(\vec{\boldsymbol{r}}, t)$ is called a **gauge function**. The transformations are called **gauge transformations**.

Definition 50.3.2 (Gauge fixing conditions). Conditions that fix a certain gauge (or class of gauge transformations) are called gauge fixing conditions. These select one of many physically equivalent configurations.

Example 50.3.3 (Lorenz gauge). A first example of a gauge fixing condition is the Lorenz gauge[2]:

$$\boxed{\nabla \cdot \vec{\boldsymbol{A}} + \varepsilon\mu \frac{\partial V}{\partial t} = 0} \tag{50.11}$$

When using this gauge fixing condition, equations 50.9 and 50.10 become uncoupled and can be rewritten as:

$$\Box \vec{\boldsymbol{A}} = -\mu \vec{\boldsymbol{J}} \tag{50.12}$$

$$\Box V = -\frac{\rho}{\varepsilon} \tag{50.13}$$

To see which gauge functions ψ are valid in this case we perform a transformation as explained above:

$$\vec{\boldsymbol{A}}' = \vec{\boldsymbol{A}} + \nabla\psi \qquad \text{and} \qquad V' = V - \frac{\partial \psi}{\partial t} \tag{50.14}$$

Substituting these transformations in equation 50.11 and using the fact that both sets of potentials $(\vec{\boldsymbol{A}}, V)$ and $(\vec{\boldsymbol{A}}', V')$ satisfiy the Lorenz gauge 50.11 gives the following condition for the gauge function ψ:

$$\Box \psi = 0 \tag{50.15}$$

Example 50.3.4 (Coulomb gauge). Apart from the Lorenz gauge 50.11, there is also the Coulomb gauge:

$$\nabla \cdot \vec{\boldsymbol{A}} = 0 \tag{50.16}$$

[2]Named after Ludvig Lorenz. Not to be confused with Hendrik Lorentz.

50.4 Energy and momentum

Definition 50.4.1 (Poynting vector).

$$\boxed{\vec{\boldsymbol{S}} = \vec{\boldsymbol{E}} \times \vec{\boldsymbol{H}}} \tag{50.17}$$

Definition 50.4.2 (Energy density).

$$\boxed{W = \frac{1}{2}\left(\vec{\boldsymbol{E}} \cdot \vec{\boldsymbol{D}} + \vec{\boldsymbol{B}} \cdot \vec{\boldsymbol{H}}\right)} \tag{50.18}$$

50.5 Differential geometry

Using the tools given (e.g. differential forms) in chapter 30 we can rewrite all of the above formulas in a more elegant form, which will also allow us to generalize them to higher dimensions and more general settings. See for example [9] for a complete derivation and interpretation. It should be noted that we used *Gaussian units* throughout this section.

Definition 50.5.1 (Field strength). Let

$$\mathbf{E} = E_1 dx^1 + E_2 dx^2 + E_3 dx^3$$

and

$$\mathbf{B} = B_1 dx^2 \wedge dx^3 + B_2 dx^3 \wedge dx^1 + B_3 dx^1 \wedge dx^2$$

be the electric and magnetic field forms respectively. Using these forms we can define the field strength as follows:

$$\mathbf{F} = \mathbf{B} - dt \wedge \mathbf{E} \tag{50.19}$$

Formula 50.5.2 (Maxwell's equations). Denote the electric 4-current by

$$\mathbf{J} = \rho dt - J_1 dx^1 - J_2 dx^2 - J_3 dx^3$$

Maxwell's equations can now be rewritten as follows:

$$d\mathbf{F} = 0 \tag{50.20}$$

$$*d(*\mathbf{F}) = 4\pi\mathbf{J} \tag{50.21}$$

where $*$ is the Hodge operator 20.57.

Definition 50.5.3 (Potential). The homogeneous equation 50.20 together with Poincaré's lemma[3] implies that there exists a differential 1-form $\mathbf{A}$ such that

$$\mathbf{F} = d\mathbf{A} \tag{50.22}$$

This 1-form is called the potential or **gauge field**. This field can be related to the ordinary scalar potential V and vectorial potential $\vec{\boldsymbol{A}}$ as follows:

$$\mathbf{A} = -V dt + A_1 dx^1 + A_2 dx^2 + A_3 dx^3 \tag{50.23}$$

Property 50.5.4 (Gauge transformation). Because $d^2 = 0$ the above equation is invariant under a transformation $\mathbf{A} \longrightarrow \mathbf{A} + df$ for any $f \in C^\infty$. This gives us exactly the gauge transformations from equation 50.14 when written out in coordinates.

[3] See theorem 30.5.8.

Part X

Relativity Theory

Chapter 51

Special Relativity

In this and the following chapters we adopt the standard Minkowskian signature $(+,-,-,-)$ unless otherwise stated. This follows the introductory literature and courses such as [17, 18]. Furthermore we also work in natural units unless stated otherwise, i.e. $\hbar = c = 1$.

51.1 Lorentz transformations

Formula 51.1.1.

$$\beta = \frac{v}{c} \tag{51.1}$$

Formula 51.1.2 (Lorentz factor).

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} \tag{51.2}$$

Formula 51.1.3 (Lorentz transformations). Let $\mathbf{V}$ be a 4-vector. A Lorentz boost along the x^1-axis is given by the following transformation:

$$\boxed{\begin{array}{lcl} V'^0 &=& \gamma\left(V^0 - \beta V^1\right) \\ V'^1 &=& \gamma\left(V^1 - \beta V^0\right) \\ V'^2 &=& V^2 \\ V'^3 &=& V^3 \end{array}} \tag{51.3}$$

Remark 51.1.4. Putting $c = +\infty$ in the previous transformation formulas gives the Galilean transformations from classical mechanics.

51.2 Energy and momentum

Formula 51.2.1 (4-velocity).

$$U^\mu = \left(\frac{dx^0}{d\tau}, \frac{dx^1}{d\tau}, \frac{dx^2}{d\tau}, \frac{dx^3}{d\tau}\right) \tag{51.4}$$

or by applying the formulas for proper time and time dilatation we obtain:

$$U^\mu = (\gamma c, \gamma \vec{\boldsymbol{u}}) \tag{51.5}$$

Formula 51.2.2 (4-momentum).

$$p^\mu = m_0 U^\mu \tag{51.6}$$

or by setting $E = cp^0$:

$$p^\mu = \left(\frac{E}{c}, \gamma m_0 \vec{\boldsymbol{u}}\right) \tag{51.7}$$

Definition 51.2.3 (Relativistic mass). The factor γm_0 in the momentum 4-factor is called the relativistic mass. By introducing this quantity (and denoting it by m), the classic formula $\vec{\boldsymbol{p}} = m\vec{\boldsymbol{u}}$ for the 3-momentum can be generalized to 4-momenta p^μ.

Formula 51.2.4 (Relativistic energy relation).

$$\boxed{E^2 = p^2c^2 + m^2c^4} \tag{51.8}$$

This formula is often called the **Einstein relation**.

Chapter 52

General Relativity

Definition 52.0.1 (Stationary spacetime). A spacetime $g_{\mu\nu}$ is called stationary if there exists a timelike Killing vector[1]. This vector can be chosen to be ∂_0 and hence we see that a spacetime is stationary if we can find a coordinate system for which the metric coefficients are time-independent.

52.1 Einstein field equations

Formula 52.1.1 (Einstein field equations). The Einstein field equations without the cosmological constant Λ read:

$$\boxed{G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}} \tag{52.1}$$

where $G_{\mu\nu}$ is the Einstein tensor 33.18 and $T_{\mu\nu}$ is the stress-energy tensor 43.11.

By taking the trace of both sides one obtains $T = -R$ (in units $c = 1, G = \frac{1}{8\pi}$) and hence we can rewrite the Einstein field equations as:

$$R_{\mu\nu} = \hat{T}_{\mu\nu} \tag{52.2}$$

where $\hat{T}_{\mu\nu} = T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T$ is the **reduced stress-energy tensor**.

Formula 52.1.2 (Einstein-Hilbert action). The (vacuum) field equations can be obtained using a variational principle from the following action:

$$S_{EH}[g_{\mu\nu}] = \int_M \sqrt{-g} R \tag{52.3}$$

When working on non-compact manifolds (such as Minkowski space) one needs an extra term to make the boundary contributions vanish. This term is due to Gibbons and Hawking:

$$S_{GH}[g_{\mu\nu}] = \oint_{\partial M} \epsilon \sqrt{h} K \tag{52.4}$$

where h_{ab} is the induced metric on the boundary, K_{ab} is the extrinsic curvature and $\epsilon = \pm 1$ is a constant depending on whether the boundary is timelike or spacelike.

[1] See definition 33.8.

52.2 Schwarzschild metric

Formula 52.2.1 (Schwarzschild metric).

$$ds^2 = \left(1 - \frac{R_s}{r}\right) c^2 dt^2 - \left(1 - \frac{R_s}{r}\right)^{-1} dr^2 - r^2 d\Omega^2 \tag{52.5}$$

where R_s is the Schwarzschild radius given by

$$R_s = \frac{2GM}{c^2} \tag{52.6}$$

Theorem 52.2.2 (Birkhoff). *The Schwarzschild metric is the unique solution of the vacuum field equation under the additional constraints of asymptotic flatness and staticity.*

Formula 52.2.3 (Reissner-Nordström metric). If we allow the black hole to have an electric charge Q, the Schwarzschild metric is modified:

$$ds^2 = \left(1 - \frac{2GM}{r} + \frac{GQ^2}{4\pi r^2}\right) c^2 dt^2 - \left(1 - \frac{2GM}{r} + \frac{GQ^2}{4\pi r^2}\right)^{-1} dr^2 - r^2 d\Omega^2 \tag{52.7}$$

Remark 52.2.4. A computation of the electric field generated by the black hole gives us:

$$E^r = \frac{Q}{4\pi r^2} \tag{52.8}$$

Although the coordinate r is not the proper distance, it still acts as a parameter for the surface of a sphere (as it does in a Euclidean or Schwarzschild metric). This explains why the above formule is the same as the one in classical electromagnetism.

52.3 Causal structure

Definition 52.3.1 (Causal cone). Let M be a Lorentzian manifold. The causal cone of a point $p \in M$ is defined as the set of points $S \subset M$ such that every point $s \in S$ is connected to p by a smooth curve that is everywhere time- or lightlike.

Definition 52.3.2 (Causal closure). Let S be a subset of a Lorentzian manifold. The causal closure of S is defined as the causal complement of the causal complement of S. A **causally closed set** is then defined as a set which is equal to its causal closure.

Part XI

Quantum Mechanics

Chapter 53

Schrödinger Equation

53.1 One dimension

53.1.1 Time independent Schrödinger equation (TISE)

Formula 53.1.1 (TISE).

$$\boxed{\hat{H}\psi(x) = E\psi(x)} \tag{53.1}$$

where $\hat{H}$ is the Hamiltonian of the system.

Property 53.1.2 (Orthonormality). Let $\{\psi_i\}$ be a set of eigenfunctions of the TISE. These functions obey the following orthogonality relations:

$$\int \psi_i^*(x)\psi_j(x)dx = \delta_{ij} \tag{53.2}$$

53.1.2 Time dependent Schrödinger equation (TDSE)

Formula 53.1.3 (TDSE).

$$\boxed{i\hbar\frac{\partial\psi}{\partial t} = \hat{H}\psi} \tag{53.3}$$

where $\hat{H}$ is the Hamiltonian of the system.

Formula 53.1.4 (Massive particle in a time-independent potential).

$$\boxed{i\hbar\frac{\partial}{\partial t}\psi(x,t) = \left(\frac{\hat{p}\cdot\hat{p}}{2m} + \hat{V}(x)\right)\psi(x,t)} \tag{53.4}$$

Formula 53.1.5 (General solution).

$$\boxed{\psi(x,t) = \sum_E c_E\psi_E(x)e^{-\frac{i}{\hbar}Et}} \tag{53.5}$$

where the functions $\psi_E(x)$ are the eigenfunctions of the TISE 53.1. The coefficients c_E can be found using the orthogonality relations:

$$c_E = \left(\int \psi_E^*(x')\psi(x',t_0)dx'\right)e^{\frac{i}{\hbar}Et_0} \tag{53.6}$$

Chapter 54

Mathematical Formalism

54.1 Postulates

54.1.1 Postulate 6: eigenfunction expansion

Definition 54.1.1 (Observable). An operator $\hat{A}$ which possesses a complete set of eigenfunctions is called an observable.

Formula 54.1.2. Let $|\Psi\rangle$ be an arbitrary wavefunction representing the system. Let the set $\{|\psi_n\rangle\}$ be a complete set of eigenfunctions of an observable of the system. The wavefunction $|\Psi\rangle$ can then be expanded as a linear combination of those eigenfunctions:

$$\boxed{|\Psi\rangle = \sum_n c_n|\psi_n\rangle + \int c_a|\psi_a\rangle da} \tag{54.1}$$

where the summation ranges over the discrete spectrum and the integral over the continuous spectrum.

Formula 54.1.3 (Closure relation). For a complete set of discrete eigenfunctions the closure relation[1] reads:

$$\sum_n |\psi_n\rangle\langle\psi_n| = \mathbb{1} \tag{54.2}$$

For a complete set of continuous eigenfunctions we have the following counterpart:

$$\int |i\rangle\langle i|di = \mathbb{1} \tag{54.3}$$

For a mixed set of eigenfunctions a similar relation is obtained by summing over the discrete eigenfunctions and integrating over the continuous eigenfunctions.

Remark. To simplify the notation we will almost always use the notation of equation 54.2 but implicitly integrate over the continuous spectrum.

54.2 Uncertainty relations

Definition 54.2.1 (Commutator). Let $\hat{A}, \hat{B}$ be two operators. We define the commutator of $\hat{A}$ and $\hat{B}$ as follows:

$$\boxed{\left[\hat{A}, \hat{B}\right] = \hat{A}\hat{B} - \hat{B}\hat{A}} \tag{54.4}$$

[1] This relation is also called the **resolution of the identity**.

Formula 54.2.2.

$$\left[\hat{A}\hat{B},\hat{C}\right] = \hat{A}\left[\hat{B},\hat{C}\right] + \left[\hat{A},\hat{C}\right]\hat{B} \tag{54.5}$$

Definition 54.2.3 (Anticommutator). Let $\hat{A},\hat{B}$ be two operators. We define the anticommutator of $\hat{A}$ and $\hat{B}$ as follows:

$$\boxed{\left\{\hat{A},\hat{B}\right\}_{+} = \hat{A}\hat{B} + \hat{B}\hat{A}} \tag{54.6}$$

Definition 54.2.4 (Compatible observables). Let $\hat{A},\hat{B}$ be two observables. If there exists a complete set of functions $|\psi_n\rangle$ that are eigenfunctions of both $\hat{A}$ and $\hat{B}$ then the two operators are said to be compatible.

Formula 54.2.5 (Heisenberg uncertainty relation). Let $\hat{A},\hat{B}$ be two observables. Let $\Delta A, \Delta B$ be the corresponding uncertainties.

$$\boxed{\Delta A \Delta B = \frac{1}{4}\left|\left\langle\left[\hat{A},\hat{B}\right]\right\rangle\right|^2} \tag{54.7}$$

54.3 Matrix representation

Formula 54.3.1. The following formula gives the (m,n)-th element of the matrix representation of $\hat{A}$ with respect to the orthonormal basis $\{\psi_n\}$:

$$\boxed{A_{mn} = \langle\psi_m|\hat{A}|\psi_n\rangle} \tag{54.8}$$

Remark 54.3.2. The basis $\{\psi_n\}$ need not consist out of eigenfunctions of $\hat{A}$.

54.4 Slater determinants

Theorem 54.4.1 (Symmetrization postulate). *Let $\mathcal{H}$ be the Hilbert space belonging to a single particle. A system of n identical particles is described by a wave function Ψ belonging to either $S^n(\mathcal{H})$ or $\Lambda^n(\mathcal{H})$.*

Remark 54.4.2. In ordinary quantum mechanics this is a postulate, but in quantum field theory this is a consequence of the spin-statistics theorem of Pauli.

Formula 54.4.3. Let $\{\sigma\}$ be the set of all permutations of the sequence $(1,...,n)$. Let $|\psi\rangle$ be the single-particle wave function. Bosonic systems are described by a wave function of the form

$$|\Psi_B\rangle = \sum_{\sigma} |\psi_{\sigma(1)}\rangle \cdots |\psi_{\sigma(n)}\rangle \tag{54.9}$$

Fermionic systems are described by a wave function of the form

$$|\Psi_F\rangle = \sum_{\sigma} \mathrm{sgn}(\sigma)|\psi_{\sigma(1)}\rangle \cdots |\psi_{\sigma(n)}\rangle \tag{54.10}$$

Definition 54.4.4 (Slater determinant). Let $\{\phi_i(\vec{\boldsymbol{q}})\}_{i\leq N}$ be a set of wave functions, called **spin orbitals**, describing a system of N identical fermions. The totally antisymmetric wave function of the system is given by:

$$\boxed{\psi(\vec{\boldsymbol{q}}_1,...,\vec{\boldsymbol{q}}_N) = \frac{1}{\sqrt{N!}}\det\begin{pmatrix} \phi_1(\vec{\boldsymbol{q}}_1) & \cdots & \phi_N(\vec{\boldsymbol{q}}_1) \\ \vdots & \ddots & \vdots \\ \phi_1(\vec{\boldsymbol{q}}_N) & \cdots & \phi_N(\vec{\boldsymbol{q}}_N) \end{pmatrix}} \tag{54.11}$$

54.5 Interaction picture

Let $\hat{H}(t) = \hat{H}_0 + \hat{V}(t)$ be the total Hamiltonian of a system where $\hat{V}(t)$ is the interaction potential. Let $|\psi(t)\rangle$ and $\hat{O}$ denote a state and operator in the Schrödinger picture.

Formula 54.5.1. In the interaction picture we define the state vector as follows:

$$|\psi(t)\rangle_I = e^{\frac{i}{\hbar}\hat{H}_0 t}|\psi(t)\rangle \tag{54.12}$$

From this it follows that the operators in the interaction picture are given by:

$$\hat{O}_I(t) = e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{O}e^{-\frac{i}{\hbar}\hat{H}_0 t} \tag{54.13}$$

Formula 54.5.2 (Schrödinger equation). Using the previous formulas, the Schrödinger equation can be rewritten as follows:

$$i\hbar\frac{d}{dt}|\psi(t)\rangle_I = \hat{V}_I(t)|\psi(t)\rangle_I \tag{54.14}$$

The time-evolution of operators in the interaction picture is given by:

$$\frac{d}{dt}\hat{O}_I(t) = \frac{i}{\hbar}\left[\hat{H}_0, \hat{O}_I(t)\right] \tag{54.15}$$

54.5.1 Adiabatic switching

Theorem 54.5.3 (Adiabatic theorem). *If a perturbation is acting slowly enough on a system such that the system can adapt its configuration at every single moment then the system will remain in the same eigenstate.*

54.6 Quantum mechanics on curved space

Using the tools of differential geometry, as presented in chapters 29 and onward, we can introduce quantum mechanics on curved space[2]. We heavily used the course material from [16].

Remark 54.6.1. A first important remark that we have to make is that the classical definition of the wave function as an element of $L^2(\mathbb{R}^d, \mathbb{C})$ is not sufficient, even in flat cartesian space. A complete description requires the introduction of the so-called *Gelfand triple*[3] where we replace the space of square-integrable functions by the Schwartz space[4] of rapidly decreasing functions. The linear functionals on this space are then given by the *tempered distributions*.

Construction 54.6.2. When working on curved spaces (or even in non-cartesian coordinates on flat space) there arise problems with the definition of the self-adjoint operators $\hat{q}^i$ and $\hat{p}_i$. The naive definition $\hat{q}^i = q^i, \hat{p}_i = -i\partial_i$ gives rise to extra terms when calculating inner products that break the canonical commutation relations and the self-adjointness of the operators.

An elegant solution to this problem is found by giving up the definition of the wave function as a function $\psi : \mathbb{R}^d \to \mathbb{C}$. Assume that we are working on a Riemannian base manifold (M, g) and that our 'naive' wave functions where living in a vector space V. We then construct a vector bundle E with typical fibre V over M. Associated to this vector bundle we then find a frame

[2] Not space-time!
[3] See also *rigged Hilbert space*.
[4] See definition 15.1.

bundle $F(E)$ and an Ehresmann connection ω that we can use to define a (local) covariant derivative ∇. The wave function is now defined as a map $\Psi : F(E) \to V$ or locally as the pullback $\psi := \varphi^*\Psi$ for some local section $\varphi : U \subseteq M \to F(E)$.

First we introduce the general inner product

$$\langle \psi, \phi \rangle = \int d^d x \sqrt{\det(g)} \psi^*(x)\phi(x) \tag{54.16}$$

Because the factor $\sqrt{\det(g)}$ transforms in the inverse manner of the measure $d^d x$, the integrand is invariant under coordinate transforms which is something we generally require of our physical laws. Using this new inner product we can check the self-adjointness of the new momentum operator $\hat{P}_i = -i\nabla_i$:

$$\begin{aligned}
\langle \psi, \hat{P}_i \phi \rangle =& \int d^d x \sqrt{\det(g)} \psi^*(x)(-i\nabla_i)\phi(x) \\
\overset{31.59}{=}& \int d^d x \sqrt{\det(g)} \psi^*(x)(-i\partial_i - i\omega_i)\phi(x) \\
=& i \int d^d x \left(\partial_i \sqrt{\det(g)}\right) \psi^*(x)\phi(x) + \int d^d x \sqrt{\det(g)} (-i\partial_i \psi)^*(x)\phi(x) \\
& - i \int d^d x \sqrt{\det(g)} \psi^*(x)\omega_i \phi(x) \\
=& \langle \hat{P}_i \psi, \phi \rangle - i \int d^d x \sqrt{\det(g)} \psi^*(x)\omega_i^* \phi(x) \\
& + i \int d^d x \left(\partial_i \sqrt{\det(g)}\right) \psi^*(x)\phi(x) - i \int d^d x \sqrt{\det(g)} \psi^*(x)\omega_i \phi(x)
\end{aligned}$$

Self-adjointness then requires that

$$\sqrt{\det(g)}(\omega_i + \omega_i^*) = \partial_i \sqrt{\det(g)} \tag{54.17}$$

or using the well known identity $(\ln f)' = \frac{f'}{f}$:

$$2\mathrm{Re}(\omega_i) = \partial_i \ln\left(\sqrt{\det(g)}\right) \tag{54.18}$$

Chapter 55

Angular Momentum

In this chapter we consider the general angular momentum operator $\hat{J} = \left(\hat{J}_x, \hat{J}_y, \hat{J}_z\right)$. This operator works on the Hilbert space spanned by the eigenbasis $\{|j, m\rangle\}$.

55.1 General operator

Property 55.1.1. The mutual eigenbasis of $\hat{J}^2$ and $\hat{J}_z$ is defined by the following two eigenvalue equations:

$$\hat{J}^2|j, m\rangle = j(j+1)\hbar^2|j, m\rangle \tag{55.1}$$

$$\hat{J}_z|j, m\rangle = m\hbar|j, m\rangle \tag{55.2}$$

Property 55.1.2. The angular momentum operators generate a Lie algebra 28.2.1. The Lie bracket is defined by following commutation relation:

$$\boxed{\left[\hat{J}_i, \hat{J}_j\right] = i\hbar\varepsilon_{ijk}\hat{J}_k} \tag{55.3}$$

Definition 55.1.3 (Ladder operators[1]). The raising and lowering operators[2] $\hat{J}_+$ and $\hat{J}_-$ are defined as:

$$\hat{J}_+ = \hat{J}_x + i\hat{J}_y \qquad \text{and} \qquad \hat{J}_- = \hat{J}_x - i\hat{J}_y \tag{55.4}$$

Corollary 55.1.4. From the commutation relations of the angular momentum operators we can derive the commutation relations of the ladder operators:

$$\left[\hat{J}_+, \hat{J}_-\right] = 2\hbar\hat{J}_z \tag{55.5}$$

Formula 55.1.5. The total angular momentum operator $\hat{J}^2$ can now be expressed in terms of $\hat{J}_z$ and the ladder operators using commutation relation 55.3:

$$\hat{J}^2 = \hat{J}_+\hat{J}_- + \hat{J}_z^2 - \hbar\hat{J}_z \tag{55.6}$$

Remark 55.1.6 (Casimir operator). From the definition of $\hat{J}^2$ it follows that this operator is a Casimir invariant[3] in the algebra generated by the operators $\hat{J}_i$.

[1] Also called the **creation** and **annihilation** operators (especially in quantum field theory).
[2] These operators will only affect the z-projection, not the total angular momentum.
[3] See definition 28.4.3.

55.2 Rotations

55.2.1 Infinitesimal rotation

Formula 55.2.1. An infinitesimal rotation $\hat{R}(\delta\vec{\varphi})$ is given by the following formula:

$$\boxed{\hat{R}(\delta\vec{\varphi}) = \mathbb{1} - \frac{i}{\hbar}\vec{\boldsymbol{J}} \cdot \delta\vec{\varphi}} \tag{55.7}$$

A finite rotation can then be produced by applying this infinitesimal rotation repeatedly, which gives:

$$\hat{R}(\vec{\varphi}) = \left(\mathbb{1} - \frac{i}{\hbar}\vec{\boldsymbol{J}} \cdot \frac{\vec{\varphi}}{n}\right)^n = \exp\left(-\frac{i}{\hbar}\vec{\boldsymbol{J}} \cdot \vec{\varphi}\right) \tag{55.8}$$

Formula 55.2.2 (Matrix elements). Applying a rotation over an angle φ around the z-axis to a state $|j, m\rangle$ gives:

$$\hat{R}(\varphi\vec{\boldsymbol{e}}_z)|j, m\rangle = \exp\left(-\frac{i}{\hbar}\hat{J}_z\varphi\right)|j, m\rangle = \exp\left(-\frac{i}{\hbar}m\varphi\right)|j, m\rangle \tag{55.9}$$

Multiplying these states with a bra $\langle j', m'|$ and using the orthonormality of the eigenstates gives the matrix elements of the rotation operator:

$$\boxed{\hat{R}_{ij}(\varphi\vec{\boldsymbol{e}}_z) = \exp\left(-\frac{i}{\hbar}m\varphi\right)\delta_{jj'}\delta_{mm'}} \tag{55.10}$$

From the expression of the angular momentum operators and the rotation operator it is clear that a general rotation has no effect on the total angular momentum number j. This means that the rotation matrix will be a block diagonal matrix with respect to j. This amounts to the following reduction of the representation of the rotation group:

$$\boxed{\langle j, m'|\hat{R}(\varphi\vec{\boldsymbol{n}})|j, m\rangle = \mathcal{D}^{(j)}_{m,m'}(\hat{R})} \tag{55.11}$$

where the values $\mathcal{D}^{(j)}_{m,m'}(\hat{R})$ are the **Wigner D-functions**.

Remark (Wigner D-functions). For every value of j there are $(2j + 1)$ values for m. The matrix $\mathcal{D}^{(j)}(\hat{R})$ is thus a $(2j + 1) \times (2j + 1)$-matrix

55.2.2 Spinor representation

Definition 55.2.3 (Pauli matrices).

$$\boxed{\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}} \tag{55.12}$$

From this definition it is clear that the Pauli matrices are Hermitian and unitary. Together with the 2×2 identity matrix[4] they form a basis for the space of 2×2 Hermitian matrices.

Formula 55.2.4. In the spinor representation ($\mathrm{J} = \frac{1}{2}$) the Wigner-D matrix reads:

$$\mathcal{D}^{(1/2)}(\varphi\vec{\boldsymbol{e}}_z) = \begin{pmatrix} e^{-i/2\varphi} & 0 \\ 0 & e^{i/2\varphi} \end{pmatrix} \tag{55.13}$$

[4] In the context of relativistic QM one often denotes the 2×2 identity matrix as σ_0.

55.3 Coupling of angular momenta

55.3.1 Total Hilbert space

Let $\mathcal{H}_i$ denote the Hilbert space of states belonging to the i^{th} particle. The Hilbert space of the total system is given by the following tensor product:

$$\mathcal{H} = \mathcal{H}_1 \otimes ... \otimes \mathcal{H}_n$$

Due to the tensor product definition above, the angular momentum operator $\hat{J}_i$ should now be interpreted as $\mathbb{1} \otimes ... \otimes \hat{J}_i \otimes ... \otimes \mathbb{1}$. This implies that the angular momentum operators $\hat{J}_{l \neq i}$ do not act on the space $\mathcal{H}_i$, so one can pull these operators through the tensor product:

$$\hat{J}_i|j_1\rangle \otimes ... \otimes |j_n\rangle = |j_1\rangle \otimes ... \otimes \hat{J}_i|j_i\rangle \otimes ... \otimes |j_n\rangle$$

The basis used above is called the **uncoupled basis**.

55.3.2 Clebsch-Gordan series

Let $\vec{\boldsymbol{J}}$ denote the total angular momentum defined as:

$$\vec{\boldsymbol{J}} = \hat{J}_1 + \hat{J}_2 \tag{55.14}$$

With this operator we can define a **coupled** state $|\mathbf{J}, \mathbf{M}\rangle$ where $\mathbf{M}$ is the total magnetic quantum number which ranges from $-\mathbf{J}$ to $\mathbf{J}$.

Formula 55.3.1 (Clebsch-Gordan coefficients). Because both bases (coupled and uncoupled) span the total Hilbert space $\mathcal{H}$ there exists a transformation between them. The transformation coefficients can be found by using the resolution of the identity:

$$\boxed{|\mathbf{J}, \mathbf{M}\rangle = \sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} |j_1, j_2, m_1, m_2\rangle\langle j_1, j_2, m_1, m_2|\mathbf{J}, \mathbf{M}\rangle} \tag{55.15}$$

These coefficients are called the Clebsch-Gordan coefficients.

Property 55.3.2. By acting with the operator $\hat{J}_z$ on both sides of equation 55.15 it is possible to proof that the CG coefficient are non-zero if and only if $\mathbf{M} = m_1 + m_2$.

Chapter 56

Dirac Equation

References for this chapter are [19]. (Note that one uses the mostly-pluses signature there.)

56.1 Dirac matrices

Definition 56.1.1 (Dirac matrices[1]). There exist multiple different representations of the Clifford generators in signature $(1,3)$. The first one is the **Dirac representation**. Here the time-like Dirac matrix γ^0 is defined as

$$\gamma^0 = \begin{pmatrix} \mathbb{1}_2 & 0 \\ 0 & -\mathbb{1}_2 \end{pmatrix} \tag{56.1}$$

The space-like Dirac matrices γ^k, $k = 1, 2, 3$ are defined using the Pauli matrices[2] σ^k:

$$\gamma^k = \begin{pmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{pmatrix} \tag{56.2}$$

The **Weyl** or **chiral** representation[3] is defined by replacing the time-like matrix γ^0 by

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1}_2 \\ \mathbb{1}_2 & 0 \end{pmatrix} \tag{56.3}$$

In signature $(3,1)$ one obtains the Weyl repreesentation by defining $\sigma^\mu \equiv (\mathbb{1}, \sigma_i)$ and $\overline{\sigma}^\mu \equiv \sigma_\mu$:

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma_\mu \\ \overline{\sigma}_\mu & 0 \end{pmatrix} \tag{56.4}$$

Remark 56.1.2. In the continuation of this paper we will adopt the Weyl representation.

Property 56.1.3. The Dirac matrices satisfy following equality:

$$\{\gamma^\mu, \gamma^\nu\}_+ = 2\eta^{\mu\nu}\mathbb{1} \tag{56.5}$$

This has the form of equation 22.4. The Dirac matrices can thus be used as the generating set of a Clifford algebra[4], called the **Dirac algebra**.

[1] Often just called the **gamma matrices**.
[2] See definition 55.12.
[3] This representation is widely used in advanced field theory and supergravity.
[4] See defiinition 22.1.1.

Notation 56.1.4 (Feynman slash notation). Let $a = a^\mu e_\mu \in V$ be a general 4-vector. The Feynman slash $\not{a}$ is defined as follows:

$$\not{a} = a^\mu \gamma_\mu \tag{56.6}$$

In fact this is just a vector space morphism:

$$/ : V \to C\ell(V, \eta) : a^\mu e_\mu \mapsto a^\mu \gamma_\mu \tag{56.7}$$

56.2 Spinors

56.2.1 Dirac equation

Formula 56.2.1 (Dirac equation). In covariant form the Dirac equation reads:

$$\boxed{(i\hbar\not{\partial} - mc)\psi = 0} \tag{56.8}$$

Definition 56.2.2 (Dirac adjoint).

$$\overline{\psi} = \psi^\dagger \gamma^0 \tag{56.9}$$

When working in the Weyl representation one should a factor i to this definition.

Definition 56.2.3 (Majorana adjoint). In the context of SUSY it is often convenient to work with a different adjoint spinor. Let $\mathcal{C} = i\gamma^3\gamma^1$ denote the charge conjugation operator. The Majorana adjoint is then defined by:

$$\overline{\psi} = \psi^t C \tag{56.10}$$

Formula 56.2.4 (Parity).

$$\hat{P}(\psi) = \gamma^0 \psi \tag{56.11}$$

56.2.2 Chiral spinors

In even dimensions one can define an additional matrix [5] which also satisfies equation 56.5:

Definition 56.2.5 (Chiral matrix).

$$\gamma_* = (-i)^{m+1} \gamma_0 \gamma_1 ... \gamma_{d-1} \tag{56.12}$$

when working in $d = 2m$ dimensions. In odd dimensions ($d = 2m + 1$) a generating set for the Clifford algebra can be obtained by taking the generating set in even dimension $d - 1$ and adjoining the element $\pm\gamma_*$. This gives two inequivalent representations of the Clifford algebra (depending on the sign).

Generally one can take the following representation for γ_*:

$$\gamma_* = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \tag{56.13}$$

[5] In $d = 4$ this matrix is often denoted by γ_5.

Definition 56.2.6 (Chiral projection). The chiral projections of a spinor ψ are defined as follows:

$$\psi_L = \frac{1+\gamma_*}{2}\psi \tag{56.14}$$

and

$$\psi_R = \frac{1-\gamma_*}{2}\psi \tag{56.15}$$

Every spinor can then be written as:

$$\psi = \psi_L + \psi_R \tag{56.16}$$

Chapter 57

Perturbation Theory

57.1 Rayleigh-Schrödinger theory

The basic of assumptions of the Rayleigh-Schrödinger perturbation theory are that the perturbation Hamiltonian is time-independent and that the eigenfunctions of the unperturbed Hamiltonian $\hat{H}_0$ also form a complete set for the perturbed Hamiltonian.

Formula 57.1.1. The perturbed eigenfunctions and eigenvalues can be expanded in the following way, where we assume that λ is a small perturbation parameter:

$$|\psi_n\rangle = \sum_{i=0}^{+\infty} \lambda^i |\psi_n^{(i)}\rangle \tag{57.1}$$

$$E_n = \sum_{i=0}^{+\infty} \lambda^i E_n^{(i)} \tag{57.2}$$

where i denotes the order of the perturbation.

57.2 Time-dependent perturbation theory

In this section we consider perturbed Hamiltoninians of the following form:

$$\hat{H}(t) = \hat{H}_0 + \lambda \hat{V}(t) \tag{57.3}$$

57.2.1 Dyson series

Formula 57.2.1 (Tomonaga-Schwinger equation). The evolution operator $\hat{U}(t)$ satisfies the following Schrödinger-type equation in the interaction picture[1]:

$$i\hbar \frac{d}{dt} \hat{U}_I |\psi(0)\rangle_I = \hat{V}_I(t) \hat{U}_I |\psi(0)\rangle_I \tag{57.4}$$

Formula 57.2.2 (Dyson series). Together with the initial value condition $\hat{U}_I(0) = \mathbb{1}$ the Tomonaga-Schwinger equation becomes an initial value problem. A particular solution is given by:

$$\hat{U}_I(t) = \mathbb{1} - \frac{i}{\hbar} \int_0^t \hat{V}_I(t') \hat{U}_I(t') dt' \tag{57.5}$$

[1] See section 54.5.

This solution can be iterated to obtain a series expansion of the evolution operator:

$$\hat{U}(t) = 1 - \frac{i}{\hbar}\int_0^t \hat{V}(t_1)dt_1 + \left(-\frac{i}{\hbar}\right)^2 \int_0^t dt_1 \int_0^{t_1} dt_2 \hat{V}(t_1)\hat{V}(t_2) + ... \tag{57.6}$$

It is clear that the integrands obey a time-ordering. By introducing the **time-ordering operator** $\mathcal{T}$:

$$\mathcal{T}\left(\hat{V}(t_1)\hat{V}(t_2)\right) = \begin{cases} \hat{V}(t_1)\hat{V}(t_2) & , \quad t_1 \geq t_2 \\ \hat{V}(t_2)\hat{V}(t_1) & , \quad t_2 > t_1 \end{cases} \tag{57.7}$$

the integrals can be rewritten in a more symmetric form:

$$\hat{U}(t) = 1 - \frac{i}{\hbar}\int_0^t \hat{V}(t_1)dt_1 + \frac{1}{2!}\left(-\frac{i}{\hbar}\right)\int_0^t dt_1 \int_0^t dt_2 \mathcal{T}\left(\hat{V}(t_1)\hat{V}(t_2)\right) + ... \tag{57.8}$$

or by comparing with the series expansion for exponential functions:

$$\boxed{\hat{U}(t) = \mathcal{T}\left(e^{-\frac{i}{\hbar}\int_0^t \hat{V}(t')dt'}\right)} \tag{57.9}$$

This expansion is called the **Dyson series**.

57.3 Variational method

Definition 57.3.1 (Energy functional).

$$E(\psi) = \frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle} \tag{57.10}$$

Property 57.3.2. The energy functional 57.10 satisfies following inequality:

$$E(\psi) \geq E_0 \tag{57.11}$$

where E_0 is the ground state energy.

Method 57.3.3. Assume that the trial function $|\psi\rangle$ depends on a set of parameters $\{c_i\}_{i\in I}$. The 'optimal' wave function (the one extremizing the energy functional) is found by solving the following system of equations:

$$\frac{\partial\psi}{\partial c_i} = 0 \qquad \forall i \in I \tag{57.12}$$

57.4 Adiabatic approximation

57.4.1 Berry phase

Consider a system for which the adiabatic approximation is valid. We then have a wavefunction of the form

$$\psi(t) = C_a(t)\psi_a(t)\exp\left[-\frac{i}{\hbar}\int_{t_0}^t E_a(t')dt'\right] \tag{57.13}$$

It follows from the orthonormality of the eigenstates $\psi_k(t)$ that the coefficient $C_a(t)$ is just a phase factor, so we can write it as

$$C_a(t) = e^{i\gamma_a(t)} \tag{57.14}$$

Substituting this ansatz in the wavefunction and the Schödinger equation gives a differential equation for the phase factor $\gamma_a(t)$. Integrating it gives:

$$\gamma_a(t) = i \int_{t_0}^{t} \left\langle \psi_a(t') \middle| \frac{\partial \psi_a(t')}{\partial t'} \right\rangle dt' \tag{57.15}$$

Due to time evolution the wavefunction accumulates a phase through the coefficient $C_a(t)$ over the period $t_0 - t_f$. This phase is called the **Berry phase**.

Lets try to apply a phase transformation to remove the Berry phase:

$$\psi'_a(t) = \psi_a(t) e^{i\eta(t)} \tag{57.16}$$

Entering this in equation 57.15 gives

$$\bar{\gamma}'_a(t) = \bar{\gamma}_a(t) - \eta(t_f) + \eta(t_0) \tag{57.17}$$

where the overhead bar denotes the integration between t_0 and t_f in equation 57.15. If the system is cyclic then $\psi_a(t_0) = \psi_a(t_f)$. Combining this with equation 57.16 gives us:

$$\eta(t_f) - \eta(t_0) = 2k\pi \qquad\qquad k \in \mathbb{N} \tag{57.18}$$

which implies that the Berry phase cannot be eliminated through a basis transformation and is thus an observable property of the system.

Definition 57.4.1 (Berry connection). The quantity

$$\mathbf{A}(\vec{\boldsymbol{x}}) = i\langle \psi_a(\vec{\boldsymbol{x}}) | \nabla_{\vec{\boldsymbol{x}}} \psi_a(\vec{\boldsymbol{x}}) \rangle \tag{57.19}$$

where $\nabla_{\vec{\boldsymbol{x}}}$ denotes the gradient in phase space, is called the Berry connection (or Berry gauge potential). Applying Stokes' theorem to 57.15 gives us:

$$\bar{\gamma}_a = \int \boldsymbol{\mathcal{B}} \cdot d\vec{\boldsymbol{S}} \tag{57.20}$$

where $\boldsymbol{\mathcal{B}} = \nabla_{\vec{\boldsymbol{x}}} \times \mathbf{A}(\vec{\boldsymbol{x}})$ is called the **Berry curvature**. Although the Berry connection is gauge dependent, the Berry curvature is gauge invariant!

Remark 57.4.2. Using the language of differential geometry one immediately finds that the accumulated phase $\bar{\gamma}_a$ is simply the holonomy associated with the Berry connection along the considered trajectory.

Chapter 58

Scattering Theory

58.1 Cross section

Formula 58.1.1 (Differential cross section).

$$\frac{d\sigma}{d\Omega} = \frac{N(\theta, \varphi)}{F} \tag{58.1}$$

where F is the incoming particle flux and N the detected flow rate[1].

58.1.1 Fermi's golden rule

Formula 58.1.2 (Fermi's golden rule). The transition probability from state i to state f is given by:

$$\boxed{\Gamma_{i\to f} = \frac{2\pi}{\hbar}|\langle f|\hat{H}|i\rangle|^2 \frac{dn}{dE_f}} \tag{58.2}$$

58.2 Lippman-Schwinger equations

In this section we consider Hamiltonians of the following form: $\hat{H} = \hat{H}_0 + \hat{V}$ where $\hat{H}_0$ is the free Hamiltonian and $\hat{V}$ the scattering potential. We will also assume that both the total Hamiltonian and the free Hamiltonian have the same eigenvalues.

Formula 58.2.1 (Lippman-Schwinger equation).

$$|\psi^{(\pm)}\rangle = |\varphi\rangle + \frac{1}{E - \hat{H}_0 \pm i\varepsilon}\hat{V}|\psi^{(\pm)}\rangle \tag{58.3}$$

where $|\varphi\rangle$ is an eigenstate of the free Hamiltonian with the same energy as $|\psi\rangle$, i.e. $\hat{H}_0|\varphi\rangle = E|\varphi\rangle$.

Remark 58.2.2. The term $\pm i\varepsilon$ is added to the denominator because otherwise it would be singular. The term has no physical meaning.

Formula 58.2.3 (Born series). If we rewrite the Lippman-Schwinger equation as:

$$|\psi\rangle = |\varphi\rangle + \hat{G}_0\hat{V}|\psi\rangle \tag{58.4}$$

[1]Because N is not defined as a flux but as a rate, the differential cross section has the dimension of area.

where $\hat{G}_0$ is the Green's operator, then we can derive the following series expansion by iterating the equation:

$$|\psi\rangle = |\varphi\rangle + \hat{G}_0\hat{V}|\varphi\rangle + \left(\hat{G}_0\hat{V}\right)^2|\varphi\rangle + ... \tag{58.5}$$

Formula 58.2.4 (Born approximation). If we cut off the Born series at the first order term in $\hat{V}$ then we obtain the Born approximation:

$$|\psi\rangle = |\varphi\rangle + \hat{G}_0\hat{V}|\varphi\rangle \tag{58.6}$$

Chapter 59

Quantum Information Theory

59.1 Bipartite systems

59.1.1 Marginal density operators

Definition 59.1.1 (Reduced density operator). Let $|\Psi\rangle_{AB}$ be the state of a bipartite system. The reduced density operator $\hat{\rho}_A$ of system A is defined as follows:

$$\hat{\rho}_A = \mathrm{Tr}_B |\Psi\rangle_{AB}\!\langle\Psi| \tag{59.1}$$

Definition 59.1.2 (Purification). Let $\hat{\rho}_A$ be the density operator of a system A. A purification of $\hat{\rho}_A$ is a pure state $|\Psi\rangle_{AB}$ of some composite system AB such that:

$$\hat{\rho}_A = \mathrm{Tr}_B |\Psi\rangle_{AB}\!\langle\Psi| \tag{59.2}$$

Property 59.1.3. Any two purifications of the same density operator $\hat{\rho}_A$ are related by a transformation $\mathbb{1}_A \otimes \hat{V}$, where $\hat{V}$ is a unitary operator on $\mathcal{H}_B$.

Part XII

Quantum Field Theory

Chapter 60

Canonical Quantization

60.1 Klein-Gordon field

60.1.1 Lagrangian and Hamiltonian

The simplest Lagrangian (density) is given by:

$$\boxed{\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2} \tag{60.1}$$

Using the principle of least action we obtain the following Euler-Lagrange equation[1]:

$$\left(\partial^\mu\partial_\mu + m^2\right)\phi = 0 \tag{60.2}$$

which can be rewritten using the **d'Alembertian** $\Box = \partial_\mu\partial^\mu$:

$$\boxed{(\Box + m^2)\phi = 0} \tag{60.3}$$

This equation is called the Klein-Gordon equation. In the limit $m \to 0$ this equation reduces to the well-known wave equation.

From the Lagrangian 60.1 we can also derive a Hamiltonian function using relations 43.5 and 43.12:

$$\boxed{H = \int d^3x \frac{1}{2}\left[\pi^2(x) + (\nabla\phi(x))^2 + m^2\phi^2(x)\right]} \tag{60.4}$$

60.1.2 Raising and lowering operators

Fourier expanding the scalar field $\phi(\vec{\boldsymbol{x}}, t)$ in momentum space and inserting it into the Klein-Gordon equation gives:

$$\left(\partial_t^2 + p^2 + m^2\right)\phi(\vec{\boldsymbol{p}}, t) = 0 \tag{60.5}$$

This is the equation for a simple harmonic oscillator with frequency $\omega_{\vec{\boldsymbol{p}}} = \sqrt{p^2 + m^2}$.

[1] See formula 43.4.

Analogous to ordinary quantum mechanics we define the raising and lowering operators $a^\dagger_{\vec{\boldsymbol{p}}}$ and $a_{\vec{\boldsymbol{p}}}$ such that:

$$\phi(\vec{\boldsymbol{x}}) = \iiint \frac{d^3p}{(2\pi)^{3/2}} \frac{1}{\sqrt{2\omega_{\vec{\boldsymbol{p}}}}} \left(a_{\vec{\boldsymbol{p}}} e^{i\vec{\boldsymbol{p}}\cdot\vec{\boldsymbol{x}}} + a^\dagger_{\vec{\boldsymbol{p}}} e^{-i\vec{\boldsymbol{p}}\cdot\vec{\boldsymbol{x}}} \right) \tag{60.6}$$

$$\pi(\vec{\boldsymbol{x}}) = \iiint \frac{d^3p}{(2\pi)^{3/2}} (-i) \sqrt{\frac{\omega_{\vec{\boldsymbol{p}}}}{2}} \left(a_{\vec{\boldsymbol{p}}} e^{i\vec{\boldsymbol{p}}\cdot\vec{\boldsymbol{x}}} - a^\dagger_{\vec{\boldsymbol{p}}} e^{-i\vec{\boldsymbol{p}}\cdot\vec{\boldsymbol{x}}} \right) \tag{60.7}$$

An equivalent definition is obtained by performing the transformation $\vec{\boldsymbol{p}} \to -\vec{\boldsymbol{p}}$ in the second term of $\phi(\vec{\boldsymbol{x}})$ and $\pi(\vec{\boldsymbol{x}})$:

$$\boxed{\begin{aligned} \phi(\vec{\boldsymbol{x}}) &= \iiint \frac{d^3p}{(2\pi)^{3/2}} \frac{1}{\sqrt{2\omega_{\vec{\boldsymbol{p}}}}} \left(a_{\vec{\boldsymbol{p}}} + a^\dagger_{-\vec{\boldsymbol{p}}} \right) e^{i\vec{\boldsymbol{p}}\cdot\vec{\boldsymbol{x}}} && (60.8) \\ \pi(\vec{\boldsymbol{x}}) &= \iiint \frac{d^3p}{(2\pi)^{3/2}} (-i) \sqrt{\frac{\omega_{\vec{\boldsymbol{p}}}}{2}} \left(a_{\vec{\boldsymbol{p}}} - a^\dagger_{-\vec{\boldsymbol{p}}} \right) e^{i\vec{\boldsymbol{p}}\cdot\vec{\boldsymbol{x}}} && (60.9) \end{aligned}}$$

When we impose the commutation relation

$$[a_{\vec{\boldsymbol{p}}}, a^\dagger_{\vec{\boldsymbol{q}}}] = \delta(\vec{\boldsymbol{p}} - \vec{\boldsymbol{q}}) \tag{60.10}$$

we obtain the following commutation relation for the scalar field and its conjugate momentum:

$$[\phi(\vec{\boldsymbol{x}}), \pi(\vec{\boldsymbol{y}})] = i\delta(\vec{\boldsymbol{x}} - \vec{\boldsymbol{y}}) \tag{60.11}$$

Now, the Hamiltonian can be calculated explicitly:

$$H = \int \frac{d^3p}{(2\pi)^3} \omega_{\vec{\boldsymbol{p}}} \left(a^\dagger_{\vec{\boldsymbol{p}}} a_{\vec{\boldsymbol{p}}} + \frac{1}{2} [a_{\vec{\boldsymbol{p}}}, a^\dagger_{\vec{\boldsymbol{p}}}] \right) \tag{60.12}$$

It is however clear from 60.10 that the second term in this integral diverges. There are two reasons for this divergence. Firstly, space is infinite, i.e. the d^3x integral in 60.4 diverges. This problem can be resolved by restricting the system to a (finite) part of space or by considering the energy density instead of the energy itself. Secondly, by including very large values for p in the integral we enter a parameter range where our theory is likely to break down. So we should introduce a "high p" cut-off[2].

A more practical solution however is to note that only energy differences are physical and so we can drop the second term altogether as it is merely a "constant".

A corollary of equation 60.12 together with the canonical commutation relations is:

$$\boxed{\begin{aligned} [H, a^\dagger_{\vec{\boldsymbol{p}}}] &= \omega_p a^\dagger_{\vec{\boldsymbol{p}}} && (60.13) \\ [H, a_{\vec{\boldsymbol{p}}}] &= -\omega_p a_{\vec{\boldsymbol{p}}} && (60.14) \end{aligned}}$$

As was the case for the quantum harmonic oscillator, the creation and annihilation operators deserve their names and we can write:

$$|\vec{\boldsymbol{k}}_1, ..., \vec{\boldsymbol{k}}_n\rangle = a^\dagger(\vec{\boldsymbol{k}}_1) \cdots a^\dagger(\vec{\boldsymbol{k}}_n)|0\rangle \tag{60.15}$$

Furthermore, this equation together with the canonical commutation relations imply that the Klein-Gordon fields are bosonic fields.

[2] See *regularization*.

60.1.3 Scalar propagator

Formula 60.1.1 (Pauli-Jordan function).

$$i[\phi(x),\phi(y)] = \underbrace{\int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega_p}\left(e^{-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} - e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}\right)}_{i\Delta(x-y)} \tag{60.16}$$

In the case that $x^0 = y^0$ (ETCR) or $(x-y)^2 < 0$ (*faster than light*) the Pauli-Jordan function is identically 0.[3]

60.1.4 Normalization constant

Under a general Lorentz boost Λ the delta function $\delta^{(3)}(\vec{\boldsymbol{p}} - \vec{\boldsymbol{q}})$ transforms[4] as $\delta^{(3)}(\Lambda\vec{\boldsymbol{p}} - \Lambda\vec{\boldsymbol{q}})\frac{\Lambda E}{E}$. Although this is clearly not Lorentz invariant, we see that the quantity $E_p\delta^{(3)}(\vec{\boldsymbol{p}} - \vec{\boldsymbol{q}})$ is an invariant.

The correct normalisation for the momentum representation thus becomes:

$$\sqrt{2E_p}a^\dagger_{\mathbf{p}}|0\rangle = |\mathbf{p}\rangle \tag{60.17}$$

and hence

$$\langle\mathbf{p}|\mathbf{q}\rangle = 2E_p(2\pi)^3\delta^{(3)}(\vec{\boldsymbol{p}} - \vec{\boldsymbol{q}}) \tag{60.18}$$

where the constants are a matter of convention (to cancel the constants in 60.6).

60.1.5 Invariant integration measure

The factor $2E_p$ does not only occur in the normalisation conditions. To find a Lorentz invariant integration measure in spacetime we consider the following integral:

$$\int \frac{d^3p}{2E_p} = \int d^4p\ \delta(p^2 - m^2)\Big|_{p^0>0} \tag{60.19}$$

By using this measure we ensure that the integral of any Lorentz invariant function $f(p)$ is again Lorentz invariant.

Example 60.1.2 (One-particle identity operator).

$$\hat{\mathbb{1}}_1 = \int \frac{d^3p}{2E_p}|\mathbf{p}\rangle\langle\mathbf{p}| \tag{60.20}$$

60.2 Dirac field

60.3 Contractions and Wick's theorem

60.3.1 Bosonic fields

In the following definitions (field) operators will be decomposed as

$$\phi = \phi^{(+)} + \phi^{(-)}$$

[3] See also the axiom of microcausality 62.1

[4] This follows from property 15.8.

where the + symbol denotes the 'positive frequency' part, i.e. the part consisting of annihilation operators[5]. The 'negative frequency' part is defined analogously.

Definition 60.3.1 (Contraction for neutral bosonic fields).

$$\overbrace{\phi(x)\phi(y)} = \begin{cases} [\phi(x)^{(+)}, \phi(y)^{(-)}] & x^0 > y^0 \\ [\phi(y)^{(+)}, \phi(x)^{(-)}] & y^0 > x^0 \end{cases} \tag{60.21}$$

Formula 60.3.2 (Feynman propagator).

$$\overbrace{\phi(x)\phi(y)} = i \underbrace{\lim_{\varepsilon \to 0^+} \int \frac{d^4k}{(2\pi)^4} \frac{e^{-i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})}}{k^2 - m^2 + i\varepsilon}}_{\Delta_F(\mathbf{x}-\mathbf{y})} \tag{60.22}$$

Definition 60.3.3 (Contraction for charged bosonic fields).

$$\overbrace{\phi(x)\overline{\phi}(y)} = \begin{cases} [\phi(x)^{(+)}, \overline{\phi}(y)^{(-)}] & x^0 > y^0 \\ [\phi(y)^{(+)}, \overline{\phi}(x)^{(-)}] & y^0 > x^0 \end{cases} \tag{60.23}$$

Definition 60.3.4 (Normal ordering). The normal ordering[6] $\mathcal{N}$ of a sequence of field operators is defined as the permuted sequence in which all annihilation operators appear on the right of the creation operators. For example:

$$\mathcal{N}\Big(\phi(x)\phi^\dagger(y)\phi(z)\Big) = \phi^\dagger(y)\phi(x)\phi(z)$$

Property 60.3.5. The vacuum state expectation value of a normal ordered sequence is 0.

Formula 60.3.6 (Wick's theorem for bosonic fields).

$$\mathcal{T}\Big(\underbrace{\phi(x_1)\phi(x_2)...\phi(x_n)}_{S}\Big) = \mathcal{N}\Big(\phi(x_1)...\phi(x_n) + \text{all possible contractions}\Big) \tag{60.24}$$

Remark 60.3.7. In the case of charged bosons, only contractions of the form $\overbrace{\phi(x)\overline{\phi}(y)}$ will remain because $[a, b^+] = 0$.

60.3.2 Fermionic fields

Definition 60.3.8 (Contraction).

$$\overbrace{\psi(x)\overline{\psi}(y)} = \begin{cases} \{\psi(x)^{(+)}, \overline{\psi}(y)^{(-)}\}_+ & x^0 > y^0 \\ -\{\psi(y)^{(+)}, \overline{\psi}(x)^{(-)}\}_+ & y^0 > x^0 \end{cases} \tag{60.25}$$

Remark 60.3.9. Only contractions of the form $\overbrace{\psi(x)\overline{\psi}(y)}$ will remain because $\{a, b^+\}_+ = 0$.

Formula 60.3.10 (Feynman propagator).

$$\overbrace{\psi(x)\overline{\psi}(y)} = i \underbrace{\lim_{\varepsilon \to 0^+} \int \frac{d^4p}{(2\pi)^4} \frac{\not{p} + m}{p^2 - m^2 + i\varepsilon} e^{-ip\cdot(x-y)}}_{S_F(x-y)} \tag{60.26}$$

[5]The classic Fourier integral is defined using an exponential $e^{-i\mathbf{k}\cdot\mathbf{x}}$. By looking at equation 60.6 and remembering that we are working in the $(1, 3)$ Minkowski signature, we see that the annihilators always occur together with a positive frequency exponential.

[6]Sometimes denoted by : :

Remark 60.3.11 (Normal ordering). One should take into account the Fermi-Dirac statistics when permuting fermionic field operators under a normal ordering. A general factor $\mathrm{sgn}(P)$, where P is the permutation of the operators, will arise in every term.

For example:

$$\mathcal{N}\Big(\psi(x)\overline{\psi}(y)\psi(z)\Big) = -\overline{\psi}(y)\psi(x)\psi(z)$$

A similar remark should be made for the time ordering $\mathcal{T}$.

Chapter 61

Yang-Mills Theory

61.1 Gauge invariance

Using the tools of differential geometry, as presented in chapters 29 and onward, we can introduce the general formulation of Yang-Mills gauge theories. Used references are [9, 15, 16, 20].

Consider a general Lie group[1] SU(n), acting on a Hilbert bundle $\mathcal{H}$ of physical states over a base space (manifold) M. A general gauge transformation has the form

$$\psi'(x) = U(x)\psi(x) \tag{61.1}$$

where $\psi, \psi' : M \to \mathcal{H}$ are sections of the physical Hilbert bundle and $U : M \to G$ encodes the local behaviour of the gauge transformation.

Theorem 61.1.1 (Local gauge principle). *The Lagrangian $\mathcal{L}[\psi]$, where $\psi(x) \in \mathcal{H}_x$, is invariant under the action of the gauge group G:*

$$\mathcal{L}[U\psi] = \mathcal{L}[\psi] \tag{61.2}$$

Generally this gauge invariance can be achieved in the following way. Denote the Lie algebra corresponding to G by $\mathfrak{g}$. Because the gauge transformation is local, the information on how it changes should be able to propagate through space. This is done by introducing a new field $B_\mu(x)$, called the **gauge field**. The most elegant formulation uses the concept of covariant derivatives:

Definition 61.1.2 (Covariant derivative).

$$\mathcal{D}_\mu = \partial_\mu + igB_\mu(x) \tag{61.3}$$

where $B_\mu : M \to \mathfrak{g}$. Here we should also note that the explicit action of the covariant derivative depends on the chosen representation of $\mathfrak{g}$ on $\mathcal{H}$. Furthermore one should note that we used the physics convention where one multiplies[2] the gauge field B by a factor ig.

So to achieve gauge invariance one should replace all derivatives by the covariant derivative. Now one could wonder if the covariant derivative itself satisfies the local gauge principle, i.e.

[1] This group is in the physics literature often called the **gauge group**.
[2] This turns anti-Hermitian fields into Hermitian fields.

$\mathcal{D}'\psi' = U\mathcal{D}\psi$. Lets write this out (from here on we will supress the coordinate dependence):

$$\begin{aligned} U^{-1}\left(\frac{\partial}{\partial x^\mu} + igB'_\mu\right)\psi' &= U^{-1}\left(\frac{\partial}{\partial x^\mu} + igB'_\mu\right)U\psi \\ &= U^{-1}\frac{\partial U}{\partial x^\mu}\psi + \frac{\partial \psi}{\partial x^\mu} + igU^{-1}B'_\mu U\psi \end{aligned} \tag{61.4}$$

This expression can only be equal to $\mathcal{D}\psi$ if

$$igB_\mu = U^{-1}\frac{\partial U}{\partial x^\mu} + igU^{-1}B'_\mu U \tag{61.5}$$

which can be rewritten as

$$B'_\mu = UB_\mu U^{-1} - \frac{1}{ig}(\partial_\mu U)U^{-1} \tag{61.6}$$

or in coordinate-independent form:

$$\boxed{\mathbf{B}' = U\mathbf{B}U^{-1} - \frac{1}{ig}dUU^{-1}} \tag{61.7}$$

Up to conventions this is exactly the content of equations 31.35 and 31.37 in differential geometry.

Example 61.1.3 (QED). For quantum electrodynamics, which has $\mathrm{U}(1) \cong S^1$ as its gauge group, we use the parametrization $U(x) = e^{ie\chi(x)}$ where $\chi : \mathbb{R}^n \to \mathbb{R}$. The relevant formulae then become:

$$\partial_\mu \longrightarrow \mathcal{D}_\mu = \partial_\mu + ieA_\mu \tag{61.8}$$

$$A_\mu \longrightarrow A'_\mu = A_\mu - \partial_\mu \chi \tag{61.9}$$

where A_μ is the classic electromagnetic potential.

61.2 Spontaneous symmetry breaking

Theorem 61.2.1 (Goldstone). *Consider a QFT with Lie group G. Denote the generators of the corresponding Lie algebra by $\mathbf{X}_a$. Generators that do not destroy the vacuum[3], i.e. $\mathbf{X}_a v \neq 0$, correspond to massless scalar (Goldstone) bosons.*

[3]This corresponds to a transformation that leaves the vacuum invariant.

Chapter 62

Axiomatic QFT

62.1 Algebraic QFT

62.1.1 Haag-Kastler axioms

Axiom 62.1 (Local net of observables). To every causally closed set[1] O one associates a C^*-algebra $\mathcal{A}(O)$. We require this assignement to satisfy the following conditions:

- Isotony: If $O_1 \subset O_2$ then $\mathcal{A}(O_1) \subset \mathcal{A}(O_2)$.
- (Causal) locality[2]: If O_1 and O_2 are spacelike separated then $[\mathcal{A}(O_1), \mathcal{A}(O_2)] = 0$ (graded commutator) within a larger algebra $\mathcal{A}(O)$ where $O_1, O_2 \subset O$.

Axiom 62.2 (Haag duality). Let $\overline{O}$ denote the spacelike complement of O and let $\mathcal{A}'$ denote the commutant of $\mathcal{A}$. Haag duality states that[3]

$$\mathcal{A}\left(\overline{O}\right)' = \mathcal{A}(O) \tag{62.1}$$

for all causally closed sets O.

Remark 62.1.1. Haag duality is known to hold for all free theories and for some less trivial theories. However it is also known to fail in the case of symmetry breaking [21].

62.1.2 Weyl systems

Definition 62.1.2 (Weyl system). Let (L, ω) be a symplectic vector space. Let K be a complex vector space and let W be a map from L to the space of unitary operators on K. The pair (K, W) is a Weyl system over (L, ω) if it satisfies:

$$W(z)W(z') = e^{\frac{i}{2}\omega(z,z')}W(z+z') \tag{62.2}$$

for all $z, z' \in K$.

Definition 62.1.3 (Heisenberg system). The generators[4] $\phi(z)$ of the map $t \mapsto W(tz)$ are said to form a Heisenberg system. These operators satisfy following properties:

[1] See definition 52.3.2.

[2] Also called **microcausality**.

[3] Here it should be understood that $\mathcal{A}\left(\overline{O}\right)$ is the algebra generated by all algebras $\infty(Q)$ where Q ranges over the causally closed sets in $\overline{O}$.

[4] Their existence is guaranteed by Stone's theorem.

- $\lambda\phi(z) = \phi(\lambda z)$ for all positive λ.
- $[\phi(z), \phi(z')] = -i\omega(z, z')$
- $\phi(z + z')$ is the closure[5] of $\phi(z) + \phi(z')$

62.2 Topological QFT

62.2.1 Atiyah-Segal axioms

Axiom 62.3 (Atiyah-Segal). A d-dimensional topological quantum field theory (TQFT) is a symmetric monoidal functor $F : \mathbf{Bord}_{d-1}^{d} \to \mathbf{FinVect}$ satisfying following axioms:

1. Normalization: $F(\emptyset) = \mathbb{C}$
2. Disjoint union: $F(M \sqcup M') = F(M) \otimes F(M')$
3. Composition: If $N = M \cup M'$ where ∂M and ∂M have opposite orientation then:

$$F(N) = F(M) \circ F(M')$$

4. Invariance: If $f : M \to M'$ is a diffeomorphism rel boundary then $F \circ f = F$.

where M, M' are d-dimensional cobordisms between $(d-1)$-dimensional (closed) smooth manifolds.

Example 62.2.1 (1D). In 1D one has the following possible smooth cobordisms and their associated TQFT operations:

Point with orientation $+$	Vector space V
Point with reversed orientation $-$	Dual space V^*
Line between points	Linear map $f : V \to V$
Cap between $\emptyset$ and points $+, -$	Coevaluation $\mathbb{C} \to V \otimes V^*$
Cup between points $-, +$ and $\emptyset$	Evaluation $V^* \otimes V \to \mathbb{C}$

Essentially this gives us the structure of a (finite-dimensional) vector space with dual.

Example 62.2.2 (2D). In 2D one can obtain a similar result by drawing all possible configurations, however the existence and combination of "*pair of pants*"-diagrams gives a richer stucture. For 2D TQFT's the corresponding object is a (finite-dimensional) commutative and cocommutative Frobenius algebra.

In dimensions 3 and higher the definition above is intractable. To allow the construction to be generalized to higher dimensions one considers the following definition:

Definition 62.2.3 (Extended TQFT). A d-dimensional extended TQFT is a symmetric monoidal functor $F : \mathbf{Bord}_{1}^{d} \to \mathbf{FinVect}$ satisfying the Atiyah-Segal axioms where the invariance axiom is required only at the highest level of k-morphisms.

[5] See definition 23.1.17.

Part XIII

Thermodynamics & Statistical Mechanics

Chapter 63

Thermodynamics

63.1 General definitions

Definition 63.1.1 (System). The part of space that we are examining.

Definition 63.1.2 (Surroundings). Everything outside the system.

Definition 63.1.3 (Immediate surrounding). The part of the surroundings that 'lies' immediately next to the system.

Definition 63.1.4 (Environment). Everything outside the immediate surroundings.

Definition 63.1.5 (Thermodynamic coordinates). Macroscopical quantities that describe the system.

Definition 63.1.6 (Intensive coordinate). Coordinate that does not depend on the total amount of material (or system size).

Definition 63.1.7 (Extensive coordinate). Coordinate that does depend on the amount of material.

Definition 63.1.8 (Thermodynamic equilibrium). A system in thermodynamic equilibrium is simultaneously in thermal, mechanical and chemical equilibrium. The system is also described by a certain set of constant coordinates.

Theorem 63.1.9. *During thermodynamic equilibrium, all intensive coordinates are uniform throughout the system.*

Definition 63.1.10 (Isolated system). An isolated system can't interact with its surroundings (due to the presence of impenetrable walls).

Definition 63.1.11 (Diathermic wall). A diathermic wall is a wall that allows heat transfer.

Definition 63.1.12 (Adiabatic wall). An adiabatic wall is a wall that does not allow heat transfer.

Definition 63.1.13 (Open system). An open system is a system that allows matter exchange.

Definition 63.1.14 (Closed system). A closed system is a system that does not allow matter exchange.

Definition 63.1.15 (Quasistatic process). A quasistatic process is a sequence of equilibrium states separated by infinitesimal changes.

Definition 63.1.16 (Path). The sequence of equilibrium states in a process is called the path.

63.2 Postulates

Theorem 63.2.1 (Zeroth law). *If two object are in thermal equilibrium with a third object then they are also in thermal equilibrium with eachother.*

Theorem 63.2.2 (First law).

$$U_f - U_i = W + Q \tag{63.1}$$

$$dU = \delta W + \delta Q \tag{63.2}$$

Remark. The δ in the heat and work differentials implies that these are 'inexact' differentials. This means that they cannot be expressed as functions of the thermodynamic coordinates. More formally a differential form dx is called inexact if the integral $\int dx$ is path dependent.

Theorem 63.2.3 (Kelvin-Planck formulation of the second law). *No machine can absorb an amount of heat and completely transform it into work.*

Theorem 63.2.4 (Clausius formulation of the second law). *Heat cannot be passed from a cooler object to a warmer object without performing work.*

Formula 63.2.5 (Clausius' inequality). In differential form, the inequality reads:

$$\frac{\delta Q}{T} \geq 0 \tag{63.3}$$

Theorem 63.2.6 (Third law). *No process can reach absolute zero in a finite sequence of operations.*

63.3 Gases

63.3.1 Ideal gases

Theorem 63.3.1 (Ideal gas law).

$$\boxed{PV = nRT} \tag{63.4}$$

Chapter 64

Statistical Mechanics

64.1 Axioms

Theorem 64.1.1 (Ergodic principle). *All microstates corresponding to the same macroscopic state are equally propable.*

Theorem 64.1.2 (Boltzmann formula). *The central axiom of statistical mechanics gives following formula for the entropy:*

$$\boxed{S = k \ln \Omega(E, V, N, \alpha)} \tag{64.1}$$

where Ω denotes the number of microstates corresponding to the system with energy E, volume V, ...

64.2 Temperature

Formula 64.2.1. The temperature of a system in contact with a heat bath is defined as:

$$\boxed{T = \left(\frac{\partial E}{\partial S}\right)_V} \tag{64.2}$$

64.3 Canonical ensemble

Formula 64.3.1 (Partition function). The partition function for discrete systems is defined as:

$$\boxed{Z = \sum_i g_i e^{-\beta \varepsilon_i}} \tag{64.3}$$

or for continuous systems:

$$Z(T) = \int \Omega(E, V, N) e^{-\beta E} dE \tag{64.4}$$

Consider a system of N indistinguishable non-interacting particles. Let ε_i be the energy associated with the i^{th} energy level and let g_i be its degeneracy. The probability p_i of finding a particle in the i^{th} energy level is given by:

$$\boxed{p_i = \frac{g_i e^{-\beta \varepsilon_i}}{Z}} \tag{64.5}$$

Definition 64.3.2 (Helmholtz free energy). A Legendre transform of the energy E gives us:

$$F = -k_B T \ln Z = E - TS \tag{64.6}$$

64.4 Grand canonical ensemble

Formula 64.4.1 (Grand canonical partition function). The partition function of the i^{th} energy level is defined as:

$$\mathcal{Z}_i = \sum_{n_k} e^{\beta n_k(\mu - \varepsilon_i)} \tag{64.7}$$

The grand canonical partition function is then given by:

$$\boxed{\mathcal{Z} = \prod_i \mathcal{Z}_i = \sum_{n_k, \varepsilon_i} e^{\beta n_k(\mu - \varepsilon_i)}} \tag{64.8}$$

Remark 64.4.2. In the case of fermions, $n_i \in \{0, 1\}$, this formula reduces to $\mathcal{Z} = e^{\beta\mu} Z$.

Definition 64.4.3 (Fugacity).

$$z = e^{\mu N} \tag{64.9}$$

Formula 64.4.4 (Quantum). In the quantum mechanical case one can rewrite the partition function as follows:

$$\mathcal{Z} = \text{tr} \exp\left(-\frac{\hat{H} - \mu\hat{N}}{T}\right) \tag{64.10}$$

which reduces to the above expressions when working in the single-particle energy eigenbasis (which is only possible for free theories).

64.5 Energy

Theorem 64.5.1 (Virial theorem).

$$\boxed{\langle T \rangle = -\frac{1}{2} \sum_k \langle \vec{\boldsymbol{r}}_k \cdot \vec{\boldsymbol{F}}_k \rangle} \tag{64.11}$$

Corollary 64.5.2. For potentials of the form $V = ar^{-n}$ this becomes:

$$2\langle T \rangle = -n \langle V \rangle \tag{64.12}$$

Theorem 64.5.3 (Equipartition theorem). *Let x be a generalized coordinate.*

$$\boxed{\left\langle x^k \frac{\partial H}{\partial x^l} \right\rangle = k_b T \delta_{kl}} \tag{64.13}$$

Corollary 64.5.4. For quadratic Hamiltonians this can be rewritten using Euler's theorem for homogeneous functions 12.10 as:

$$\langle T \rangle = \frac{1}{2} k_b T \tag{64.14}$$

64.6 Black-body radiation

Formula 64.6.1 (Planck's law).

$$\boxed{B_\nu(\nu, T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{\frac{h\nu}{kt}} - 1}} \tag{64.15}$$

Formula 64.6.2 (Wien's displacement law).

$$\boxed{\lambda_{max} T = b} \tag{64.16}$$

where $b = 2.897\,772\,9(17) \times 10^{-3}$ Km is **Wien's displacement constant**.

Part XIV

Condensed Matter Physics

Chapter 65

Material Physics

65.1 Crystals

Theorem 65.1.1 (Steno's law). *The angles between crystal faces of the same type are constant and do not depend on the total shape of the crystal.*

Definition 65.1.2 (Zone). The collection of faces parallel to a given axis, is called a zone. The axis itself is called the zone axis.

65.1.1 Analytic representation

Definition 65.1.3 (Miller indices). Let a, b, c be the lengths of the (not necessarily orthogonal) basis vectors of the crystal lattice. The lattice plane intersecting the axes at $\left(\frac{a}{h}, \frac{b}{k}, \frac{c}{k}\right)$ is denoted by the Miller indices $(h\ k\ l)$.

Notation 65.1.4. Negative numbers are written as $\overline{a}$ instead of $-a$.

Formula 65.1.5 (Coordinates of axes). Let a, b, c denote the lengths of the basis vectors. The axis formed by the intersection of the planes $(h_1\ k_1\ l_1)$ and $(h_2\ k_2\ l_2)$, pointing in the direction of the point (au, bv, cw) is denoted by $[u\ v\ w]$. Where

$$u = \begin{vmatrix} k_1 & l_1 \\ k_2 & l_2 \end{vmatrix} \qquad v = \begin{vmatrix} l_1 & h_1 \\ l_2 & h_2 \end{vmatrix} \qquad w = \begin{vmatrix} h_1 & k_1 \\ h_2 & k_2 \end{vmatrix} \tag{65.1}$$

Theorem 65.1.6 (Hauy's law of rational indices). *The Miller indices of every natural face of a crystal will always have rational proportions.*

65.2 Symmetries

Definition 65.2.1 (Equivalent planes/axes). When applying certain symmetries to a plane or axis, it often occurs that we obtain a set of equivalent planes/axes. These equivalence classes are denoted respectively by $\{h\ k\ l\}$ and $< h\ k\ l >$.

Property 65.2.2 (Rotational symmetry). Only $1, 2, 3, 4$ and 6-fold rotational symmetries can occur.

65.3 Crystal lattice

Formula 65.3.1. For an orthogonal crystal lattice, the distance between planes of the family $(h\ k\ l)$ is given by:

$$d_{hkl} = \frac{1}{\sqrt{\left(\frac{h}{a}\right)^2 + \left(\frac{k}{b}\right)^2 + \left(\frac{l}{c}\right)^2}} \tag{65.2}$$

65.3.1 Bravais lattice

Definition 65.3.2 (Bravais lattice). A crystal lattice generated by a certain point group symmetry is called a Bravais lattice. There are 14 different Bravais lattices in 3 dimensions. These are the only possible ways to place (infinitely) many points in 3D space by applying symmetry operations to a point group.

Definition 65.3.3 (Wigner-Seitz cell). The part of space consisting of all points closer to a given lattice point than to any other.

65.3.2 Reciprocal lattice

Formula 65.3.4 (Reciprocal basis vectors). The reciprocal lattice corresponding to a given Bravais lattice with primitive basis $\{\vec{\boldsymbol{a}}, \vec{\boldsymbol{b}}, \vec{\boldsymbol{c}}\}$ is defined by the following reciprocal basis vectors

$$\vec{\boldsymbol{a}}^* = 2\pi \frac{\vec{\boldsymbol{b}} \times \vec{\boldsymbol{c}}}{\vec{\boldsymbol{a}} \cdot (\vec{\boldsymbol{b}} \times \vec{\boldsymbol{c}})} \tag{65.3}$$

The vectors $\vec{\boldsymbol{b}}^*$ and $\vec{\boldsymbol{c}}^*$ are obtained by permutation of (a, b, c). These vectors satisfy the relations

$$\begin{aligned} \vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{a}}^* &= 2\pi \\ \vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{b}}^* &= 2\pi \\ \vec{\boldsymbol{c}} \cdot \vec{\boldsymbol{c}}^* &= 2\pi \end{aligned} \tag{65.4}$$

$$\tag{65.5}$$

Notation 65.3.5 (Reciprocal lattice vector). The reciprocal lattice vector $\vec{\boldsymbol{r}}^*_{hkl}$ is defined as follows:

$$\vec{\boldsymbol{r}}^*_{hkl} = h\vec{\boldsymbol{a}}^* + k\vec{\boldsymbol{b}}^* + l\vec{\boldsymbol{c}}^* \tag{65.6}$$

Property 65.3.6. The reciprocal lattice vector $\vec{\boldsymbol{r}}^*_{hkl}$ has the following properties:

- $\vec{\boldsymbol{r}}^*_{hkl}$ is perpendicular to the family of planes $(h\ k\ l)$ of the direct lattice.
- $||\vec{\boldsymbol{r}}^*_{hkl}|| = \frac{2\pi n}{d_{hkl}}$

65.4 Diffraction

65.4.1 Constructive interference

Formula 65.4.1 (Laue conditions). Suppose that an incident beam makes angles α_0, β_0 and γ_0 with the lattice axes. The diffracted beam making angles α, β and γ with the axes will be

observed if following conditions are satisfied:

$$a(\cos\alpha - \cos\alpha_0) = h\lambda$$
$$b(\cos\beta - \cos\beta_0) = k\lambda$$
$$c(\cos\gamma - \cos\gamma_0) = l\lambda$$

If these conditions have been met then we observe a diffracted beam of order hkl.

Remark 65.4.2. Further conditions can be imposed on the angles, such as the pythagorean formula for orthogonal axes. This has the consequence that the only two possible ways to obtain a diffraction pattern are:

- a fixed crystal and a polychromatic beam
- a rotating crystal and a monochromatic beam

Formula 65.4.3 (Vectorial Laue conditions). Let $\vec{k}_0, \vec{k}$ denote the wave vector of respectively the incident and diffracted beams. The Laue conditions can be reformulated in the following way:

$$\boxed{\vec{k} - \vec{k}_0 = \vec{r}^*_{hkl}} \tag{65.7}$$

Formula 65.4.4 (Bragg's law). Another equivalent formulation of the Laue conditions is given by following formula:

$$\boxed{2d_{hkl}\sin\theta = n\lambda} \tag{65.8}$$

where

λ : wavelength of the incoming beam

θ : the **Bragg angle**

d_{hkl} : distance between neighbouring planes

Remark 65.4.5. The angle between the incident and diffracted beams is given by 2θ.

Construction 65.4.6 (Ewald sphere). A simple construction to determine if Bragg difraction will occur is the Ewald sphere: Put the origin of the reciprocal lattice at the tip of the incident wave vector $\vec{\boldsymbol{k}}_i$. Now construct a sphere with radius $\frac{2\pi}{\lambda}$ centered on the start of $\vec{\boldsymbol{k}}_i$. All points on the sphere that coincide with a reciprocal lattice point satisfy the vectorial Laue condition 65.7. Therefore Bragg diffraction will occur in the direction of all the intersections of the Ewald sphere and the reciprocal lattice.

65.4.2 Intensity of diffracted beams

Definition 65.4.7 (Systematic extinctions). Every particle in the motive emits its own waves. These waves will interfere and some will cancel out which leads to the absence of certain diffraction spots. These absences are called systematic extinctions.

Definition 65.4.8 (Atomic scattering factor). The waves produced by the individual electrons of an atom, which can have a different phase, can be combined into a resulting wave. The amplitude of this wave is called the atomic scattering factor.

Definition 65.4.9 (Structure factor). The waves coming from the individual atoms in the motive can also be combined, again taking into account the different phases, into a resulting wave. The amplitude of this wave is called the structure factor and it is given by:

$$F(hkl) = \sum_j f_j \exp\left[2\pi i(hx_j + ky_j + lz_j)\right] \tag{65.9}$$

where f_j is the atomic scattering factor of the j^{th} atom in the motive.

Example 65.4.10. A useful example of systematic extinctions is the structure factor of an FCC or BCC lattice for the following specific situations:

If $h+k+l$ is odd, then $F(hkl)=0$ for a BCC lattice. If h,k and l are not all even or all odd then $F(hkl)=0$ for an FCC lattice.

Definition 65.4.11 (Laue indices). Higher order diffractions can be rewritten as a first order diffraction in the following way:

$$2d_{nhnknl}\sin\theta=\lambda \qquad \text{with} \qquad d_{nhnknl}=\frac{d_{hkl}}{n} \tag{65.10}$$

Following from the interpretation of the Bragg law as diffraction being a reflection at the lattice plane $(h\ k\ l)$ we can introduce the (fictitious) plane with indices $(nh\ nk\ nl)$. These indices are called Laue indices.

Remark. In contrast to Miller indices which cannot possess common factors, the Laue indices obviously can.

65.5 Alloys

Theorem 65.5.1 (Hume-Rothery conditions). An element can be dissolved in a metal (forming a solid solution) if the following conditions are met:

- The difference between the atomic radii is $\leq 15\%$.
- The crystal structures are the same.
- The elements have a similar electronegativity.
- The valency is the same.

65.6 Lattice defects

Definition 65.6.1 (Vacancy).

A lattice point where an atom is missing. Also called a **Schottky defect**.

Formula 65.6.2 (Concentration of Schottky defects†). Let N denote the number of lattice points and n the number of vacancies. The following relation gives the temperature dependence of Schottky defects:

$$\frac{n}{n+N}=e^{-E_v/kT} \tag{65.11}$$

where T is the temperature and E_v the energy needed to create a vacancy.

Remark. A similar relation holds for interstitials.

Definition 65.6.3 (Interstitial). An atom placed at a position which is not a lattice point.

Definition 65.6.4 (Frenkel pair). An atom displaced from a lattice point to an interstitial location (hereby creating a vacancy-interstitial pair) is called a Frenkel defect.

Formula 65.6.5 (Concentration of Frenkel pairs). Let n_i denote the number of atoms displaced from the bulk of the lattice to any N_i possible interstitial positions and thus creating n_i vacancies. The following relation holds:

$$\frac{n_i}{\sqrt{NN_i}} = e^{-E_{fr}/2kT} \tag{65.12}$$

where E_{fr} denotes the energy needed to create a Frenkel pair.

Remark 65.6.6. In compounds the number of vacancies can be much higher than in monoatomic lattices.

Remark 65.6.7. The existence of these defects creates the possibility of diffusion.

65.7 Electrical properties

65.7.1 Charge carriers

Formula 65.7.1 (Conductivity). Definition 48.2 can be modified to account for both positive and negative charge carriers:

$$\sigma = n_n q_n \mu_n + n_p q_p \mu_p \tag{65.13}$$

Remark. The difference between the concentration of positive and negative charge carriers can differ by orders of magnitude across different materials. It can differ by up to 20 orders of magnitude.

65.7.2 Band structure

Definition 65.7.2 (Valence band). The energy band corresponding to the outermost (partially) filled atomic orbital.

Definition 65.7.3 (Conduction band). The first unfilled energy band.

Definition 65.7.4 (Band gap). The energy difference between the valence and conduction bands (if they do not overlap). It is the energy zone[1] where no electron states can exist.

Definition 65.7.5 (Fermi level). The energy level having a 50% chance of being occupied at thermodynamic equilibrium.

Formula 65.7.6 (Fermi function). The following distribution gives the probability of a state with energy E_i being occupied by an electron:

$$\boxed{f(E_i) = \frac{1}{e^{(E_i - E_f)/kT} + 1}} \tag{65.14}$$

where E_f is the Fermi level as defined above.

65.7.3 Intrinsic semiconductors

Formula 65.7.7. Let n denote the charge carrier density as before. We find the following temperature dependence:

$$n \propto e^{-E_g/2kt} \tag{65.15}$$

where E_g is the band gap. This formula can be directly derived from the Fermi function by noting that for intrinsic semiconductors the Fermi level sits in the middle of the band gap, i.e. $E_c - E_f = E_g/2$, and that for most semiconductors $E_g \gg kT$.

[1] For a basic derivation see [22].

65.7.4 Extrinsic semiconductors

Definition 65.7.8 (Doping). Intentionally introducing impurities to modify the (electrical) properties.

Definition 65.7.9 (Acceptor). Group III element added to create an excess of holes in the valence band. The resulting semiconductor is said to be a **p-type semiconductor**.

Definition 65.7.10 (Donor). Group IV element added to create an excess of electrons in the valence band. The resulting semiconductor is said to be an **n-type semiconductor**.

65.7.5 Ferroelectricity

Some materials can exhibit certain phase transitions between a paraelectric and ferroelectric state.

Paraelectric materials have the property that the polarisation $\vec{P}$ and the electric field $\vec{E}$ are proportional. Ferroelectric materials have the property that they exhibit permanent polarization, even in the absence of an electric field. This permanent behaviour is the result of a symmetry breaking, i.e. the ions in the lattice have been shifted out of their 'central' positions and induce a permanent dipole moment.

The temperature at which this phase transition occurs is called the **ferroelectric Curie temperature**. Above this temperature the material will behave as a paraelectric material.

Remark 65.7.11. Ferroelectricity can only occur in crystals with unit cells that do not have a center of symmetry. This would rule out the possiblity of having the asymmetry needed for the dipole moment.

Definition 65.7.12 (Saturation polarization). The maximum polarization obtained by a ferroelectric material. It it obtained when the domain formation also reaches a maximum.

Definition 65.7.13 (Remanent polarization). The residual polarization of the material when the external electric field is turned off.

Definition 65.7.14 (Coercive field). The electric field needed to cancel out the remanent polarization.

Definition 65.7.15 (Piezoelectricity). Materials that obtain a polarization when exposed to mechanical stress are called piezoelectric materials.

Remark 65.7.16. All ferroelectric materials are piezoelectric, but the converse is not true. All crystals without a center of symmetry are piezoelectric. This property is however only a necessary (and not a sufficient) condition for ferroelectricity, as mentioned above.

Example 65.7.17 (Transducer). A device that converts electrical to mechanical energy (and vice versa).

65.8 Magnetic properties

Definition 65.8.1 (Diamagnetism). In diamagnetic materials, the magnetization is oriented opposite to the applied field, so $B < H$. The susceptibility is small, negative and independent of the temperature.

Remark 65.8.2. All materials exhibit a diamagnetic character.

Definition 65.8.3 (Paramagnetism). The susceptibility is small, positive and inversely proportional to the temperature.

Definition 65.8.4 (Ferromagnetism). Spontaneous magnetization can occur. The susceptibility is large and dependent on the applied field and temperature. Above a certain temperature, the **ferromagnetic Curie temperature**, the materials will behave as if they were only paramagnetic.

65.8.1 Paramagnetism

Formula 65.8.5 (Curie's law). If the interactions between the particles can be neglected, we obtain the following law:

$$\chi = \frac{C}{T} \tag{65.16}$$

Materials that satisfy this law are called **ideal paramagnetics**.

Formula 65.8.6 (Curie-Weiss law). If the interactions between particles cannot be neglected, we obtain the following law:

$$\chi = \frac{C}{T - \theta} \tag{65.17}$$

where $\theta = CN_W$ with N_W the **Weiss-constant**. This deviation of the Curie law is due to the intermolecular interactions that induce an internal magnetic field $H_m = N_W M$.

Formula 65.8.7 (Brillouin function B_J).

$$\boxed{B_J(y) = \frac{2J+1}{2J}\coth\left(\frac{2J+1}{2J}y\right) - \frac{1}{2J}\coth\left(\frac{y}{2J}\right)} \tag{65.18}$$

where $y = \dfrac{g\mu_B JB}{kT}$

Remark 65.8.8. Because $\coth(y \to \infty) \approx 1$ we have:

$$\text{if } T \to 0 \quad \text{then} \quad M = Ng\mu_B JB_J(y \to \infty) = Ng\mu_B J \tag{65.19}$$

This value is called the **absolute saturation magnetization**.

65.8.2 Ferromagnetism

Ferromagnetics are materials that have strong internal interactions which lead to large scale (with respect to the lattice constant) parallel ordering of the atomic magnetic (dipole) moments. This leads to the spontaneous magnetization of the material and consequently a nonzero total dipole moment.

Remark. In reality, ferromagnetic materials do not always spontaneously possess a magnetic moment in the absence of an external field. When stimulated by a small external field, they will however display a magnetic moment, much larger than paramagnetic materials would.

Definition 65.8.9 (Domain). The previous remark is explained by the existence of Weiss domains. These are spontaneously magnetized regions in a magnetic material. The total dipole moment is the sum of the moments of the individual domains. If not all the domains have a parallel orientation then the total dipole moment can be 0, a small external field is however sufficient to change the domain orientation and produce a large total magnetization.

Definition 65.8.10 (Bloch walls). A wall between two magnetic domains.

Definition 65.8.11 (Ferromagnetic Curie temperature). Above this temperature the material loses its ferromagnetic properties and it becomes a paramagnetic material following the Curie-Weiss law.

Remark 65.8.12. For ferromagnetic (and ferrimagnetic) materials it is impossible to define a magnetic susceptibility as the magnetization is nonzero even in the absence of a magnetic field.[2] Above the critical temperature (Curie/Néel) it is however possible to define a susceptiblity as the materials become paramagnetic in this region.

65.8.3 Antiferromagnetism

When the domains in a magnetic material have an antiparallel ordering[3], the total dipole moment will be small. If the temperature rises, the thermal agitation however will disturb the orientation of the domains and the magnetic susceptibility will rise.

Definition 65.8.13 (Néel temperature). At the Néel temperature, the susceptibility will reach a maximum. Above this temperature ($T > T_N$) the material will become paramagnetic, satisfying the following formula:

$$\chi = \frac{C}{T + \theta} \tag{65.20}$$

This resembles a generalization of the Curie-Weiss law with a negative and therefore virtual critical temperature.

65.8.4 Ferrimagnetism

Materials that are not completely ferromagnetic nor antiferromagnetic, due to an unbalance between the sublattices, will have a nonzero dipole moment even in the absence of an external field. The magnitude of this moment will however be smaller than that of a ferromagnetic material. These materials are called ferrimagnetic materials.

Formula 65.8.14 (Néel hyperbola). Above the Néel temperature it is possible to define a susceptibility given by:

$$\frac{1}{\chi} = \frac{T}{C} - \frac{1}{\chi_0} - \frac{\sigma}{T - \theta'} \tag{65.21}$$

65.9 Mathematical description

Theorem 65.9.1 (Neumann's principle). *The symmetry elements of the physical properties of a crystal should at least contain those of the point group of the crystal.*

[2] This can be seen from equation 49.1: $M = \chi H$. The susceptibility should be infinite.
[3] This will occur if it is energetically more favourable.

Chapter 66

Tensor Networks

66.1 Matrix Product States

66.1.1 Finite-dimensional lattices

Definition 66.1.1 (Matrix product state). Let $\mathcal{H}_n$ be the local Hilbert spaces of dimension d_n where $n \in \{1, ..., N\}$. A state $|\psi\rangle$ in the total Hilbert space $\bigotimes_i \mathcal{H}_i$ is a matrix product state (with periodic boundary conditions) if there exist matrices $A^{i_n}(n) \in \mathcal{L}(\mathbb{C}^{D_n}, \mathbb{C}^{D_{n-1}})$ with $i_n \leq d_n$ such that:

$$|\psi\rangle = \sum_{\{i_k\}} \mathrm{tr}\Big(\prod_n^N A^{i_n}(n)\Big)|i_1...i_N\rangle \tag{66.1}$$

For each lattice site n the set of matrices $\{A^{i_n}_{\alpha\beta}(n)\}$ can be regarded as the content of one rank-3 tensor. The periodic boundary condition requires that $D_0 = D_N$ (otherwise the trace would be ill-defined). Different boundary conditions can be implemented by adding an additional factor X at the end of the trace.

Notation 66.1.2. In the continuation of this chapter we will abbreviate matrix product states as **MPS**.

Remark 66.1.3 (Physical and virtual spaces). For each *physical* index i_n one can regard the matrix $A^{i_n}(n)$ as a linear map between *virtual* (or ancilla) spaces $\mathbb{C}^{D_n}$.

Formula 66.1.4 (MPS projector). Consider an MPS given by tensors $\{A(n)\}_{n\leq N}$. The associated MPS projector is defined as:

$$\mathcal{P}(A) = \sum_{i,\alpha,\beta} A^i_{\alpha\beta}(n)|i\rangle\langle\alpha\beta| \tag{66.2}$$

Formula 66.1.5 (Transfer operator). Give the MPS tensors $\{A(n)\}_{n\leq N}$ one can define a transfer operator:

$$\mathbb{E}(n) = \sum_{i=1}^{d_n} A^i(n) \otimes \overline{A^i}(n) \tag{66.3}$$

Formula 66.1.6 (Superoperator). More generally we can define for every local observable $\hat{O}_n$ a superoperator in $\mathcal{L}(\mathbb{C}^{D_n} \otimes \overline{\mathbb{C}^{D_n}}, \mathbb{C}^{D_{n-1}} \otimes \overline{\mathbb{C}^{D_{n-1}}})$:

$$\mathbb{E}_{O_n}(n) = \sum_{i,i'=1}^{d_n} \langle i|\hat{O}_n|i'\rangle A^{i'}(n) \otimes \overline{A^i}(n) \tag{66.4}$$

Comparing with the definition of the transfer operator we see that $\mathbb{E}$ is given by the superoperator associated to the unit operator. Given two sets of MPS tensors $\{A(n)\}, \{B(n)\}$ we define a generalized superoperator by:

$$\mathbb{E}_B^A(n) = \sum_{i=1}^{d_n} A^i(n) \otimes \overline{B^i}(n) \tag{66.5}$$

Example 66.1.7. Using these definitions we can rewrite the formulas for expectation values more efficiently. Given a product operator $\hat{O} = \bigotimes_i^N \hat{O}_i$ we find that:

$$\langle \psi[A] | \hat{O} | \psi[A] \rangle = \mathrm{tr}\Big(\mathbb{E}_{O_1}(1) \cdots \mathbb{E}_{O_N}(N) \Big) \tag{66.6}$$

Formula 66.1.8. Associated to the superoperator $\mathbb{E}_O(n)$ one can define a map acting on the virtual operators:

$$\mathcal{E}_{O_n}^{(n)}(\phi) = \sum_{i,i'=1}^{d_n} \langle s | \hat{O}_n | s' \rangle A^{i'}(n) \phi A^i(n)^\dagger \tag{66.7}$$

$$\tilde{\mathcal{E}}_{O_n}^{(n)}(\phi) = \sum_{i,i'=1}^{d_n} \langle s | \hat{O}_n | s' \rangle A^i(n)^\dagger \sigma A^{i'}(n) \tag{66.8}$$

where $\phi \in \mathcal{L}(\mathbb{C}^{D_n}), \sigma \in \mathcal{L}(\mathbb{C}^{D_{n-1}})$.

Property 66.1.9. The map $\mathcal{E}_{\mathbb{1}}^{(n)}$ associated to the transfer operator is a CP map[1] and the associated Kraus operators are the MPS matrices $A^i(n)$.

66.1.2 Translation-invariant states

Definition 66.1.10 (Uniform MPS). By setting all MPS tensors $A(n) = B$ for a given tensor B one obtains a translation-invariant (TI) state, i.e. a state invariant under a shift of the index n. These MPS form the variational class of uniform MPS.

Remark 66.1.11 (TIMPS). Not every TIMPS should be uniform, there should only exist a local gauge transformation $A'(n) = U(n)A(n)U(n)^{-1}$ such that $A'(n)$ is uniform (in certain cases this is only possible by enlarging the bond dimension).

[1] See definition 23.3.6.

Part XV

Appendices

Appendix A

Derivations: Mathematics

A.1 Group theory

A.1.1 Explanation for property 3.1.65

Pick an element $x \in X$. The stabilizer of x with respect to G is the set

$$S_x = \{g \in G | g \cdot x = x\}$$

Due to the transitivity of the group action we have that

$$\forall x, y \in X : \exists h \in G : h \cdot x = y$$

So for every $z \in X$ we can choose a group element g_z such that $g_z \cdot x = z$. For all elements in the coset $g_z S_x = \{g_z s \in G | s \in S_x\}$ the following equality is satisfied:

$$(g_z s) \cdot x = g_z \cdot (s \cdot x) = g_z \cdot x = z$$

This implies that the map $\Phi : G/S_x \to X$ is surjective.

Now we need to prove that Φ is also injective. We give a proof by contradiction. Choose two distinct cosets gS_x and hS_x. Then there exist two elements $G, H \in X$ such that $g \cdot x = G$ and $h \cdot x = H$. Now assume that $G = H$. This means that

$$\begin{aligned} & g \cdot x = h \cdot x \\ \iff & (h^{-1}g) \cdot x = x \\ \iff & h^{-1}g \in S_x \\ \iff & hS_x \ni h(h^{-1}g) = g \end{aligned}$$

This would imply that $gS_x = hS_x$ which is in contradiction to our assumption. It follows that $G \neq H$ such that Φ is injective. □

A.2 Calculus

A.2.1 Proof of method 12.10.14

The function $F(x)$ is defined as follows:

$$F(x) = \sum_{n=0}^{+\infty} \frac{a_n}{n!} x^n \tag{A.1}$$

We now perform a Borel transform:

$$\begin{aligned}\int_0^{+\infty} F(xt)e^{-t}dt &= \sum_{n=0}^{N}\int_0^{+\infty}\frac{a_n}{n!}x^n t^n e^{-t}dt \\ &= \sum_{n=0}^{N}\frac{a_n}{n!}x^n\int_0^{+\infty} t^n e^{-t}dt \\ &= \sum_{n=0}^{N}\frac{a_n}{n!}x^n\Gamma(n+1) \\ &= \sum_{n=0}^{N}a_n x^n\end{aligned} \tag{A.2}$$

where we used the definition of the Gamma function 12.17 on line 3 and the relation between the factorial function and the Gamma function 12.18 on line 4. □

A.3 Linear algebra

A.3.1 Proof of equivalence of definitions 20.8.14 and 20.8.15

$$(u+v)\otimes(u+v) - u\otimes u - v\otimes v = u\otimes v + v\otimes u \tag{A.3}$$

The LHS is an element of the ideal I generated by $\{v\otimes v | v\in V\}$. Using the ideal generated by elements such as in the RHS gives the usual definition of the exterior algebra based on the wedge product as defined in 20.48 because it imposes the relation $u\wedge v = -v\wedge u$.

We do however have to pay attention to one little detail. As mentioned in 20.8.15 the general definition uses the ideal I to construct the quotient space. The other construction is only equivalent when working over a field with a characteristic different from 2. This follows from the fact that we have to divide by 2 when trying to obtain the ideal I from the RHS by setting $u = v$.

A.4 Manifolds and bundles

A.4.1 Proof of equivalence of definitions 27.2.3 and 27.2.7

Let (U,φ) be a chart around the point $p\in M$. Using the first definition of a tangent vector (27.2.3), i.e.:

$$\left.\frac{\partial}{\partial q^i}\right|_p : \mathcal{F}_p(M,\mathbb{R})\to\mathbb{R} : f\mapsto \frac{\partial}{\partial q^i}\left(f\circ\varphi^{-1}\right)(\varphi(p))$$

we can rewrite equation 27.7:

$$v_p(f) = \frac{\partial(f\circ\varphi^{-1})}{\partial q^i}(\varphi(p))\frac{dq^i}{dt}(0)$$

as follows:

$$v_p(f) = \left.\frac{\partial f}{\partial q^i}\right|_p\frac{dq^i}{dt}(0)$$

Because the partial derivatives as defined in 27.2.3 form a basis for the tangent space (by construction), we see that this equation is in fact an expansion of the tangent vector v_p in terms

of that basis. It follows that vectors tangent to curves[1] are also tangent vectors according to the first definition.

To prove the other direction we have to show that the partial derivative operators can be constructed as vectors tangent to curves.

A tangent vector can be expanded, according to the first construction, in the following way:

$$v_p = v^i \left.\frac{\partial}{\partial q^i}\right|_p$$

where we also define $v = (v^1, ..., v^n)$. We can then construct the curve $\gamma : t \mapsto \varphi^{-1}(q_0 + vt)$. It is obvious that the tangent vector v_p is tangent to the curve γ. From this it follows that we have an isomorphism between the tangent vectors from to the first definition and the equivalence classes of vectors tangent to curves from the second definition. These definitions are thus equivalent.□

Although the previous equivalence implies that the tangent space construction using germs of curves gives us a vector space we could also check the vector space axioms directly. First we prove that the sum of vectors tangent to the curves γ and δ is again a vector tangent to some curve $\chi : \mathbb{R} \to M$. For this let us define the curve

$$\chi(t) \equiv \varphi^{-1} \circ \Big(\varphi \circ \gamma(t) + \varphi \circ \delta(t) - \varphi(p)\Big)$$

where φ is again the coordinate map in some chart (U, φ) around $p \in M$. Using equation 27.7 we find:

$$\begin{aligned}
v_{p,\chi}(f) =& \frac{\partial(f \circ \varphi^{-1})}{\partial q^i}(\varphi(p)) \frac{d(\varphi^i \circ \chi)}{dt}(0) \\
=& \frac{\partial(f \circ \varphi^{-1})}{\partial q^i}(\varphi(p)) \frac{d}{dt}\left(\varphi^i \circ \gamma + \varphi^i \circ \delta - \varphi^i(p)\right) \\
=& \frac{\partial(f \circ \varphi^{-1})}{\partial q^i}(\varphi(p)) \left(\frac{d(\varphi^i \circ \gamma)}{dt} + \frac{d(\varphi^i \circ \delta)}{dt}\right) \\
=& v_{p,\gamma}(f) + v_{p,\delta}(f)
\end{aligned}$$

The constant term $-\varphi(p)$ in the definition of $\chi(t)$ is necessary to make sure that $\chi(0) = \gamma(0) = \delta(0) = p$. The scalar multiplication by a number $\lambda \in K$ can be proven by defining the curve $\chi(t) = \varphi^{-1} \circ \Big[\lambda\Big(\varphi \circ \gamma(t)\Big)\Big]$.

A.4.2 Explanation for example 30.3.13

In this derivation we use the Landau little-o notation $o(t)$, i.e.:

$$\lim_{t \to 0} \frac{o(t)}{t} = 0 \tag{A.4}$$

Now assume that X is a smooth vector field and f is a smooth function. Because the Lie derivative is a local operation we can work in a local chart such that γ is (again locally) equivalent

[1] More precisely: representatives of equivalence classes of vectors tangent to curves.

to a curve[2] $\beta_p : U \to \mathbb{R}^n$ and such that we can expand $\beta_p(t)$ around $p \in U$:

$$\begin{aligned}
\mathcal{L}_X f(p) &= \lim_{t\to 0}\left[\frac{f(\beta_p(0) + t\beta_p'(0) + o(t)) - f(p)}{t}\right] \\
&= \lim_{t\to 0}\left[\frac{f(p + tX(p) + o(t)) - f(p)}{t}\right] \\
&= \lim_{t\to 0}\left[\frac{f(p) + tDf(p)\cdot X(p) + o(t) - f(p)}{t}\right] \\
&= \sum_k \frac{\partial f}{\partial x^k}(p)X_k(p) + \lim_{t\to 0}\frac{o(t)}{t} \\
&= \sum_k \frac{\partial f}{\partial x^k}(p)X_k(p)
\end{aligned} \tag{A.5}$$

where we used the defining condition 30.16 for integral curves on line 2. If we now rewrite this equation as an operator equality, we obtain:

$$\boxed{\mathcal{L}_X = \sum_k X_k \frac{\partial}{\partial x^k}} \tag{A.6}$$

A.4.3 Explanation for formula 30.3.14

For vector fields we cannot just take the difference at two different points because the tangent spaces generally do not coincide. We can solve this by using the flow 30.3.7:

$$\mathcal{L}_X Y = \lim_{t\to 0}\frac{(T\sigma_t)^{-1}[X(\gamma_p(t))] - X(p)}{t} \tag{A.7}$$

where the $T\sigma_t$ is the differential 30.1.7 of the flow which satisfies $(T\sigma)^{-1} = T\sigma_{-t}$. To see that this definition makes sense we have to show that $(T\sigma_t)^{-1}[X(\gamma_p(t))] \in T_pM$. This goes as follows:

$$\begin{aligned}
(T\sigma_t)^{-1}[X(\gamma_p(t))](f) &= T\sigma_{-t}[X(\gamma_p(t))](f) \\
&= X(\sigma_{-t}\circ\gamma_p(t))(f\circ\sigma_{-t}) \\
&= X(\sigma_{-t}\circ\sigma_t(p))(f\circ\sigma_{-t}) \\
&= X(p)(f\circ\sigma_{-t}) \\
&\in T_pM
\end{aligned}$$

for all $f \in C^k(M, \mathbb{R})$. On line 3 we used the definition of the flow 30.3.7.

We can also rewrite the second term in the numerator of A.7 using the flow:

$$X(p) = X(\sigma_0(p)) = T\sigma_0(X)$$

Using the definition of the pushforward of vector fields 30.13 the Lie derivative can be rewritten as:

$$\begin{aligned}
\mathcal{L}_X Y &= \lim_{t\to 0}\frac{\sigma_{-t*}X(\gamma_p(t)) - \sigma_{0*}X(\gamma_p(0))}{t} \\
&= \frac{d}{dt}(\sigma_{-t*}X)(\gamma_p(t))\Big|_{t=0}
\end{aligned}$$

[2]The vector field $X(p) = (p, Y(p))$ where Y is a smooth vector field on $\mathbb{R}^n$ can also be identified with Y itself. This is implicitly done in the derivation by using the notation X for both vector fields.

Or finally by using the relation between pushforward and pullback 30.14 this becomes:

$$\boxed{\mathcal{L}_X Y = \frac{d}{dt}(\sigma_t^* X)(\gamma_p(t))\Big|_{t=0}} \tag{A.8}$$

A.4.4 Explanation of remark 30.4.16

Looking at formula 30.37 for the exterior derivative of a smooth function and remembering the definition of the gradient 20.2 we see that these two definitions appear very similar. The major difference lies in the fact that ∇f is a vector in $\mathbb{R}^3$ and df is a covector in $\mathbb{R}^{*3}$. However there exists an isomorphism between these spaces and so we can identify ∇f and df.

Similar relations hold for the rotor 20.9 and divergence 20.7, however here we have to use a different construction as we will be working with the spaces Λ^1 and Λ^2. However we can use the Hodge star 20.56 to obtain the correct dimensions.

Consider a vector $\vec{\boldsymbol{f}} = (f_1, f_2, f_3)$ where f_i is smooth. Using these functions f_i we can construct a 1-form $\alpha = f_1 dx_1 + f_2 dx_2 + f_3 dx_3$ and a 2-form $\omega = f_1 dx_2 \wedge dx_3 + f_2 dx_3 \wedge dx_1 + f_3 dx_1 \wedge dx_2$. After applying the exterior derivative (in the corresponding spaces) we obtain:

$$d\alpha = \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3}\right) dx_2 \wedge dx_3 + \left(\frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1}\right) dx_3 \wedge dx_1 + \left(\frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2}\right) dx_1 \wedge dx_2$$

$$d\omega = \left(\frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} + \frac{\partial f_3}{\partial x_3}\right) dx_1 \wedge dx_2 \wedge dx_3$$

Using result 20.8.23 and the isomorphism $\sim : \mathbb{R}^{3*} \to \mathbb{R}^3$ we can rewrite this as:

$$\boxed{\begin{aligned} \sim df &= \nabla f \qquad (A.9)\\ \sim (*d\alpha) &= \nabla \times \vec{\boldsymbol{f}} \qquad (A.10)\\ *d\omega &= \nabla \cdot \vec{\boldsymbol{f}} \qquad (A.11) \end{aligned}}$$

Appendix B

G-Structures

In the following table we give an overview of the more common G-structures one can define on a smooth (simply-connected) manifolds M^n.

Geometric structure	Structure group	Remarks
Orientation	$\mathrm{SL}_n(\mathbb{R})$	$\mathrm{GL}_n^+(\mathbb{R})$ is sufficient for orientability. The special linear group gives rise to a volume form.
Riemannian metric	$O(n)$	
Almost-symplectic structure*	$\mathrm{Sp}_{2k}(\mathbb{R})$	Integrability (in the form of a closed form) gives a symplectic manifold.
Almost-complex structure*	$\mathrm{GL}_k(\mathbb{C})$	Integrability (in the form of Newlander-Nirenberg) gives a complex manifold.
Almost-Hermitian structure*	$\mathrm{U}(k)$	Integrability gives a Kähler manifold.
Calabi-Yau*	$\mathrm{SU}(k)$	
Hyperkähler**	$\mathrm{Sp}(k)$	It follows that every hyperkähler implies Calabi-Yau.
Quaternionic-Kähler**[1]	$(\mathrm{Sp}(k) \times \mathrm{Sp}(1))/\mathbb{Z}_2$	These manifolds are not strictly Kähler since the structure group is not a subgroup of $\mathrm{U}(2k)$.

Structures marked with $*$ require the real dimension $n = 2k$ to be even. Structures marked with $**$ require the real dimension $n = 4k$ to be a multiple of 4.

Remark. This table is strongly related to the classification of (irreducible simply-connected non-symmetric) Riemannian manifolds by *Berger*.

[1] This definition additionally requires that $k > 1$.

Appendix C

Derivations: Lagrangian formalism

C.1 d'Alembert's principle

In the following derivation we assume a constant mass.

$$
\begin{aligned}
&\sum_k \left(\vec{\boldsymbol{F}}_k - \dot{\vec{\boldsymbol{p}}}_k\right) \dot{\vec{\boldsymbol{r}}}_k = 0 \\
\iff &\sum_k \left(\vec{\boldsymbol{F}}_k - \dot{\vec{\boldsymbol{p}}}_k\right) \cdot \left(\sum_l \frac{\partial \vec{\boldsymbol{r}}}{\partial q_l} \dot{q}_l\right) = 0 \\
\iff &\sum_l \left(\sum_k \vec{\boldsymbol{F}}_k \cdot \frac{\partial \vec{\boldsymbol{r}}}{\partial q_l} - \sum_k m \ddot{\vec{\boldsymbol{r}}} \cdot \frac{\partial \vec{\boldsymbol{r}}}{\partial q_l}\right) \dot{q}_l = 0 \\
\iff &\sum_l \left(Q_l - \sum_k m \ddot{\vec{\boldsymbol{r}}} \cdot \frac{\partial \vec{\boldsymbol{r}}}{\partial q_l}\right) \dot{q}_l = 0
\end{aligned} \tag{C.1}
$$

Now we look at the following derivative:

$$
\begin{aligned}
&\frac{d}{dt}\left(\dot{\vec{\boldsymbol{r}}} \cdot \frac{\partial \vec{\boldsymbol{r}}}{\partial q_l}\right) = \ddot{\vec{\boldsymbol{r}}} \cdot \frac{\partial \vec{\boldsymbol{r}}}{\partial q_l} + \dot{\vec{\boldsymbol{r}}} \cdot \frac{d}{dt}\left(\frac{\partial \vec{\boldsymbol{r}}}{\partial q_l}\right) \\
\iff &\ddot{\vec{\boldsymbol{r}}} \cdot \frac{\partial \vec{\boldsymbol{r}}}{\partial q_l} = \frac{d}{dt}\left(\dot{\vec{\boldsymbol{r}}} \cdot \frac{\partial \vec{\boldsymbol{r}}}{\partial q_l}\right) - \dot{\vec{\boldsymbol{r}}} \cdot \frac{d}{dt}\left(\frac{\partial \vec{\boldsymbol{r}}}{\partial q_l}\right) \\
\iff &\ddot{\vec{\boldsymbol{r}}} \cdot \frac{\partial \vec{\boldsymbol{r}}}{\partial q_l} = \frac{d}{dt}\left(\dot{\vec{\boldsymbol{r}}} \cdot \underbrace{\frac{\partial \vec{\boldsymbol{r}}}{\partial q_l}}_{A}\right) - \dot{\vec{\boldsymbol{r}}} \cdot \left(\frac{\partial \dot{\vec{\boldsymbol{r}}}}{\partial q_l}\right)
\end{aligned} \tag{C.2}
$$

To evaluate A we can take a look at another derivative:

$$
\begin{aligned}
\frac{\partial \dot{\vec{\boldsymbol{r}}}}{\partial \dot{q}_l} &= \frac{\partial}{\partial \dot{q}_l}\left(\sum_k \frac{\partial r}{\partial q_k} \dot{q}_k\right) \\
&= \sum_k \frac{\partial r}{\partial q_k} \delta_{kl} \\
&= \frac{\partial \vec{\boldsymbol{r}}}{\partial q_l} \\
&= A
\end{aligned}
$$

Substituting this in formula C.2 gives:

$$\begin{aligned}\ddot{\vec{\boldsymbol{r}}} \cdot \frac{\partial \vec{\boldsymbol{r}}}{\partial q_l} &= \frac{d}{dt}\left(\dot{\vec{\boldsymbol{r}}} \cdot \frac{\partial \dot{\vec{\boldsymbol{r}}}}{\partial \dot{q}_l}\right) - \dot{\vec{\boldsymbol{r}}} \cdot \left(\frac{\partial \dot{\vec{\boldsymbol{r}}}}{\partial q_l}\right) \\ &= \frac{d}{dt}\left(\frac{1}{2}\frac{\partial \dot{\vec{\boldsymbol{r}}}^2}{\partial \dot{q}_l}\right) - \frac{1}{2}\frac{\partial \dot{\vec{\boldsymbol{r}}}^2}{\partial q_l}\end{aligned} \tag{C.3}$$

If we multiply this by the mass m and sum over all particles we get :

$$\begin{aligned}\sum_k m_k \ddot{\vec{\boldsymbol{r}}}_k \cdot \frac{\partial \vec{\boldsymbol{r}}_k}{\partial q_l} &= \frac{d}{dt}\frac{\partial}{\partial \dot{q}_l}\left(\sum_k \frac{1}{2} m \dot{\vec{\boldsymbol{r}}}_k^2\right) - \frac{\partial}{\partial q_l}\left(\sum_k \frac{1}{2} m \dot{\vec{\boldsymbol{r}}}_k^2\right) \\ &= \frac{d}{dt}\frac{\partial T}{\partial \dot{q}_l} - \frac{\partial T}{\partial q_l}\end{aligned} \tag{C.4}$$

Where we have denoted the total kinetic energy in the last line as T.
Plugging this result into formula C.1 gives us:

$$\sum_l \left(Q_l - \frac{d}{dt}\frac{\partial T}{\partial \dot{q}_l} - \frac{\partial T}{\partial q_l}\right)\dot{q}_l = 0 \tag{C.5}$$

As all the q_l are independent the following relation should hold for all l:

$$\begin{aligned}& Q_l - \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_l}\right) - \frac{\partial T}{\partial q_l} = 0 \\ \iff & \boxed{\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_l}\right) - \frac{\partial T}{\partial q_l} = Q_l}\end{aligned} \tag{C.6}$$

This last equation is known as a **Lagrange equation of the first kind**.
If we have a system with only conservative forces acting on it, we can write the force on the i-th particle as:

$$F_i = -\nabla_i V \tag{C.7}$$

With this in mind, lets take a look at the derivative of the potential V with respect to the l-th generalized coordinate:

$$\begin{aligned}\frac{\partial V}{\partial q_l} &= \sum_i (\nabla_i V) \cdot \frac{\partial \vec{\boldsymbol{r}}_i}{\partial q_l} \\ &= -Q_l\end{aligned} \tag{C.8}$$

The differentiation of V with respect to any generalized velocity $\dot{q}_l$ is trivially zero. This combined with the last formula C.8 and with formula C.6 gives:

$$\begin{aligned}& \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_l}\right) - \frac{\partial T}{\partial q_l} = Q_l \\ \iff & \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_l}\right) - \frac{\partial T}{\partial q_l} = -\frac{\partial V}{\partial q_l} + \frac{\partial V}{\partial \dot{q}_l} \\ \iff & \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_l} - \frac{\partial V}{\partial \dot{q}_l}\right) - \frac{\partial}{\partial q_l}(T - V) = 0\end{aligned} \tag{C.9}$$

If we introduce a new variable L, called the **Lagrangian**, we get the **Lagrangian equation of the second kind**:

$$\boxed{\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_l}\right) - \frac{\partial L}{\partial q_l} = 0} \tag{C.10}$$

C.2 Hamilton's principle

In this part we start from the principle of least action. First we the define the **action** as following:

$$\boxed{I = \int_{t_1}^{t_2} L\left(y(t), \dot{y}(t), t\right) dt} \tag{C.11}$$

Then we require that this action is minimal for the physically acceptable path. To do this we define a family of paths:

$$y(t, \alpha) = y(t) + \alpha\eta(t) \tag{C.12}$$

Where $\eta(t)$ is an arbitrary function with the following boundary conditions:

$$\begin{cases} \eta(t_1) = 0 \\ \eta(t_2) = 0 \end{cases} \tag{C.13}$$

If we define the action integral over this family of paths, the integral C.11 becomes a function of α:

$$I(\alpha) = \int_{t_1}^{t_2} L\left(y(t, \alpha), \dot{y}(t, \alpha), t\right) dt \tag{C.14}$$

Requiring that the action integral is stationary for $y(t)$ (thus $\alpha = 0$) is equivalent to:

$$\left(\frac{dI}{d\alpha}\right)_{\alpha=0} = 0 \tag{C.15}$$

This condition combined with formula C.14 gives us:

$$\frac{dI}{d\alpha} = \int_{t_1}^{t_2} \frac{d}{d\alpha} L\left(y(t, \alpha), \dot{y}(t, \alpha), t\right) dt \tag{C.16}$$

As we evaluate this derivative in $\alpha = 0$ we can replace $y(t, \alpha)$ by $y(t)$ due to definition C.12.

$$\begin{aligned} \frac{dI}{d\alpha} &= \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial L}{\partial \dot{y}}\frac{\partial \dot{y}}{\partial \alpha}\right] dt \\ &= \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial y}\eta(t) + \frac{\partial L}{\partial \dot{y}}\dot{\eta}(t)\right] dt \end{aligned} \tag{C.17}$$

If we substitute $\frac{\partial L}{\partial \dot{y}} := h(t)$ and apply integration by parts to the second term in this integral, we get:

$$\begin{aligned} \frac{dI}{d\alpha} &= \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial y}\eta(x) + h(t)\dot{\eta}(t)\right] dt \\ &= \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial y}\eta(t) + h(t)\frac{d\eta}{dt}\right] dt \\ &= \int_{t_1}^{t_2} \frac{\partial L}{\partial y}\eta(t)dt + \eta(t_2)h(t_2) - \eta(t_1)h(t_1) - \int_{t_1}^{t_2} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}}\right)\eta(t)dt \end{aligned} \tag{C.18}$$

Due to the initial conditions C.15 for the function $\eta(t)$, the two terms in the middle vanish and we obtain:

$$\frac{dI}{d\alpha} = \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial y} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}}\right)\right]\eta(t)dt \tag{C.19}$$

Furthermore, as the function $\eta(t)$ was arbitrary, the only possible way that this derivative can become zero is when the integrand is identically zero:

$$\boxed{\frac{\partial L}{\partial y} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}}\right) = 0} \tag{C.20}$$

If we compare this result with formula C.10 we see that we can also obtain the **Lagrangian equations of the second kind** by starting from the principle of least action. (Where the variable y represents the generalized coordinates q_l and the variable $\dot{y}$ represents the generalized velocities $\dot{q}_l$)

Remark C.2.1. Differential equations of the form

$$\boxed{\frac{\partial f}{\partial y}(y, \dot{y}, x) = \frac{d}{dx}\left(\frac{\partial f}{\partial \dot{y}}(y, \dot{y}, x)\right)} \tag{C.21}$$

are known as **Euler-Lagrange equations**.

C.3 Explanation for Noether's theorem 43.3.1

The general transformation rule for the Lagrangian is:

$$\mathcal{L}(x) \to \mathcal{L}(x) + \alpha\delta\mathcal{L}(x) \tag{C.22}$$

To have a symmetry, i.e. keep the action invariant, the deformation factor has to be a 4-divergence:

$$\mathcal{L}(x) \to \mathcal{L}(x) + \alpha\partial_\mu \mathcal{J}^\mu(x) \tag{C.23}$$

To obtain formula 43.8 we vary the Lagrangian explicitly:

$$\begin{aligned}
\delta\mathcal{L} &= \frac{\partial \mathcal{L}}{\partial \phi}\delta\phi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\delta(\partial_\mu \phi) \\
&= \frac{\partial \mathcal{L}}{\partial \phi}\delta\phi + \partial_\mu\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\delta\phi\right) - \partial_\mu\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\right)\delta\phi \\
&= \partial_\mu\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\delta\phi\right) + \left[\frac{\partial \mathcal{L}}{\partial \phi} - \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\right]\delta\phi
\end{aligned}$$

The second term vanishes due to the Euler-Lagrange equation C.20. Combining these formulas gives us:

$$\partial_\mu\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\delta\phi\right) - \partial_\mu \mathcal{J}^\mu(x) = 0 \tag{C.24}$$

From this equation we can conclude that the current

$$\boxed{j^\mu(x) = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\delta\phi - \mathcal{J}^\mu(x)} \tag{C.25}$$

is conserved.

Appendix D

Derivations: Optics and material physics

D.1 Optics

D.1.1 Law of Lambert-Beer 46.9

From formula 46.8 we now that the complex refractive index can be written as

$$\widetilde{n} = n + ik$$

Where k is called the **extinction coefficient**.
From classical optics we also know that in a material the speed of light obeys the following relation:

$$c = \widetilde{n}v$$

Where we have used the complex refractive index. It readily follows that the wavenumber (sadly also given the letter k) can be written as:

$$k = \frac{\omega}{v} = \widetilde{n}\frac{\omega}{c}$$

From classical electromagnetism we know that a plane wave can be written as:

$$E(x,t) = Re\left\{A\,\exp\left[i(kx - \omega t + \phi)\right]\right\}$$

So everything put together we get:

$$E(x,t) = Re\left\{A\,\exp\left[i\left((n+ik)\frac{\omega}{c}x - \omega t + \phi\right)\right]\right\}$$

or also:

$$E(x,t) = Re\left\{A\,\exp\left[in\frac{\omega}{c}x\right]\cdot\exp\left[-k\frac{\omega}{c}x\right]\cdot\exp\left[-i\omega t\right]\cdot\exp\left[i\phi\right]\right\}$$

We also know that the intensity is given by the following relation:

$$I(x) = |E(x)|^2 = E^*(x)\cdot E(x)$$

So only the second factor will remain. Dividing this by its value for $x = 0$ we get:

$$\frac{I(x)}{I(0)} = \frac{E(x)\cdot E^*(x)}{E(0)\cdot E^*(0)} = \exp\left[-\frac{2k\omega}{c}x\right] = \exp[-\alpha x]$$

Where α is the absorption coefficient as defined in formula 46.10. □

Appendix E

Derivations: Classical and Statistical Mechanics

E.1 Moments of inertia

In this section we will always use formula 42.8 to calculate the moment of inertia.

E.1.1 Disk

The volume of a (solid) disk is given by:

$$V_{disk} = \pi R^2 d \tag{E.1}$$

where R is the radius and d is the thickness. The mass density is then given by:

$$\rho = \frac{M}{\pi R^2 d} \tag{E.2}$$

Using cylindrical coordinates the moment of inertia then becomes:

$$I = \frac{M}{\pi R^2 d} \int_0^{2\pi} d\varphi \int_0^d dz \int_0^R r^3 dr \tag{E.3}$$

$$= \frac{M}{\pi R^2 d} 2\pi d \frac{R^4}{4} \tag{E.4}$$

$$= \frac{1}{2} M R^2 \tag{E.5}$$

E.1.2 Solid sphere

The volume of a solid sphere is given by:

$$V_{sphere} = \frac{4}{3}\pi R^3 \tag{E.6}$$

where R is the radius. The mass density is then given by:

$$\rho = \frac{M}{\frac{4}{3}\pi R^3} \tag{E.7}$$

We will use spherical coordinates to derive the moment of inertia, but we have to be carefull. The r in formula 42.8 is the distance between a point in the body and the axis of rotation. So it is not the same as the r in spherical coordinates which is the distance between a point and the origin. However the relation between these two quantities is easily found using basic geometry to be:

$$r = r' \sin\theta \tag{E.8}$$

where r' is the spherical coordinate. Now we can calculate the moment of inertia as follows:

$$I = \frac{M}{\frac{4}{3}\pi R^3} \int_0^{2\pi} d\varphi \int_0^R r'^4 dr' \int_0^\pi \sin^3\theta d\theta \tag{E.9}$$

$$= \frac{M}{\frac{4}{3}\pi R^3} 2\pi \frac{R^5}{5}\frac{4}{3} \tag{E.10}$$

$$= \frac{2}{5} M R^2 \tag{E.11}$$

E.2 Schottky defects

Let E_v be the energy needed to remove a particle from its lattice point and move it to the surface. We will neglect any surface effects and assume that the the energy E_v is independent of the distance to the surface.

The total energy of all vacancies is then given by $E = nE_v$. The number of possible microstates is

$$\Omega = \frac{(N+n)!}{N!n!} \tag{E.12}$$

where we used the fact that the removal of n particles creates n more lattice points at the surface. Using Boltzmann's entropy formula 64.1 and Stirling's formula we obtain

$$S(N,n) = k\ln\Omega = k\big[(N+n)\ln(N+n) - n\ln n - N\ln N\big] \tag{E.13}$$

Using 64.2 we can find the temperature:

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{N,V} = \frac{dS}{dn}\frac{dn}{dE} = \frac{k}{E_v}\ln\frac{N+n}{n} \tag{E.14}$$

which can be rewritten as

$$\boxed{\frac{n}{N+n} = \exp\left(-\frac{E_v}{kT}\right)} \tag{E.15}$$

The density of Frenkel defects can be derived analogously.

Appendix F

Units and symbols

pk	745.7 W

Table F.1: Units

List of Symbols

The following symbols are used throughout the summary:

Abbreviations

CR	Cauchy-Riemann
TQFT	Topological quantum field theory
TVS	Topological vector space

Operations

Ad_g	Adjoint representation of a Lie group G.
ad_X	Adjoint representation of a Lie algebra $\mathfrak{g}$.
arg	Argument of a complex number.
e	Identity element of a group.
Im	Imaginary part of a complex number.
$\mathrm{Ind}_f(z)$	Index of a point $z \in \mathbb{C}$ with respect to a function f.
Par_t^γ	Parallel transport map with respect to the curve γ.
Re	Real part of a complex number.
Res	Residue of a complex function.
$\{\cdot,\cdot\}$	Poisson bracket
$\sphericalangle(v,w)$	Angle between the vectors v and w.
$X \times Y$	Cartesian product of the sets X and Y.
$X + Y$	Sum of the vector spaces X and Y.
$X \oplus Y$	Direct sum of the vector spaces X and Y.
$X \otimes Y$	Tensor product of the vector spaces X and Y.
$\mathbb{1}_X$	Identity map on the set X.
$\approx$	is approximately equal to
$\hookrightarrow$	is included in
$\cong$	is isomorphic to
$\mapsto$	mapsto

Collections

Ab	Category of Abelian groups.
$\mathrm{Aut}(V)$	Set of automorphisms (invertible endomorphisms) on a set V.
$\mathcal{B}_0(X,Y)$	Space of compact bounded oeprators between Banach spaces.
$\mathcal{B}(V,W)$	Space of bounded linear maps from the space X to the space Y.
$C_p^\infty(M)$	Ring of all smooth functions $f : M \to \mathbb{R}$ defined on a neighbourhood of $p \in M$.

$\mathrm{C}^\omega(V)$	Set of all analytic function on a set V.
Diff	Category of smooth manifolds.
D^n	Standard n-disk
$\mathrm{End}(V)$	Ring of endomorphisms on a set V.
$\mathcal{E}\mathrm{nd}$	Endomorphism operad
$\mathrm{GL}(V)$	General linear group: group of all automorphisms on a vector space V.
$\mathrm{GL}_n(K)$	General linear group: group of all invertible n-dimensional matrices over the field K.
Grp	Category of groups and group homomorphisms.
Grpd	Category of groupoids.
$\mathrm{Hol}_p(\omega)$	Holonomy group at p with respect to the connection ω.
$\mathrm{Hom}(V, W)$	Set of morphisms from a set V to a set W.
hTop	Homotopy category
$K^0(X)$	K-theory over a (compact Hausdorff) space X.
LX	Free loop space on X.
Man^p	Category of C^p-manifolds.
$\mathrm{Open}(X)$	Category of open subsets of a topological space X.
$P(S), 2^S$	Power set of S.
$\mathrm{Pin}(V)$	Pin group of the Clifford algebra $C\ell(V, Q)$.
$\mathrm{Sh}(X)$	Category of sheafs over a topological space X.
$\mathrm{SL}(2, \mathbb{C})$	Special linear group of dimension 2 over the field of complex numbers.
S^n	Standard n-sphere
$S^n(V)$	Space of symmetric rank n tensors over a vector space V.
$\mathrm{Spec}(R)$	Spectrum of a commutative ring R.
$\mathrm{Sp}_n(K)$	Symplectic group: Group of matrices preserving the canonical symplectic form over the field K.
$\mathrm{Sp}(n)$	Compact symplectic group: $\mathrm{Sp}_{2n}(\mathbb{C}) \cap \mathrm{U}(2n)$.
$\mathrm{TL}_n(\delta)$	Temperley-Lieb algebra with $n-1$ generators and parameter δ.
T^n	Standard n-torus. Cartesian product of n times S^1.
Top	Category of topological spaces.
$U(\mathfrak{g})$	Universal enveloping algebra of a Lie algebra $\mathfrak{g}$.
$\mathrm{Vect}(X)$	Category of vector bundles over a topological space X.
Y^X	Set of functions from a set X to a set Y.
$\emptyset$	Empty set
$\pi_n(X, x_0)$	n^{th} homotopy space over X with basepoint x_0.
$[a, b]$	Closed interval
$]a, b[$	Open interval
$\Lambda^n(V)$	Space of antisymmetric rank n tensors over a vector space V.
ΩX	(Based) loop space on X.
$\Omega^k(M)$	$C^\infty(M)$-module of differential k-forms on the manifold M.
$\rho(A)$	Resolvent set of a bounded linear operator A.

$\mathfrak{X}(M)$	$C^\infty(M)$-module of vector fields on the manifold M.

Bibliography

[1] Charles A. Weibel. *An Introduction to Homological Algebra*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1994.

[2] P.J. Hilton and U. Stammbach. *A Course in Homological Algebra*. Springer.

[3] W. S. Massey. *A Basic Course in Algebraic Topology*. Springer.

[4] W. J. Merry. Algebraic Topology. `https://www.merry.io/algebraic-topology`.

[5] J. L. Brylinski. *Loop Spaces, Characteristic Classes and Geometric Quantization*. Birkhauser.

[6] W. Greiner. *Quantum Mechanics*. Springer.

[7] J. H. Heinbockel. Branch Cuts. `http://www.math.odu.edu/~jhh/ch45.PDF`.

[8] N. Jeevanjee. *An Introduction to Tensors and Group Theory for Physicists*. Birkhauser.

[9] S. B. Sontz. *Principal Bundles: The Classical Case*. Springer.

[10] W. Fulton and J. Harris. *Representation Theory: A First Course*. Springer.

[11] S. B. Sontz. *Principal Bundles: The Quantum Case*. Springer.

[12] Y. Choquet-Bruhat, C. DeWitt-Morette, and M. Dillard-Bleick. *Analysis, Manifolds and Physics, Part 1: Basics*. North-Holland.

[13] Y. Choquet-Bruhat and C. DeWitt-Morette. *Analysis, Manifolds and Physics, Part 2*. North-Holland.

[14] I. Kolar, P. W. Michor, and J. Slovak. *Normal Operations in Differential Geometry*. Springer.

[15] C. Nash and S. Sen. *Topology and Geometry for Physicists*. Dover Publications.

[16] F. Schuller. Lectures on the Geometric Anatomy of Theoretical Physics. `https://www.youtube.com/channel/UC6SaWe7xeOp31Vo8cQG1oXw`.

[17] M. E. Peskin and D. V. Schroeder. *An Introduction to Quantum Field Theory*. Westview Press.

[18] W. Greiner and J. Reinhardt. *Field Quantization*. Springer.

[19] A. Van Proeyen and Daniel Freedman. *Supergravity*. Cambridge University Press.

[20] F. Belgun. Gauge Theory. `http://www.math.uni-hamburg.de/home/belgun/Gauge4.pdf`.

[21] John E. Roberts. Spontaneously Broken Gauge Symmetries and Superselection Rules. 1974.

[22] B. H. Bransden and C. J. Joachain. *Quantum Mechanics*. Prentice Hall.

[23] Safko Goldstein and Poole. *Classical Mechanics*. Pearson.

[24] F. Cardin. *Elementary Symplectic Topology and Mechanics*. Springer.

[25] H. Radjavi and P. Rosenthal. *Invariant Subspaces*. Dover Publications.

[26] M. Karoubi. *K-Theory: An Introduction*. Springer.

[27] P. Petersen. *Riemannian Geometry*. Springer.

[28] I. M. Anderson. *The Variational Bicomplex*.

[29] S. Cordes, G. Moore, and S. Ramgoolam. Lectures on 2D Yang-Mills Theory, Equivariant Cohomology and Topological Field Theories. arXiv:hep-th/9411210v2.

[30] D. C. Ferguson. A Theorem of Looman-Menchoff. `http://digitool.library.mcgill.ca/thesisfile111406.pdf`.

[31] H. Lyre. Berry Phase and Quantum Structure. arXiv:1408.6867.

[32] V. Itskov, P. J. Olver, and F. Valiquette. Lie Completion of Pseudogroups. *Transformation Groups*, 16:161--173, 2011.

[33] R. Borcherds. Lie Groups. `https://math.berkeley.edu/~reb/courses/261/`.

[34] A. Losev. From berezin integral to batalin-vilkovisky formalism: A mathematical physicist's point of view. 04 2007.

[35] B. Iliev. Normal Frames for General Connections on Differentiable Fibre Bundles. arXiv:math/0405004.

[36] P. Stachura. Short and Biased Introduction to Groupoids. arXiv:1311.3866.

[37] O. Caramello. Lectures on Topos Theory at the University of Insubria. `https://www.oliviacaramello.com/Teaching/Teaching.htm`.

[38] Stacks Project. `https://stacks.math.columbia.edu/`.

[39] The nLab. `https://ncatlab.org/nlab`.

[40] Wikipedia. `https://www.wikipedia.org/`.

[41] E. Jamsin and J. Palmkvist. Modave lectures V: Lie and Kac-Moody Algebras. `http://www.ulb.ac.be/sciences/ptm/pmif/Rencontres/ModaveV/ModavelecturesFinalElla.pdf`.

Index