Summary: Mathematics & Physics

Nicolas Dewolf

November 26, 2017

Contents

Chapter 1

Introduction

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 **bold text**. 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/crucial 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) 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.

Part I

Set Theory

Chapter 2

Set theory

2.1 Collections

Definition 2.1.1 (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 2.1.2. $S \subset P(S)$

Definition 2.1.3 (Collection). Let A be a set. A collection of elements in A is a subset of A.

Definition 2.1.4 (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 $f : I \rightarrow A$. 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 2.1.5 (Helly family). A Helly family of order k is a pair (X, F) with $F \subset 2^X$ such that for every finite $G \subset F$:

$$\bigcap_{V\in G} V = \emptyset \implies \exists H \subset G : \bigcap_{V\in H} V = \emptyset \wedge |H| \leq k \tag{2.1}$$

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

Definition 2.1.6 (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 2.1.7 (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 2.1.8 (Cover). A cover of S is a collection of sets $\mathcal{F} \subseteq 2^S$ such that

$$\bigcup_{V\in\mathcal{F}} V = S \tag{2.2}$$

2.2 Set operations

Definition 2.2.1 (Symmetric difference).

$$A\Delta B = (A\backslash B) \cup (B\backslash A) \tag{2.3}$$

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

$$E^c = \Omega\backslash E \tag{2.4}$$

Formula 2.2.3 (de Morgan's laws).

$$\left(\bigcup_i A_i\right)^c = \bigcap_i A_i^c \tag{2.5}$$

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

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.*

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

2.3.2 Lattices

Definition 2.3.8 (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 2.3.9. The join of $\{a, b\}$ is denoted by $a \wedge b$. The meet of $\{a, b\}$ is denoted by $a \vee b$.

Definition 2.3.10 (Lattice). A poset $(P, \leq)$ is called a lattice if it is both a join- and a meet-semilattice.

Definition 2.3.11 (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 2.3.12 (Net). A net on a topological space X is a subset of X indexed by a directed set I.

2.3.3 Bounded sets

Definition 2.3.13 (Supremum). The supremum $\sup(X)$ of a set X is the smallest upper bound of X.

Definition 2.3.14 (Infimum). The infimum $\inf(X)$ of a set X is the greatest lower bound of X.

Definition 2.3.15 (Maximum). If $\sup(X) \in X$ the supremum is called the maximum of X. This is denoted by $\max(X)$.

Definition 2.3.16 (Minimum). If $\inf(X) \in X$ the supremum is called the minimum of X. This is denoted by $\min(X)$.

2.3.4 Real numbers

Property 2.3.17 (First axiom). The set of real numbers is an ordered field $(\mathbb{R}, +, \cdot, <)$

Property 2.3.18 (Completeness axiom[3]). Every non-empty subset of $\mathbb{R}$ that is bounded above has a supremum.

Property 2.3.19. $\mathbb{Q} \subset \mathbb{R}$

[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**.

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

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 2.3.20 (Extended real line).

$$\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\} = [-\infty, \infty] \tag{2.7}$$

2.3.5 Filter

Definition 2.3.21 (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[4] 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.

[4]See equations 2.5 and 2.6.

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

$$\mathcal{G} = \bigcap\{\mathcal{F} : \mathcal{F} \text{ is a } \sigma\text{-algebra that contains } \mathcal{A}\} \tag{2.8}$$

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 $\sigma(\mathcal{A})$.

Definition 2.4.8 (Borel set). Let $\mathcal{B}$ be the σ-algebra generated by all open[5] 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 9.1.32):
$$\mathrm{Pr}_1 : \Omega \rightarrow \Omega_1 : (\omega_1, \omega_2) \mapsto \omega_1$$
$$\mathrm{Pr}_2 : \Omega \rightarrow \Omega_2 : (\omega_1, \omega_2) \mapsto \omega_2$$

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 $\mathcal{A}$.*

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

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.

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:

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

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

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:

$$L_c(f) = \{x \in X : f(x) = c\} \tag{2.9}$$

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**.

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 $\star$

Definition 3.1.3 (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 3.1.4 (Commutative group[1]). Let $(G, \star)$ be a group. If $\star$ is commutative, then G is called a commutative group.

[1] Also called an Abelian group.

3.1.1 Cosets

Definition 3.1.5 (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

$$gH = \{gh : h \in H\} \tag{3.1}$$

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**. The sets of left and right cosets are denoted by G/H and $H\backslash G$ respectively.

Definition 3.1.6 (Quotient group). Let G be a group and N a normal subgroup. The quotient group G/N is defined as the set of cosets of N in G. This set can be turned into a group itself 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 3.1.7 (Center). The center of a group is defined as follows:

$$Z(G) = \{z \in G : \forall g \in G, zg = gz\} \tag{3.2}$$

This set is a normal subgroup of G.

3.1.2 Order

Definition 3.1.8 (Order of a group). The number of elements in the group. It is denoted by $|G|$ or $\mathrm{ord}(G)$.

Definition 3.1.9 (Order of an element). The order of an element $a \in G$ is the smallest integer n such that

$$a^n = e \tag{3.3}$$

where e is the identity element of G.

Definition 3.1.10 (Torsion group). A torsion group is a group for which all element have finite order. The torsion set $\mathrm{Tor}(G)$ of a group G is the set of all elements $a \in G$ that have finite order. For Abelian groups, $\mathrm{Tor}(G)$ is a subgroup.

3.1.3 Symmetric and alternating groups

Definition 3.1.11 (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 $\mathrm{Sym}(X)$ of a finite set X is analogously defined.

Definition 3.1.12 (Alternating group). The alternating group A_n is the subgroup of S_n containing all even permutations.

Definition 3.1.13 (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 3.1.14. Let τ be a k-cycle. Then τ is k-cyclic (hence the name *cycle*):

$$\tau^k = \mathbb{1}_G \tag{3.4}$$

Example 3.1.15. 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 3.1.16 (Transposition). A permutation which exchanges two elements but lets the other ones unchanged.

3.1.4 Direct product

Definition 3.1.17 (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:

$$(g_1, h_1) \cdot (g_2, h_2) = (g_1 g_2, h_1 h_2) \tag{3.5}$$

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.

Notation 3.1.18. When the groups are Abelian, the direct product is sometimes called the **direct sum** and is denoted by $\oplus$.

Definition 3.1.19 (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 $N \cap H = \{\mathbb{1}\}$.
- 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 homomorphism $\rho : G \to H$ which satisfies $\rho|_H = \mathbb{1}$ 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.

Definition 3.1.20 (Outer semidirect product). Let G, H be two groups and let $\varphi : H \to \mathrm{Aut}(G)$ be a group homomorphism. 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:

$$(g_1, h_1) \cdot (g_2, h_2) = (g_1 \varphi(h_1)(g_1), h_1 h_2) \tag{3.6}$$

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

By noting that the set $N = \{(g, \mathbb{1}_H)|g \in G\}$ is a normal subgroup isomorphic to G and that the set $B = \{(\mathbb{1}_G, h)|h \in H\}$ is a subgroup isomorphic to H we can also construct the outer semidirect product $G \rtimes_\varphi H$ as the inner semidirect product $N \rtimes H$.

Remark 3.1.21. The direct product of groups is a special case of the outer semidirect product where the group homomorphism is given by the trivial map $\varphi : h \mapsto \mathbb{1}_G$.

3.1.5 Free groups

Definition 3.1.22 (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:

$$G = \left\{ \sum_i a_i g_i \,\middle|\, a_i \in \mathbb{Z} \right\} \tag{3.7}$$

The set of generators $\{g_i\}_{i\in I}$ is then called a **basis**[2] of G. The number of elements in the basis is called the **rank** of G.

Property 3.1.23. Consider a free group G. Let $H \subset G$ be a subgroup. Then H is also free.

Theorem 3.1.24. *Let G be a finitely generated Abelian group of rank n, i.e. its basis has n elements. This group can be constructed in two different ways:*

$$G = F/H \tag{3.8}$$

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

$$G = A \oplus T \qquad \text{where} \qquad T = Z_{h_1} \oplus \cdots \oplus Z_{h_m} \tag{3.9}$$

where A is a freely 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[3]*. The rank $n - m$ and the numbers h_i are unique.*

3.1.6 Group presentations

Definition 3.1.25 (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 3.1.26 (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.

[2]In analogy with the basis of a vector space.

[3]See also definition 3.1.10.

Definition 3.1.27 (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.

3.1.7 Group actions

Definition 3.1.28 (Group homomorphism). A group homomorphism $\Phi : G \to H$ is a map satisfying $\forall g, h \in G$

$$\Phi(gh) = \Phi(g)\Phi(h) \tag{3.10}$$

Definition 3.1.29 (Kernel). The kernel of a group homomorphism $\Phi : G \to H$ is defined as the set

$$K = \{g \in G : \Phi(g) = \mathbb{1}_H\} \tag{3.11}$$

Theorem 3.1.30 (First isomorphism theorem). *Let G, H be a groups and let $\varphi : G \to H$ be a group homomorphism. If φ is surjective than $G/\ker\varphi \cong H$.*

Definition 3.1.31 (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 3.1.32. A group action can alternatively be defined as a group homomorphism from G to $\mathrm{Sym}(V)$. It assigns a permutation of V to every element $g \in G$.

Notation 3.1.33. The action $\rho(g, v)$ is often denoted by $g \cdot v$ or even gv.

Definition 3.1.34 (Orbit). The orbit of an element $x \in X$ with respect to a group G is defined as the set:

$$x \cdot G = \{x \cdot g | g \in G\} \tag{3.12}$$

Definition 3.1.35 (Stabilizer). The stabilizer group or **isotropy group** of an element $x \in X$ with respect to a group G is defined as the set:

$$G_x = \{g \in G | g \cdot x = x\} \tag{3.13}$$

This is a subgroup of G.

Definition 3.1.36 (Free action). A group action is free if $g \cdot x = x$ implies $g = e$ for every $x \in X$.

Definition 3.1.37 (Faithful action[4]). A group action is faithful if the homomorphism $G \to \mathrm{Sym}(X)$ 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 $g \cdot x \neq h \cdot x$.

[4] A faithful action is also called an **effective** action.

Definition 3.1.38 (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 3.1.39 (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 3.1.40 (†). Let X be a set and let G be a group such that the action of G on X is transitive. Then their exists a bijection $X \cong G/G_x$ where G_x is the stabilizer of any element $x \in X$.

Definition 3.1.41 (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):

$$g \cdot (a + b) = g \cdot a + g \cdot b \tag{3.14}$$

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

Definition 3.1.42 (G-module homomorphism). A G-module homomorphism is a map $f : V \to W$ satisfying

$$g \cdot f(v) = f(g \cdot v) \tag{3.15}$$

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**.

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.

3.2.1 Ideals

Definition 3.2.3 (Ideal). Let $(R, +, \cdot)$ be a ring with $(R, +)$ its additive group. A subset $I \subseteq R$ is called an ideal[5] 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 3.2.4 (Unit ideal). Let $(R, +, \cdot)$ be a ring. R itself is called the unit ideal.

Definition 3.2.5 (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 3.2.6 (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:

$$ab \in I \implies \mathrm{a} \in I \vee b \in I \tag{3.16}$$

Definition 3.2.7 (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 3.2.8 (Minimal ideal). A proper ideal is said to be minimal if it contains no other nonzero ideal.

Construction 3.2.9 (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:

$$I = \left\{ \sum_{i=1}^{n} l_i x_i r_i \;\middle|\; \forall i \leq n : l_i, r_i \in R \text{ and } x_i \in X \right\} \tag{3.17}$$

Left and right ideals are generated in a similar fashion.

3.2.2 Modules

Definition 3.2.10 (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 15.2.1 but where the scalars are only elements of a ring instead of a field.

Property 3.2.11. For a general R-module the existence of a basis is not guaranteed unless R is a division ring. See construction 15.2.8 to see how this basis can be constructed.

Corollary 3.2.12. As every field is in particular a division ring, the existence of a basis follows from the above property for R-modules.

Definition 3.2.13 (Free module). A module is said to be free if it admits a basis.

Definition 3.2.14 (Projective module). A module P is said to be projective if:

$$P \oplus M = F \tag{3.18}$$

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

[5]More generally: two-sided ideal

3.2.3 Graded rings

Definition 3.2.15 (Graded ring). Let R be a ring that can be written as the direct sum of Abelian groups A_k:

$$R = \bigoplus_{k\in\mathbb{N}} A_k \tag{3.19}$$

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

$$vw = (-1)^{mn} wv \tag{3.20}$$

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[6]. 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 homomorphisms 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). Consider a direct system (A_i, f_{ij}) over a (directed) set I. The direct limit A of these direct systems is defined as follows:

$$\varinjlim A_i = \bigsqcup_{i\in I} A_i \Big/ \sim \tag{3.21}$$

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.

[6]See definition 2.3.11.

3.3.2 Exact sequences

Definition 3.3.3 (Exact sequence). Consider a sequence (finite or infinite) of algebraic structures and their corresponding homomorphisms:

$$A_0 \xrightarrow{\Phi_1} A_1 \xrightarrow{\Phi_2} \cdots \xrightarrow{\Phi_n} A_n \tag{3.22}$$

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 $h \in \mathbb{N}$.

Definition 3.3.4 (Short exact sequence). A short exact sequence is an exact sequence of the form:

$$0 \to A_0 \xrightarrow{\Phi_1} A_1 \xrightarrow{\Phi_2} A_3 \to 0 \tag{3.23}$$

A long exact sequence is an infinite exact sequence.

Property 3.3.5. 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).

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.

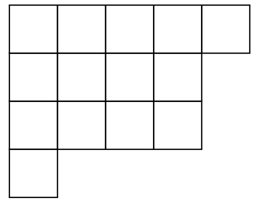

Figure 3.1: A Young diagram representing the partition $(5, 4, 4, 1)$ of 14.

Definition 3.4.3 (Young diagram[7]). 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.

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. Using the diagram 3.1 we obtain the conjugate partition $(4, 3, 3, 3, 1)$ represented by

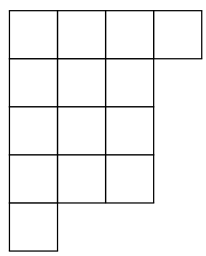

Figure 3.2: A Young diagram representing the partition $(4, 3, 3, 3, 1)$ of 14.

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:

$$\Lambda = (\Lambda_1, ..., \Lambda_m; \Lambda_{m+1}, ..., \Lambda_N) \tag{3.24}$$

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 $n \equiv |\Lambda| = \sum_{i=1}^{N}$.

[7] 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

Topology

Chapter 4

Topology

4.1 Topological spaces

Definition 4.1.1 (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.

Definition 4.1.2 (Relative topology). 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, called the relative topology:

$$\tau_{\text{rel}} = \{U_i \cap Y : U_i \in \tau_X\} \tag{4.1}$$

Definition 4.1.3 (Disjoint union). Let $\{X_i\}_{i \in I}$ be a family of topological spaces. Now consider the disjoint union

$$X = \bigsqcup_{i \in I} X_i \tag{4.2}$$

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:

$$\tau_X = \{U \subseteq X | \forall i \in I : \phi_i^{-1}(U) \text{ is open in } X_i\} \tag{4.3}$$

Definition 4.1.4 (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:

$$\tau_\sim = \{U \subseteq X/_\sim | \pi^{-1}(U) \text{ is open in } X\} \tag{4.4}$$

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

Example 4.1.5 (Discrete topology). The discrete topology is the topology such that every subset is open (and thus also closed).

Example 4.1.6 (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:

$$\mathcal{B} = \left\{ \prod_{i\in I} U_i \,\middle|\, U_i \in \tau_i \right\} \tag{4.5}$$

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

Definition 4.1.8 (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.

Definition 4.1.9 (Suspension). Let X be a topological space. The suspension of X is defined as the following quotient space:

$$SX = (X \times [0,1])/\{(x,0) \sim (y,0) \text{ and } (x,1) \sim (y,1) | x, y \in X\} \tag{4.6}$$

4.2 Neighbourhoods

4.2.1 Neighbourhoods

Definition 4.2.1 (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 $a \in U \subseteq V$.

Definition 4.2.2 (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:

$$U = \bigcup_{V \in \mathcal{F}} V \tag{4.7}$$

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

Definition 4.2.3 (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 4.2.4 (First-countability). A topological space (Ω, τ) is first-countable if for every point $x \in \Omega$ there exists a countable local basis.

Property 4.2.5 (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 4.2.6 (Second-countability). A topological space (Ω, τ) is second-countable if there exists a countable global basis.

Property 4.2.7. Let X be a topological space. The closure of a subset V is given by:

$$\overline{V} = \{x \in X | \exists \text{ a net } (x_\lambda)_{\lambda \in I} \text{ in } X : x_\lambda \to x\} \tag{4.8}$$

This implies that the topology on X is completely determined by the convergence of nets[1].

Corollary 4.2.8. In first-countable spaces we only have to consider the convergence of sequences.

Definition 4.2.9 (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 4.2.10. 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.

4.2.2 Separation axioms

Definition 4.2.11 (T_0 axiom[2]). 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 4.2.12 (T_1 axiom[3]). 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.

Definition 4.2.13 (Hausdorff space). A topological space is a Hausdorff space or T_2 space if it satisfies the following axiom:

$$(\forall x, y \in \Omega)(\exists \text{ neighbourhoods } U, V)(x \in U, y \in V, U \cap V = \emptyset) \tag{4.9}$$

This axiom is called the **Hausdorff separation axiom** or T_2 axiom. The points are said to be separated by neighbourhoods.

[1] See definition 2.3.12.
[2] T_1 spaces are also said to carry the **Kolmogorov topology**.
[3] T_1 spaces are also said to carry the **Fréchet topology**.

Property 4.2.14. Every singleton (and thus also every finite set) is closed in a Hausdorff space.

Definition 4.2.15 (Urysohn space[4]). A topological space is an Urysohn space if every two distcinct points are separated by closed neighbourhoods.

Definition 4.2.16 (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 4.2.17 (T_3 axiom). A space that is both regular and T_0 is T_3.

Definition 4.2.18 (Normal space). A topological space is said to be normal if every two closed subsets have disjoint neighbourhoods.

Definition 4.2.19 (T_4 axiom). A space that is both normal and T_1 is T_4.

Property 4.2.20. A space satisfying the separation axiom T_k also satisfies all separation axioms $T_{i\leq k}$.

4.3 Convergence and continuity

4.3.1 Convergence

Definition 4.3.1 (Convergence). A sequence $(x_n)_{n\in\mathbb{N}}$ in X is said to converge to a point $a \in X$ if:

$$(\forall \text{ neighbourhoods } V \text{ of } a)(\exists\ N > 0)(\forall\ n > N)(x_n \in V) \tag{4.10}$$

Property 4.3.2. Every subsequence of a converging sequence converges to the same point[5].

Property 4.3.3. Let X be a Hausdorff space. The limit of a converging sequence in X is unique.

4.3.2 Continuity

Definition 4.3.4 (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 4.3.5. *Let X be a first-countable space. Consider a function $f : X \to Y$. The following statements are equivalent:*

- *f is continuous*

[4] Sometimes called a $T_{2\frac{1}{2}}$ space.

[5] This limit does not have to be unique. See the next property for more information.

- *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* $a \in X$.

Corollary 4.3.6. 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 4.3.3).

Remark 4.3.7. If the space X is not first-countable, we have to consider the convergence of nets 2.3.12.

Theorem 4.3.8 (Urysohn's lemma). *A topological space X is normal*[6] *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.*

$$f(a) = 0, \forall a \in A \qquad\qquad f(b) = 1, \forall b \in B \tag{4.11}$$

Theorem 4.3.9 (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.

4.3.3 Homeomorphisms

Definition 4.3.10 (Homeomorphism). A function f is called a homeomorphism if both f and f^{-1} are continuous and bijective.

Definition 4.3.11 (Embedding). A function is an embedding if it is homeomorphic onto its image.

4.4 Connectedness

Definition 4.4.1 (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 4.4.2. 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 4.4.3 (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* $c \in f(X)$.

[6] See definition 4.2.18.

Definition 4.4.4 (Path-connected space[7]). 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 4.4.5. Every path-connected space is connected.

The converse does not hold. There exists however the following (stronger) relation:

Property 4.4.6. A connected and locally path-connected space is path-connected.

Remark 4.4.7. 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.

4.5 Compact spaces

4.5.1 Compactness

Definition 4.5.1 (Sequentially compact). A topological space is sequentially compact if every sequence[8] has a convergent subsequence.

Definition 4.5.2 (Finite intersection property). A family $\mathcal{F} \subseteq 2^X$ of subsets has the finite intersection property[9] if every finite subfamily has a non-zero intersection:

$$\bigcap_{i\in I} V_i \neq \emptyset \tag{4.12}$$

for all finite index sets I.

Definition 4.5.3 (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 4.5.4. A first-countable space is sequentially compact if and only if every countable open cover has a finite subcover.

Definition 4.5.5 (Lindelöf space). A space for which every open cover has a countable subcover.

Property 4.5.6. Every second-countable space is also a Lindelöf space.

Definition 4.5.7 (Compact space). A topological space X is compact if every open cover of X has a finite subcover.

[7] A similar notion is that of **arcwise-connectedness** where the function φ is required to be a homeomorphism.

[8] The sequence itself does not have to converge.

[9] The family is then called a FIP-family.

Theorem 4.5.8 (Heine-Borel[10]). *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 4.5.9 (Heine-Borel on $\mathbb{R}^n$). *A subset of $\mathbb{R}^n$ is compact if and only if it is closed and bounded.*

Theorem 4.5.10 (Tychonoff's theorem). *Any product[11] of compact topological spaces is again compact when equipped with the (Tychonoff) product topology 4.1.6.*

Definition 4.5.11 (Locally compact). A topological space is locally compact if every point $x \in X$ has a compact neighbourhood.

Theorem 4.5.12 (Dini's theorem). *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 4.5.13 (Paracompact space). A topological space is paracompact if every open cover has a locally finite open refinement.

Property 4.5.14 (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 4.5.15 (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 4.5.16. 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.

4.5.2 Compactifications

Definition 4.5.17 (Dense). A subset $V \subseteq X$ is dense in a topological space X if $\overline{V} = X$.

Definition 4.5.18 (Separable space). A topological space is separable if it contains a countable dense subset.

Property 4.5.19. Every second-countable space is separable.

Definition 4.5.20 (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 4.5.21. 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.

[10] Also Borel-Lebesgue.

[11] Finite, countably infinite or even uncountably infinite.

Remark. It is important to note that compactifications are not unique.

Definition 4.5.22 (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.

4.6 Homotopy theory

4.6.1 Homotopy

Definition 4.6.1 (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 4.6.2 (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 4.6.3. Every homeomorphism is a homotopy equivalence.

Definition 4.6.4 (Null-homotopic). A continuous function is null-homotopic if it is homotopic to a constant function.

Definition 4.6.5 (Contractible space). A topological space X is said to be contractible if the identity map $\mathbb{1}_X$ is null-homotopic. Equivalently, the space is homotopy-equivalent to a point.

4.6.2 Fundamental group

In this subsection we will always assume to be working with pointed spaces 4.1.8. The base point will be denoted by x_0.

Definition 4.6.6 (Loop space). The set of all **loops** in X, i.e. all continuous functions $\delta : [0,1] \to X$ for which $\delta(0) = \delta(1)$. It is denoted by ΩX. This set can be equipped with a multiplication operation corresponding to the concatenation of loops[12].

Definition 4.6.7 (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.

[12] 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 $[0, 1]$.

Property 4.6.8. 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 4.6.9 (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[13]:

Definition 4.6.10 (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 4.6.11. For $n \geq 1$ the sets $\pi_n(X, x_0)$ are groups.

Property 4.6.12. For $n \geq 2$ the homotopy groups $\pi_n(X, x_0)$ are abelian.

Property 4.6.13. 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 4.6.14. Homeomorphic spaces have the same homotopy groups π_n.

Formula 4.6.15. 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:

$$\pi_n(X \times Y, (x_0, y_0)) = \pi_n(X, x_0) \otimes \pi_n(Y, y_0) \tag{4.13}$$

where $\otimes$ denotes the direct product of groups 3.1.17.

4.7 Homology

4.7.1 Simplexes

Definition 4.7.1 (Simplex). A k-simplex σ^k is defined as the following set:

$$\sigma^k = \left\{ \sum_{i=0}^{k} \lambda_i t_i \middle| \sum_{i=0}^{k} \lambda_i = 1 \text{ and } \lambda_i \geq 0 \right\} \tag{4.14}$$

where the points (vertices) t_i are linearly independent, i.e. the vectors $t_i - t_0$ are linearly independent. Equivalently it is the convex hull of the $k+1$ vertices $\{t_0, ..., t_k\}$.

[13]Note however that we replace the interval $[0, 1]$ by the sphere S^1. This is nonrestrictive as we can construct S^n by mapping (identifying) the boundary of $[0, 1]^n$ to the basepoint x_0.

Definition 4.7.2 (Barycentric coordinates). The coordinates λ_i from previous defintion are called barycentric coordinates. This comes 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.

Definition 4.7.3 (Simplicial complex). A simplicial complex $\mathcal{K}$ is a set of simplexes 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 we have $\sigma_1 \cap \sigma_2 = \emptyset$ 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 4.7.4 (Polyhedron). Let $\mathcal{K}$ be a simplicial complex. The polyhedron associated with $\mathcal{K}$ is the topological spaces constructed by equipping $\mathcal{K}$ with the Euclidean subspace topology.

Definition 4.7.5 (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 4.7.6 (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 4.7.7. *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 vertexes of $\mathcal{K}$. Let G be the free group generated by the elements g_{ij} corresponding to the ordered 1-simplexes $[v_i, v_j] \in \mathcal{C}$. 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 $[v_i, v_j] \in \mathcal{C}$.*

then the group G is isomorphic to fundamental group $\pi_1(\mathcal{K}, a_0)$.

Corollary 4.7.8. 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.

4.7.2 Simplicial homology

Definition 4.7.9 (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 (additive) group generated by the k-simplexes in $\mathcal{K}$:

$$C_k(\mathcal{K}) = \left\{ \sum_i a_i \sigma_i \;\middle|\; \sigma_i \text{ is a } k\text{-simplex in } \mathcal{K} \text{ and } a_i \in \mathbb{Z} \right\} \tag{4.15}$$

For $k > n$ we define $C_k(\mathcal{K})$ to be $\{0\}$.

Definition 4.7.10 (Boundary operator). The boundary operator $\partial_k : C_k(\mathcal{K}) \to C_{k-1}(\mathcal{K})$ is the homomorphism defined by following properties:

- ∂_k is linear, i.e. $\partial_k(\sum_i a_i\sigma_i) = \sum_i a_i\partial_k(\sigma_i)$
- For every oriented k-simplex $[v_0, ..., v_k]$ we have that

$$\partial_k[v_0, ..., v_k] = \sum_{i=0}^{k}(-1)^i[v_0, ..., \hat{v}_i, ..., v_k] \tag{4.16}$$

- The boundary of every 0-chain is 0.

where $[v_0, ..., \hat{v}_i, ..., v_k]$ denotes the $(k-1)$-simplex obtained by removing the vertex v_i.

Property 4.7.11. The boundary operators satisfy following relation:

$$\partial_k \circ \partial_{k-1} = 0 \tag{4.17}$$

This property turns the system (C_k, ∂_k) into a so-called **chain complex**.

Definition 4.7.12 (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 also called **cycles**.

Definition 4.7.13 (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 4.7.14 (Homology group). From property 4.17 it follows that $B_k(\mathcal{K})$ is a subgroup of $Z_k(\mathcal{K})$. We can thus define the k^{th} homology group $H_k(\mathcal{K})$ as the following quotient:

$$H_k(\mathcal{K}) = Z_k(\mathcal{K})/B_k(\mathcal{K}) \tag{4.18}$$

Theorem 3.1.24 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 4.7.15. If two topological spaces are of the same homotopy type then they have isomorphic homology groups. It follows that homeomorphic spaces have isomorphic homology groups.

Corollary 4.7.16. It follows from the definition of a triangulation that we can (easily) construct the homology groups for a given triangulable space by looking at one of its triangulations.

Definition 4.7.17 (Betti numbers). The numbers $R_k(\mathcal{X})$ from the definition of homology groups are called the Betti numbers of $\mathcal{X}$.

Formula 4.7.18 (Euler characteristic). The Euler characteristic of a triangulable space X is defined as follows[14]:

$$\boxed{\chi(X) = \sum_i(-1)^i R_i(X)} \tag{4.19}$$

[14]This formula is sometimes called the *Euler-Poincaré* or *Poincaré* formula.

Definition 4.7.19. The definition of homology groups can be generalized by letting the (formal) linear combinations used in the definition of the chain group (see 4.7.9) be of the following form:

$$c^k = \sum_i g_i \sigma_i^k \tag{4.20}$$

where $G = \{g_i\}$ is an Abelian group and σ_i^k are k-simplexes. The k^{th} homology group of X with coefficients in G is denoted by $H_k(X; G)$. In case of G being 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 (or field) is given by the *Universal coefficient theorem*.

Formula 4.7.20 (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:

$$H_k(X \times Y; F) = \bigoplus_{k=i+j} H_i(X; F) \otimes H_j(Y; F) \tag{4.21}$$

When F is replaced by the set of integers the torsion subgroups have to be taken into account. This will not be done here.

4.7.3 Relative homology

In this section we use a simplicial complex K and a subcomplex L.

Definition 4.7.21 (Relative chain group). The k-chain group of K modulo L is defined as the following quotient group:

$$C_k(K, L) = C_k(K)/C_k(L) \tag{4.22}$$

Definition 4.7.22 (Relative boundary operator). The relative boundary operator $\overline{\partial}_k$ is defined as follows:

$$\overline{\partial}_k(c_k + C_k(L)) = \partial_k c_k + C_{k-1}(L) \tag{4.23}$$

where $c_k \in C_k(K)$. This operator is, just like the ordinary boundary operator ∂_k, a homomorphism.

Definition 4.7.23 (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:

$$H_k(K, L) = \frac{\ker \overline{\partial}_k}{\operatorname{im} \overline{\partial}_{k+1}} \tag{4.24}$$

Elements $h_k \in H_k(K, L)$ can thus be written as $h_k = z_k + C_p(L)$ where z_k does not have to be a relative k-cycle. We merely require that $\partial_k z_k$ is a chain in $C_{k-1}(L)$.

Definition 4.7.24 (Homology sequence). Using the relative homology groups we obtain following (long) exact sequence:

$$\cdots \rightarrow H_k(L) \xrightarrow{i_*} H_k(K) \xrightarrow{j_*} H_k(K,L) \xrightarrow{\partial_k} H_{k-1}(L) \rightarrow \cdots \tag{4.25}$$

where i_* and j_* are the homology homomorphisms induced by the inclusions $i : L \rightarrow K$ and $j : K \rightarrow (K, L)$.

Theorem 4.7.25 (Excision theorem). *Let X, U, V be a triangulable spaces such that $U \subset V \subset X$. If the closure $\overline{U}$ is contained in the interior V° then:*

$$H_k(X,V) = H_k(X\backslash U, V\backslash U) \tag{4.26}$$

4.7.4 Singular homology

Definition 4.7.26 (Singular simplex). Consider the **standard** k-simplex Δ^k:

$$\Delta^k = \{(x_0, ..., x_k) \in \mathbb{R}^{k+1} | \sum_i x_i = 1 \text{ and } x_i \geq 0\} \tag{4.27}$$

A singular k-simplex in a topological space X is defined as a continuous map $\sigma^k : \Delta^k \rightarrow X$. The name singular comes from the fact that the maps σ^k need not be invertible.

Definition 4.7.27 (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 freely generated group is in most cases infinite as there are multiple ways to map Δ^k to X.

Definition 4.7.28 (Singular boundary operator). The singular boundary operator ∂ (we use the same notation as for simplicial boundary operators) is defined by its linear action of the singular chain group $S_k(X)$. It follows that we only have to know the action on the singular simplexes σ^k.

We first introduce the notation $[v_0, ..., v_k] := [\sigma^k(e_0), ..., \sigma^k(e_k)]$ where e_i is the i^{th} vertex of the standard simplex Δ^k. The action on the singular simplex σ^k is then given by:

$$\partial\sigma^k = \sum_i^k (-1)^k [v_0, ..., \hat{v}_i, ..., v_k] \tag{4.28}$$

The singular boundary operators satisfy the same relation as in the simplicial case:

$$\partial_k \circ \partial_{k-1} = 0 \tag{4.29}$$

Definition 4.7.29 (Singular homology group). The singular homology groups are defined as follows:

$$H_k(X; G) = \frac{\ker \partial_k}{\text{im } \partial_{k+1}} \tag{4.30}$$

Theorem 4.7.30. *Let X be a triangulable space. The k^{th} singular homology group of X is isomorphic to the k^{th} simplicial homology group of X.*

Remark 4.7.31. When X is not triangulable the previous theorem is not valid. The singular approach to homology is thus a more general construction, but it is often more difficult to compute the homology groups (even in the case of triangulable spaces).

4.7.5 Examples

Example 4.7.32. Let X be a contractible space. We then find that:

$$H_k(X) = \begin{cases} \mathbb{Z} & k = 0 \\ \{0\} & k > 0 \end{cases} \tag{4.31}$$

Example 4.7.33. Let P be a connected polyhedron. We then find that:

$$H_0(P) = \mathbb{Z} \tag{4.32}$$

Example 4.7.34. The homology groups of the n-sphere S^n are given by:

$$H_k(S^n) = \begin{cases} \mathbb{Z} & k = 0 \text{ or } k = n \\ \{0\} & \text{otherwise} \end{cases} \tag{4.33}$$

4.7.6 Axiomatic approach

Definition 4.7.35 (Eilenberg-Steenrod axioms). All homology theories have the following properties in common. By treating these properties as axioms we can construct homology theories as a sequence of functors H_k. The axioms are as follows:

1. **Homotopy**: If f, g are homotopic maps then their induced homology maps are the same.
2. **Excision**[15]: 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 $k \geq 1$.
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
$$\cdots \to H_k(A) \xrightarrow{i_*} H_k(X) \xrightarrow{j_*} H_k(X, A) \xrightarrow{\partial_k} H_{k-1}(A) \to \cdots \tag{4.34}$$
where i_* and j_* are the homology homomorphisms induced by the inclusions $i : A \to X$ and $j : X \to (X, A)$.

Let X be a singleton. The group $H_0(X)$ is called the **coefficient group** and gives the coefficients used in the construction of the free Abelian chain groups C_k.

Remark 4.7.36. If the dimension axiom is removed from the set of axioms, then we obtain a so-called *extraordinary homology theory*.

[15]See also theorem 4.7.25.

4.8 Sheaf theory

4.8.1 Sheafs

Definition 4.8.1 (Sheaf). Let X be a topological space. A sheaf over X is a tuple (S, X, π), where S is a topological space and $\pi : S \to X$ a continuous surjection, such that the following two conditions are satisfied:

- For every point $s \in S$ there exists a neighbourhood U such that $\pi|_U$ is homeomorphism onto some open neighbourhood of $\pi(s) \in X$. This map induces the discrete topology on S
- For every $x \in X$, the set $\pi^{-1}(x)$ is an algebraic structure such that the corresponding algebraic operation is continuous.

Definition 4.8.2 (Stalk). The preimage $\pi^{-1}(x)$ is called the stalk over x and is often denoted by S_x.

Definition 4.8.3 (Homomorphism of sheaves). Let S, S' be two sheaves over the same space with projections π and π'. A homomorphism of sheaves is a map Φ satisfying the following conditions:

- $\Phi : S \to S'$ is continuous.
- $\pi = \pi' \circ \Phi$, i.e. Φ maps stalks in S to corresponding stalks in S'.
- For each $x \in X$, the restriction $\Phi|_x : S_x \to S'_x$ is a homomorphism of the algebraic structures corresponding to the stalks.

4.8.2 Presheafes

Definition 4.8.4 (Presheaf). Let X be a topological space. A presheaf over X consists of an algebraic structure S_U for every open set $U \subseteq X$ and a homomorphism $\Phi^U_V : S_U \to S_V$ for every two open sets $U, V \subseteq X$ with $V \subseteq U$ such that the following conditions are satisfied:

- If $U = \emptyset$ then $S_U = 0$, where 0 is the zero object in the category corresponding to the algebraic structure of S_U.
- $\Phi^U_U = \mathbb{1}_X$
- If $W \subseteq V \subseteq U$ then $\Phi^U_W = \Phi^V_W \circ \Phi^U_V$.

Definition 4.8.5 (Homomorphism of presheaves). Let S, S' be two presheaves. A homomorphism $S \to S'$ is a set of homomorphisms $\Psi_U : S_U \to S'_U$ that commute with the maps Φ^U_V.

Construction 4.8.6. For every presheaf over X we can construct a sheaf (S, X, π). For every $x \in X$ we set the stalk S_x to be the direct limit 3.21 of the direct system (S_U, Φ^U_V). The set S is then defined as the union of all sets S_x and π maps every element of $S_x \subset S$ to x.

The topology on S is defined by means of the following basis. For every $U \in X$ and every element $f \in S_U$ we construct a subset $f_U \subset S$ given by $\{f_x \in S_x : x \in U\}$ where f_x is called the **germ**[16] of f at x. The basis for our topology is then given by the set $\{f_U : U \subset X, f \in S_U\}$.

4.8.3 Sections

Definition 4.8.7 (Section). A section of a sheaf (S, X, π) over an open set U is a continuous map $s : U \to S$ such that $\pi \circ s = \mathbb{1}_U$. The set of all sections carry the same algebraic structure as S.

Remark. A **global** section is a section $s : X \to S$.

[16] This is a generalization of definition 4.2.9.

Chapter 5

Metric spaces

5.1 General definitions

Definition 5.1.1 (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 5.1.2 (Metric space). A set M equipped with a metric d is called a metric space and is denoted by (M, d).

Definition 5.1.3 (Diameter). The diamater of a subset $U \subset M$ is defined as

$$\mathrm{diam}(U) = \sup_{x,y \in U} d(x, y) \tag{5.1}$$

Definition 5.1.4 (Bounded). A subset $U \subseteq M$ is bounded if $\mathrm{diam}(U) < +\infty$.

Property 5.1.5. 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 5.1.6 (Open ball). An open ball centered on a point $x_0 \in M$ with radius $R > 0$ is defined as the set:

$$\boxed{B(x_0, R) = \{x \in M : d(x, x_0) < R\}} \tag{5.2}$$

Definition 5.1.7 (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. $\overline{B}(x_0, R) = \{x \in M : d(x, x_0) \leq R\}$.

[1] See next chapter.

Definition 5.1.8 (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 5.1.9 (Open set). A subset $N \subset M$ is said to be open if every point $x \in N$ is an interior point of N.

Definition 5.1.10 (Closed set). A subset $V \subset M$ is said to be closed if its complement is open.

Definition 5.1.11 (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 5.1.12 (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 5.1.13 (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:

$$\forall \varepsilon > 0 : \exists N_0 \in \mathbb{N} : \forall n \geq N_0 : d(x_n, a) < \varepsilon \tag{5.3}$$

Definition 5.1.14 (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:

$$\forall \varepsilon > 0 : \exists \delta_\varepsilon : \forall x \in \mathrm{dom}(f) : d(a, x) < \delta_\varepsilon \implies d'(f(a), f(x)) < \varepsilon \tag{5.4}$$

Property 5.1.15. Let (M, d) be a metric space. The distance function $d : M \times M \to \mathbb{R}$ is a continuous function.

Definition 5.1.16 (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:

$$\forall \varepsilon > 0 : \exists \delta_\varepsilon : \forall x, y \in \mathrm{dom}(f) : d(x, y) < \delta_\varepsilon \implies d'(f(x), f(y)) < \varepsilon \tag{5.5}$$

This is clearly a stronger notion than that of continuity as the number ε is equal for all points $y \in \mathrm{dom}(f)$.

5.2 Examples of metrics

Definition 5.2.1 (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 5.2.2. The projection associated with the set M_j is defined as:

$$\mathrm{pr}_j : M \to M_j : (a_1, ..., a_n) \mapsto a_j \tag{5.6}$$

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 5.2.3 (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 $\mathcal{C}(K, \mathbb{C})$:

$$d_\infty(f, g) = \sup_{x\in K} |f(x) - g(x)| \tag{5.7}$$

Example 5.2.4 (p-metric). We can define following set of metrics on $\mathbb{R}^n$:

$$\boxed{d_p(x, y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p\right)^{1/p}} \tag{5.8}$$

Example 5.2.5 (Chebyshev distance).

$$d_\infty(x, y) = \max_{1\leq i\leq n} |x_i - y_i| \tag{5.9}$$

It is also called the **maximum metric** or L_∞ metric.

Remark 5.2.6. 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:

$$d_\infty(x, y) = \lim_{p\to\infty} d_p(x, y) \tag{5.10}$$

which is also the origin of the name L_∞ metric.

5.3 Metrizable spaces

Definition 5.3.1 (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 5.3.2 (Urysohn's metrization theorem). *Every second-countable T_3 space is metrizable.*

5.4 Compactness in metric spaces

Theorem 5.4.1 (Stone). *Every metric space is paracompact.*

Definition 5.4.2 (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 $\forall F \in \mathcal{F} : \mathrm{diam}(F) \leq \varepsilon$.
- For every $\varepsilon > 0$ there exists a finite subset $E \subset M$ such that $M \subseteq \bigcup_{x \in E} B(x, \varepsilon)$.

Property 5.4.3. 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 5.4.4. *For a metric space M the following statements are equivalent:*

- *M is compact.*
- *M is sequentially compact.*
- *M is complete and totally bounded.*

Theorem 5.4.5 (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 5.4.6 (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:

$$(\forall f \in \mathcal{F})(\forall x \in U)(d(f(x), f(a)) \leq \varepsilon) \tag{5.11}$$

for all $\varepsilon \geq 0$.

Property 5.4.7. 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 5.4.8 (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 $x \in K$.*

[2] See formula 5.7.

5.5 Complete metric spaces

Definition 5.5.1 (Cauchy sequence). A sequence $(x_n)_{n\in\mathbb{N}}$ in a metric space (M,d) is Cauchy (or has the Cauchy property) if

$$(\forall\varepsilon>0)(\exists N\in\mathbb{N})(\forall m,n\geq N)(d(x_m,x_n)<\varepsilon) \tag{5.12}$$

Property 5.5.2.

- Every closed subset of a complete metric space is complete.
- Every complete subset of a metric space is closed.

Property 5.5.3 (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 5.5.4 (Completeness). A metric space is complete if it satisfies the Cauchy criterion.

5.6 Injective metric spaces

Definition 5.6.1 (Metric retraction). Let (M,d) be a metric space. A function $f:X\rightarrow X$ 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 $x,y\in M$:

$$d(f(x),f(y))\leq d(x,y) \tag{5.13}$$

The image of f is called a (metric) retract of M.

Definition 5.6.2 (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 retract of X.

Property 5.6.3. Every injective metric space is complete.

5.7 Convex spaces

Definition 5.7.1 (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:

$$d(x,z)=d(x,y)+d(y,z) \tag{5.14}$$

Definition 5.7.2 (Hyperconvex space). A convex space for which the set of closed balls has the Helly property[3] is called a hyperconvex space.

[3]See definition 2.1.5.

Theorem 5.7.3 (Aronszajn & Panitchpakdi). *A metric space is injective if and only if it is hyperconvex.*

Part III

Calculus

Chapter 6

Calculus

6.1 Sequences

Definition 6.1.1 (Limit superior). Let $(x_i)_{i\in\mathbb{N}}$ be a sequence of real numbers. The limit superior is defined as follows:

$$\limsup_{i\to+\infty} x_i = \inf_{i\geq 1}\left\{\sup_{k\geq i} x_k\right\} \tag{6.1}$$

Definition 6.1.2 (Limit inferior). Let $(x_i)_{i\in\mathbb{N}}$ be a sequence of real numbers. The limit superior is defined as follows:

$$\liminf_{i\to+\infty} x_i = \sup_{i\geq 1}\left\{\inf_{k\geq i} x_k\right\} \tag{6.2}$$

Theorem 6.1.3. *A sequence $(x_i)_{i\in\mathbb{N}}$ converges pointwise if and only if $\limsup_{i\to+\infty} x_i = \liminf_{i\to+\infty} x_i$.*

6.2 Continuity

Definition 6.2.1 (Lipschitz continuity). A function $f:\mathbb{R}\to\mathbb{R}$ is Lipschitz continuous if there exists a constant $C>0$ such that

$$|f(x)-f(x')|\leq C|x-x'| \tag{6.3}$$

for all $x, x'\in\mathbb{R}$.

Theorem 6.2.2 (Darboux's theorem). *Let f be a differentiable function on a closed interval I. Then f' has the intermediate value property*[1].

[1] This means that the function satisfies the conclusion of the intermediate value theorem 4.4.3.

Remark 6.2.3 (Darboux function). Functions that have the intermediate value property are called Darboux functions.

Corollary 6.2.4 (Bolzano's theorem). 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 6.2.5. The image of a compact set is also a compact set.

Theorem 6.2.6 (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.*

6.3 Convergence

Definition 6.3.1 (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

$$\forall x \in \text{dom}(f_n) : \lim_{n\to+\infty} f_n(x) = f(x) \tag{6.4}$$

Definition 6.3.2 (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

$$\sup_{x\in\text{dom}(f_n)} \left\{ |\lim_{n\to+\infty} f_n(x) - f(x)| \right\} = 0 \tag{6.5}$$

6.4 Derivative

6.4.1 Single variable

Formula 6.4.1 (Derivative).

$$\boxed{f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}} \tag{6.6}$$

Theorem 6.4.2 (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:*

$$f'(c) = \frac{f(b) - f(a)}{b - a} \tag{6.7}$$

Definition 6.4.3 (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 $i = 1, \ldots, n$) then f is said to be of class $\mathbf{C}^n(I)$.

Definition 6.4.4 (Smooth function). A function f is said to be smooth if it is of class $\mathbf{C}^\infty$.

Definition 6.4.5 (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 class of analytic functions defined on I is denoted by $\mathbf{C}^\omega(I)$.

Theorem 6.4.6 (Schwarz's theorem). *Let $f \in C^2(\mathbb{R}^n, \mathbb{R})$, then:*

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right) = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right) \tag{6.8}$$

for all indices $i, j \leq n$.

Method 6.4.7 (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:

First we take a look at the natural logarithm of the function:

$$\ln[u(x)] = g(x)\ln[f(x)]$$

Then we look at the derivative of the natural logarithm:

$$\frac{d\ln[u(x)]}{dx} = \frac{1}{u(x)}\frac{du(x)}{dx} \implies \frac{du(x)}{dx} = u(x)\frac{d\ln[u(x)]}{dx}$$

But according to the first equation we also have:

$$\frac{d\ln[u(x)]}{dx} = \frac{d}{dx}g(x)\ln[f(x)] = \frac{dg(x)}{dx}\ln[f(x)] + \frac{g(x)}{f(x)}\frac{df(x)}{dx}$$

Combining these two equations gives:

$$\boxed{\frac{d}{dx}\left[f(x)^{g(x)}\right] = f(x)^{g(x)}\left[\frac{dg}{dx}(x)\ln[f(x)] + \frac{g(x)}{f(x)}\frac{df}{dx}(x)\right]} \tag{6.9}$$

Theorem 6.4.8 (Euler's homogeneous function theorem). *Let f be a homogeneous function, i.e. $f(ax_1, ..., ax_n) = a^n f(x_1, ..., x_n)$. Then f satisfies following equality:*

$$\sum_k x_k \frac{\partial f}{\partial x_k} = nf(x_1, ..., x_n) \tag{6.10}$$

6.5 Riemann integral

Definition 6.5.1 (Improper Riemann integral).

$$\boxed{\int_{-\infty}^{+\infty} f(x)dx = \lim_{\substack{a\to-\infty \\ b\to+\infty}} \int_a^b f(x)dx} \tag{6.11}$$

6.6 Fundamental theorems

Theorem 6.6.1 (First fundamental theorem of calculus). *Let $f(x)$ be a continuous function defined on the open interval I. Let $c \in I$. The following theorem establishes a link between integration and differentiation:*

$$\boxed{\exists F(x) = \int_c^x f(x')dx' : F'(x) = f(x)} \tag{6.12}$$

Furthermore this function $F(x)$ is uniformly continuous on I.

Remark 6.6.2. 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 6.6.3 (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:*

$$\boxed{\int_a^b f'(x)dx = f(b) - f(a)} \tag{6.13}$$

Theorem 6.6.4 (Differentiation under the integral sign[2]).

$$\boxed{\frac{d}{dx}\int_{a(x)}^{b(x)} f(x,y)dy = f(x,b(x)) \cdot b'(x) - f(x,a(x)) \cdot a'(x) + \int_{a(x)}^{b(x)} \frac{\partial f(x,y)}{\partial x}dy} \tag{6.14}$$

6.7 Taylor expansion

Formula 6.7.1 (Exponential function).

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \tag{6.15}$$

6.8 Euler integrals

6.8.1 Euler integral of the first kind

Formula 6.8.1 (Beta function).

$$\boxed{B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}dt} \tag{6.16}$$

[2]This is a more general version of the so called 'Leibnitz integral rule'.

6.8.2 Euler integral of the second kind

Formula 6.8.2 (Gamma function).

$$\boxed{\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt} \tag{6.17}$$

Formula 6.8.3. $\mathrm{B(x,\ y)} = \dfrac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$

Formula 6.8.4 (Recursion formula). Let $n!$ denote the factorial for integer numbers.

$$\Gamma(n) = (n-1)! \tag{6.18}$$

Chapter 7

Series

7.1 Convergence tests

Theorem 7.1.1. *A series* $\sum_{i=1}^{+\infty} a_i$ *can only converge if* $\lim_{i\to+\infty} a_i = 0$.

Property 7.1.2 (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 7.1.3 (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 7.1.4 (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 7.1.5 (MacLaurin-Cauchy integral test). Let f be a continuous non-negative 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 7.1.6. 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 7.1.7. If the integral in the previous theorem converges, then the series has following lower and upper bounds:

$$\int_n^{+\infty} f(x)dx \leq \sum_{i=n}^{+\infty} a_i \leq f(n) + \int_n^{+\infty} f(x)dx \tag{7.1}$$

Method 7.1.8 (d'Alembert's ratio test).

$$R = \lim_{n\to+\infty} \left|\frac{a_{n+1}}{a_n}\right| \tag{7.2}$$

Following cases arise:

- $R < 1$: the series converges absolutely
- $R > 1$: the series does not converge
- $R = 1$: the test is inconclusive

Method 7.1.9 (Cauchy's root test).

$$R = \limsup_{n\to+\infty} \sqrt[n]{|a_n|} \tag{7.3}$$

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 7.1.10 (Radius of convergences). The number $\frac{1}{R}$ is called the radius of convergence.

Remark 7.1.11. 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 7.1.12 (Gauss's test). If $u_n > 0$ for all n then we can write the ratio of successive terms as follows:

$$\left|\frac{u_n}{u_{n+1}}\right| = 1 + \frac{h}{n} + \frac{B(n)}{n^k} \tag{7.4}$$

where $k > 1$ and $B(n)$ is a bounded function when $n \to \infty$. The series converges if $h > 1$ and diverges otherwise.

7.2 Asymptotic expansions

Definition 7.2.1 (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:

$$f(x) - \sum_{n=0}^{N} = O(x^{N+1}) \tag{7.5}$$

Method 7.2.2 (Borel transform†). Define the function $F(x) = \sum_{n=0}^{+\infty} \frac{a_n}{n!} x^n$. If the integral

$$\int_0^{+\infty} e^{-t} F(xt) dt < +\infty \tag{7.6}$$

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 7.2.3 (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 8

Complex calculus

8.1 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 8.1.1 (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 8.1.2 (Real/imaginary part). A complex number z can also be written as $\mathrm{Re}(z) + i\mathrm{Im}(z)$ where

$$\mathrm{Re}(z) = \frac{z + \overline{z}}{2} \tag{8.1}$$

$$\mathrm{Im}(z) = \frac{z - \overline{z}}{2i} \tag{8.2}$$

Definition 8.1.3 (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 8.1.4 (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:

$$\begin{aligned} z + \infty &= \infty \\ z * \infty &= \infty \\ \frac{z}{\infty} &= 0 \end{aligned} \tag{8.3}$$

for all non-zero $z \neq \infty$. As there exists no multiplicative inverse for ∞ the Riemann sphere does not form a field.

[1]See definition 4.5.22.

8.2 Holomorphic functions

Definition 8.2.1 (Holomorphic). A function f is holomorphic on an open set U if it is complex differentiable at every point $z_0 \in U$.

Definition 8.2.2 (Biholomorphic). A complex function f is said to be biholomorphic if both f and f^{-1} are holomorphic.

Property 8.2.3 (Cauchy-Riemann conditions). A holomorphic function $f(z)$ satisfies the following conditions:

$$\boxed{\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \qquad \text{and} \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}} \tag{8.4}$$

or equivalently:

$$\boxed{\frac{\partial f}{\partial \bar{z}} = 0} \tag{8.5}$$

Theorem 8.2.4 (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 8.2.5. Functions u, v satisfying the CR-conditions are harmonic functions, i.e. they satisfy Laplace's equation.

Property 8.2.6. Functions u, v satisfying the CR-conditions have orthogonal level curves 2.9.

8.3 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 8.3.1 (Contour). A contour is a curve $z(t)$ that can be parametrized by

$$\left.\begin{array}{l} x = x(t) \\ y = y(t) \end{array}\right\} \rightarrow z(t) = z = x + iy \tag{8.6}$$

Formula 8.3.2 (Complex contour integral). The complex contour integral of a function $f(z) = u(z) + iv(z)$ is defined as the following line integral:

$$\int_{z_1}^{z_2} f(z)dz = \int_{(x_1,y_1)}^{(x_2,y_2)} [u(x,y) + iv(x,y)](dx + idy) \tag{8.7}$$

[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.

Theorem 8.3.3 (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 contour C in Ω:*

$$\boxed{\oint_C f(z)dz = 0} \tag{8.8}$$

Corollary 8.3.4. The contour integral of a holomorphic function depends only on the limits of integration and not on the contour connecting them.

Formula 8.3.5 (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:

$$\boxed{f(z_0) = \frac{1}{2\pi i}\oint_C \frac{f(z)}{z - z_0}dz} \tag{8.9}$$

Corollary 8.3.6 (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[4] on Ω and:

$$\boxed{f^{(n)}(z_0) = \frac{1}{2\pi i}\oint_C f(z)\frac{n!}{(z - z_0)^{n+1}}dz} \tag{8.10}$$

Furthermore, the derivatives are also holomorphic on Ω.

Theorem 8.3.7 (Morera's Theorem). *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.*

Definition 8.3.8 (Meromorphic). A function f is called meromorphic when it is analytic on the whole complex plane with exception of isolated poles and removable singularities.

Theorem 8.3.9 (Sokhotski-Plemelj[5]). *Let $f(x)$ be a continuous complex-valued function defined on the real line and let $a < 0 < b$.*

$$\lim_{\varepsilon\to 0^+}\int_a^b \frac{f(x)}{x \pm i\varepsilon}dx = \mp i\pi f(0) + \mathcal{P}\int_a^b \frac{f(x)}{x}dx \tag{8.11}$$

where $\mathcal{P}$ denotes the Cauchy principal value.

8.4 Laurent series

Definition 8.4.1 (Laurent series). If f is function, analytic on an annulus A, then f can be expanded as the following series:

$$f(z) = \sum_{n=-\infty}^{\infty} a_n(z - z_0)^n \qquad \text{with} \qquad a_n = \frac{1}{2\pi i}\oint \frac{f(z')}{(z' - z_0)^{n+1}}dz' \tag{8.12}$$

---

[3]Also called the *Cauchy-Goursat theorem*.
[4]See definition 6.4.5.
[5]See for example [15], page 104.

Remark 8.4.2. 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 8.4.3 (Principal part). The principal part of a Laurent series is defined as the sum:

$$\sum_{n=-\infty}^{-1} a_n (z - z_0)^n \tag{8.13}$$

8.5 Singularities

8.5.1 Poles

Definition 8.5.1 (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 8.5.2 (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 8.5.3 (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 8.5.4 (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 $\lim_{h\to 0} f(h)$.

8.5.2 Branch cuts

Formula 8.5.5 (Roots). Let $z \in \mathbb{C}$. The n^{th} roots[6] of $z = re^{i\theta}$ are given by:

$$z^{1/n} = \sqrt[n]{r} \exp\left(i\frac{\theta + 2\pi k}{n}\right) \tag{8.14}$$

where $k \in \{0, 1, ..., n\}$.

Formula 8.5.6 (Complex logarithm). We parametrize z as $z = re^{i\theta}$.

$$\mathrm{LN}(z) = \ln(r) + i(\theta + 2\pi k) \tag{8.15}$$

[6] Also see the *fundamental theorem of algebra* 15.1.3.

Definition 8.5.7 (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 $\arg(\text{LN}) \in]\alpha, \alpha + 2\pi]$ is often denoted by LN_α or $\log_\alpha$.

Definition 8.5.8 (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 8.5.9 (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 8.5.10. Consider the complex function

$$f(z) = \frac{1}{\sqrt{(z - z_1)...(z - z_n)}}$$

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 [6] for the proof.)

Definition 8.5.11 (Principal value). The principal value of a multi-valued complex function is defined as the choice of branch such that $\arg(f) \in] - \pi, \pi]$.

8.5.3 Residue theorem

Definition 8.5.12 (Residue). By applying formula 8.7 to a polynomial function we find:

$$\int_C (z - z_0)^n dz = 2\pi i \delta_{n,-1} \tag{8.16}$$

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 8.5.13. The residue of a complex function $f(z)$ at a pole z_0 is denoted by $\text{Res}[f(z)]_{z=z_0}$.

Formula 8.5.14. For a pole of order m, the residue is calculated as follows:

$$\text{Res}\left[f(z)\right]_{z=z_j} = a_{-1} = \lim_{z \to z_0} \frac{1}{(m-1)!} \left(\frac{\partial}{\partial z}\right)^{m-1} (f(z)(z - z_0)) \tag{8.17}$$

For essential singularities the residue can be found by writing out the Laurent series explicitly.

Theorem 8.5.15 (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:*

$$\boxed{\oint_C f(z)dz = 2\pi i \sum_j \mathrm{Res}\left[f(z)\right]_{z=z_j}} \tag{8.18}$$

Remark 8.5.16. For poles on the contour C, only half of the residue contributes to the integral.

Formula 8.5.17 (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:

$$\frac{1}{2\pi i}\oint_C \frac{f(z)}{f'(z)}dz = Z_f - P_f \tag{8.19}$$

Formula 8.5.18 (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 or **index** of a with respect to the function f is defined as:

$$\mathrm{Ind}_f(a) = \frac{1}{2\pi i}\oint_C \frac{f'(z)}{f(z)-a}dz \tag{8.20}$$

This number will always be an integer.

8.6 Limit theorems

Theorem 8.6.1 (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 8.6.2 (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 8.6.3 (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

$$\int_C g(z)dz = 0$$

8.7 Analytic continuation

Theorem 8.7.1 (Schwarz' reflection principle). *Let $f(z)$ be analytic on the upper half plane. If $f(z)$ is real when z is real then*

$$f(\overline{z}) = \overline{f(z)} \tag{8.21}$$

Chapter 9

Measure theory and Lebesgue integration

9.1 Measure

9.1.1 General definitions

Definition 9.1.1 (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 9.1.2 (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.

Definition 9.1.3 (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:

$$\mu\left(\{x \in X : \neg P(x)\}\right) = 0 \tag{9.1}$$

Definition 9.1.4 (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 $A \subset E$:

$$A \in \Sigma \quad \text{and} \quad \mu(A) = 0$$

Definition 9.1.5 (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.

[1] also called σ-additivity

[2] In probability theory this is foten often called **almost surely**.

Definition 9.1.6 (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:

$$\begin{aligned} \mu(B) &= \inf\{\mu(O) : O \text{ open}, O \supset B\} \\ \mu(B) &= \sup\{\mu(F) : F \text{ closed}, F \subset B\} \end{aligned} \tag{9.2}$$

Definition 9.1.7 (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 $\forall A_i : P(A_i) < +\infty$.

Method 9.1.8. 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.

9.1.2 Lebesgue measure

Formula 9.1.9 (Length of an interval). The length of an open interval $I = (a, b)$ is defined as:

$$l\,(I) = b - a \tag{9.3}$$

Definition 9.1.10 (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

$$A \subseteq \bigcup_{n=1}^{+\infty} I_n \tag{9.4}$$

with

$$\sum_{i=1}^{+\infty} l(I_n) < \varepsilon \tag{9.5}$$

Theorem 9.1.11. *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 9.1.12. Any countable set is null.

Definition 9.1.13 (Outer measure). Let $X \subseteq \mathbb{R}$ be an open set. The (Lebesgue) outer measure is defined as:

$$\boxed{m^*(X) = \inf\left\{\sum_{i=1}^{+\infty} l(I_i) \text{ with } (I_i)_{i\in\mathbb{N}} \text{ a sequence of open intervals that covers } X\right\}} \tag{9.6}$$

Property 9.1.14. Let I be an interval. The outer measure equals the length: $m^*(I) = l(I)$.

Property 9.1.15. The outer measure is translation invariant: $m^*(A + t) = m^*(A) \quad, \forall A, t$

Property 9.1.16. $m^*(A) = 0$ if and only if A is null.

Property 9.1.17. If $A \subset B$ then $m^*(A) \leq m^*(B)$.

Property 9.1.18 (Countable subadditivity). For every sequence of sets $(E_i)_{i\in\mathbb{N}}$ the following inequality holds:

$$m^*\left(\bigcup_{i=1}^{+\infty} E_i\right) \leq \sum_{i=1}^{+\infty} m^*(E_i) \tag{9.7}$$

Theorem 9.1.19 (Carathéodory's criterion / Lebesgue measure). *Let X be a set. If X satisfies the following equation, it is said to be Lebesgue measurable:*

$$\forall E \subseteq \mathbb{R} : m^*(E) = m^*(E \cap X) + m^*(E \cap X^c) \tag{9.8}$$

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 9.1.20. All null sets and intervals are measurable.

Property 9.1.21 (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:

$$\boxed{m\left(\bigcup_{i=1}^{+\infty} E_i\right) = \sum_{i=1}^{+\infty} m(E_i)} \tag{9.9}$$

Remark. Previous property, together with the properties of the outer measure, implies that the Lebesgue measure is indeed a proper measure as defined in 9.1.1.

Property 9.1.22. $\mathcal{M}$ is a σ-algebra[3] over $\mathbb{R}$.

Theorem 9.1.23. *For every $A \subset \mathbb{R}$ there exists a sequence $(O_i)_{i\in\mathbb{N}}$ of open sets such that:*

$$A \subset \bigcap_i O_i \qquad \text{and} \qquad m\left(\bigcap_i O_i\right) = m^*(A) \tag{9.10}$$

Theorem 9.1.24. *For every $E \in \mathcal{M}$ there exists a sequence $(F_i)_{i\in\mathbb{N}}$ of closed sets such that:*

$$\bigcup_i F_i \subset E \qquad \text{and} \qquad m\left(\bigcup_i F_i\right) = m(E) \tag{9.11}$$

Remark. The previous 2 theorems imply that the Lebesgue measure is a regular Borel measure 9.2.

Theorem 9.1.25. *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 $m^*(E\backslash F) < \varepsilon$.*

[3]See definition 2.4.2.

Property 9.1.26. 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:

$$\forall i : A_i \subseteq A_{i+1} \implies m\left(\bigcup_{i=1}^{+\infty} A_i\right) = \lim_{i\to+\infty} m(A_i) \tag{9.12}$$

$$\forall i : A_i \supseteq A_{i+1} \wedge m(A_1) < +\infty \implies m\left(\bigcap_{i=1}^{+\infty} A_i\right) = \lim_{i\to+\infty} m(A_i) \tag{9.13}$$

Remark 9.1.27. This property is not only valid for the Lebesgue measure but for every countably additive set function.

Property 9.1.28. The Lebesgue measure $m(X)$ is continuous at $\emptyset$, i.e. if $(A_i)_{i\in\mathbb{N}} \to \emptyset$ then $\lim\limits_{i\to+\infty} m(A_i) = 0$.

Theorem 9.1.29. *$\mathcal{M}$ is the completion of $\mathcal{B}$.*

Corollary 9.1.30. $\mathcal{B} \subset \mathcal{M} \subset \mathcal{F}_\mathbb{R}$

Definition 9.1.31 (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:

$$\mathcal{M}_B = \{A \cap B : A \in \mathcal{M}\} \qquad \text{and} \qquad \forall E \in \mathcal{M}_B : m_B(E) = m(E) \tag{9.14}$$

Furthermore, the measure space $(B, \mathcal{M}_B, m_B)$ is complete.

9.1.3 Measurable functions

Definition 9.1.32 (Measurable function). A function f is (Lebesgue) measurable if for every interval $I \subset \mathbb{R} : f^{-1}(I) \in \mathcal{M}$.

Definition 9.1.33 (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 9.1.34. Inclusion 9.1.30 implies that every Borel function is also Lebesgue measurable.

Theorem 9.1.35. *The class of Lebesgue measurable[5] functions defined on $E \in \mathcal{M}$ is closed under multiplication and it forms a vector space.*

Property 9.1.36. Following types of functions are measurable:

- monotone functions

[4]These functions are often simply called 'Borel functions'.
[5]This property is also valid for Borel functions.

- continuous functions
- indicator functions

Corollary 9.1.37. 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 9.1.38. Let f be a measurable function. The set[6] $\{x : f(x) = a\}$ is also measurable for all $a \in \mathbb{R}$.

Theorem 9.1.39. *Define following functions, which are measurable if f is measurable as a result of previous properties:*

$$f^+(x) = \begin{cases} f(x) & \text{if } \ f(x) > 0 \\ 0 & \text{if } \ f(x) \leq 0 \end{cases} = \max(f, 0) \tag{9.15}$$

$$f^-(x) = \begin{cases} 0 & \text{if } \ f(x) > 0 \\ -f(x) & \text{if } \ f(x) \leq 0 \end{cases} = \max(-f, 0) \tag{9.16}$$

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.

9.1.4 Limit operations

Property 9.1.40. 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.

Property 9.1.41. Let f be a measurable function. Let g be a function such that $f = g$ almost everywhere. The function g is measurable.

Corollary 9.1.42. 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 9.1.43 (Essential supremum).

$$\text{ess sup} f = \sup\{z : f \geq z \text{ a.e.}\} \tag{9.17}$$

[6] This set is called the 'level set' of f.
[7] This property is also valid for Borel functions.

Definition 9.1.44 (Essential infimum).

$$\text{ess inf} f = \inf\{z : f \leq z \text{ a.e.}\} \tag{9.18}$$

Property 9.1.45. 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 9.1.46. 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.

9.2 Lebesgue integral

9.2.1 Simple functions

Definition 9.2.1 (Indicator function). An important function when working with sets is the following one:

$$\boxed{\mathbb{1}_A(x) = \begin{cases} 1 & \text{if} \quad x \in A \\ 0 & \text{if} \quad x \notin A \end{cases}} \tag{9.19}$$

Definition 9.2.2 (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:

$$f(x) = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}(x) \tag{9.20}$$

with $A_i = f^{-1}(a_i) \in \mathcal{M}$

Remark 9.2.3 (Step function). If the sets A_i are intervals, the simple function is often called a 'step function'.

Formula 9.2.4 (Lebesgue integral of simple functions). Let φ be a simple function as defined in equation 9.20. Let $\mu : \mathcal{M} \rightarrow \mathbb{R}$ 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:

$$\int_E \varphi d\mu = \sum_{i=1}^{n} a_i \mu(E \cap A_i) \tag{9.21}$$

Example 9.2.5. 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:

$$\int_{\mathbb{R}} \mathbb{1}_{\mathbb{Q}} dm = 1 \times m(\mathbb{Q}) + 0 \times m(\mathbb{R} \backslash \mathbb{Q}) = 0 \tag{9.22}$$

where the measure of the rational numbers is 0 because it is a countable set (see corollary 9.1.12.

9.2.2 Measurable functions

Formula 9.2.6 (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:

$$\int_E fdm = \sup\left\{\int_E \varphi dm : \varphi \text{ a simple function such that } \varphi \leq f\right\} \tag{9.23}$$

Property 9.2.7. The Lebesgue integral $\int_E fdm$ of a measurable function f is always non-negative.

Notation 9.2.8. The following notation is frequently used (both in the sense of Riemann and Lebesgue):

$$\int fdm = \int_{\mathbb{R}} fdm \tag{9.24}$$

Formula 9.2.9. The following equality is easily proved as for every set $A \subseteq \mathbb{R} : A \cup A^c = \mathbb{R}$.

$$\int_A fdm = \int f\mathbb{1}_A dm \tag{9.25}$$

Theorem 9.2.10. *Let f be a non-negative measurable function. Then $f = 0$ a.e. if and only if $\int_{\mathbb{R}} fdm = 0$.*

Property 9.2.11. The Lebesgue integral over a null set is 0.

Property 9.2.12. Let f, g me measurable functions. The Lebesgue integral has the following properties:

- $f \leq g$ a.e. implies $\int fdm \leq \int gdm$.
- Let A be a measurable set. Let $B \subset A$. Then $\int_B fdm \leq \int_A fdm$.
- The Lebesgue integral is linear.
- For every two disjoint measurable sets A and B we have that $\int_{A\cup B} fdm = \int_A fdm + \int_B fdm$.
- **Mean value theorem**: If $a \leq f(x) \leq b$, then $am(A) \leq \int_A fdm \leq bm(A)$.

Theorem 9.2.13. *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 9.2.14. *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 $\int_a^b |f - h|dm < \varepsilon$.*

[8]See remark 9.2.3.

9.2.3 Integrable functions

Definition 9.2.15 (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:

$$\int_E f dm = \int_E f^+ dm - \int_E f^- dm \tag{9.26}$$

Remark. The difference between the integral 9.23 and the integral of an integrable function is that with the latter f does not have to be non-negative.

Theorem 9.2.16. *f is integrable if and only if $|f|$ is integrable. Furthermore, $\int_E |f| dm = \int_E f^+ dm + \int_E f^- dm$.*

Property 9.2.17. 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 9.2.18. *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 9.2.19. 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 9.2.20. 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 9.6 for further information.

9.2.4 Convergence theorems

Theorem 9.2.21 (Fatou's lemma). *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of non-negative measurable functions.*

$$\int_E \left(\liminf_{n\to\infty} f_n \right) dm \leq \liminf_{n\to\infty} \int_E f_n dm \tag{9.27}$$

Theorem 9.2.22 (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:*

$$\boxed{\int_E f dm = \lim_{n\to\infty} \int_E f_n(x) dm} \tag{9.28}$$

Method 9.2.23. 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 9.2.24 (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*

$$\int_E f dm = \lim_{n\to\infty} \int_E f_n(x) dm \tag{9.29}$$

Property 9.2.25. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of non-negative measurable functions. The following equality applies:

$$\int \sum_{n=1}^{+\infty} f_n(x) dm = \sum_{n=1}^{+\infty} \int f_n(x) dm \tag{9.30}$$

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 9.2.26 (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

$$\int \sum_{i=1}^{\infty} f_n(x) dm = \sum_{i=1}^{\infty} \int f_n(x) dm \tag{9.31}$$

Theorem 9.2.27 (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.

9.2.5 Relation to the Riemann integral

Theorem 9.2.28 (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 9.2.29. *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 9.2.30. *If $f \geq 0$ and the improper Riemann integral 6.11 exists, then the Lebesgue integral $\int f dm$ exists and the two integrals coincide.*

9.3 Examples

Definition 9.3.1 (Dirac measure[9]). We define the Dirac measure as follows:

$$\delta_a(X) = \begin{cases} 1 & \text{if } a \in X \\ 0 & \text{if } a \notin X \end{cases} \tag{9.32}$$

The integration with respect to the Dirac measure has the following nice property[10]:

$$\int g(x) d\delta_a = g(a) \tag{9.33}$$

Example 9.3.2. Let $\mu = \delta_2, X = (-4; 1)$ and $Y = (-2; 17)$. The following two integrals are easily computed:

$$\int_X d\mu = 0$$
$$\int_Y d\mu = 1$$

9.4 Space of integrable functions

9.4.1 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 now do know how to integrate

[9]Compare to 11.6.
[10]This equality can be proved by applying formula 29.14 with $X \equiv a$.

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' on a different 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 9.4.1. $L^1(E)$ is a Banach space[11].

Formula 9.4.2. A norm on $L^1(E)$ is given by:

$$||f||_1 = \int_E |f|dm \tag{9.34}$$

9.4.2 Hilbert space L^2

Property 9.4.3. L^2 is a Hilbert space[12].

Formula 9.4.4. A norm on $L^2(E)$ is given by:

$$||f||_2 = \left(\int_E |f|^2 dm\right)^{\frac{1}{2}} \tag{9.35}$$

This norm is induced by the following inner product:

$$\boxed{\langle f|g\rangle = \int_E f\overline{g}dm} \tag{9.36}$$

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 9.35 is finite.

Definition 9.4.5 (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:

$$\langle f|g\rangle = 0 \implies \text{f and g are orthogonal} \tag{9.37}$$

Furthermore it is also possible to introduce the angle between functions in the same way as equation 15.43.

Formula 9.4.6 (Cauchy-Schwarz inequality). Let $f, g \in L^2(E, \mathbb{C})$. We have that $fg \in L^1(E\mathbb{C})$ and:

$$\boxed{\left|\int_E f\overline{g}dm\right| \leq ||fg||_1 \leq ||f||_2||g||_2} \tag{9.38}$$

Remark. This follows immediately from formula 9.40.

Property 9.4.7. If E has finite Lebesgue measure then $L^2(E) \subset L^1(E)$.

[11] See definition 17.1.4.
[12] See definition 17.2.1.

9.4.3 L^p spaces

Generalizing the previous two Lebesgue function classes leads us to the notion of L^p spaces with the following norm:

Property 9.4.8. For all $1 \leq p \leq +\infty$ $L^p(E)$ is a Banach space with a norm given by:

$$||f||_p = \left(\int_E |f|^p \, dm\right)^{\frac{1}{p}} \tag{9.39}$$

Remark 9.4.9. 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 9.4.10 (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:

$$||fg||_1 \leq ||f||_p ||g||_q \tag{9.40}$$

Formula 9.4.11 (Minkowski's inequality). For every $p \geq 1$ and $f, g \in L^p(E)$ we have

$$||f + g||_p \leq ||f||_p + ||g||_p \tag{9.41}$$

Property 9.4.12. If E has finite Lebesgue measure then $L^q(E) \subset L^p(E)$ when $1 \leq p \leq q < +\infty$.

9.4.4 L^∞ space of essentially bounded measurable functions

Definition 9.4.13 (Essentially bounded function). Let f be a measurable function satisfying $\text{ess sup}|f| < +\infty$. The function f is said to be essentially bounded and the set of all such functions is denoted by $L^\infty(E)$.

Formula 9.4.14. A norm on L^∞ is given by:

$$||f||_\infty = \text{ess sup}|f| \tag{9.42}$$

This norm is called the **supremum norm** and it induces the supremum metric 5.7.

Property 9.4.15. L^∞ is a Banach space.

9.5 Product measures

9.5.1 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 9.5.1 (Hypercube). Let $I_1, ..., I_n$ be a sequence of intervals.

$$\mathbf{I} = I_1 \times ... \times I_n \tag{9.43}$$

Definition 9.5.2 (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:

$$l(\mathbf{I}) = \prod_{i=1}^{n} l(I_i) \tag{9.44}$$

9.5.2 Construction of the product measure

Property 9.5.3 (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 $A_2 \in \mathcal{F}_2$:

$$\boxed{P(A_1 \times A_2) = P_1(A_1)P_2(A_2)} \tag{9.45}$$

Definition 9.5.4 (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 9.5.5. 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 9.5.6. 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 9.5.7 (Product measure). From previous property it follows that we can write the product measure $P(A)$ in the following way:

$$\boxed{P(A) = \int_{\Omega_2} P_1(A_{\omega_2}) dP_2(\omega_2)} \tag{9.46}$$

Property 9.5.8. 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

$$\boxed{\int_{\Omega_2} P_1(A_{\omega_2}) dP_2(\omega_2) = \int_{\Omega_1} P_2(A_{\omega_1}) dP_1(\omega_1)} \tag{9.47}$$

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.

9.5.3 Fubini's theorem

Property 9.5.9. 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 9.5.10 (Section of a function). The functions $\omega_1 \mapsto f(\omega_1, \omega_2)$ and $\omega_2 \mapsto f(\omega_1, \omega_2)$ are called sections of f.

Theorem 9.5.11 (Tonelli's theorem). *Let $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$ be a non-negative function. The following equalities apply:*

$$\begin{aligned}\int_{\Omega_1\times\Omega_2} f(\omega_1,\omega_2)d(P_1\times P_2)(\omega_1,\omega_2) &= \int_{\Omega_1}\left(\int_{\Omega_2} f(\omega_1,\omega_2)dP_2(\omega_2)\right)dP_1(\omega_1)\\ &= \int_{\Omega_2}\left(\int_{\Omega_1} f(\omega_1,\omega_2)dP_1(\omega_1)\right)dP_2(\omega_2)\end{aligned} \tag{9.48}$$

Corollary 9.5.12 (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 9.48 holds.

Remark 9.5.13. 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.

9.6 Radon-Nikodym theorem

Definition 9.6.1. 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

$$\forall A \in \mathcal{F} : \mu(A) = 0 \implies \nu(A) = 0 \tag{9.49}$$

Notation 9.6.2. This relation is denoted by $\nu \ll \mu$.

Theorem 9.6.3 (Absolute continuity). *Let μ, ν be finite measures on a measurable space $(\Omega, \mathcal{F})$. Then $\nu \ll \mu$ if and only if*

$$\forall \varepsilon > 0 : \exists \delta > 0 : \forall A \in \mathcal{F} : \mu(A) < \delta \implies \nu(A) < \varepsilon \tag{9.50}$$

Property 9.2.20 can be generalized to arbitrary measure spaces as follows:

Property 9.6.4. 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 9.6.5 (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 9.6.6 (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 9.6.7 (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 $\dfrac{d\nu}{d\mu}$.

Theorem 9.6.8 (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 9.6.9. The function g in the previous theorem is unique up to a μ-null (or ν-null) set.

Property 9.6.10. 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

$$\int_\Omega f d\nu = \int_\Omega f h_\nu d\mu \tag{9.51}$$

Remark 9.6.11. This property also holds for all functions $f \in L^1(\mu)$.

Property 9.6.12. Let λ, ν, μ be σ-finite measures. If $\lambda \ll \mu$ and $\nu \ll \mu$ then we have:

- $\dfrac{d(\lambda + \nu)}{d\mu} = \dfrac{d\lambda}{d\mu} + \dfrac{d\lambda}{d\mu}$ a.e.
- Chain rule: if $\lambda \ll \nu$ then $\dfrac{d\lambda}{d\mu} = \dfrac{d\lambda}{d\nu}\dfrac{d\nu}{d\mu}$ a.e.

9.7 Lebesgue-Stieltjes measure

Chapter 10

Integral transforms

10.1 Fourier series

Definition 10.1.1 (Dirichlet kernel). The Dirichlet kernel is the collection of functions of the form

$$D_n(x) = \frac{1}{2\pi}\sum_{k=-n}^{n} e^{ikx} \tag{10.1}$$

Formula 10.1.2 (Sieve property). If $f \in C^1[-\pi,\pi]$ then

$$\lim_{n\to+\infty}\int_{-\pi}^{\pi} f(x)D_n(x)dx = 0 \tag{10.2}$$

Formula 10.1.3. For 2π-periodic functions, the n-th degree Fourier approximation is given by following convolution:

$$s_n(x) = \sum_{k=-n}^{n}\widetilde{f}(k)e^{ikx} = (D_n * f)(x) \tag{10.3}$$

Theorem 10.1.4 (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 10.1.5 (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:

$$\boxed{f(x) = \sum_{n=-\infty}^{+\infty}\left(\frac{1}{2l}\int_{-l}^{l} e^{-i\frac{n\pi x'}{l}} f(x')dx'\right)e^{i\frac{n\pi x}{l}}} \tag{10.4}$$

Formula 10.1.6 (Fourier coefficients). As seen in the general formula, the Fourier coefficient $\widetilde{f}(n)$ can be calculated by taking the inner product 17.5 of $f(x)$ and the n-th eigenfunction e_n:

$$\widetilde{f}(n) = \langle e_n|f\rangle = \int_{-l}^{l} e_n^*(x)f(x)dx \qquad \text{with} \qquad e_n = \sqrt{\frac{1}{2l}}e^{i\frac{n\pi x}{l}} \tag{10.5}$$

Definition 10.1.7 (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

$$f^{L,\nu}(x) = \frac{f^L(x+) + f^L(x-)}{2} \tag{10.6}$$

Theorem 10.1.8. *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 $x \in \mathbb{R}$.*

10.2 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:

$$\boxed{\mathcal{F}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt} \tag{10.7}$$

$$f(t) = \mathcal{F}^{-1}(t) = \frac{1}{\sqrt{2\pi}} ⨍_{-\infty}^{\infty} \mathcal{F}(\omega) e^{i\omega t} d\omega \tag{10.8}$$

Equation 10.7 is called the (forward) Fourier transform of $f(t)$ and equation 10.8 is called the inverse Fourier transform.

Notation 10.2.1. The Fourier transform of a function $f(t)$, as seen in equation 10.7, is often denoted by $\widetilde{f}(\omega)$.

Theorem 10.2.2 (Convergence of the Fourier integral). *If $f : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous (see 6.3) 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 10.2.3 (Fourier inversion theorem). *If both $f(t), \mathcal{F}(\omega) \in \mathcal{L}^1(\mathbb{R})$ are continuous then the Cauchy principal value in 10.8 can be replaced by a normal integral.*

Remark 10.2.4. Schwartz functions (see 11.1) 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 10.2.5. From the Riemann-Lebesgue lemma 9.2.27 it follows that

$$\mathcal{F}(\omega) \to 0 \qquad \text{if} \qquad |\omega| \to 0 \tag{10.9}$$

Property 10.2.6 (Parceval's theorem). Let $(f, \widetilde{f})$ and $(g, \widetilde{g})$ be two Fourier transform pairs.

$$\int_{-\infty}^{+\infty} f(x)g(x)dx = \int_{-\infty}^{+\infty} \widetilde{f}(k)\widetilde{g}(k)dk \tag{10.10}$$

Corollary 10.2.7 (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:

$$\int_{-\infty}^{+\infty} |f(x)|^2 dx = \int_{-\infty}^{+\infty} |\widetilde{f}(k)|^2 dk \tag{10.11}$$

10.2.1 Convolution

Formula 10.2.8 (Convolution).

$$\boxed{(f * g)(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(\tau) g(t-\tau) d\tau} \tag{10.12}$$

Property 10.2.9 (Commutativity).

$$f * g = g * f \tag{10.13}$$

Theorem 10.2.10 (Convolution Theorem).

$$\widetilde{f * g} = \widetilde{g}\widetilde{f} \tag{10.14}$$

10.3 Laplace transform

Formula 10.3.1 (Laplace transform).

$$\mathcal{L}\{F(t)\}_{(s)} = \int_0^{\infty} f(t) e^{-st} dt \tag{10.15}$$

Formula 10.3.2 (Bromwich integral).

$$f(t) = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} \mathcal{L}\{F(t)\}_{(s)} e^{st} ds \tag{10.16}$$

Notation 10.3.3. The Laplace transform as defined in equation 10.15 is sometimes denoted by $f(s)$.

10.4 Mellin transform

Formula 10.4.1 (Mellin transform).

$$\mathcal{M}\{f(x)\}(s) = \int_0^{+\infty} x^{s-1} f(x) dx \tag{10.17}$$

Formula 10.4.2 (Inverse Mellin transform).

$$f(x) = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} \mathcal{M}\{f(x)\}_{(s)} x^{-s} ds \tag{10.18}$$

10.5 Integral representations

Formula 10.5.1 (Heaviside step function).

$$\theta(x) = \frac{1}{2\pi i}\int_{-\infty}^{\infty} \frac{e^{ikx}}{x - i\varepsilon} dk \tag{10.19}$$

Formula 10.5.2 (Dirac delta function).

$$\delta^{(n)}(\vec{\boldsymbol{x}}) = \frac{1}{(2\pi)^n}\int_{-\infty}^{\infty} e^{i\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{x}}} d^n k \tag{10.20}$$

Chapter 11

Distributions

11.1 Generalized function

Definition 11.1.1 (Schwartz space). The Schwartz space or **space of rapidly decreasing functions**[1] $S(\mathbb{R})$ is defined as:

$$S(\mathbb{R}) = \left\{ f(x) \in C^\infty(\mathbb{R}) : \forall i, j \in \mathbb{N} : \forall x \in \mathbb{R} : |x^i f^{(j)}(x)| < +\infty \right\} \tag{11.1}$$

Remark 11.1.2. This definition can be generalized to functions of the class $C^\infty(\mathbb{R}^n)$ or functions $f : \mathbb{R} \to \mathbb{C}$. The Schwartz space is then denoted by $S(\mathbb{R}, \mathbb{C})$.

Definition 11.1.3 (Functions of slow growth). The set of functions of slow growth $N(\mathbb{R})$ is defined as:

$$N(\mathbb{R}) = \{ f(x) \in C^\infty(\mathbb{R}) : \forall i \in \mathbb{N}, \exists M_i > 0 : |f^{(i)}(x)| = O(|x|^i) \text{ for } |x| \to +\infty \} \tag{11.2}$$

Remark. It is clear that all polynomials belong to $N(\mathbb{R})$ but not to $S(\mathbb{R})$.

Property 11.1.4. If $f(x) \in S(\mathbb{R})$ and $a(x) \in N(\mathbb{R})$ then $a(x)f(x) \in S(\mathbb{R})$.

Definition 11.1.5 (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.

[1]These functions are said to be rapidly decreasing because every derivative $f^{(j)}(x)$ decays faster than any polynomial x^i for $x \to +\infty$.

Notation 11.1.6. Let ψ be a generalized function. Let $f \in S(\mathbb{R})$. The inner product 9.36 is generalized by following functional:

$$\langle \psi | f \rangle = \lim_{n \to +\infty} \int_{-\infty}^{+\infty} \psi_n(x) f(x) dx \tag{11.3}$$

Property 11.1.7. 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 11.1.8 (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 $\lim\limits_{n \to +\infty} f_n(x) = f(x)$ converges pointwise on that interval.

11.2 Dirac Delta distribution

Definition 11.2.1 (Heaviside function). Define the generalized function $H \sim \{H_n(x) \in S(\mathbb{R})\}$ as:

$$H(x) = \begin{cases} 0 & if \quad x < 0 \\ 1 & if \quad x \geq 0 \end{cases} \tag{11.4}$$

From this definition it follows that for every $f \in S(\mathbb{R})$:

$$\langle H | f \rangle = \int_0^{+\infty} f(x) dx \tag{11.5}$$

Remark 11.2.2. 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 11.2.3 (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 11.1.7 and 11.5 we have for every $f \in S(\mathbb{R})$:

$$\begin{aligned} \langle \delta | f \rangle &= \langle H' | f \rangle \\ &= -\langle H | h \rangle \\ &= -\int_0^{+\infty} f'(x) dx \\ &= f(0) \end{aligned} \tag{11.6}$$

[2]The space of Lebesgue integrable functions 9.2.18.

Property 11.2.4 (Sampling property). The result from previous definition can be generalized in the following way:

$$\boxed{f(x_0) = \int_{\mathbb{R}} f(x)\delta(x - x_0)dx} \tag{11.7}$$

Example 11.2.5 (Dirac comb).

$$III_b(x) = \sum_n \delta(x - nb) \tag{11.8}$$

Property 11.2.6. 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:

$$\delta[f(x)] = \sum_{i=1}^{n} \frac{1}{|f'(x_i)|}\delta(x - x_i) \tag{11.9}$$

Property 11.2.7 (Convolution with delta function). Let $f(x) \in S(\mathbb{R})$. Let $\otimes$ denote the convolution.

$$\delta(x) \otimes f(x) = \int_{-\infty}^{+\infty} \delta(x - \alpha)f(\alpha)d\alpha = f(x) \tag{11.10}$$

Formula 11.2.8 (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 function

$$f_c(x) = f(x) - \sum_{i=1}^{n} \sigma_i H(x - x_i)$$

which is a continuous function. Differentiation gives

$$f'(x) = f_c'(x) + \sum_{i=1}^{n} \sigma_i \delta(x - x_i)$$

It follows that the derivative in a generalized sense of a piecewise continuous function equals the derivative in the classical sense plus a summation of delta functions at every jump discontinuity.

11.3 Fourier transform

Theorem 11.3.1. *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 10.10 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 functions is also a generalized functions.*

Property 11.3.2. 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:

$$\langle \psi | F \rangle = \langle \Psi | f \rangle \tag{11.11}$$

Formula 11.3.3 (Fourier representation of delta function).

$$\boxed{\delta(x-a) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{ik(x-a)} dk} \tag{11.12}$$

Chapter 12

Ordinary differential equations

12.1 Boundary conditions

Unigue solutions of a differential equation are obtained by supplying additional conditions. These are called boundary conditions.

12.1.1 Periodic boundary conditions

Periodic boundary conditions are conditions of the following form:

$$y(x) = y(x + \varphi) \tag{12.1}$$

By induction it follows that for every n:

$$y(x) = y(x + n\varphi) \tag{12.2}$$

12.1.2 Dirichlet boundary conditions

Dirichlet boundary conditions are conditions of the following form:

$$y(x) = f(x) \qquad , \qquad x \in \partial\Omega \tag{12.3}$$

where Ω is the domain of the problem.

Remark 12.1.1. When y is a function of multiple variables, α can be a function as well. For example (in spherical coordinates: ρ, ϕ, θ):

$$y(x, \phi, \theta) = \alpha(\phi, \theta) \tag{12.4}$$

12.1.3 Neumann boundary conditions

Neumann boundary conditions are conditions of the following form:

$$y'(a) = \alpha \tag{12.5}$$

Remark 12.1.2. 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):

$$\frac{\partial y}{\partial \hat{n}}(\vec{x}) = f(\vec{x}) \qquad , \qquad \vec{x} \in S \tag{12.6}$$

12.2 First order ODE's

Formula 12.2.1 (First order ODE).

$$\boxed{y'(t) + a(t)y(t) = R(t)} \tag{12.7}$$

If the function $R(t)$ is identically zero, then the ODE is said to be **homogenous**.

Theorem 12.2.2. *Let $U \subseteq \mathbb{R}$ be an open set. Let the functions $a(t), R(t) : U \rightarrow \mathbb{R}$ be continuous. The solutions $\varphi(t) : U \rightarrow \mathbb{R}$ of equation 12.7 are given by:*

$$\boxed{\varphi(t) = e^{-\int a(t)dt}\left(c + \int R(t)e^{\int a(t)dt}dt\right)} \tag{12.8}$$

where c is a constant.

12.3 Second order ODE's

Formula 12.3.1 (Second order ODE).

$$\boxed{y''(t) + a(t)y'(t) + b(t)y(t) = R(t)} \tag{12.9}$$

Formula 12.3.2 (Homogeneous second order ODE).

$$y''(t) + a(t)y'(t) + b(t)y(t) = 0 \tag{12.10}$$

12.3.1 General solution

Formula 12.3.3. Let $\varphi : U \rightarrow \mathbb{R}$ be a nowhere zero solution of the homogeneous equation 12.10. The general solution of equation 12.9 is then given by:

$$\boxed{y(t) = c_1\ \varphi + c_2\ \varphi \int \frac{e^{-\int a}}{\varphi^2} + \psi_0} \tag{12.11}$$

where ψ_0 is a particular solution of equation 12.9.

Theorem 12.3.4. *Let ψ_0 be a solution of equation 12.9. The set of all solutions is given by the affine space:*

$$\{\psi_0 + \chi : \chi \text{ is a solution of the homogeneous equation 12.10}\} \tag{12.12}$$

Theorem 12.3.5. *Two solutions of the homogeneous equation 12.10 are independent if the wronskian is nonzero:*

$$W(\varphi_1(x), \varphi_2(x)) = \begin{vmatrix} \varphi_1(x) & \varphi_2(x) \\ \varphi_1'(x) & \varphi_2'(x) \end{vmatrix} \neq 0 \tag{12.13}$$

Formula 12.3.6 (Abel's identity). An explicit formula for the wronskian is given by:

$$\boxed{W(x) = W(x_0)\exp\left(-\int_{x_0}^{x} a(x')dx'\right)} \tag{12.14}$$

12.3.2 Constant coefficients

Theorem 12.3.7. *A map $\varphi : U \to \mathbb{C}$ is a complex solution of equation 12.10 if and only if $Re\{\varphi\}$ and $Im\{\varphi\}$ are real solutions of equation 12.10.*

Formula 12.3.8 (Characteristic equation). When having an ODE of the form[1] :

$$y''(t) + py'(t) + qy(t) = 0 \tag{12.15}$$

where p and q are constants, we define the characteristic equation as follows:

$$\lambda^2 + p\lambda + q = 0 \tag{12.16}$$

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 12.15 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 $\lambda_1 = a + ib$: $y(t) = c_1\ e^{at}\cos(bt) + c_2\ e^{at}\sin(bt)$

[1] Any other form of homogeneous second order ODE's with constant coefficients can be rewritten in this form.

[2] See theorem 15.1.3 ("Fundamental theorem of algebra").

12.3.3 Method of Frobenius

Formula 12.3.9 (Method of Frobenius). To find a solution of the homogeneous equation 12.10 we assert a solution of the form:

$$\boxed{y(x) = \sum_{i=0}^{\infty} a_i (x - x_0)^{i+k}} \tag{12.17}$$

where k is a constant.

Definition 12.3.10 (Indicial equation). After inserting the solution 12.17 into the homogeneous equation 12.10 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 12.3.11. *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 12.11)*
- $k_1 - k_2 \in \mathbb{Z}$*: 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 12.3.12 (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.*

12.4 Sturm-Liouville theory

Definition 12.4.1 (Sturm-Liouville boundary value problem). The following ODE, subject to mixed boundary conditions, is called a Sturm-Liouville boundary value problem:

$$\boxed{\frac{d}{dx}\left[p(x)\frac{dy}{dx}\right] + \left[g(x) + \lambda r(x)\right] y(x) = 0} \tag{12.18}$$

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

$$\begin{cases} \alpha_1 y(a) + \beta_1 y'(a) = 0 \\ \alpha_2 y(b) + \beta_2 y'(b) = 0 \end{cases} \tag{12.19}$$

where at least one of the constants $\alpha_1, \alpha_2, \beta_1$ or β_2 is non-zero.

[3] It is important to 'sync' the power of all terms in order to obtain one 'large' coefficient.

Formula 12.4.2. 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:

$$\begin{vmatrix} \alpha_1 u_1(a;\lambda) + \beta_1 u_1'(a;\lambda) & \alpha_1 u_2(a;\lambda) + \beta_1 u_2'(a;\lambda) \\ \alpha_1 u_1(b;\lambda) + \beta_1 u_1'(b;\lambda) & \alpha_1 u_2(b;\lambda) + \beta_1 u_2'(b;\lambda) \end{vmatrix} = 0 \tag{12.20}$$

The independent eigenfunctions can be found by substituting the found eigenvalues in the ODE 12.18.

Definition 12.4.3 (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

$$\left[a_2(x)\frac{d^2}{dx^2} + a_1(x)\frac{d}{dx} + a_0(x)\right] y(x) = 0 \tag{12.21}$$

This equation can be rewritten in a self-adjoint form by setting:

$$p(x) = e^{\int \frac{a_1}{a_2}dx} \qquad \text{and} \qquad g(x) = \frac{a_0}{a_2}e^{\int \frac{a_1}{a_2}dx} \tag{12.22}$$

Property 12.4.4. The eigenfunctions corresponding to distinct eigenvalues are orthogonal with respect to the weight function $r(x)$.

Theorem 12.4.5 (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 13

Partial differential equations

13.1 General linear equations

Formula 13.1.1 (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;

$$x_i = \frac{\det(A_i)}{\det(A)} \tag{13.1}$$

where A_i is the matrix obtained by replacing the i^{th} column of A by the column matrix b.

Definition 13.1.2 (Characteristic curve). Curve along which the highest order partial derivatives are not uniquely defined.

13.2 First order PDE

Formula 13.2.1 (First order quasilinear PDE).

$$\boxed{P(x,y,z)\frac{\partial z}{\partial x} + Q(x,y,z)\frac{\partial z}{\partial y} = R(x,y,z)} \tag{13.2}$$

Formula 13.2.2 (Characteristic curve). The PDE will have no unique solution if

$$\begin{vmatrix} P & Q \\ dx & dy \end{vmatrix} = 0 \tag{13.3}$$

and will have a non-unique solution if

$$\begin{vmatrix} P & R \\ dx & dz \end{vmatrix} = 0 \tag{13.4}$$

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 13.2.3. *The general solution of 13.2 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*

$$\frac{dx}{P} = \frac{dy}{Q} = \frac{dz}{R} \tag{13.5}$$

where c_1, c_2 are constants which are fixed by boundary conditions.

Remark 13.2.4. Looking at the defining equations of the characteristic curve, it is clear that these fix the general solution of the PDE.

13.3 Characteristics

Formula 13.3.1 (Second order quasilinear PDE). Consider the following pseudolinear differential equation for the function $u(x, y)$:

$$R(x, y)u_{xx} + S(x, y)u_{xy} + T(x, y)u_{yy} = W(x, y, u, p, q) \tag{13.6}$$

where $p = u_x$ and $q = u_y$.

Formula 13.3.2 (Equation of characteristics). Consider the following two differential equations

$$\begin{cases} u_{xx}dx + u_{xy}dy = dp \\ u_{xy}dx + u_{yy}dy = dq \end{cases} \tag{13.7}$$

According to Cramer's rule 13.1 these equations, together with the PDE 13.6, give the following condition for the characteristic curves:

$$\begin{vmatrix} R(x, y) & S(x, y) & T(x, y) \\ dx & dy & 0 \\ 0 & dx & dy \end{vmatrix} = 0 \tag{13.8}$$

which is equivalent to following equation:

$$\boxed{R\left(\frac{dy}{dx}\right)^2 - S\left(\frac{dy}{dx}\right) + T = 0} \tag{13.9}$$

Definition 13.3.3 (Types of characteristics). Equation 13.9 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 13.3.4 (Canonical form). Consider the general change of variables $\xi = \xi(x, y)$, $\eta = \eta(x, y)$ and $z = \zeta$. With this change, the PDE 13.6 becomes:

$$A(\xi_x, \xi_y)\frac{\partial^2\zeta}{\partial\xi^2} + 2B(\xi_x, \xi_y, \eta_x, \eta_y)\frac{\partial^2\zeta}{\partial\xi\partial\eta} + A(\eta_x, \eta_y)\frac{\partial^2\zeta}{\partial\eta^2} = F(\xi, \eta, \zeta, \zeta_\xi, \zeta_\eta) \tag{13.10}$$

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 13.9 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

$$\boxed{\frac{\partial^2\zeta}{\partial\xi\partial\eta} = H(\xi,\eta,\zeta,\zeta_\xi,\zeta_\eta)} \tag{13.11}$$

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

$$\boxed{\frac{\partial^2\zeta}{\partial\eta^2} = G(\xi,\eta,\zeta,\zeta_\xi,\zeta_\eta)} \tag{13.12}$$

where $G = \frac{F}{A(\eta_x,\eta_y)}$.

- **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

$$\boxed{\frac{\partial^2\zeta}{\partial\alpha^2} + \frac{\partial^2\zeta}{\partial\beta^2} = K(\alpha,\beta,\zeta,\zeta_\alpha,\zeta_\beta)} \tag{13.13}$$

Theorem 13.3.5 (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.*

13.3.1 D'Alemberts method

Consider the wave equation

$$\frac{\partial^2 u}{\partial x^2}(x,t) = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2}(x,t) \tag{13.14}$$

By applying the method from previous subsection, it is clear that the characteristics are given by

$$\xi = x + ct \qquad \text{and} \qquad \eta = x - ct \tag{13.15}$$

Furthermore, it follows that the wave equation is a hyperbolic equation which can be rewritten in the canonical form:

$$\frac{\partial^2 u}{\partial\xi\partial\eta}(\xi,\eta) = 0 \tag{13.16}$$

Integration with respect to ξ and η and rewriting the solution in terms of x and t gives

$$u(x,t) = f(x+ct) + g(x-ct) \tag{13.17}$$

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

$$u(x,0) = v(x) \qquad \text{and} \qquad \frac{\partial u}{\partial t}(x,0) = q(x) \tag{13.18}$$

By applying these conditions to the general solution 13.17 it can be shown that the general solution subject to the given boundary conditions is given by:

$$\boxed{u(x,t) = \frac{1}{2}\left[v(x+ct) + v(x-ct)\right] + \frac{1}{2c}\int_{x-ct}^{x+ct} q(z)dz} \tag{13.19}$$

Remark 13.3.6. Because x is not bounded, this solution is only valid for infinite strings.

13.4 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.

13.4.1 Cartesian coordinates

Method 13.4.1 (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 13.4.2. Consider following PDE:

$$\frac{\partial u}{\partial t} - a\frac{\partial^2 u}{\partial x^2} = 0 \tag{13.20}$$

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:

$$\begin{cases} X''(x) & = & \lambda X(x) \\ T'(t) & = & a\lambda T(t) \end{cases}$$

13.4.2 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 13.3.5 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.

13.5 Non-homogeneous boundary conditions

Formula 13.5.1 (Non-homogeneous boundary condition).

$$\alpha u(a,t) + \beta\frac{\partial u}{\partial x}(a,t) = h(t) \tag{13.21}$$

When $h(t)$ is identically zero, the boundary condition becomes homogeneous.

[2]Think of the Dirichlet boundary condition 12.3.

[2]The Dirichlet boundary problem originated with the Laplace equation.

Method 13.5.2 (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 13.5.3. 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 13.5.4. 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 13.5.5 (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 12.8.
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 13.5.6. 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.

[4]Non-homogeneous boundary conditions can be turned homogeneous by the previous two methods.

Chapter 14

Bessel functions

14.1 Bessel's differential equation (BDE)

A Bessel's differential equation is an ordinary differential equation of the following form:

$$\boxed{z^2y'' + zy' + (z^2 - n^2)y = 0} \tag{14.1}$$

The solutions of this ODE are the Bessel functions of the first and second kind (also called respectivelly Bessel and Neumann functions).

$$J_n(z) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!(m+n)!}\left(\frac{z}{2}\right)^{2m+n} \tag{14.2}$$

$$N_n(z) = \lim_{\nu\to n} \frac{cos(\nu\pi)J_N(z) - J_{-n}(z)}{sin(\nu\pi)} \tag{14.3}$$

Remark. Solution 14.2 can be found by using Frobenius' method.

Property 14.1.1. For $n \notin \mathbb{N}$ the solutions $J_n(z)$ and $J_{-n}(z)$ are independent.

Remark 14.1.2. For $n \notin \mathbb{N}$ the limit operation in function 14.3 is not necessary as $\sin(n\pi)$ will never become 0 in this case.

14.2 Generating function

Define the following function:

$$g(x,t) = exp\left[\frac{x}{2}\left(t - \frac{1}{t}\right)\right] \tag{14.4}$$

If we expand this function as a Laurent series, we obtain the following formula:

$$g(x,t) = \sum_{n=-\infty}^{+\infty} J_n(x)t^n \tag{14.5}$$

By applying the residue theorem 8.5.15, we can express the functions $J_n(x)$ as follows:

$$J_n(x) = \frac{1}{2\pi i}\oint_C \frac{g(x,t)}{t^{n+1}}dt \tag{14.6}$$

The function $g(x,t)$ is called the generating function of the Bessel functions.

14.3 Applications

14.3.1 Laplace equation

When solving the Laplace equation in cylindrical coordinates we obtain a BDE with integer n, which has the **cylindrical Bessel functions** 14.2 and 14.3 as solutions.

14.3.2 Helmholtz equation

When solving the Helmholtz equation in spherical coordinates we obtain a variant of the BDE for the radial part:

$$z^2y'' + 2zy' + [z^2 - n(n+1)]y = 0 \tag{14.7}$$

where n is an integer. The solutions, called **spherical Bessel functions**, are related to the cylindrical Bessel functions in the following way:

$$j_n(r) = \sqrt{\frac{\pi}{2x}}J_{n+\frac{1}{2}}(r) \tag{14.8}$$

and similarly for the Neumann functions.

Part IV

Linear Algebra

Chapter 15

Linear Algebra

15.1 General

15.1.1 Polynomes

Definition 15.1.1 (Degree). The exponent of the highest order power in x. It is often denoted by $\deg(f)$.

Definition 15.1.2 (Monic polynome). A polynome of which the highest order term has coefficient 1.

Theorem 15.1.3 (Fundamental theorem of algebra). *Let $f(x) \in K[x]$ with $\deg(f) \geq 1$. Then $f(x)$ has at least 1 root in $\mathbb{C}$.*

Theorem 15.1.4. *If $f(x) \in \mathbb{C}[x]$ is a monic polynome 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 $n_1, ..., n_k \in \mathbb{N}$.

15.2 Vector spaces

In this and coming sections all vector spaces can be finite- or infinite-dimensional. If necessary, the dimension will be specified.

Definition 15.2.1 (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 is 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: $a(\vec{v}+\vec{w}) = a\vec{v}+a\vec{w}$

Definition 15.2.2 (Linear combination). The vector w is a linear combination of elements in the set $\{v_n\}$ if it can be written as:

$$w = \sum_n \lambda_n v_n \tag{15.1}$$

for some subset $\{\lambda_n\}$ of the field K.

Definition 15.2.3 (Linear independence). A set finite $\{v_n\}_{n\leq N}$ is said to be linearly independent if the following relation holds:

$$\sum_{n=0}^{N} \lambda_n v_n = 0 \iff \forall n : \lambda_n = 0 \tag{15.2}$$

A general set $\{w_i\}_{i\in I}$ is linearly independent if every finite subset of it is linearly independent.

Definition 15.2.4 (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 15.2.5 (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 15.2.6. Every set T that spans V contains a basis of V.

Remark 15.2.7. Here it is time for a little side note. 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.

It follows that for finite-dimensional spaces we do not have to worry. In infinite-dimensional spaces however we have to keep this in mind. An alternative, which allows infinitely many elements is given by the concept of a *Schauder* basis.

We continue by constructing this peculiar type of basis:

Construction 15.2.8 (Hamel basis). Consider the set of all linearly independent subsets of V. Under the relation of inclusion this set becomes a partially ordered set[1]. From Zorn's lemma 2.3.7 it follows that there exists at least one maximal linearly independent set.

Now we have to show that this maximal subset S is also a generating set of V. For this let us choose a vector $v \in V$ that is not already in S. From the maximality of S it follows that

[1] See definition 2.3.2.

$S \cup v$ 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:

$$\sum_{i=0}^{n} a^i e_i + bv = 0 \tag{15.3}$$

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]:

$$v = -\frac{1}{b} \sum_{i=0}^{n} a^i e_i \tag{15.4}$$

Because v was randomly chosen we conclude that S is a generating set for V. This set is called a Hamel basis of V.

Remark. 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 is equivalent to the axiom of choice (and thus to Zorn's lemma).

15.2.1 Subspaces

Definition 15.2.9 (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:

$$W \leq V \iff \forall w_1, w_2 \in W : \forall \lambda, \mu \in K : \lambda w_1 + \mu w_2 \in W \tag{15.5}$$

Definition 15.2.10 (Grassmannian). Let V be a K-vector space. The set consisting of all k-dimensional subspaces of V is denoted by $\mathrm{Gr}(k, V)$.

Property 15.2.11. $GL(V)$ acts transitively[3] on all k-dimensional subspaces of V. From property 3.1.40 it follows that the coset space $GL(V)/H_W$ for any stabilizer H_W of some $W \in \mathrm{Gr}(k, V)$ is isomorphic (as a set) to $\mathrm{Gr}(k, V)$.

15.2.2 Algebra

Definition 15.2.12 (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[4]:

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 turn the binary operation into a bilinear operation.

[2] It is this step that requires R to be a division ring in property 3.2.11 because else we would not generally be able to divide by $b \in R$.

[3] See definition 3.1.38

[4] These conditions imply that the binary operation is a bilinear map.

Definition 15.2.13 (Unital algebra). An algebra V is said to be unital if it contains an identity element with respect to the bilinear map $\star$.

Definition 15.2.14 (Clifford algebra). Let V be a unital associative algebra over the field K. If the bilinear map is a quadratic form then V is called a Clifford algebra.

Notation 15.2.15. Let V be an algebra and Q a quadratic form. The Clifford algebra is denoted by $C\ell(V,Q)$.

15.2.3 Sum and direct sum

Definition 15.2.16 (Sum). Let V be a K-vector space. Let $W_1, W_2, ..., W_k$ be subspaces of V. The 'sum' of the subspaces $W_1, ..., W_k$ is defined as follows:

$$W_1 + ... + W_k := \left\{ \sum_{i=1}^k w_i : w_i \in W_i \right\} \tag{15.6}$$

Definition 15.2.17 (Direct sum). If every element v of the sum as defined in definition 15.2.16 can be written as a unique linear combination, then the sum is called a direct sum.

Notation 15.2.18 (Direct sum).

$$W_1 \oplus ... \oplus W_k = \bigoplus_{i=1}^k W_i$$

Theorem 15.2.19. *Let V be a K-vector space. Let W, W_1, W_2 be three subspaces of V such that $W = W_1 \oplus W_2$. We have the following properties:*

- *If $\mathcal{B}_1$ is a basis of W_1 and if $\mathcal{B}_2$ is a basis of W_2, $\mathcal{B}_1 \cup \mathcal{B}_2$ is a basis of W.*
- $\dim(W) = \dim(W_1) + \dim(W_2)$

Theorem 15.2.20. *Let V be a finite-dimensional K-vector space. Let W_1, W_2 be two subspaces of V. Then the following relation holds:*

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2) \tag{15.7}$$

The second item in previous property is a direct consequence of this property.

Definition 15.2.21 (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 $V = W \oplus W'$.

Theorem 15.2.22. *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.*

15.2.4 Graded vector space

Similar to definition 3.2.15 we can define the following:

Definition 15.2.23 (Graded vector space). Let V_n be a vector space for all $n \in \mathbb{N}$. The vector space

$$V = \bigoplus_{n\in\mathbb{N}} V_n \tag{15.8}$$

is called a graded vector space.

Definition 15.2.24 (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 15.2.25 (Superalgebra). Let A be a $\mathbb{Z}_2$-graded algebra, i.e.:

$$A = A_0 \oplus A_1 \tag{15.9}$$

such that for all $i, j \mod 2$:

$$A_i \star A_j \subseteq A_{i+j} \tag{15.10}$$

15.3 Linear maps[5]

Definition 15.3.1 (Zero map). Let $f : A \rightarrow B$ be a (linear) map. The map f is called a zero map if:

$$\forall a \in A : f(a) = 0 \tag{15.11}$$

Definition 15.3.2 (Restriction). Let $f : A \rightarrow B$ be a (linear) map. Let $C \subset A$. The (linear) map $f|_C : C \rightarrow B : c \rightarrow f(c)$ is called the restriction of f to C.

Definition 15.3.3 (Injective). A map $f : A \rightarrow B$ is called injective if the following condition is satisfied:

$$\forall a, a' \in A : f(a) = f(a') \implies a = a' \tag{15.12}$$

Notation 15.3.4 (Injective map).

$$f : A \hookrightarrow B$$

Definition 15.3.5 (Surjective). A map $f : A \rightarrow B$ is called surjective if the following condition is satisfied:

$$\forall b \in B, \exists a \in A : f(a) = b \tag{15.13}$$

Notation 15.3.6 (Surjective map).

$$f : A \twoheadrightarrow B$$

[5]Other names are **linear mapping** and **linear transformation**.

Definition 15.3.7 (Bijective). A map is called bijective if it is both injective and surjective.

Notation 15.3.8 (Bijective map).

$$f : A \xrightarrow{\sim} B$$

Definition 15.3.9 (Isomorphism). A linear bijective map f between two K-vector spaces is called an isomorphism.

Notation 15.3.10 (Isomorphic). If two K-vector spaces V, W are isomorphic we denote it as following:

$$V \cong W$$

Definition 15.3.11 (Automorphism). An isomorphism from V to V is called an automorphism. The set of all automorphisms on V, which is in fact a group, is denoted by $\mathrm{Aut}(V)$.

Definition 15.3.12 (C^r-diffeomorphism). An isomorphism of class $C^r(K)$ with an inverse that is also of class $C^r(K)$ is called a C^r-diffeomorphism.

Theorem 15.3.13. *Let $f : A \rightarrow B$ be a map. The following statements are equivalent:*

1. *There exists a map*[6] $g : B \rightarrow A$ *such that* $f \circ g = \mathbf{1}_B$ *and* $g \circ f = \mathbf{1}_A$*.*
2. f *is bijective.*

Corollary 15.3.14. From theorem 15.3.13 and the definition of isomorphisms we can conclude that isomorphisms are precisely those maps that are invertible.

Definition 15.3.15 (General linear group[7]). The set of all automorphisms $f : V \rightarrow V$ is called the general linear group $GL_K(V)$ of $GL(V)$.

Definition 15.3.16 (Rank). The dimension of the image of a linear map is called the rank.

Definition 15.3.17 (Kernel). The kernel of a linear map $f : V \rightarrow W$ is the following subset of V:

$$\ker(f) = \{v \in V \mid f(v) = 0\} \tag{15.14}$$

Definition 15.3.18 (Nullity). The dimension of the kernel is called the nullity.

Theorem 15.3.19. *A linear map $f : V \rightarrow W$ is injective if and only if $ker(f) = \{0\}$.*

Property 15.3.20. Let $f : V \rightarrow W$ be a linear map. Let $U \leq V$. We have the following two properties of the restriction $f|_U$ of f to U:

- $\ker(f|_U) = \ker(f) \cap U$
- $\mathrm{im}(f|_U) \leq \mathrm{im}(f)$

[6] The map g is called the **inverse** of f.

[7] This group is isomorphic to the general linear group of invertable matrices, hence the similar name and notation. (See definition 15.5.7)

15.3.1 Linear operator

Definition 15.3.21 (Linear operator). A linear automorphism $f : V \to V$ is called a linear operator. It is also more generally known as an **endomorphism** on V.

Property 15.3.22. Let $\lambda, \mu \in K$. An operator $f : V \to V$ is called linear if it satisfies the following condition:

$$f(\lambda v_1 + \mu v_2) = \lambda f(v_1) + \mu f(v_2) \tag{15.15}$$

Theorem 15.3.23. *Let V be finite-dimensional K-vector space. Let $f : V \to V$ be a linear operator. The following statements are equivalent:*

- *f is injective*
- *f is surjective*
- *f is bijective*

15.3.2 Dimension

Definition 15.3.24 (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:

$$\boxed{\dim(V) = n} \tag{15.16}$$

Property 15.3.25. Let V be a finite-dimensional K-vector space. Every basis of V has the same number of elements.[8]

Theorem 15.3.26 (Dimension theorem[9]). *Let $f : V \to W$ be a linear map.*

$$\dim(\mathrm{im}(f)) + \dim(\ker(f)) = \dim(\mathrm{V}) \tag{15.17}$$

Theorem 15.3.27. *Two K-vector spaces are isomorphic if and only if they have the same dimension.*

15.3.3 Homomorphisms

Definition 15.3.28 (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 of V to W, or shorter: the 'hom-space' of V to W.

$$\mathrm{Hom}_K(V, W) = \{f : V \to W \mid \text{f is a linear map}\} \tag{15.18}$$

[8]This theorem can be generalized to infinite-dimensional spaces by stating that all bases have the same *cardinality*.

[9]Also called the **rank-nullity theorem**.

Theorem 15.3.29. *If V, W are two finite-dimensional K-vector spaces we have:*

$$\dim\left(Hom_K(V, W)\right) = \dim(V) \cdot \dim(W) \tag{15.19}$$

Definition 15.3.30 (Endomorphism ring). The space $\mathrm{Hom}_K(V, V)$ with the composition as multiplication forms a ring, the endomorphism ring. It is denoted as $\mathrm{End}_K(V)$ or $\mathrm{End}(V)$.

Definition 15.3.31 (Minimal polynomial). Let $f \in \mathrm{End}(V)$ and V a finite-dimensional K-vector space. The monic polynomial $\mu_f(x)$ of the lowest order such that $\mu_f(f) = 0$ is called the minimal polynomial of f.

Property 15.3.32. Let $f \in \mathrm{End}(V)$. Let $\mu_f(x)$ be the minimal polynomial of f. Let $\varphi(x) \in K[x]$. If $\varphi(f) = 0$, then the minimal polynomial $\mu_f(x)$ divides $\varphi(x)$.

15.3.4 Dual space

Definition 15.3.33 (Dual space). Let V be a K-vector space. The dual space V^* of V is the following vector space:

$$V^* := \mathrm{Hom}_K(V, K) = \{f : V \to K \ : \ f \text{ is a linear map}\} \tag{15.20}$$

Definition 15.3.34 (Linear form). The elements of V^* are called *linear forms*.

Property 15.3.35. From theorem 15.3.29 it follows that $\dim(V^*) = \dim(V)$.

Remark 15.3.36. If V is infinite-dimensional, theorem 15.3.35 is not valid. In the infinite-dimensional case we **always** have $|V^*| > |V|$ (where we now use the cardinality instead of the dimension).

Definition 15.3.37 (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:

$$\boxed{\varepsilon_i : V \to K : \sum_{j=1}^{n} a_i e_i \mapsto a_i} \tag{15.21}$$

The relation between the basis and dual basis can also be written as:

$$\varepsilon^i(e_j) = \delta^i_j \tag{15.22}$$

Definition 15.3.38 (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 15.3.39 (Transpose). When $V = W$ the dual map f^* is often denoted by f^T.

Definition 15.3.40 (Natural pairing). The natural pairing of V and its dual V^* is defined as the following bilinear map:

$$\langle v, v^* \rangle = v^*(v) \tag{15.23}$$

15.3.5 Convex functions

Definition 15.3.41 (Convex function). Let X be a convex subset of V. A function $f : X \to \mathbb{R}$ is convex if for all $x, y \in X$ and $t \in [0,1]$:

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) \tag{15.24}$$

Remark 15.3.42. For a concave function we have to turn the inequality around.

Corollary 15.3.43. A linear map $f : X \to \mathbb{R}$ is both convex and concave.

Theorem 15.3.44 (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. $\forall k \leq n$*

$$\sum_{i=1}^n x_i = \sum_{i=1}^n y_i \tag{15.25}$$

$$x_{(1)} + ... + x_{(k)} \geq y_{(1)} + ... + y_{(k)} \tag{15.26}$$

where $x_{(i)}$ denotes the ordering[10] of the tuple $(x_1, ..., x_n)$. Then

$$\sum_{i=1}^n f(x_i) \geq \sum_{i=1}^n f(y_i) \tag{15.27}$$

15.4 Inner product

In the following section all vector spaces V will be $\mathbb{R}$- or $\mathbb{C}$-vector spaces.

Notation 15.4.1 (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 3 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 15.4.2. Inner products are special cases of **non-degenerate Hermitian forms** which do not possses the positive-definiteness property.

Corollary 15.4.3. The first two properties have the result of conjugate linearity in the second argument:

$$\langle f|\lambda g + \mu h\rangle = \overline{\lambda}\langle f|g\rangle + \overline{\mu}\langle f|h\rangle \tag{15.28}$$

[10] In decreasing order: $x_{(1)} \geq ... \geq x_{(n)}$.

15.4.1 Inner product space

Definition 15.4.4 (Inner product space[11]). A vector space equipped with an inner product $\langle\cdot|\cdot\rangle$ is called an inner product space.

Definition 15.4.5 (Metric dual[12]). 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:

$$L : V \to V^* : v \mapsto \langle v|\cdot\rangle \tag{15.29}$$

Definition 15.4.6 (Adjoint operator). Let A be a linear operator on V. Let v, w be two vectors in V. The *Hermitian* adjoint of A is defined as the linear operator $A^\dagger$ that satisfies:

$$\langle A^\dagger v, w\rangle = \langle v, Aw\rangle \tag{15.30}$$

Alternatively one can define the adjoint using the metric dual $L(\cdot)$ as follows:

$$\boxed{A^\dagger = L^{-1} \circ A^T \circ L} \tag{15.31}$$

If $A = A^\dagger$ then A is said to be **Hermitian** or **self-adjoint**.

Corollary 15.4.7. The Hermitian adjoint of a complex matrix $A \in \mathbb{C}^{m\times n}$ is given by:

$$A^\dagger = \overline{A}^T \tag{15.32}$$

where $\overline{A}$ denotes the complex conjugate of A and A^T the transpose of A.

The definition of an adjoint operator 15.4.6 can be generalized to the case where $A^\dagger$ is not unique (for example when A is not globally defined) in the following way:

Definition 15.4.8 (Conjugate operators). Two operators B and C are said to be conjugate if:

$$\langle Bx, y\rangle = \langle x, Cy\rangle \tag{15.33}$$

Example 15.4.9. The Lie algebra associated with the group of isometries $\mathrm{Isom}(V)$ of a non-degenerate Hermitian form satisfies following condition:

$$\langle Xv, w\rangle = -\langle v, Xw\rangle \tag{15.34}$$

for all Lie algebra elements X. It follows that the Lie algebra consists of all anti-hermitian operators.

[11] Sometimes called a **prehilbert space**.

[12] See also definition 27.1.

15.4.2 Orthogonality

Definition 15.4.10 (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:

$$\langle v|w\rangle = 0 \tag{15.35}$$

An orthogonal **system** is a set of vectors, none of them the null vector, that are mutually orthogonal.

Property 15.4.11. Orthogonal systems are linearly independent.

Definition 15.4.12 (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:

$$\langle v|v\rangle = 1 \tag{15.36}$$

Definition 15.4.13 (Orthogonal complement[13]). Let W be a subspace of V. The following subspace is called the orthogonal complement of W:

$$W^\perp = \{v \in V \mid \forall w \in W : \langle v|w\rangle = 0\} \tag{15.37}$$

Property 15.4.14. The inner-product is invariant under transformations between orthonormal bases.

Property 15.4.15.

$$W \cap W^\perp = \{0\} \tag{15.38}$$

Property 15.4.16. Let V be a finite-dimensional K-vector space. The orthogonal complement $W^\perp$ is a complementary subspace[14] to W, i.e. $W \leq V$: $W \oplus W^\perp = V$.

Corollary 15.4.17. Let $W \leq V$ where V is a finite-dimensional K-vector space. We have the following relation:

$$(W^\perp)^\perp = W \tag{15.39}$$

Definition 15.4.18 (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:

$$\mathrm{proj}_W(v) = \sum_{i=1}^k \langle v|w_i\rangle w_i \tag{15.40}$$

$$\mathrm{proj}_w(v) = \frac{\langle v|w\rangle}{\langle w|w\rangle} w \tag{15.41}$$

[13] $W^\perp$ is pronunciated as 'W-perp'.
[14] hence the name

Property 15.4.19.

1. $\forall w \in W : \mathrm{proj}_W(w) = w$
2. $\forall u \in W^\perp : \mathrm{proj}_W(u) = 0$

Method 15.4.20 (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:

$$\begin{aligned} w_1 &= u_1 & e_1 &= \frac{w_1}{||w_1||} \\ w_2 &= u_2 - \frac{\langle u_2|w_1\rangle}{||u_2||^2}w_1 & e_2 &= \frac{w_2}{||w_2||} \\ &\vdots & &\vdots \\ w_n &= u_n - \sum_{k=1}^{n-1}\frac{\langle u_n|w_k\rangle}{||u_n||^2}w_k & e_n &= \frac{w_n}{||w_n||} \end{aligned} \tag{15.42}$$

15.4.3 Angle

Definition 15.4.21 (Angle). Let v, w be elements of an inner product space. The angle θ between v and w is defined as:

$$\boxed{\cos\theta = \frac{\langle v|w\rangle}{||v||||w||}} \tag{15.43}$$

15.5 Matrices

Notation 15.5.1. The set of all $m \times n$-matrices defined over the field K is denoted as $M_{m,n}(K)$ or $\mathrm{Mat}_{m,n}(K)$. If $m = n$, the set is denoted as $M_n(K)$ or $Mat_n(K)$.

Property 15.5.2 (Dimension). The dimension of $M_{m,n}(K)$ is mn.

Definition 15.5.3 (Trace). Let $A = (a_{ij}) \in M_n(K)$. We define the trace of A as follows:

$$\boxed{\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}} \tag{15.44}$$

Property 15.5.4. 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 15.5.5 (Hilbert-Schmidt norm). Also called the **Frobenius norm**. This matrix norm is given by following formula

$$||A||_{HS}^2 = \sum_{i,j} |A_{ij}|^2 = \mathrm{tr}(A^\dagger A) \tag{15.45}$$

If one identifies $M_n(\mathbb{C})$ with $\mathbb{C}^{2n}$ then this norm equals the standard Hermitian norm.

Formula 15.5.6 (Hadamard product). Let A, B be two matrices in $M_{m\times n}(K)$. The Hadamard product is defined as the entry-wise product:

$$(A \circ B)_{ij} = A_{ij} B_{ij} \tag{15.46}$$

Definition 15.5.7 (General linear group). The set of invertable matrices is called the general linear group and is denoted by $GL_n(K)$.

Property 15.5.8. For all $A \in GL_n(K)$ we have:

- $A^T \in GL_n(K)$
- $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$

Property 15.5.9. Let $A \in M_{m,n}(K)$. Denote the set of columns of A as $\{A_1, A_2, ..., A_n\}$ and the set of rows of A as $\{R_1, R_2, ..., R_m\}$. The set of columns is a subspace of K^m and the set of rows is a subspace of K^n. Furthermore we have:

$$\dim(\mathrm{span}(A_1, ..., A_n)) = \dim(\mathrm{span}(R_1, ..., R_m))$$

Definition 15.5.10 (Rank of a matrix). We can define the rank of matrix $A \in M_{m,n}(K)$ as follows:

$$\mathrm{rk}(A) := \dim(\mathrm{span}(A_1, ..., A_n)) \overset{15.5.9}{=} \dim(\mathrm{span}(R_1, ..., R_m)) \tag{15.47}$$

Property 15.5.11. 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 $A \in GL_n(K)$ and $B \in M_{n,r}(K)$. We have $\mathrm{rk}(AB) = \mathrm{rk}(B)$.
3. Let $A \in GL_n(K)$ and $B \in M_{r,n}(K)$. We have $\mathrm{rk}(BA) = \mathrm{rk}(B)$.

Property 15.5.12. Let $A \in M_{m,n}(K)$. First define the following linear map:

$$\boxed{L_A : K^n \to K^m : v \mapsto Av} \tag{15.48}$$

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 15.47.

15.5.1 System of equations

Theorem 15.5.13. *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 15.48. 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 15.5.14 (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.*

15.5.2 Coordinates and matrix representations

Definition 15.5.15 (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 15.5.16 (Coordinate isomorphism). With the previous definition in mind we can define the coordinate isomorphism of v with respect to $\mathcal{B}$ as follows:

$$\boxed{\beta : V \to K^n : \sum_{i=1}^n \lambda_i b_i \mapsto (\lambda_1, ..., \lambda_n)^T} \tag{15.49}$$

Definition 15.5.17 (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 15.5.18 (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 15.5.19 (Construction of a matrix representation). From definition 15.5.17 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 15.5.20. *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:*

$$\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \end{pmatrix} = A_{f,\mathcal{B},\mathcal{C}} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \tag{15.50}$$

Theorem 15.5.21. *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 15.5.22. *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 15.5.23. *Let β and γ be the coordinate isomorphisms with respect to respectively $\mathcal{B}$ and $\mathcal{C}$. From theorem 15.5.20 it follows that:*

$$\gamma(f(v)) = A_f \cdot \beta(v) \tag{15.51}$$

or alternatively

$$\gamma \circ f = L_{A_f} \circ \beta \tag{15.52}$$

Theorem 15.5.24. *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:*

$$A_{g \circ f} = A_g A_f \tag{15.53}$$

Theorem 15.5.25. *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:*

$$A_{g \circ f} = A_g A_f \tag{15.54}$$

Theorem 15.5.26. *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[15]*:*

$$GL_K(V) \to GL_n(K) : f \mapsto A_f \tag{15.55}$$

Remark 15.5.27. The sets $GL_K(V)$ and $GL_N(K)$ are groups. So the previous theorem states that the map $f \mapsto A_f$ is a group isomorphism.

Theorem 15.5.28. *Let $V = K^n$. Let $f \in V^*$. From construction 15.5.19 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 15.5.20 gives:*

$$f(\lambda_1, ..., \lambda_n)^T = (f(e_1), ..., f(e_n))(\lambda_1, ..., \lambda_n)^T = \sum_{i=1}^{n} f(e_i)\lambda_i \tag{15.56}$$

or alternatively with $\{\varepsilon_1, ..., \varepsilon_n\}$ the dual basis to the standard basis of V:

$$\boxed{f = \sum_{i=1}^{n} f(e_i)\varepsilon_i} \tag{15.57}$$

[15] Follows from theorem 15.5.25.

Theorem 15.5.29. *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 $\mathcal{B}$.*

15.5.3 Coordinate transforms

Definition 15.5.30 (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 $\mathcal{B}$:

$$b'_j = q_{1j}b_1 + ... + q_{nj}b_n \tag{15.58}$$

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 15.5.31. *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 15.5.32. *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:*

$$A' = P^{-1}AQ \tag{15.59}$$

Corollary 15.5.33. Let $f \in \text{End}_K(V)$ and let Q be the transition matrix. From theorem 15.5.32 it follows that:

$$A' = Q^{-1}AQ \tag{15.60}$$

Definition 15.5.34 (Matrix conjugation). Let $A \in M_n(K)$. The set

$$\{Q^{-1}AQ \mid Q \in GL_n(K)\} \tag{15.61}$$

is called the conjugacy class[16] of A. Another name for conjugation is **similarity transformation**.

Remark 15.5.35. 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 15.5.36. From property 15.5.4 it follows that the trace of a matrix is invariant under similarity transformations:

$$\boxed{\mathrm{tr}(Q^{-1}AQ) = \mathrm{tr}(A)} \tag{15.62}$$

Definition 15.5.37 (Matrix congruence). Let $A, B \in M_n(K)$. If there exists a matrix P such that

$$A = P^T BP \tag{15.63}$$

then the matrices are said to be congruent.

Property 15.5.38. Every matrix congruent to a symmetric matrix is also symmetric.

Theorem 15.5.39. *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. We can find the following result:*

$$Q^T Q = \mathbb{1}_n$$

15.5.4 Determinant

Definition 15.5.40 (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 15.5.41 (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 15.5.42 (Adjugate matrix). The adjugate matrix of $A \in M_n(K)$ is defined as follows:

$$\mathrm{adj}(A) := \begin{pmatrix} \alpha_{11} & \alpha_{21} & \cdots & \alpha_{n1} \\ \alpha_{12} & \alpha_{22} & \cdots & \alpha_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_{1n} & \alpha_{2n} & \cdots & \alpha_{nn} \end{pmatrix} \tag{15.64}$$

or shorter: $\mathrm{adj}(A) = (\alpha_{ij})^T$.

[16] This is the general definition of conjugacy classes for groups. Furthermore, these classes induce a partitioning of the group.

Remark. It is important to notice that we have to transpose the matrix after the elements have been replaced by their cofactor.

Property 15.5.43. 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. $\det(A_{i_1}, \ldots, A_{i_n}) = \operatorname{sgn}(i_1, \ldots, i_n)\det(A_1, \ldots, A_n)$
6. $\det(A_1, \ldots, A_n) = \det(A_1, \ldots, A_i + \lambda A_k, \ldots, A_n)$ where A_i, A_k are columns of A.
7. The determinant can be evaluated as follows:

$$\det(A) = \sum_{i=1}^{n}(-1)^{i+k} a_{ik} \det(A_{ik}) \tag{15.65}$$

Theorem 15.5.44. *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 15.5.45. *For all $A \in M_n(K)$ we find $A\,adj(A) = adj(A)A = \det(A)I_n$.*

Theorem 15.5.46. *For all $A \in GL_n(K)$ we find $A^{-1} = \det(A)^{-1}\ adj(A)$.*

An alternative definition of a $k \times k$-minor is:

Definition 15.5.47. 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 15.5.48. *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 $k \times k$-minor different from 0.*

Theorem 15.5.49. *Let $f \in End_K(V)$. The determinant of the matrix representation of f is invariant under basis transformations.*

Definition 15.5.50 (Determinant of a linear operator). The previous theorem allows us to unambiguously define the determinant of f as follows:

$$\det(f) := \det(A)$$

where A is some matrix representation of f.

15.5.5 Characteristic polynomial

Definition 15.5.51 (Characteristic polynomial[17]). 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:

$$\chi_f(x) := \det(x\mathbb{1}_n - A) \in K[x] \tag{15.66}$$

is a monic polynomial of degree n in the variable x and the polynomial does not depend on the choice of basis.

Definition 15.5.52 (Characteristic equation[18]). The following equation is called the characteristic equation of f:

$$\boxed{\chi_f(x) = 0} \tag{15.67}$$

Formula 15.5.53. 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:

$$\begin{cases} c_0 = (-1)^n \det(A) \\ c_{n-1} = -\mathrm{tr}(A) \end{cases}$$

Theorem 15.5.54 (Cayley-Hamilton).

1. *Let $A \in M_n(K)$ with characteristic polynomial $\chi_A(x)$. We find the following relation:*

$$\chi_A(A) = A^n + \sum_{i=1}^{n-1} c_i A^i = 0 \tag{15.68}$$

2. *Let $f \in End_K(V)$ with characteristic polynomial $\chi_f(x)$. We find that*

$$\chi_f(f) = f^n + \sum_{i=1}^{n-1} c_i f^i = 0 \tag{15.69}$$

Corollary 15.5.55. From theorem 15.3.32 and the Cayley-Hamilton theorem it follows that the minimal polynomial $\mu_f(x)$ is a divisor of the characteristic polynomial $\chi_f(x)$.

[17] This polynomial can also be used directly for a matrix A as theorem 15.5.21 matches every matrix A with some linear operator f.

[18] This equation is sometimes called the **secular equation**.

15.5.6 Linear groups

Definition 15.5.56 (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 15.5.57 (Elementary matrix). $E_{ij}(c)$ is the elementary matrix with element c on the i,j-th position.

Property 15.5.58. We have the following property:

$$\det(E_{ij}(c)) = 1 \tag{15.70}$$

which implies that $E_{ij}(c) \in GL_n(K)$.

Property 15.5.59. 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 15.5.60. *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 15.5.61 (Special linear group). The following subset of $GL_n(K)$ is called the special linear group:

$$SL_n(K) = \{A \in GL_n(K) \mid \det(A) = 1\} \tag{15.71}$$

Theorem 15.5.62. *Every $A \in SL_n(K)$ can be written as a product of elementary matrices.*[19]

Definition 15.5.63 (Orthogonal group). The orthogonal and special orthogonal group are defined as follows:

$$\begin{aligned} O_n(K) &= \{A \in GL_n(K) \mid AA^T = A^TA = I_n\} \\ SO_n(K) &= O_n(K) \cap SL_n(K) \end{aligned}$$

Property 15.5.64. For orthogonal matrices, conjugacy 15.5.34 and congruency 15.5.37 are equivalent.

[19]Follows readily from theorem 15.5.60.

Definition 15.5.65 (Unitary group). The unitary and special unitary group are defined as follows:

$$U_n(K,\sigma) = \{A \in GL_n(K) \mid A\overline{A}^T = \overline{A}^T A = I_n\}$$
$$SU_n(K,\sigma) = U_n(K) \cap SL_n(K)$$

where σ denotes the *involution*[20] $a^\sigma \equiv \overline{a}$.

Remark. If $K = \mathbb{C}$ where the involution is taken to be the complex conjugate, the σ is often ommited in the definition: $U_n(K)$ and $SU_n(K)$.

Definition 15.5.66 (Unitary equivalence). Let A, B be two matrices in $M_n(K)$. 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**.

15.6 Eigenvectors

Definition 15.6.1 (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:

$$f(v) = \lambda v \tag{15.72}$$

Where $\lambda \in K$ is the **eigenvalue** belonging to v.

Definition 15.6.2 (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:

$$\ker(\lambda \mathbf{1}_V - f) \tag{15.73}$$

Theorem 15.6.3 (Characteristic equation[21]). *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 15.67.*

Theorem 15.6.4. *A linear operator $f \in End_K(V)$ defined over an n-dimensional K-vector space V has at most n different eigenvalues.*[22]

Method 15.6.5 (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 15.6.3.
2. Then we find the eigenvector v_i belonging to the eigenvalue λ_i by using the following equation:

$$(\mathbf{A} - \lambda_i \mathbf{1}_V)\, v_i = 0 \tag{15.74}$$

[20] An involution is an operator that is its own inverse: $f(f(x)) = x$.
[21] This theorem also holds for the eigenvalues of a matrix $A \in M_n(K)$.
[22] This theorem also holds for a matrix $A \in M_n(K)$.

15.6.1 Diagonalization

Definition 15.6.6 (Diagonalizable operator). An operator $f \in \text{End}_K(V)$ on a finite-dimensional K-vector space V is diagonalizable if there exists a matrix representation $A \in M_n(K)$ of f such that A is a diagonal matrix.

Theorem 15.6.7. *A linear operator f defined on a finite-dimensional K-vector space V has a diagonal matrix as matrix representation if and only if the set of eigenvectors of f is a basis of V.*

Theorem 15.6.8. *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 15.6.9. Using the fact that the trace of a linear operator is invariant under similarity transformations (see property 15.62) we get following useful formula:

$$\boxed{\text{tr}(f) = \sum_i \lambda_i} \tag{15.75}$$

where $\{\lambda_i\}_{0\leq i\leq n}$ are the eigenvalues of f.

Property 15.6.10. Let V be an n-dimensional K-vector space. Let $f \in \text{End}_K(V)$ 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.

Definition 15.6.11 (Multiplicity). Let V be a K-vector space. Let $f \in \text{End}_K(V)$ be a linear operator with characteristic polynomial[23]:

$$\chi_f(x) = \prod_{i=1}^{n} (x - \lambda_i)^{n_i} \tag{15.76}$$

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 15.6.12. The geometric multiplicity is always at least 1.

[23] 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.

Property 15.6.13. The algebraic multiplicity is always greater than or equal to the geometric multiplicity.

Theorem 15.6.14. *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 15.6.15. Every Hermitian operator $f \in \text{End}_K(\mathbb{C}^n)$ 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.[24]

Property 15.6.16. Let $A, B \in \text{End}_K(V)$ be two linear operators. If the commutator $[A, B] = 0$, then the two operators have a common eigenbasis.

Theorem 15.6.17 (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*[25]*.*

15.7 Euclidean space $\mathbb{R}^n$

A finite-dimensional $\mathbb{R}$-vector space is called a **Euclidean space**.

15.7.1 Angle

Definition 15.7.1 (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[26]:

$$\angle(u, v) = \text{acos}\frac{\langle u|v\rangle}{||u|| \cdot ||v||} \tag{15.77}$$

where we set the range of acos as $[0, \pi]$.

Notation 15.7.2. When working in a Euclidean space the inner product $\langle v|w\rangle$ is often written as $v \cdot w$ or even vw.

[24] This implies that the matrix P diagonalizing the Hermitian operator is unitary, i.e. $P^{-1} = P^\dagger$.

[25] Also with respect to conjugation, which are equivalent to similarity transformations according to property 15.5.64.

[26] This formula follows readily from the Cauchy-Schwarz inequality (see theorem 17.2.6).

15.7.2 Vector product

Definition 15.7.3 (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 15.7.4 (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 15.7.5. It is convenient to take the standard basis $(e_1, \ldots, e_n)$ to be positively oriented.

Formula 15.7.6 (Cross product).

$$\boxed{(v \times w)_i = \varepsilon_{ijk} v_j w_k} \tag{15.78}$$

where ε_{ijk} is the 3-dimensional Levi-Civita symbol.

Remark 15.7.7. It is important to note that the previous construction is only valid in 3 dimenensions.

Chapter 16

Vector calculus

16.1 Nabla-operator

Definition 16.1.1 (Nabla).

$$\nabla \equiv \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \tag{16.1}$$

Following formulas can be found by using basic properties of (vector) calculus.

Formula 16.1.2 (Gradient).

$$\nabla V = \left(\frac{\partial V_x}{\partial x}, \frac{\partial V_y}{\partial y}, \frac{\partial V_z}{\partial z}\right) \tag{16.2}$$

Formula 16.1.3. Let $\varphi(\vec{\boldsymbol{x}})$ be a scalar field. The total differential $d\varphi$ can be rewritten as

$$d\varphi = \nabla\varphi \cdot d\vec{r} \tag{16.3}$$

Property 16.1.4. The gradient of a scalar function V is perpendicular to the level sets 2.9 of V.

Definition 16.1.5 (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 $\vec{a}$:

$$\nabla_{\vec{\boldsymbol{a}}}V \equiv (\vec{a} \cdot \nabla)V \tag{16.4}$$

Example 16.1.6. 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

$$\frac{\partial\varphi}{\partial s} = \frac{d\vec{r}}{ds} \cdot \nabla\varphi \tag{16.5}$$

Definition 16.1.7 (Conservative vector field). A vector field obtained as the gradient of a scalar function.

Property 16.1.8. A vector field is conservative if and only if its line integral is path independent.

Formula 16.1.9 (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:

$$\nabla T = \frac{\partial T}{\partial x^i} \otimes \vec{\boldsymbol{e}}^i \tag{16.6}$$

Formula 16.1.10 (Divergence).

$$\nabla \cdot \vec{\boldsymbol{A}} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \tag{16.7}$$

Definition 16.1.11 (Solenoidal vector field). A vector field $\vec{\boldsymbol{V}}(\vec{\boldsymbol{x}})$ is said to be solenoidal if it satisfies:

$$\nabla \cdot \vec{\boldsymbol{V}} = 0 \tag{16.8}$$

It is also known as a **divergence free vector field**.

Formula 16.1.12 (Rotor / curl).

$$\nabla \times \vec{\boldsymbol{A}} = \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}, \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}, \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \tag{16.9}$$

Definition 16.1.13 (Irrotational vector field). A vector field $\vec{\boldsymbol{V}}(\vec{\boldsymbol{x}})$ is said to be irrotational if it satisfies:

$$\nabla \times \vec{\boldsymbol{V}} = 0 \tag{16.10}$$

Remark 16.1.14. All conservative vector fields are irrotational but irrotational vector fields are only conservative if the domain is simply-connected[2]

16.1.1 Laplacian

Definition 16.1.15 (Laplacian).

$$\triangle V \equiv \nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} \tag{16.11}$$

$$\nabla^2 \vec{\boldsymbol{A}} = \nabla \left(\nabla \cdot \vec{\boldsymbol{A}} \right) - \nabla \times \left(\nabla \times \vec{\boldsymbol{A}} \right) \tag{16.12}$$

Remark 16.1.16. Equation 16.12 is called the **vector laplacian**.

Formula 16.1.17 (Laplacian in different coordinate systems).

[1] See definition 25.3.21.
[2] See definition 4.6.9.

- Cylindrical coordinates (ρ, ϕ, z):

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial}{\partial\rho}\right) + \frac{1}{\rho^2}\frac{\partial^2}{\partial\phi^2} + \frac{\partial^2}{\partial z^2} \tag{16.13}$$

- Spherical coordinates (r, ϕ, θ):

$$\frac{1}{r^2}\left[\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)\right] \tag{16.14}$$

16.1.2 Mixed properties[3]

$$\nabla \times (\nabla V) = 0 \tag{16.15}$$

$$\nabla \cdot \left(\nabla \times \vec{\boldsymbol{V}}\right) = 0 \tag{16.16}$$

In Cartesian coordinates equation 16.12 can be rewritten as follows:

$$\nabla^2\vec{\boldsymbol{A}} = (\triangle A_x, \triangle A_y, \triangle A_z) \tag{16.17}$$

16.1.3 Helmholtz decomposition

Formula 16.1.18 (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:

$$\vec{\boldsymbol{P}} = \nabla \times \vec{\boldsymbol{A}} + \nabla V \tag{16.18}$$

16.2 Line integrals

Formula 16.2.1 (Line integral of a continuous scalar field). Let f be a continuous scalar 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:

$$\int_\Gamma f(s)ds = \int_a^b f(\vec{\boldsymbol{\varphi}}(t))||\vec{\boldsymbol{\varphi}}'(t)||dt \tag{16.19}$$

Formula 16.2.2 (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:

$$\int_\Gamma \vec{\boldsymbol{F}}(\vec{\boldsymbol{s}}) \cdot d\vec{\boldsymbol{s}} = \int_a^b \vec{\boldsymbol{F}}(\vec{\boldsymbol{\varphi}}(t)) \cdot \vec{\boldsymbol{\varphi}}'(t)dt \tag{16.20}$$

[3]See remark 25.5.10 for a differential geometric approach.

16.3 Integral theorems[4]

Theorem 16.3.1 (Fundamental theorem of calculus for line integrals).
Let $\vec{\Gamma}:\mathbb{R}\to\mathbb{R}^3$ be a smooth curve.

$$\int_{\Gamma(a)}^{\Gamma(b)} \nabla f(\vec{\boldsymbol{r}})\cdot d\vec{\boldsymbol{r}} = \varphi(\Gamma(b)) - \varphi(\Gamma(a)) \tag{16.21}$$

Theorem 16.3.2 (Kelvin-Stokes' theorem).

$$\oint_{\partial S} \vec{\boldsymbol{A}}\cdot d\vec{\boldsymbol{l}} = \iint_S \left(\nabla\times\vec{\boldsymbol{A}}\right) dS \tag{16.22}$$

Theorem 16.3.3 (Divergence theorem[5]).

$$\oiint_{\partial V} \vec{\boldsymbol{A}}\cdot d\vec{\boldsymbol{S}} = \iiint_V \left(\nabla\cdot\vec{\boldsymbol{A}}\right) dV \tag{16.23}$$

Corollary 16.3.4 (Green's identity).

$$\oiint_{\partial V} (\psi\nabla\phi - \phi\nabla\psi)\cdot d\vec{\boldsymbol{S}} = \iiint_V \left(\psi\nabla^2\phi - \phi\nabla^2\psi\right) dV \tag{16.24}$$

16.4 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 21.14. Also there is no Einstein summation used, all summations are written explicitly.

Formula 16.4.1 (Unit vectors).

$$\frac{\partial\vec{r}}{\partial q^i} = h_i\hat{e}_i \tag{16.25}$$

Formula 16.4.2 (Gradient).

$$\nabla V = \sum_{i=1}^{3}\frac{1}{h_i}\frac{\partial V}{\partial q^i}\hat{e}_i \tag{16.26}$$

Formula 16.4.3 (Divergence).

$$\nabla\cdot\vec{\boldsymbol{A}} = \frac{1}{h_1h_2h_3}\left(\frac{\partial}{\partial q^1}(A_1h_2h_3) + \frac{\partial}{\partial q^2}(A_2h_3h_1) + \frac{\partial}{\partial q^3}(A_3h_1h_2)\right) \tag{16.27}$$

Formula 16.4.4 (Rotor).

$$(\nabla\times\vec{\boldsymbol{A}})_i = \frac{1}{h_jh_k}\left(\frac{\partial}{\partial q^j}(A_kh_k) - \frac{\partial}{\partial q^k}(A_jh_j)\right) \tag{16.28}$$

where $i\neq j\neq k$.

[4] These theorems follow from the more general Stokes' theorem 26.3.
[5] Also known as *Gauss's theorem* or the *Gauss-Ostrogradsky theorem*.

Chapter 17

Banach spaces and Hilbert spaces

In this chapter the term "linear operator", which is normally reserved for maps of the form $f : V \to V$, is 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 15.4.

17.1 Banach spaces

Definition 17.1.1 (Norm). Let V be a K-vector space. 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 17.1.2. A norm $||\cdot||$ clearly induces a metric[1] by setting $d(x, y) = ||x - y||$.

Definition 17.1.3 (Normed vector space). A K-vector space equipped with a norm $||\cdot||$.

Definition 17.1.4 (Banach space). A normed vector space that is complete[2] with respect to the norm $||\cdot||$.

Definition 17.1.5 (Reflexive space). A Banach space V for which its dual coincides with the dual of its dual, i.e. $V^* = (V^*)^*$.

Property 17.1.6. Every finite-dimensional Banach spaces is reflexive. This follows from property 15.3.35.

[1] See definition 5.1.1.
[2] See condition 5.12.

Property 17.1.7. 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 17.1.8. The topological (continuous) dual of a Banach space is also a Banach space.

17.1.1 Bounded operators

Definition 17.1.9 (Bounded operator). Let $L : V \to W$ be a linear operator between two Banach spaces. The operator is said to be bounded if there exists a scalar M that satisfies the following condition:

$$\boxed{\forall v \in V : ||Lv||_W \leq M||v||_V} \tag{17.1}$$

Notation 17.1.10. The space of bounded linear operators from V to W is denoted by $\mathcal{B}(V, W)$.

Property 17.1.11. If V is a Banach space then $\mathcal{B}(V)$ is also a Banach space.

Definition 17.1.12 (Operator norm). The operator norm of L is defined as follows:

$$||L||_{op} = \inf\{M|M \text{ satisfies condition 17.1}\} \tag{17.2}$$

As the name suggests it is a norm on $\mathcal{B}(V, W)$. The topology induced by this norm is called the norm topology.

Equivalent definitions of the operator norm are:

$$||L||_{op} = \sup_{||x||\leq 1} ||L(x)|| = \sup_{||x||=1} ||L(x)|| = \sup_{x\neq 0} \frac{||L(x)||}{||x||} \tag{17.3}$$

Following property reduces the problem of continuity to that of boundedness.

Property 17.1.13. 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 17.1.14. Let A be a bounded linear operator with eigenvalue λ. We then have:

$$|\lambda| \leq ||A||_{op} \tag{17.4}$$

Property 17.1.15. Let A be a bounded linear operator. Let $A^\dagger$ denote its adjoint[3]. Then $A^\dagger$ is bounded and $||A||_{op} = ||A^\dagger||_{op}$.

[3]See definition 15.4.6.

17.1.2 Theorems

Property 17.1.16. Let X be a general TVR. Every linear map $\varphi : \mathbb{K}^n \to X$ is continuous.

Property 17.1.17. Let X be a finite-dimensional normed vector space. Every linear bijection $\varphi : \mathbb{K}^n \to X$ is a homeomorphism.

Corollary 17.1.18. 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 17.1.19 (Open mapping theorem[4]). *Let $f : V \to W$ be a continuous linear operator between two Banach spaces. If f is surjective then it also open.*

17.1.3 Spectrum

Definition 17.1.20 (Resolvent set). Let A be a bounded linear operator on a normed space V. The resolvent set $\rho(A)$ consists of all scalar $\lambda \in \mathbb{C}$ such that $(A - \lambda\mathbb{1})^{-1}$ is a bounded linear operator, called the resolvent of A, on a dense subset of V. These scalars λ are called **regular values** of A.

Definition 17.1.21 (Spectrum). The set of scalars $\mu \notin \rho(A)$ is called the spectrum of A.

Remark 17.1.22. 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 17.1.23 (Point spectrum). The set of scalars $\mu \in \mathbb{C}$ for which the resolvent of A fails to be injective is called the point spectrum of A. This set contains exactly the eigenvalues of A.

Definition 17.1.24 (Continuous spectrum). The set of scalars $\mu \in \mathbb{C}$ for which the resolvent of A fails to be surjective but for which the range of the resolvent is dense in V is called the continuous spectrum of A. The scalars for which the range is not dense is called the **residual spectrum** $\sigma_r(A)$.

Definition 17.1.25 (Compression spectrum). The set of scalars $\mu \in \mathbb{C}$ for which the resolvent of A fails to have a dense range in V is called the compression spectrum $\sigma(A)$. It follows that $\sigma_r(A) \subseteq \sigma(A)$.

17.2 Hilbert space

Definition 17.2.1 (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).

[4] Sometimes called the *Banach-Schauder* theorem.

Example 17.2.2. Let $f, g \in \mathcal{L}^2([a,b],\mathbb{C})$, the inner product of f and g is defined as:

$$\boxed{\langle f|g\rangle = \int_a^b f^*(x)\overline{g(x)}dx} \tag{17.5}$$

Remark 17.2.3. See section 9.4.2 for a more formal treatment of this subject.

Formula 17.2.4. It is also possible to define an inner product with respect to a weight function $\phi(x)$:

$$\int_a^b f^*(x)g(x)\phi(x)dx \tag{17.6}$$

Using this formula it is possible to define orthogonality with respect to a weight function.

17.2.1 Inner products and norms

Formula 17.2.5. Let V be an inner product space. A norm on V can be induced by the inner product in the following way:

$$||v||^2 = \langle v|v\rangle \tag{17.7}$$

However not every norm induces an inner product. Only norms that satisfy the parallellogram law 17.9 induce an inner product. This inner product can be recovered through the polarization identity 17.10 (see below).

Property 17.2.6 (Cauchy-Schwarz inequality).

$$\boxed{|\langle v|w\rangle| \leq ||v||\ ||w||} \tag{17.8}$$

where the equality holds if and only if v and w are linearly dependent.

Corollary 17.2.7. 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 17.2.8 (Parallellogram law).

$$||v+w||^2 + ||v-w||^2 = 2(||v||^2 + ||w||^2) \tag{17.9}$$

Formula 17.2.9 (Polarization identity).

$$4\langle v|w\rangle = ||v+w||^2 - ||v-w||^2 + i\left(||v+iw||^2 - ||v-iw||^2\right) \tag{17.10}$$

Formula 17.2.10 (Pythagorean theorem). In an inner product space the triangle equality reduces to the well-known Pythagorean theorem for orthogonal vectors v, w:

$$||v+w||^2 = ||v||^2 + ||w||^2 \tag{17.11}$$

This formula can be extended to any set of orthogonal vectors $x_1, ..., x_n$:

$$\boxed{\left\|\sum_{i=1}^n x_i\right\|^2 = \sum_{i=1}^n ||x_i||^2} \tag{17.12}$$

17.2.2 Generalized Fourier series

Property 17.2.11 (Bessel's inequality). First of all we have following general equality for orthonormal vectors $x_1, ..., x_n$ and complex scalars $a_1, ..., a_n$:

$$\left\|x - \sum_{i=1}^{n} a_i x_i\right\|^2 = ||x||^2 - \sum_{i=1}^{n} |\langle x, x_i\rangle|^2 + \sum_{i=1}^{n} |\langle x, x_i\rangle - a_i|^2 \tag{17.13}$$

This expression becomes minimal for $a_i = \langle x, x_i\rangle$ (last term becomes 0). This leads to Bessel's inequality:

$$\boxed{\sum_{i=1}^{n} |\langle x, x_i\rangle|^2 \leq ||x||^2} \tag{17.14}$$

Corollary 17.2.12. The sum in 17.14 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 17.2.13. *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:*

$$\left\|\sum_{i=1}^{+\infty} a_i x_i\right\|^2 = \sum_{i=1}^{+\infty} |a_i|^2 \tag{17.15}$$

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 17.2.14. 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 $\mathcal{H}$.

17.2.3 Complete sets

Definition 17.2.15 (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:

$$x = \sum_{i\in I}\langle x, x_i\rangle x_i \tag{17.16}$$

This implies that a complete set is a basis for the vector space.

Another characterization is the following.

Alternative Definition 17.2.16. A complete set of orthonormal vectors is a set $S \subset V$ such that we cannot add another vector w to it satisfying:

$$\forall v_i \in S : \langle v_i, w\rangle = 0 \qquad \wedge \qquad w \neq 0 \tag{17.17}$$

Property 17.2.17. For complete sequences (x_n) the inequality of Bessel 17.14 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 17.2.18 (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 17.2.19. 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 $x = 0$.

17.2.4 Orthogonality and projections

The basic notions on orthogonality in inner product space can be found in section 15.4.2.

Property 17.2.20. 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 17.2.21. The previous property implies that the orthogonal complemement of some arbitrary subset of a Hilbert space is a Hilbert space itself.

Theorem 17.2.22 (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 17.2.23. An equivalent definition for the unique $h' \in K$ is $||h-h'|| = \inf\{||h-k|| : k \in K\}$.

Corollary 17.2.24. It follows that given a complete (or closed) subspace S the Hilbert space $\mathcal{H}$ can be decomposed as $\mathcal{H} = S \oplus S^\perp$.

17.2.5 Separable Hilbert spaces

The definition of separable spaces in the sense of point-set topology is given in 4.5.18. An equivalent definition for Hilbert spaces is the following.

Alternative Definition 17.2.25 (Separable Hilbert space). A Hilbert space is separable if it contains a complete sequence of orthonormal vectors.

Corollary 17.2.26. 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 17.2.27. *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 17.2.28. Every orthogonal subset of a separable Hilbert space is countable.

17.2.6 Compact operators

Definition 17.2.29 (Compact operator). Let A be a linear operator on a Hilbert space $\mathcal{H}$. A is said to be compact if for every sequence (x_n) in $\mathcal{H}$ the sequence $(A[x_n])$ has a convergent subsequence.

Property 17.2.30. Every compact operator on a Hilbert space is bounded.

Corollary 17.2.31. Every linear operator on a finite-dimensional Hilbert space is bounded.

17.2.7 Linear functionals

Property 17.2.32. Let f be a continuous linear functional. Then $\dim(\ker f)^\perp$ is 0 or 1 where the former case only arises when $f \equiv 0$.

Theorem 17.2.33 (Riesz' representation theorem). *Let $\mathcal{H}$ be a Hilbert space. For every continuous linear functional $\rho : \mathcal{H} \to \mathbb{R}$ there exists a unique element $x_0 \in \mathcal{H}$ such that*

$$\rho(h) = \langle h, x_0 \rangle \tag{17.18}$$

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 17.2.34. This theorem justifies the bra-ket notation used in quantum mechanics where one associates to every ket $|\psi\rangle \in \mathcal{H}$ a bra $\langle\psi| \in \mathcal{H}^*$.

Chapter 18

Operator algebras

18.1 Involutive algebras

Definition 18.1.1 (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 18.1.2 (Involutive algebra[1]). An involutive algebra is an associative algebra A over a commutative ring R with involution $^-$ together with an operator $^* : A \rightarrow A$ such that:

- $(a+b)^* = a^* + b^*$
- $(ab)^* = b^* a^*$
- $(\lambda a)^* = \overline{\lambda} a^*$

where $\lambda \in R$.

18.2 C*-algebras

Definition 18.2.1 (C*-algebra). A C*-algebra is a involutive Banach algebra[2] A such that the **C*-identity**

$$||a^* a|| = ||a||\ ||a^*|| \tag{18.1}$$

is satisfied.

Definition 18.2.2 (Positive). An element of a C*-algebra is called positive if it is self-adjoint and if its spectrum is contained in $[0, +\infty[$. A linear functional on a C*-algebra is called positive if every positive element is mapped to a positive number.

[1]Also called a ***-algebra**.

[2]See definition 17.1.4.

Definition 18.2.3 (State). Let A be a C*-algebra. A state ψ on A is a positive linear functional of unit norm.

Chapter 19

Tensor calculus

19.1 Tensor product

19.1.1 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 19.1.1. The tensor product of vector spaces V and W is defined as[1] 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:

$$\boxed{(v \otimes w)(g, h) = v(g)w(h)} \tag{19.1}$$

Definition 19.1.2 (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 $\mathbf{T}(e_i, ..., e_j, e^k, ..., e^l) = T_{i...j}{}^{k...l}$.

Property 19.1.3 (Universal property). A set X together with a bilinear map $\mathcal{T} : V \times W \to X$ is said to have the universal property if for every bilinear map $f : V \times W \to Z$, where Z is some other vector space, there exists a unique linear map $f' : X \to Z$ such that $f = f' \circ \mathcal{T}$.

Corollary 19.1.4. The tensor product is unique up to a linear isomorphism. This results in

$$V \otimes W \cong W \otimes V \tag{19.2}$$

[1] "isomorphic to" would be a better terminology. See the "universal property" 19.1.3. For a complete proof and explanation, see [13].

The isomorphism is given by:

$$v(f) \equiv f(v) \tag{19.3}$$

where $v \in V$ and $f \in V^*$.

Notation 19.1.5 (Tensor power).

$$V^{\otimes n} = \underbrace{V \otimes ... \otimes V}_{n \text{ copies}} \tag{19.4}$$

Remark 19.1.6. More generally, the tensor product of r copies of V and s copies of V^* is the vector space $\mathcal{T}_s^r(V) = V^{\otimes r} \otimes V^{*\otimes s}$. These tensors are said to be of **type** (r, s).

Remark 19.1.7. 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 $\mathcal{T}_1^0 V$ 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 15.3.35). The space $\mathcal{T}_0^1 V$ will also be isomorphic to V itself.

Definition 19.1.8. The scalars (elements of the base field K) are by definition the $(0, 0)$ tensors.

Alternative Definition 19.1.9. The tensor space $\mathcal{T}_s^r(V)$ is spanned by the basis

$$\underbrace{e_i \otimes ... \otimes e_j}_{r \text{ basis vector}} \otimes \underbrace{\varepsilon^k \otimes ... \otimes \varepsilon^l}_{s \text{ dual basis vectors}}$$

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 19.1.10 (Dimension of tensor product). From the previous construction it follows that the dimension of $\mathcal{T}_s^r(V)$ 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 19.1.1 and using the definition of the dual vectors 15.22 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. $\square$

19.2 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 transformations A and B should be well-defined, they are each other's inverse: $B = A^{-1}$.

Definition 19.2.1 (Contravariant). A tensor component that transforms by the following rule is called contravariant:

$$v^i = A^i{}_j v'^j \tag{19.5}$$

Definition 19.2.2 (Covariant). A tensor component that transforms by the following rule is called covariant:

$$p_i = B^j{}_i\, p'_j \tag{19.6}$$

Example 19.2.3 (Mixed tensor). As an example of a mixed tensor we give the transformation formula for the mixed third-order tensor T^k_{ij}:

$$T^k_{ij} = A^k{}_w B^u{}_i B^v{}_j T'^w{}_{uv}$$

Theorem 19.2.4 (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*[2]*. 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.

19.3 Tensor operations

19.3.1 General operations

Definition 19.3.1 (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

$$V \otimes V^* : e_i \otimes e^j \mapsto e^j(e_i) \tag{19.7}$$

Definition 19.3.2 (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 19.3.3. Let $A^i{}_k$ and $B^j{}_{lm}$ be two tensors. The direct product is equal to:

$$C^{i\ j}_{\ k\ lm} = A^i{}_k B^j{}_{lm}$$

[2] This rule does not necessarily hold when $B = 0$ as transformations rules are not defined for the null-tensor.

Formula 19.3.4 (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:

$$\boxed{(\hat{A}\otimes\hat{B})(v\otimes w) = (\hat{A}v)\otimes(\hat{B}w)} \tag{19.8}$$

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.

19.3.2 Determinant

Definition 19.3.5 (n-form). An n-form is a totally anti-symmetric element $\omega\in\mathcal{T}_n^0V$. $\dim V$-forms are also called **top forms** or **volume forms**.

Definition 19.3.6 (Determinant). Let φ be an element in $\mathcal{T}_1^1V\cong\mathrm{End}(V)$. Let ω be a volume form and let $\{e_i\}_{i\leq n}$ be a basis for V. The determinant of φ is then defined as:

$$\det\varphi = \frac{\omega(\varphi(e_1),...,\varphi(e_n))}{\omega(e_1,...,e_n)} \tag{19.9}$$

This definition is well-defined, i.e. it is independent of the choice of volume form and basis. Furthermore it coincides with definition 15.5.50.

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 15.5.50 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 25.2).

19.3.3 Differentiation

Property 19.3.7.

$$\vec{\boldsymbol{\nabla}}\cdot(\vec{\boldsymbol{A}}\otimes\vec{\boldsymbol{B}}) = (\vec{\boldsymbol{\nabla}}\cdot\vec{\boldsymbol{A}})\vec{\boldsymbol{B}} + (\vec{\boldsymbol{A}}\cdot\vec{\boldsymbol{\nabla}})\vec{\boldsymbol{B}} \tag{19.10}$$

19.3.4 Levi-Civita tensor

Definition 19.3.8 (Levi-Civita tensor). Let e^i be the dual vector to e_i. In n dimensions, we define the Levi-Civita tensor as follows:

$$\boldsymbol{\varepsilon} = \varepsilon_{12...n}e^1\otimes e^2\otimes...\otimes e^n \tag{19.11}$$

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 19.3.9. 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).

Formula 19.3.10 (Cross product). By using the Levi-Civita symbol, we can define the i-th component of the cross product[3] of two vectors $\vec{\boldsymbol{v}}, \vec{\boldsymbol{w}}$ as follows:

$$\boxed{(\vec{\boldsymbol{v}} \times \vec{\boldsymbol{w}})_i = \varepsilon_{ijk} v_j w_k} \tag{19.12}$$

19.3.5 Complexification

Definition 19.3.11 (Complexification). Let V be a real vector space. The complexification of V is defined as the following tensor product:

$$V^{\mathbb{C}} = V \otimes \mathbb{C} \tag{19.13}$$

On its own this remains a real vector space. However we can turn this space into a complex vector space by generalizing the scalar product as follows:

$$\alpha(v \otimes \beta) = v \otimes (\alpha\beta) \tag{19.14}$$

for all $\alpha, \beta \in \mathbb{C}$.

Property 19.3.12. By noting that every element $\overline{v} \in V^{\mathbb{C}}$ can be written as

$$\overline{v} = (v_1 \otimes 1) + (v_2 \otimes i)$$

we can decompose the complexification as follows:

$$V^{\mathbb{C}} \cong V \oplus iV \tag{19.15}$$

19.4 (Anti)symmetric tensors

19.4.1 Symmetric tensors

Notation 19.4.1. The space of symmetric $(0, n)$ tensors is denoted by $S^n(V^*)$. The space of symmetric $(n, 0)$ tensors is denoted by $S^n(V)$.

[3] Following from remark 19.3.9 we can see that the cross product is in fact not a vector, but a pseudovector.

19.4.2 Antisymmetric tensors

Definition 19.4.2 (Antisymmetric tensor). Tensors that change sign under the interchange of any two indices.

Notation 19.4.3. 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 19.4.4. Let $n = \dim(V)$. $\Lambda^r(V)$ equals the null-space for all $r \geq n$.

19.4.3 Wedge product

Definition 19.4.5 (Wedge product).

$$\boxed{f \wedge g = f \otimes g - g \otimes f} \tag{19.16}$$

From this definition it immediately follows that the wedge product is antisymmetric.

Formula 19.4.6. Let $\{P_i\}_i$ be the set of all permutations of the sequence $(1, ..., k)$.

$$e_1 \wedge ... \wedge e_k = \sum_i \operatorname{sgn}(P_i) e_{P_i(1)} \otimes ... \otimes e_{P_i(k)} \tag{19.17}$$

Construction 19.4.7. Let $\{e_i\}_{1 \leq i \leq n}$ be a basis for V. It is clear from the definition 19.16 that a basis for $\Lambda^r(V)$ is given by

$$\{e_{i_1} \wedge ... \wedge e_{i_r} \mid \forall k : 1 \leq i_k \leq \dim(V)\}$$

The dimension of this space is given by:

$$\dim \Lambda^k(V) = \binom{n}{k} \tag{19.18}$$

Remark 19.4.8. For $k = 0$, the above construction is not useful, so we just define $\Lambda^0(V) = \mathbb{R}$.

Formula 19.4.9 (Levi-Civita symbol). The Levi-Civita tensor in n dimensions as introduced in 19.11 can now be rewritten more concisely as:

$$\boldsymbol{\varepsilon} = e_1 \wedge ... \wedge e_n \tag{19.19}$$

Formula 19.4.10. In 3 dimensions there exists an important isomorphism $J : \Lambda^2(\mathbb{R}^3) \to \mathbb{R}^3$:

$$J(\lambda)^i = \frac{1}{2} \varepsilon^i{}_{jk} \lambda^{jk} \tag{19.20}$$

where $\lambda \in \Lambda^2(\mathbb{R}^3)$.

Looking at the definition of the cross product 15.78, 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 19.4.11. 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 15.29). Evaluating this formula using the properties of the wedge and tensor products leads to the well known BAC-CAB rule of triple cross products:

$$(C \cdot A)B - (B \cdot A)C$$

Remark 19.4.12. The wedge product can also be defined without prior knowledge of tensor products. The wedge product is then defined by letting $V \wedge W$ be the smallest space such that for all elements $v \in V, w \in W$ the following property holds:

$$v \wedge w = -w \wedge v \tag{19.21}$$

Now, let $\{e_i\}_{i \leq m}$ be an ordered basis for V and let $\{d_j\}_{j \leq n}$ be an ordered basis for W. The wedge product $V \wedge W$ is spanned by the basis $\{e_i \wedge d_j\}_{i \leq m, j \leq n}$.

19.4.4 Exterior algebra

Definition 19.4.13 (Exterior power). In the theory of exterior algebras, the space $\Lambda^k(V)$ is often called the k^{th} exterior power of V.

Definition 19.4.14 (Exterior algebra). We can define a graded vector space[4] $\Lambda^*(V)$ 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:

$$\wedge : \Lambda^k(V) \times \Lambda^l(V) \rightarrow \Lambda^{k+l}(V)$$

This algebra is called the exterior algebra or **Grassmann algebra** of V.

Alternative Definition 19.4.15 (†). Let $T(V)$ be the (free) tensor algebra over the vector space V, i.e.

$$T(V) = \bigoplus_{k \geq 0} T^{\otimes k}(V) \tag{19.22}$$

where $T^{\otimes k}(V)$ is the k^{th} tensor power of V. 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 19.4.16. The exterior algebra is both an associative algebra and a unital algebra with unit element $1 \in \mathbb{R}$. Furthermore it is also commutative in the graded sense (see 3.20).

[4]See definition 15.8.

19.4.5 Hodge star

It follows from equation 19.18 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 15.4.2.

Definition 19.4.17 (Volume element). Let V be an n-dimensional vector space with ordered basis $\{e_i\}_{i\leq n}$. The volume element on V is defined as:

$$\mathrm{Vol}(V) := e_1 \wedge ... \wedge e_n \tag{19.23}$$

It is clear that this is an element of $\Lambda^n(V)$.

Definition 19.4.18 (Orientation). Let $\omega \in \Lambda(V)$ be an element of degree n. From the previous definition it follows that this k-blade is a scalar multiple of Vol because $\Lambda^n(V)$ is one-dimensional:

$$\omega = r\mathrm{Vol}(V)$$

The k-blade ω induces an orientation on V in the following way. If the scalar $r > 0$ then the orientation is said to be **positive**. If $r < 0$ then the orientation is **negative**.

Formula 19.4.19 (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:

$$\boxed{\langle v_1 \wedge ... \wedge v_k | w_1 \wedge ... \wedge w_k\rangle_k = \det(\langle v_i, w_j\rangle)} \tag{19.24}$$

For an orthogonal basis, this formula factorises into:

$$\langle v_1 \wedge ... \wedge v_k | w_1 \wedge ... \wedge w_k\rangle_k = \langle v_1|w_1\rangle \cdots \langle v_k|w_k\rangle \tag{19.25}$$

Definition 19.4.20 (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:

$$\omega \wedge \rho = \langle *\omega, \rho\rangle_{n-k}\mathrm{Vol}(V) \tag{19.26}$$

where $\langle\cdot,\cdot\rangle$ is the inner product 19.24 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 $\mathrm{Vol}(V)$. This implies that it can be written as

$$c(\rho)\mathrm{Vol}(V)$$

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 19.4.21. Let $\{e_i\}_{i\leq n}$ be a positively oriented ordered orthonormal 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 $\omega = e_{i_1}\wedge...\wedge e_{i_k}$.

$$\boxed{*\omega = \mathrm{sgn}(\tau)\prod_{m=1}^{n-k}\langle e_{j_m}|e_{j_m}\rangle e_{j_1}\wedge ... \wedge e_{j_{n-k}}} \tag{19.27}$$

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 $\mathrm{Vol}(V)$

Corollary 19.4.22. Consider three vectors $u, v, w \in \mathbb{R}^3$.

$$*(v\wedge w) = v\times w \tag{19.28}$$

$$*(v\times w) = v\wedge w \tag{19.29}$$

$$*(u\wedge v\wedge w) = u\cdot(v\times w) \tag{19.30}$$

Remark 19.4.23. Formula 19.20 is an explicit evaluation of the first equation 19.28.

Proof. The sign $\mathrm{sgn}(\tau)$ can be written using the Levi-Civita symbol ε_{ijk} as defined in 19.11. 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 19.4.24. Consider an inner product space V, then

$$\boxed{**\ \omega = (-1)^{k(n-k)}\omega} \tag{19.31}$$

In $n = 4$ this leads to $**\omega = \omega$ which means that the Hodge star is an involution in 4-dimensional inner product spaces.

Definition 19.4.25 (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.

19.4.6 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 19.4.26 (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 19.4.27. Furthermore, from the anti-commutativity it follows that we can regard the Grassmann variables as being non-zero square-roots of zero.

Property 19.4.28. 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

$$p(\theta) = a + b\theta \tag{19.32}$$

Definition 19.4.29. We can equip the exterior algebra Λ with Grassmann variables θ_i with an involution similar to that on $\mathbb{C}$:

$$(\theta_i\theta_j...\theta_k)^* = \theta_k...\theta_j\theta_i \tag{19.33}$$

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 20

Clifford Algebra

20.1 Clifford algebra

Definition 20.1.1 (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:

$$v \cdot v = Q(v)1 \tag{20.1}$$

where 1 is the unit element in V. This condition implies that the square of a vector is a scalar.

Notation 20.1.2. The Clifford algebra corresponding to V and Q is denoted by $C\ell(V, Q)$.

Construction 20.1.3. The previous definition can be given an explicit construction. First we construct the tensor algebra of V:

$$T(V) = \bigoplus_{k\in\mathbb{N}} V^{\otimes k} \tag{20.2}$$

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 20.1.4. Looking at definition 19.4.15 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 linearly isomorphic when $\text{char}(V) \neq 2$.

Property 20.1.5 (Dimension). If V has dimension n then $C\ell(V, Q)$ has dimension 2^n.

20.2 Geometric algebra

Definition 20.2.1 (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

[1] See definition 3.2.9.

the Clifford algebra $C\ell(V, g)$. If $\mathrm{char}(V) \neq 2$ then the bilinear form uniquely determines a quadratic form $Q : v \mapsto g(v, v)$ as required in definition 20.1.1.

Definition 20.2.2 (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.

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:

$$ab = \frac{1}{2}(ab + ba) + \frac{1}{2}(ab - ba) \tag{20.3}$$

We can then define the inner product as the symmetric part:

$$a \cdot b := \frac{1}{2}(ab + ba) = \frac{1}{2}\left((a + b)^2 - a^2 - b^2\right) = g(a, b) \tag{20.4}$$

Analogously we define the exterior (outer) product as the antisymmetric part:

$$a \wedge b := \frac{1}{2}(ab - ba) \tag{20.5}$$

These definitions allow us the rewrite formula 20.3 as:

$$\boxed{ab = a \cdot b + a \wedge b} \tag{20.6}$$

Looking at the last equality in the definition of the inner product 20.4 we see that condition 20.1 is satisfied when $a = b$. Furthermore we see that when g is fully degenerate, i.e. $g(v, v) = 0$ for all $v \in V$, the inner product is identically zero for all vectors and the geometric algebra coincides with exterior algebra[2] over V. For general forms g the exterior algebra is a subalgebra of the GA.

Definition 20.2.3 (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 19.4.3. Sums of multivectors of different grades are called mixed multivectors[3].

Let $n = \dim(V)$. Multivectors of grade n are also called **pseudoscalars** and multivectors of grade $n - 1$ are also called **pseudovectors**.

Definition 20.2.4 (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.

[2]See definition 19.4.14.

[3]These elements do not readily represent a geometric structure.

Formula 20.2.5. Let A, B be two multivectors of respectively grade m and n. Their inner product is defined as:

$$A \cdot B = \langle AB \rangle_{|m-n|} \tag{20.7}$$

Their exterior product is defined as:

$$A \wedge B = \langle AB \rangle_{m+n} \tag{20.8}$$

20.3 Pin group

20.3.1 Clifford group

Definition 20.3.1 (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 operator $\hat{}$ defined by:

$$\hat{v} = \begin{cases} v & v \in V_0 \\ -v & v \in V_1 \end{cases} \tag{20.9}$$

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[4].

Formula 20.3.2 (Twisted conjugation). Let $v \in V$ be a vector and let $s \in C\ell(V,Q)$ be an invertible element of the Clifford algebra over V. The twisted conjugation of v by s is given by the map:

$$\chi : C\ell(V,Q) \times V : \chi(s)v = sv\hat{s}^{-1} \tag{20.10}$$

Definition 20.3.3 (Clifford group). The Clifford group $\Gamma(V,Q)$ is defined as follows:

$$\Gamma(V,Q) = \{s \in C\ell(V,Q) : sv\hat{s}^{-1} \in V, v \in V\} \tag{20.11}$$

It is the set of Clifford algebra elements that stabilize V under twisted conjugation. Furthermore this set is closed under multiplication and forms a group. It contains all invertible elements $s \in C\ell(V,Q)$.

Property 20.3.4. If V is finite-dimensional the map

$$\chi : \Gamma(V,Q) \to O(V,Q) : s \mapsto \chi(s) \tag{20.12}$$

defines a representation[5] called the *vectorial representation*. Furthermore, from the first isomorphism theorem 3.1.30 it follows that $O(V,Q)$ is isomorphic to $\Gamma(V,Q)/\ker\chi$ where $\ker\chi = \mathbb{R}\backslash\{0\}$. This isomorphism[6] also implies that the Clifford group is given by the the set of finite products of invertible elements $v \in V$:

$$\Gamma(V,Q) = \left\{\prod_i^k s_i : s_i \text{ invertible in } V, n \in \mathbb{N}\right\} \tag{20.13}$$

[4]See definition 15.2.25.

[5]The surjectiveness of the map χ follows from the *Cartan-Dieudonné theorem*.

[6]Together with the *Cartan-Dieudonné theorem*.

Corollary 20.3.5. By nothing that pure rotations can be decomposed as an even number of reflections we find that:

$$\Gamma^+(V,Q)/\mathbb{R}_0 = SO(V,Q) \tag{20.14}$$

where Γ^+ is the intersection of the even Clifford algebra and the Clifford group.

20.3.2 Pin and Spin groups

Formula 20.3.6 (Spinor norm). On $\Gamma(V,Q)$ (and in fact on al of $C\ell(V,Q)$) one can define the spinor norm:

$$\mathcal{N}(x) = \hat{x}x \tag{20.15}$$

where $\hat{}$ is the main involution. The map $|\mathcal{N}|$ gives a group homomorphism from $\Gamma(V,Q)$ to $\mathbb{R}_0^+$.

Definition 20.3.7 (Pin and spin groups). Using the spinor norm $\mathcal{N}$ we can now define the pin and spins groups as follows:

Part V

Differential Geometry

Chapter 21

Curves and Surfaces

21.1 Curves

Property 21.1.1 (Regular curve). Let $\vec{c}(t)$ be a curve defined on an interval I. $\vec{c}(t)$ is said to be regular if $\frac{d\vec{c}}{dt} \neq \vec{0}$ for all $t \in I$.

Definition 21.1.2 (C^r-parameter transformation). A transformation $\varphi :]c, d[\rightarrow]a, b[: u \mapsto t$ such that $\varphi(u)$ is a C^r-diffeomorphism[1].

Definition 21.1.3 (Geometric property). A geometric property is a property that is invariant under:

1. parameter transformations
2. positive orthonormal changes of basis

Theorem 21.1.4. *Let $\vec{c}(t), \vec{d}(t)$ be two curves with the same image. We then have the following relation:*

$$\vec{c}(t) \text{ regular} \iff \vec{d}(t) \text{ regular} \tag{21.1}$$

21.1.1 Arc length

Definition 21.1.5 (Natural parameter). Let $\vec{c}(t)$ be a curve. The parameter t is said to be a natural parameter if:

$$\left|\left|\frac{d\vec{c}}{dt}\right|\right| \equiv 1 \tag{21.2}$$

Formula 21.1.6 (Arc length). The following function $\phi(t)$ is a bijective map and a natural parameter of $\vec{c}(t)$:

$$\phi(t) = \int_{t_0}^{t} ||\dot{\vec{c}}(t)||dt \tag{21.3}$$

[1] See definition 15.3.12

Remark. The arc length as defined above is often denoted by 's'.

Theorem 21.1.7. *Let $\vec{c}(t)$ be a curve. Let u be an alternative parameter of $\vec{c}(t)$. 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 21.3.

Remark. As the last theorem implies, no unique natural parameter or arc length exists.

21.1.2 Frenet-Serret frame

Definition 21.1.8 (Tangent vector). Let $\vec{c}(s)$ be parametrized by arc length. The tangent vector $\vec{t}(s)$ is defined as:

$$\vec{t}(s) = \vec{c}\,'(s) \tag{21.4}$$

Property 21.1.9. From the definition of the natural parametrization 21.2 and the previous definition it follows that the tangent vector is a unit vector:

$$\vec{t}(s) \cdot \vec{t}(s) = \left|\left|\frac{d\vec{c}(s)}{ds}\right|\right|^2 = 1$$

Definition 21.1.10 (Principal normal vector). Let $\vec{c}(s)$ be parametrized by arc length. The principal normal vector is defined as:

$$\vec{n}(s) = \frac{\vec{t}\,'(s)}{||\vec{t}\,'(s)||} \tag{21.5}$$

Property 21.1.11. From property 21.1.9 and the definition of the principal normal vector it follows that the tangent vector and principal normal vector are orthogonal:

$$\vec{t}(s) \cdot \vec{t}(s) = 1 \implies \vec{t}(s) \cdot \vec{t}\,'(s) = 0 \implies \vec{t}(s) \cdot \vec{n}(s) = 0$$

Definition 21.1.12 (Binormal vector). Let $\vec{c}(s)$ be parametrized by arc length. The binormal vector is defined as:

$$\vec{b}(s) = \vec{t}(s) \times \vec{n}(s) \tag{21.6}$$

Definition 21.1.13 (Frenet-Serret frame). As the vectors $\vec{t}(s), \vec{n}(s)$ and $\vec{b}(s)$ are mutually orthonormal and linearly independent, we can use them to construct a positive orthonormal basis. The ordered basis $\left(\vec{t}(s), \vec{n}(s), \vec{b}(s)\right)$ is called the **Frenet-Serret** frame.

Remark. This basis does not have to be the same in every point of $\vec{c}(s)$.

Definition 21.1.14 (Curvature). Let $\vec{c}(s)$ be parametrized by arc length. The curvature of $\vec{c}(s)$ is defined as:

$$\frac{1}{\rho(s)} = ||\vec{t}\,'(s)|| \tag{21.7}$$

Definition 21.1.15 (Torsion). Let $\vec{\boldsymbol{c}}(s)$ be a curve parametrized by arc length. The torsion of $\vec{\boldsymbol{c}}(s)$ is defined as:

$$\tau(s) = \rho(s)^2(\vec{\boldsymbol{t}}\ \ \vec{\boldsymbol{t}}'\ \ \vec{\boldsymbol{t}}'') \tag{21.8}$$

Formula 21.1.16 (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:

$$\begin{cases} \vec{\boldsymbol{t}}'(s) & = & & & \frac{1}{\rho(s)}\vec{\boldsymbol{n}}(s) & & \\ \vec{\boldsymbol{n}}'(s) & = & -\frac{1}{\rho(s)}\vec{\boldsymbol{t}}(s) & & + & & \tau(s)\vec{\boldsymbol{b}}(s) \\ \vec{\boldsymbol{b}}'(s) & = & & & -\tau(s)\vec{\boldsymbol{n}}(s) & & \end{cases} \tag{21.9}$$

Theorem 21.1.17 (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.*

21.2 Surfaces

Notation 21.2.1. Let $\vec{\boldsymbol{\sigma}}$ be a surface[2]. The derivative of $\vec{\boldsymbol{\sigma}}$ with respect to the coordinate q^i is written as follows:

$$\frac{\partial\vec{\boldsymbol{\sigma}}}{\partial q^i} = \vec{\boldsymbol{\sigma}}_i \tag{21.10}$$

21.2.1 Tangent vectors

Definition 21.2.2 (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:

$$\forall\vec{\boldsymbol{r}} \in T_P\Sigma : \left[\vec{\boldsymbol{r}} - \vec{\boldsymbol{\sigma}}(q_0^1, q_0^2)\right] \cdot \left[\vec{\boldsymbol{\sigma}}_1(q_0^1, q_0^2) \times \vec{\boldsymbol{\sigma}}_2(q_0^1, q_0^2)\right] = 0 \tag{21.11}$$

Definition 21.2.3 (Normal vector). The cross product in equation 21.11 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:

$$\vec{\boldsymbol{N}}(q_0^1, q_0^2) = \frac{1}{||\vec{\boldsymbol{\sigma}}_1 \times \vec{\boldsymbol{\sigma}}_2||}\left(\vec{\boldsymbol{\sigma}}_1 \times \vec{\boldsymbol{\sigma}}_2\right) \tag{21.12}$$

21.2.2 First fundamental form

Definition 21.2.4 (Metric coefficients). Let $\vec{\boldsymbol{\sigma}}$ be a surface. The metric coefficients g_{ij} are defined as follows:

$$\boxed{g_{ij} = \vec{\boldsymbol{\sigma}}_i \cdot \vec{\boldsymbol{\sigma}}_j} \tag{21.13}$$

[2] $\vec{\boldsymbol{\sigma}}$ denotes the surface as a vector field. Σ denotes the geometric image of $\vec{\boldsymbol{\sigma}}$.

Definition 21.2.5 (Scale factor). The following factors are often used in vector calculus:

$$g_{ii} = h_i^2 \tag{21.14}$$

Definition 21.2.6 (First fundamental form). Let $\vec{\boldsymbol{\sigma}}$ be a surface. Define a bilinear form $I_P(\vec{\boldsymbol{v}}, \vec{\boldsymbol{w}}) : T_P\Sigma \times T_P\Sigma \rightarrow \mathbb{R}$ that restricts the inner product to $T_P\Sigma$:

$$I_P(\vec{\boldsymbol{v}}, \vec{\boldsymbol{w}}) = \vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{w}} \tag{21.15}$$

This bilinear form is called the first fundamental form or **metric**.

Corollary 21.2.7. 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 21.13:

$$I_P(\vec{\boldsymbol{v}}, \vec{\boldsymbol{w}}) = v^i\vec{\boldsymbol{\sigma}}_i \cdot w^j\vec{\boldsymbol{\sigma}}_j = g_{ij}v^iw^j$$

Notation 21.2.8. The length can be written as

$$s = \int \sqrt{||\dot{\vec{\boldsymbol{c}}}(t)||}dt = \int \sqrt{ds^2}$$

where the second equality is formally defined. The two equalities together can be combined into the following notation for the metric:

$$\boxed{ds^2 = g_{ij}dq^idq^j} \tag{21.16}$$

Formula 21.2.9. Let (g_{ij}) be the metric tensor. We define the matrix (g^{ij}) as its inverse:

$$(g^{ij}) = \frac{1}{\det(g_{ij})}\begin{pmatrix} g_{22} & -g_{12} \\ -g_{12} & g_{11} \end{pmatrix} \tag{21.17}$$

21.2.3 Isometries

Definition 21.2.10 (Isometry). An isometry is a distance-preserving map, i.e. a diffeomorphism $\Phi : \Sigma \rightarrow \Sigma'$ that maps arc segments in Σ to arc segments with the same length in Σ'.

Property 21.2.11. A diffeomorphism Φ is an isometry if and only if the metric coefficients of σ and σ' are the same.

Definition 21.2.12 (Conformal map). A diffeomorphism $\Phi : \Sigma \rightarrow \Sigma'$ 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 21.2.13. A diffeomorphism Φ is conformal if and only if the metric coefficients of σ and σ' are proportional.

Definition 21.2.14 (Surface preserving map). A diffeomorphism $\Phi : \Sigma \rightarrow \Sigma'$ is sadi to be surface-preserving if it maps a segment of Σ to a segment of Σ' with the same surface.

Property 21.2.15. A diffeomorphism Φ is surface-preserving if and only if the metric coefficients of σ and σ' satisfy:

$$g'_{11}g'_{22} - (g'_{12})^2 = g_{11}g_{22} - g_{12}^2 \tag{21.18}$$

for all points (q^1, q^2).

Corollary 21.2.16. A map that is surface-preserving and conformal is also isometric.

21.2.4 Second fundamental form

Definition 21.2.17 (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:

$$\boxed{II_P(\vec{\boldsymbol{v}}, \vec{\boldsymbol{w}}) = L_{ij}(q^1, q^2)v^i w^j} \tag{21.19}$$

where $L_{ij} = \vec{\boldsymbol{N}} \cdot \vec{\boldsymbol{\sigma}}_{ij}$.

Definition 21.2.18 (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 $(q^1(s), q^2(s))$ is defined as:

$$\boxed{\frac{1}{\rho_n(s)} = \vec{\boldsymbol{c}}\,''(s) \cdot \vec{\boldsymbol{N}}(s)} \tag{21.20}$$

From the definition of the second fundamental form it follows that the normal curvature can be written as:

$$\frac{1}{\rho_n(s)} = II(\vec{\boldsymbol{t}}, \vec{\boldsymbol{t}}) = \frac{II\left(\dot{\vec{\boldsymbol{c}}}(t), \dot{\vec{\boldsymbol{c}}}(t)\right)}{I\left(\dot{\vec{\boldsymbol{c}}}(t), \dot{\vec{\boldsymbol{c}}}(t)\right)} \tag{21.21}$$

where the last equality holds for any given parameter t.

Theorem 21.2.19 (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 $(q^1(t_0), q^2(t_0))$ if $\vec{\boldsymbol{c}}(t_0) = \vec{\boldsymbol{d}}(t_0)$ and if $\dot{\vec{\boldsymbol{c}}}(t_0) \parallel \dot{\vec{\boldsymbol{d}}}(t_0)$. Furthermore, the osculating circles of all curves with the same normal curvature at a given point form a sphere.*

Property 21.2.20. 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 21.2.21 (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 a point $(q^1(s), q^2(s))$ is defined as:

$$\boxed{\frac{1}{\rho_g(s)} = \left(\vec{\boldsymbol{N}}(s)\ \vec{\boldsymbol{t}}(s)\ \vec{\boldsymbol{t}}\,'(s)\right)} \tag{21.22}$$

Formula 21.2.22. 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:

$$\boxed{\frac{1}{\rho^2} = \frac{1}{\rho_n^2} + \frac{1}{\rho_g^2}} \tag{21.23}$$

21.2.5 Curvature of a surface

Definition 21.2.23 (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:

$$L_P(\vec{\boldsymbol{\sigma}}_1) = -\vec{\boldsymbol{N}}_1 \qquad \text{and} \qquad L_P(\vec{\boldsymbol{\sigma}}_2) = -\vec{\boldsymbol{N}}_2 \tag{21.24}$$

Formula 21.2.24. Let $\vec{\boldsymbol{v}}, \vec{\boldsymbol{w}} \in T_P\Sigma$. The following equalities hold:

$$L_P(\vec{\boldsymbol{v}}) \cdot \vec{\boldsymbol{w}} = L_P(\vec{\boldsymbol{w}}) \cdot \vec{\boldsymbol{v}} = II_P(\vec{\boldsymbol{v}}, \vec{\boldsymbol{w}}) \tag{21.25}$$

Formula 21.2.25 (Matrix elements of L_P). Let (g^{ij}) be the inverse of the metric tensor. The matrix elements of L_P are defined as:

$$L^k_j = g^{ki} L_{ij}$$

Formula 21.2.26 (Weingarten formulas).

$$\vec{\boldsymbol{N}}_j = -L^k_j \vec{\boldsymbol{\sigma}}_k \tag{21.26}$$

Theorem 21.2.27. *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 eigenfunctions 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 21.2.28 (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**.

Remark. 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 21.2.29. For the principal directions we have $L_P(\vec{\boldsymbol{h}}_1) \cdot \vec{\boldsymbol{h}}_2 = 0$[3]. 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 21.2.30 (Gaussian curvature). The Gaussian curvature K of a surface is defined as the determinant of the Weingarten map, i.e.:

$$K = \frac{1}{R_1 R_2} \tag{21.27}$$

Definition 21.2.31 (Mean curvature). The mean curvature H of a surface is defined as the trace of the Weingarten map, i.e.:

$$H = \frac{1}{2}\left(\frac{1}{R_1} + \frac{1}{R_2}\right) \tag{21.28}$$

[3] Tangent vectors that satisfy this equation are called **adjoint** tangent vectors.

Property 21.2.32. The principal curvatures are the solutions of the following equation:

$$x^2 - 2Hx + K = 0$$

This is the characteristic equation (15.67) of the Weingarten map.

Definition 21.2.33. 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.
- 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.

Theorem 21.2.34. *A surface Σ containing only umbilics is a part of a sphere or a part of a plane.*

Theorem 21.2.35. *In the neighbourhood of a point P of a surface with principal curvatures $1/R_1$ and $1/R_2$ is locally the same as the following quadric:*

$$x_3 = \frac{1}{2}\left(\frac{x_1^2}{R_1} + \frac{x_2^2}{R_2}\right) \tag{21.29}$$

if we ignore terms of order > 2.

Theorem 21.2.36 (Euler's formula). *The normal curvature of a couple $(P, \vec{\boldsymbol{e}})$ where is $\vec{\boldsymbol{e}} = \vec{\boldsymbol{h}}_1 \cos\theta + \vec{\boldsymbol{h}}_2 \sin\theta \in T_P\Sigma$ is given by:*

$$\frac{1}{\rho_n} = \frac{\cos^2\theta}{R_1} + \frac{\sin^2\theta}{R_2} \tag{21.30}$$

Definition 21.2.37 (Asymptotic curve). An asymptotic curve is a curve which is in every point P tangent to a direction with zero normal curvature.

Formula 21.2.38 (Differential equation for asymptotic curves).

$$L_{11}\left(\dot{q}^1(t)\right)^2 + 2L_{12}\dot{q}^1(t)\dot{q}^2(t) + L_{22}\left(\dot{q}^2(t)\right)^2 = 0 \tag{21.31}$$

Property 21.2.39. 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.

Definition 21.2.40 (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 21.2.41 (Rodrigues' formula). A curve is a line of curvature if and only if it satisfies the following formula:

$$\frac{d\vec{\boldsymbol{N}}}{dt}(t) = -\frac{1}{R(t)}\frac{d\vec{\boldsymbol{c}}}{dt}(t) \tag{21.32}$$

If the curve satisfies this formula, then the scalar function $1/R(t)$ coincides with the principal curvature along the curve.

Formula 21.2.42 (Differential equation for curvature lines).

$$\begin{vmatrix} (\dot{q}^2)^2 & -\dot{q}^1\dot{q}^2 & (\dot{q}^1)^2 \\ g_{11} & g_{12} & g_{22} \\ L_{11} & L_{12} & L_{22} \end{vmatrix} = 0 \tag{21.33}$$

Property 21.2.43. From theorem 21.2.27 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.

21.2.6 Christoffel symbols and geodesics

Formula 21.2.44 (Gauss' formulas).

$$\vec{\boldsymbol{\sigma}}_{ij} = L_{ij}\vec{\boldsymbol{N}} + \Gamma^k{}_{ij}\vec{\boldsymbol{\sigma}}_k \tag{21.34}$$

where the **Christoffel symbols** $\Gamma^k{}_{ij}$ are defines as:

$$\boxed{\Gamma^k{}_{ij} = g^{kl}\vec{\boldsymbol{\sigma}}_l \cdot \vec{\boldsymbol{\sigma}}_{ij}} \tag{21.35}$$

Corollary 21.2.45. From the expression of the Christoffel symbols we can derive an alternative expression using only the metric tensor g_{ij}:

$$\Gamma^k{}_{ij} = \frac{1}{2}g^{kl}\left(\frac{\partial g_{il}}{\partial q^j} - \frac{\partial g_{ij}}{\partial q^l} + \frac{\partial g_{jl}}{\partial q^i}\right) \tag{21.36}$$

Definition 21.2.46 (Geodesic). A geodesic is a curve with zero geodesic curvature.

Theorem 21.2.47. *A curve on a surface is an geodesic if and only if the tangent plane and the osculation plane are orthogonal in every point P of the surface.*

Formula 21.2.48 (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:

$$\boxed{q''^k + \Gamma^k{}_{ij}q'^i q'^j = 0} \tag{21.37}$$

21.2.7 Theorema Egregium

Formula 21.2.49 (Codazzi-Mainardi equations).

$$\frac{\partial L_{ij}}{\partial q^k} - \frac{\partial L_{ik}}{\partial q^j} = \Gamma^l{}_{ik} L_{lj} - \Gamma^l{}_{ij} L_{lk} \tag{21.38}$$

Definition 21.2.50 (Riemann curvature tensor).

$$\boxed{R^l{}_{ijk} = \frac{\partial \Gamma^l{}_{ik}}{\partial q^j} - \frac{\partial \Gamma^l{}_{ij}}{\partial q^k} + \Gamma^s{}_{ik}\Gamma^l{}_{sj} - \Gamma^s{}_{ij}\Gamma^l{}_{ks}} \tag{21.39}$$

Formula 21.2.51 (Gauss' equations).

$$R^l{}_{ijk} = L_{ik} L^l_j - L_{ij} L^l_k \tag{21.40}$$

Theorem 21.2.52 (Theorema Egregium). *The Gaussian curvature K (formula 21.27) is completely determined by the metric tensor g_{ij} and its derivatives:*

$$\boxed{K = \frac{R^l{}_{121} g_{l2}}{g_{11} g_{22} - g_{12}^2}} \tag{21.41}$$

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 21.2.53. From the condition of isometries 21.2.11 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 21.2.54. 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.

Chapter 22

Manifolds

22.1 Charts

Definition 22.1.1 (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 22.1.2 (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 22.1.3 (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 22.1.4 (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 22.1.5 (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.

Definition 22.1.6 (C^k-manifold). If all transition maps are C^k-diffeomorphisms than the manifold is called a C^k-manifold. A C^∞-manifold is also called a smooth manifold.

Theorem 22.1.7 (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 22.1.8 (Radó-Moise). *In the dimensions $1, 2$ and 3 there exists for every topological manifold a unique smooth structure.*

Theorem 22.1.9. *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 22.1.10 (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

$$f_{\varphi\psi} = \psi \circ f \circ \varphi^{-1} \tag{22.1}$$

is smooth on $\mathbb{R}^n$.

Remark. The function $f_{\varphi\psi}$ in equation 22.1 is called the **local representation** of f.

Notation 22.1.11. The set of all C^∞ functions on a manifold M defined on a neighbourhood of $m \in M$ is denoted by $C^\infty_m(M)$. This set forms a commutative unital ring when equipped with the usual sum and product (composition) of functions.

22.2 Tangent vectors

Definition 22.2.1 (Tangent vector). Let M be a smooth manifold and $p \in M$. Let $f, g : M \to \mathbb{R} \in C^\infty_p(M)$. 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**[2].

Property 22.2.2. For every constant function $c : p \mapsto c$ we have:

$$v_p(c) = 0 \tag{22.2}$$

[1] In this definition one can replace 'smooth' by 'C^k-differentiable'.

[2] Generally, every operation that satisfies the Leibniz property is called a derivation.

Definition 22.2.3 (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:

$$\boxed{\left.\frac{\partial}{\partial q^i}\right|_p : C_p^\infty(M,\mathbb{R}) \to \mathbb{R} : f \mapsto \frac{\partial}{\partial q^i}\left(f \circ \varphi^{-1}\right)(\varphi(p))} \tag{22.3}$$

where (U,φ) is a coordinate chart such that $p \in U$ and $(q^1,...,q^n)$ are local coordinates.

Remark 22.2.4. 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 22.2.5. From the above tangent space construction it follows that:

$$\boxed{\dim(T_pM) = \dim(M)} \tag{22.4}$$

This also implies that the tangent spaces over two distinct points $p, q \in M$ are isomorphic.

Definition 22.2.6 (Curve). A smooth function $\gamma : \mathbb{R} \to M$ with $\gamma(0) = m$ is called a smooth curve through $m \in M$.

Alternative Definition 22.2.7 (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:

$$\frac{d(\varphi \circ \gamma_1)}{dt}(0) = \frac{d(\varphi \circ \gamma_2)}{dt}(0) \tag{22.5}$$

or equivalently, if their local representatives are tangent in 0. This relation imposes an
def equivalence relation[3] 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:

$$\boxed{v_p(f) = \left.\frac{d(f \circ c)}{dt}\right|_{t=0}} \tag{22.6}$$

Applying the chain rule gives us

$$v_p(f) = \frac{\partial(f \circ \varphi^{-1})}{\partial q^i}(\varphi(p))\frac{dq^i}{dt}(0) \tag{22.7}$$

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.

[3]The relation is well-defined (under a change of chart) because the transition maps (and their Jacobian matrices) are invertible and thus non-singular.

22.3 Curvature

Formula 22.3.1 (Riemann Curvature Tensor). Let $V \in TM$. Let D_μ be the covariant derivative.

$$\boxed{[D_\mu, D_\nu]V^\rho = R^\rho{}_{\kappa\mu\nu}V^\kappa} \tag{22.8}$$

Formula 22.3.2 (Ricci tensor).

$$R_{\mu\nu} = R^\lambda{}_{\mu\lambda\nu} \tag{22.9}$$

Formula 22.3.3 (Ricci scalar).

$$R = R^\mu{}_\mu \tag{22.10}$$

This scalar quantity is also called the **scalar curvature**.

Formula 22.3.4 (Einstein tensor).

$$\boxed{G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R} \tag{22.11}$$

Theorem 22.3.5. *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:*

$$\nabla_\mu G^{\mu\nu} = 0 \tag{22.12}$$

22.4 Submanifolds

Definition 22.4.1 (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 22.4.2 (Immersion). Let $f : M \to N$ be a differentiable function between smooth manifolds. f is called an immersion if its derivative[4] is everywhere injective, or equivalently if its derivative has maximal rank[5] everywhere:

$$\text{rk}_p(f) = \dim(M), \forall p \in M \tag{22.13}$$

Definition 22.4.3 (Submersion). Let $f : M \to N$ be a differentiable map between smooth manifolds. A **regular point**[6] of f is a point $x \in M$ such that $T_x f$ is surjective. f is called a submersion if its derivative is everywhere surjective, or equivalently if

$$\text{rk}_p(f) = \dim(N), \forall p \in M \tag{22.14}$$

[4]This is formally defined in 25.20. For now it is the map represented by the Jacobian matrix.
[5]See definition 25.3.8.
[6]$f(x)$ is then called a **regular value**.

Definition 22.4.4 (Embedding). A differentiable function between smooth manifolds is called a smooth embedding if its both an injective immersion and an embedding in the topological sense 4.3.11. This implies that the submanifold topology coincides with the subspace topology 4.1.

Definition 22.4.5 (Embedded submanifold). Let M be a manifold. A subset N is an embedded[7] submanifold if the inclusion map $f : M \hookrightarrow N$ is a smooth embedding.

Definition 22.4.6 (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:

$$\mathbb{R}^m \cong \mathbb{R}^m \times \{\underbrace{0, ..., 0}_{n-m}\} \overset{\iota}{\hookrightarrow} \mathbb{R}^m \times \mathbb{R}^{n-m} \cong \mathbb{R}^n \tag{22.15}$$

where $\iota : (x_1, ..., x_m) \mapsto (x_1, ..., x_m, \underbrace{0, ..., 0}_{n-m})$ is the canonical inclusion map.

Alternative Definition 22.4.7. A k-dimensional embedded manifold N of M can now 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) with

$$\varphi(U \cap N) = \varphi(U) \cap (\mathbb{R}^k \times \{\underbrace{0, ..., 0}_{n-k}\}) \tag{22.16}$$

where $n = \dim(M)$. The set $U \cap N$ is called a slice of (U, φ) in analogy with the previous definition of a (standard) slice.

Theorem 22.4.8 (Submersion theorem[8]). *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* $\dim(M_2)$.

22.5 Manifolds with boundary

Definition 22.5.1 (Manifold with boundary). Let $\mathbb{H}^n$ denote the upper half space, i.e.:

$$\mathbb{H}^n = \{(x_1, ..., x_n) | x_n \geq 0\} \subset \mathbb{R}^n \tag{22.17}$$

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] 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.

[8] Also called the **regular value theorem**.

Remark 22.5.2 (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 $\mathrm{Int}(M) = M$ ∂M, in the sense of manifolds, should not be confused with the topological interior.

Property 22.5.3. Let M be an n-dimensional manifold with boundary. Let (U, φ) be a chart for $p \in \partial M$. Then

$$\varphi(p) \in \partial\mathbb{H}^n = \{(x_1, ..., x_n) | x_n = 0\} \tag{22.18}$$

Chapter 23

Lie groups and Lie algebras

23.1 Lie groups

Definition 23.1.1 (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 23.1.2 (Lie subgroup). A subset of a Lie group is a Lie subgroup if it is both a subgroup and a closed submanifold.

Theorem 23.1.3 (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 23.1.4. 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.

23.1.1 Left invariant vector fields

Definition 23.1.5 (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 $g \in G$:

$$L_{g,*}X(h) = X(g \cdot h) \tag{23.1}$$

where L_g denotes the left action map associated with g.

Property 23.1.6. The set $\mathcal{L}(G)$ of LIVF's on a Lie group G is a vector space over $\mathbb{R}$.

Property 23.1.7. 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.

23.1.2 One-parameter subgroups

Definition 23.1.8 (One-parameter subgroup). A one-parameter (sub)group is a continuous group homomorphism $\Phi : \mathbb{R} \to G$ from the additive group of real numbers to a Lie group G.

Property 23.1.9. Let $\Phi : \mathbb{R} \to G$ be a one-parameter subgroup of G. Let $\Psi : G \to H$ be a continuous group homomorphism. Then $\Psi \circ \Phi : \mathbb{R} \to H$ is a one-parameter subgroup of H.

Property 23.1.10. Let X be a LIVF on a Lie group G. Let γ_X be the integral curve of X through $e \in G$. The maximal flow domain $D(X)$ is $]-\infty, +\infty[$ and the flow[3] σ_t determines a one-parameter subgroup on G. Furthermore, for every one-parameter subgroup $\phi(t)$ we can construct a LIVF $X = \phi'(0)$. This correspondence is a bijection.

23.1.3 Cocycles

Definition 23.1.11 (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:

$$g_{ij} = g_{ik} \circ g_{kj} \tag{23.2}$$

Property 23.1.12. 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 $x \in M$.

23.2 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 23.1.1) and real Lie algebras.

23.2.1 Definitions

Definition 23.2.1 (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]See definition 25.32.

3. Jacobi identity: $[a,[b,c]]+[b,[c,a]]+[c,[a,b]]=0$

Property 23.2.2. Let G be a Lie group. The tangent space T_eG has the structure of a Lie algebra where the Lie bracket is given by the commutator of vector fields 25.37.

The following definition gives the equality of the set of LIVF's on a Lie group G and the Lie algebra $\mathfrak{g} := T_eG$:

Alternative Definition 23.2.3. From the second part of property 23.1.7 it follows that the Lie algebra $\mathfrak{g}$ associated to G is isomorphic to the set of LIVF's on G. Using property 25.37 we can show that the Lie bracket also defines a LIVF on G. It follows that $\mathfrak{g}$ is closed under Lie brackets.

Notation 23.2.4. Lie algebras are denoted by fraktur symbols. For example, the Lie algebra associated with the Lie group G is mostly denoted by $\mathfrak{g}$.

Theorem 23.2.5 (Ado's theorem). *Every finite-dimensional Lie algebra can be embedded as a subalgebra of $\mathfrak{gl}_n$.*

Definition 23.2.6 (Lie algebra homomorphism). A map $\Phi : \mathfrak{g} \rightarrow \mathfrak{h}$ is a Lie algebra homomorphism if it satisfies following condition

$$\Phi([X,Y]) = [\Phi(X), \Phi(Y)] \tag{23.3}$$

for all $X, Y \in \mathfrak{g}$.

Property 23.2.7. Let G, H be Lie groups with G simply-connected. A linear map $\Phi : \mathfrak{g} \rightarrow \mathfrak{h}$ is the differential of a Lie group homomorphism $\phi : G \rightarrow H$ if and only if Φ is a Lie algebra homomorphism.

23.2.2 Examples

Example 23.2.8. The cross product $\times : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$ turns $\mathbb{R}^3$ into a Lie algebra.

Example 23.2.9. An interesting example is the Lie algebra associated to the Lie group of invertible complex[4] 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:

$$L_{A,*}(B) = L_A(B) \tag{23.4}$$

Following two examples of Lie algebras can be checked using condition 15.34:

Example 23.2.10 (Lie algebra of $O(3)$). The set of 3×3 anti-symmetric matrices. It is also important to note that $\mathfrak{o}(3) = \mathfrak{so}(3)$.

Example 23.2.11 (Lie algebra of $SU(2)$). The set of 2×2 traceless anti-Hermitian matrices. This result can be generalized to arbitrary $n \in \mathbb{N}$.

[4] As usual, this result is also valid for real matrices.

23.2.3 Exponential map

Formula 23.2.12 (Exponential map). Let $X \in \mathfrak{g}$ be a LIVF on G. We define the exponential map $\exp : \mathfrak{g} \to G$ as:

$$\boxed{\exp(X) := \gamma_X(1)} \tag{23.5}$$

where γ_X is the corresponding one-parameter subgroup.

Property 23.2.13. The exponential map is the unique map $\mathfrak{g} \to G$ such that $\exp(0) = e$, whose differential at the origin in $\mathfrak{g}$ is given by the identity $\mathbb{1}_{\mathfrak{g}}$ and for which the restrictions to the lines through the origin in $\mathfrak{g}$ are one-parameter subgroups of G.

Corollary 23.2.14. Because the identity element $\mathbb{1}_{\mathfrak{g}} = \exp_{e,*}$ is an isomorphism, the inverse function theorem 25.3.9 implies that the image of *exp* will contain a neighbourhood of the identity $e \in G$. If G is connected then property 23.1.4 implies that exp generates all of G.

Together with the property that $\psi \circ \exp = \exp \circ\, \psi_*$ for every Lie group homomorphism $\psi : G \to H$ it follows that *if G is connected, a Lie group homomorphism $\psi : G \to H$ is completely determined by its differential ψ_* at the identity $e \in G$.*

Example 23.2.15 (Matrix Lie groups). For matrix Lie groups we define the classic matrix exponential:

$$e^{tX} = \sum_{k=0}^{+\infty} \frac{(tX)^k}{k!} \tag{23.6}$$

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 $\exp(-X)$.

23.2.4 Structure

Definition 23.2.16 (Structure constants). As Lie algebras are closed under Lie brackets, every Lie bracket can be expanded in term of a chosen basis $\{X_k\}_{k\in I}$ as follows:

$$[X_i, X_j] = \sum_{k\in I} {c_{ij}}^k X_k \tag{23.7}$$

where the factors ${c_{ij}}^k$ are called the structure constants[5] of the Lie algebra.

Property 23.2.17. 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.

[5]Note that these constants are basis-dependent.

Formula 23.2.18 (Baker-Campbell-Hausdorff formula). This formula is the solution of the equation

$$Z = \log(\exp(X)\exp(X)) \tag{23.8}$$

for $X, Y \in \mathfrak{g}$. The solution is given by following formula

$$e^X e^Y = \exp\left(X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}[X, [X, Y]] - \frac{1}{12}[Y, [X, Y]] + \cdot\right) \tag{23.9}$$

However this formula will only converge if X, Y are sufficiently small (for matrix Lie algebras this means, using the Hilbert-Schmidt norm 15.45: $||X|| + ||Y|| < \frac{\ln(2)}{2}$). 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 23.2.19 (Lie product formula[6]). Let $\mathfrak{g}$ be a Lie algebra. The following formula applies to any $X, Y \in \mathfrak{g}$:

$$e^{X+Y} = \lim_{n\to+\infty}\left(e^{\frac{X}{n}} e^{\frac{Y}{n}}\right)^n \tag{23.10}$$

23.2.5 Killing form

Definition 23.2.20 (Killing form). Let $\mathfrak{g}$ be a finite-dimensional Lie algebra. Define the symmetric bilinear form[7]

$$\boxed{K(X, Y) = \mathrm{tr}(\mathrm{ad}_X \mathrm{ad}_Y)} \tag{23.11}$$

Theorem 23.2.21 (Cartan's criterion). *A Lie algebra is semisimple if and only if its Killing form is non-degenerate.*

Corollary 23.2.22. If a Lie algebra is semisimple its Killing form induces a metric

$$g : (X, Y) \mapsto -\ \mathrm{tr}(\mathrm{Ad}_X, \mathrm{Ad}_Y) \tag{23.12}$$

which turns the corresponding Lie group G into a Riemannian manifold.

Property 23.2.23. The Killing-form is Ad-invariant, i.e.

$$K(\mathrm{Ad}_g(X), \mathrm{Ad}_g(Y)) = K(X, Y) \tag{23.13}$$

for all $g \in G$. From this it follows that $\mathrm{Ad} : G \to \mathrm{Isom}(\mathfrak{g})$.

[6] Also called the Lie-Trotter formula.

[7] This is a symmetric $(0, 2)$-tensor in $\mathfrak{g}^* \otimes \mathfrak{g}^*$. (See 19.1.9)

23.2.6 Universal enveloping algebra

Definition 23.2.24 (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 $g \otimes h - h \otimes g - [g,h]$.

Definition 23.2.25 (Casimir invariant[8]). Let $\mathfrak{g}$ be a Lie algebra. A Casimir invariant J is an element of the center of $U(\mathfrak{g})$, the universal enveloping algebra of $\mathfrak{g}$.

23.2.7 Poisson algebras and Lie superalgebras

Definition 23.2.26 (Poisson algebra). Let V be a vector space equipped with two bilinear operation $\star$ 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** acts as a derivation[9] with respect to the operation $\star$, i.e.

$$\{x, y \star z\} = \{x,y\} \star z + y \star \{x,z\}$$

Definition 23.2.27 (Gerstenhaber algebra).

[8] Also known as a *Casimir operator* or *Casimir element*.
[9] See definition 22.2.1.

Chapter 24

Representation Theory

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].

Definition 24.1.2 (Subrepresentation). A subrepresentation of a representation V is a subspace of V invariant under the action of the group G.

Example 24.1.3 (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 3.1.32 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 24.1.4. 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 15.3.38. 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 24.1.5. A representation ρ which acts on spaces V, W can also be extended to the tensor product $V \otimes W$ in the following way:

$$g(v \otimes w) = g(v) \otimes g(w) \tag{24.4}$$

[1] See definition 15.3.11.

[2] See definition 3.1.31.

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

$$W = \{a_1v_1 + a_2v_2 + ... + a_nv_n | a_1 + a_2 + ... + a_n = 0\} \tag{24.5}$$

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 irreducible representations of a finite group G. Let $\varphi : V \to W$ be a G-module homomorphism. 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 under the action of G and complementary to W.

Property 24.2.5. Let G be a finite group. A representation 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.

24.3 Lie group representations

For more information on Lie groups and Lie algebras see chapter 23.

Definition 24.3.1 (Adjoint representation). 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 $T_e\Psi_g$:

$$\mathrm{Ad}_g : T_eG \to T_eG : X \mapsto gXg^{-1} \tag{24.8}$$

It is a representation of G on its own tangent space $T_eG \equiv \mathfrak{g}$.

24.4 Lie algebra representations

Formula 24.4.1 (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:

$$\mathrm{ad}_X := T_e(\mathrm{Ad}_g) \tag{24.9}$$

where $g = e^{tX}$. Explicitly, let $\mathfrak{g}$ be a Lie algebra. For every element $X \in \mathfrak{g}$ we define the Lie bracket as follows:

$$[X, Y] := \mathrm{ad}_X(Y) \tag{24.10}$$

Property 24.4.2. Given the antisymmetry of the Lie bracket the Jacobi identity is equivalent to ad: $\mathfrak{g} \rightarrow \mathrm{Aut}(\mathfrak{g})$ being a Lie algebra homomorphism, i.e. $\mathrm{ad}_{[X,Y]} = [\mathrm{ad}_X, \mathrm{ad}_Y]$.

Using the exponential map we can give property 23.2.7 an explicit form:

Formula 24.4.3 (Induced homomorphism). Let $\phi : G \rightarrow H$ be a Lie group homomorphism[3] with G connected and simply-connected. This homomorphism induces a a Lie algebra homomorphism $\Phi : \mathfrak{g} \rightarrow \mathfrak{h}$ given by:

$$\Phi(X) = \frac{d}{dt}\phi\left(e^{tX}\right)\bigg|_{t=0} \tag{24.11}$$

or equivalently:

$$\phi\left(e^{tX}\right) = e^{t\Phi(X)} \tag{24.12}$$

Corollary 24.4.4 (Commutator). For the general linear group GL_n the Lie bracket is given by the commutator:

$$\boxed{[X, Y] = XY - YX} \tag{24.13}$$

This follows from definition 24.10: $[X, Y] = \frac{d}{dt}\mathrm{Ad}_{\gamma(t)}(Y)\big|_{t=0}$ with $\gamma(0) = e$ and $\gamma'(0) = X$.

Remark 24.4.5. The homomorphism induced by $\mathrm{Ad} : G \rightarrow H$ is precisely $\mathrm{ad} : \mathfrak{g} \rightarrow \mathfrak{h}$. Informally we can thus say that the infinitesimal version of the similarity transformation is given by the commutator (in case of $G = \mathrm{GL}_n$).

[3]Continuity (inherent to the definition of a Lie group homomorphism) is needed to ensure that $\phi(e^{tX})$ is also a one-parameter subgroup (see 23.1.9).

Chapter 25

Bundle theory

25.1 Fibre bundles

Definition 25.1.1 (Bundle). A bundle is a triple (E, B, π) where E, B are topological spaces and π is a continuous surjective map.

Definition 25.1.2 (Fibered manifold). A fibered manifold is a surjective submersion[1] $\pi : E \to B$ 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 25.1.3 (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 smooth 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:

$$\begin{tikzcd}
\pi^{-1}(U_i) \arrow[rr, "\varphi_i"] \arrow[dr, "\pi"'] & & U_i \times F \arrow[dl, "\mathrm{pr}_1"] \\
& U_i &
\end{tikzcd}$$

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**[2], (U_i, φ_i) a **bundle chart**[3] and the set $\{(U_i, \varphi_i)\}_{i \in I}$ a **trivializing cover**.

[1] See definition 22.4.3.

[2] This name follows from the fact that the bundle is locally homeomorphic to a (trivial) product space: $E \cong U \times F$.

[3] 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[4] $g_{ji} : U_i \cap U_j \to G$, associated to the (left) action (which we require to be faithful[5]) of G on every fibre, by the following relation:

$$\varphi_j \circ \varphi_i^{-1}(b, x) = (b, g_{ji}(b) \cdot x) \tag{25.1}$$

Remark 25.1.4. One should pay attention that the bundle charts are not coordinate charts in the original sense 22.1.1 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 25.1.5. A fibre bundle (E, B, π, F, G) is often indicated by the following diagram:

$$\begin{tikzcd} F \arrow[r, hook] & E \arrow[d, "\pi"] \\ & B \end{tikzcd}$$

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 25.1.6 (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 $\pi^{-1}(b)$.

Definition 25.1.7 (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 25.1.8. A smooth fibre bundle is also a smooth manifold.

Construction 25.1.9 (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 23.1.11. 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[6].

[4]See definition 23.1.11.
[5]See definition 3.1.37.
[6]See definition 4.1.3.

3. From this disjoint union we construct a quotient space[7] (equipped with the quotient space topology) by applying following equivalence relation for every i, j:

$$(p, f) \sim (p, g_{ji}(x) \cdot f) \tag{25.2}$$

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 $x \in M$.

4. Local trivializations are given by the maps $\varphi_i : \pi^{-1}(U_i) \to U_i \times F$ that satisfy:

$$\varphi_i^{-1} : (x, f) \mapsto [(x, f)] \tag{25.3}$$

where $[A]$ means the equivalence class of A.

Definition 25.1.10 (Compatible[8] bundle charts). A bundle chart (V, ψ) is compatible with a trivializing cover $\{(U_i, \varphi_i)\}_{i\in I}$ if whenever $V \cap U_i \neq \emptyset$ their exists a map $h_i : V \cap U_i \to G$ such that:

$$\psi \circ \varphi_i^{-1}(b, x) = (b, h_i(b)x) \tag{25.4}$$

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.

Definition 25.1.11 (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 25.1 commute:

$$\begin{tikzcd}
E_1 \arrow[r, "f_E"] \arrow[d, "\pi_1"'] & E_2 \arrow[d, "\pi_2"] \\
B_1 \arrow[r, "f_B"'] & B_2
\end{tikzcd}$$

Figure 25.1: Bundle map between fibre bundles.

Definition 25.1.12 (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 25.1.13 (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[9] $\{(U_i, \varphi_i)\}_{i\in I}$ 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:

$$g'_{ji}(b) = \rho_j(b) \circ g_{ji}(b) \circ \rho_i^{-1}(b) \tag{25.5}$$

[7] See definition 4.4.

[8] Also called an **admissible chart**.

for every $b \in U_i \cap U_j$. An explicit form of these functions is given by:

$$\rho_i = \varphi_i' \circ \varphi_i^{-1} \tag{25.6}$$

This transformation is called a **gauge transformation** (especially in physics) .

Property 25.1.14. 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 25.1.15 (Trivial bundle). A fibre bundle (E, B, π, F) is trivial if $E = B \times F$.

Definition 25.1.16 (Trivialization). A trivialization of a fibre bundle ξ is an equivalence $\xi \to B \times F$. Bundles for which a trivialization can be found are also called *trivial bundles*.

Definition 25.1.17 (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 25.1.18 (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:

$$f^*E = \{(b', e) \in B' \times E : f(b') = \pi(e)\} \tag{25.7}$$

The topology on f^*E is given by the subspace topology.

Definition 25.1.19 (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:

$$F_1 \diamond F_2 = \{f \times g \in F_1 \times F_2 : \pi_1(f) = \pi_2(g)\} \tag{25.8}$$

25.1.1 Sections

Definition 25.1.20 (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 25.1.21. 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 25.1.22. The sections on a fibre bundle E pullback to the pullback bundle f^*E by setting $f^*s = s \circ f$.

[9]Remark that the collection $\{U_i\}_{i \in I}$ is the same for both trivializing covers.

25.1.2 Jet bundles

Definition 25.1.23 (Jet). Consider a fibre bundle (E, B, π) with its sections $\Gamma(E)$. Two sections $\sigma, \xi \in \Gamma(E)$ with local coordinates (σ^i) and (ξ^i) define the same r-jet at a point $p \in B$ if and only if:

$$\left.\frac{\partial^\alpha \sigma^i}{\partial x^\alpha}\right|_p = \left.\frac{\partial^\alpha \xi^i}{\partial x^\alpha}\right|_p \tag{25.9}$$

for all $0 \leq i \leq \dim E$ 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 $p \in B$ with representative σ is denoted by $j^r_p\sigma$. The number r is called the **order** of the jet.

Definition 25.1.24 (Jet manifold). Consider a fibre bundle (E, B, π). The r-jet manifold $J^r(\pi)$ of the projection π is defined as:

$$J^r(\pi) = \{j^r_p\sigma : \sigma \in \Gamma(E), p \in B\} \tag{25.10}$$

The set $J^0(\pi)$ is identified with the total space E.

Definition 25.1.25 (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

$$\pi_r : J^r(\pi) \to B : j^r_p\sigma \mapsto p \tag{25.11}$$

$$\pi_r : J^r(\pi) \to E : j^r_p\sigma \mapsto \sigma(p) \tag{25.12}$$

These projections satisfy $\pi_r = \pi \circ \pi_{r,0}$. We can also define a **k-jet projection** $\pi_{r,k}$ as the map

$$\pi_{r,k} : J^r(\pi) \to J^k(\pi) : j^r_p\sigma \mapsto j^k_p\sigma \tag{25.13}$$

where $k \leq r$. The k-jet projections satisfy a transitivity property $j_{k,m} = j_{r,m} \circ j_{k,r}$.

Definition 25.1.26 (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:

$$j^r\sigma : B \to J^r(\pi) : p \mapsto j^r_p\sigma \tag{25.14}$$

Definition 25.1.27 (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)[10] on E define *induced* bundle charts on $J^r(\pi)$ in the following way:

$$U^r_i = \{j^r_p\sigma : \sigma(p) \in U_i\} \tag{25.15}$$

$$\varphi^r_i = \left(x^k, u^\alpha, \left.\frac{\partial^I u^\alpha}{\partial x^I}\right|_p\right) \tag{25.16}$$

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 $J^r(\pi)$.

[10] Where $\varphi_i = (x^k, u^\alpha)$ with x^k the base space coordinates and u^α the total space coordinates.

25.2 Principal bundles

Definition 25.2.1 (Principal bundle). A principal bundle is a fibre bundle (E, B, π, G, G) such that the structure group and the typical fibre are the same, i.e. we identify the structure group with the group of left translations of G.

Remark 25.2.2. 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**. 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 25.2.3. The dimension of P is given by:

$$\dim P = \dim M + \dim G \tag{25.17}$$

Property 25.2.4. Let $\pi : P \to B$ be a principal G-bundle with local trivializations $\{(U_i, \varphi_i)\}_{i\in I}$. There exists a (faithful) right action of G on P given by:

$$z \cdot g = \varphi_i^{-1}(b, hg) \tag{25.18}$$

for all $g, h \in G$ and $z \in \pi(U_i)$. This action preserves fibres ($y \cdot g \in F_b$ for all $y \in F_b, g \in G$). Furthermore, it is free[11] and it is transitive. It follows that the fibres over B are exactly the orbits of the right action on P.

Every local trivialization φ_i is also G-equivariant:

$$\varphi_i(z \cdot g) = \varphi_i(z) \cdot g \tag{25.19}$$

Definition 25.2.5 (Bundle map). A bundle map $F : P_1 \to P_2$ between principal G-bundles is a pair of smooth maps (f_B, f_E) such that:

1. The diagram 25.2 below commutes.
2. f_E is G-equivariant[12].

The map f_E is said to **cover** f_B.

Example 25.2.6 (Associated principal bundle). For every fibre bundle (E, B, π, F, G) we can construct an associated principal G-bundle by replacing the fibre F by G itself.

Property 25.2.7. 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 25.2.8 (Frame bundle). Let V be an n-dimensional vector space. Denote the set of ordered bases, also called **frames**, of V by $F(V)$. This set is isomorphic to the group $GL(\mathbb{R}^n)$ which follows from the fact that every basis transformation is given by the action of an element of the general linear group. We can thus construct a principal bundle associated to the vector bundle E by replacing every fibre $\pi^{-1}(b)$ by $F(\pi^{-1}(b)) \cong \mathrm{GL}(\mathbb{R}^n)$.

[11] See definition 3.1.36.
[12] See definition 3.1.42.

$$\begin{tikzcd}
E_1 \arrow[r, "f_E"] \arrow[d, "\pi_1"'] & E_2 \arrow[d, "\pi_2"] \\
B_1 \arrow[r, "f_B"'] & B_2
\end{tikzcd}$$

Figure 25.2: Bundle map between principal G-bundles.

25.2.1 Reduction of the structure group

Construction 25.2.9. Consider a fibre bundle $\mathcal{F} = (E, B, \pi, F, G)$. Let H be a subgroup of G. If there exists a fibre bundle with structure group H equivalent to $\mathcal{F}$ then we say that the structure group G can be reduced to H.

Property 25.2.10. An n-dimensional manifold is orientable if and only if the structure group $GL(\mathbb{R}^n)$ of its frame bundle $F(M)$ is reducible to $GL^+(\mathbb{R}^n)$, i.e. the group of invertible matrices with positive determinant.

Definition 25.2.11 (G-structure). Consider a manifold M. A G-structure on M is the reduction of the structure group $GL(\mathbb{R}^n)$ of the frame bundle $F(M)$ to a subgroup $G \subset GL(\mathbb{R}^n)$.

Example 25.2.12. An $O(n)$-structure on M turns the manifold into a Riemannian manifold[13]. Because the cotangent bundle T^*M transforms[14] using the transpose inverse of the transition maps of the tangent bundle TM, which for maps in $O(n)$ is equal to the oringinal maps, these two bundles are equivalent. The isomorphism is given by the *musical isomorphisms*[15].

25.2.2 Sections

Where every vector bundle has at least one global section, the **zero section**[16], a general principal bundle does not necessarily have a global section. This is made clear by the following property:

Property 25.2.13. 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 $\mathrm{Triv}(P)$.

[13]See definition 27.1.3.
[14]See example 25.3.16.
[15]See definition 27.1.2.
[16]This is the map $s : b \to \vec{\mathbf{0}}$ for all $b \in B$.

25.3 Vector bundles

The tangent space, as introduced in subsection 22.2, can also be introduced in a more topological way:

25.3.1 Tangent bundle

Construction 25.3.1 (Tangent bundle). Let M be an n-dimensional manifold with atlas $\{(U_i, \varphi_i)\}_{i\leq n}$. Construct for every open set U an associated set $TU = U \times \mathbb{R}^n$ and construct for every smooth function f an associated smooth function on TU, called the **differential** or **derivative** of f, by:

$$Tf : U \times \mathbb{R}^n \to f(U) \times \mathbb{R}^n : (x, v) \mapsto (f(x), Df(x)v) \tag{25.20}$$

where $Df(x) : \mathbb{R}^n \to \mathbb{R}^n$ is the linear operator associated with the Jacobian matrix of f in x. Applying this definition to the transition functions ψ_{ji} we obtain a new set of functions $\widetilde{\psi}_{ji} := T\psi_{ji}$ given by:

$$\widetilde{\psi}_{ji}(\varphi_i(x), v) = \big(\varphi_j(x), D(\varphi_j \circ \varphi_i^{-1})(\varphi_i(x))v\big) \tag{25.21}$$

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 25.1.9 to the triple $(M, \mathbb{R}^n, GL(\mathbb{R}^n))$ together with the base cover $\{U_i\}_{i\leq n}$ and the cocycle $\{\widetilde{\psi}_{ji}\}_{i,j\in I}$.

Remark. The charts in the atlas of the constructed bundle are sometimes called **natural charts** because the first n coordinates are equal to the coordinates of the base space.

Alternative Definition 25.3.2. The above construction eventually comes down to the following (easier) definition of the tangent bundle:

$$TM = \bigsqcup_{p\in M} T_pM \tag{25.22}$$

equipped with the disjoint union topology 4.1.3 and the projection map[17]

$$\pi : TM \to M : (p, X) \mapsto p \tag{25.23}$$

where X is a tangent vector in T_pM. An atlas on TM is then given by the charts $(\pi^{-1}(U), \theta)$ with

$$\theta : TM \to \mathbb{R}^{2n} : (p, X) \mapsto (\varphi(p), X^1, ..., X^n) \tag{25.24}$$

where $X = X^i \left.\frac{\partial}{\partial x^i}\right|_p \in T_pM$ and where (U, φ) is a chart on M around the point $p \in M$.

[17] The single-valuedness of π is ensured because the tangent bundle is defined as the disjoint union of the tangent spaces.

Property 25.3.3. 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 the tangent bundle is a manifold of dimension $2n$.

Definition 25.3.4 (Tangent space). Let $x \in M$. The topological definition of the tangent space is given by the fibre

$$T_xM := \tau_M^{-1}(x) \tag{25.25}$$

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:

$$\begin{aligned}(x, v_1) + (x, v_2) &:= (x, v_1 + v_2)\\ r(x, v) &:= (x, rv)\end{aligned}$$

Remark 25.3.5. 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 25.3.6 (Differential). The map T from 25.20 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 $T_{f(x)}V$ (where $V = f(U)$) linearly. So T_xf is a linear map on fibres.

Property 25.3.7. The map $Tf : TM \to TN$ (see 25.20) has following properties[18]:

- $T(\mathbb{1}_M) = \mathbb{1}_{TM}$
- Let f, g be two smooth functions on smooth manifolds. Then $T(f \circ g) = Tf \circ Tg$.

Remark. We can also use a construction similar to that of the tangent bundle to reconstruct the original manifold M from the sets $\varphi_i(U_i)$.

Definition 25.3.8 (Rank). Let $f : M \to N$ be a differentiable map between smooth manifolds. Using the fact that Tf is a linear map of fibres[19], we define the rank of f at $p \in M$ as the rank (as in 15.3.16) of the differential $Tf : T_pM \to T_{f(p)}N$.

Theorem 25.3.9 (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.*

[18] This turns the map T into a functor on the category of smooth manifolds. We can view T as a functorial derivative.

[19] See definition 25.3.6.

25.3.2 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 25.3.10 (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.

Now we can construct a new topological space E, similar to the construction of the tangent bundle, by taking the disjoint union of the sets $\varphi_i(U_i) \times V$ and quotienting out using the functions $\widetilde{g}_{ji} : (\varphi_i(x), v) \mapsto (\varphi_j(x), g_{ji}(x) \cdot v)$, where $g \cdot v \equiv \rho(g)v$. This gives us a set of natural charts[20] $\{(\widetilde{U}_i, \widetilde{\varphi}_i)\}_{i\leq n}$, a projection map $\pi : E \to M$ induced by the local projection $\varphi_i(U_i) \times V \to \varphi_i(U_i)$ and a naturally defined vector space on every fibre $V_x := \pi^{-1}(x)$. Furthermore every fibre V_x is (although not necessarilly canonically) isomorphic to V.

This set E is called a **smooth vector bundle** over M with *typical fibre* V and *projection map* π.

Remark 25.3.11. 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.

Remark 25.3.12. Vector bundles are smooth fibre bundles where the typical fibre is a vector space V and the structure group is given by $GL(V)$.

Definition 25.3.13 (Associated vector bundle). Consider a representation $\rho : GL(\mathbb{R}^n) \to GL(\mathbb{R}^l)$ and the cocycle $t_{ji} := D(\psi_{ji}) \circ \varphi_i$ as defined for tangent bundles. 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. The vector bundle $E = \rho(TM)$ so obtained is called the associated bundle of the tangent bundle induced by ρ.

Remark 25.3.14. It should also be noted that every vector bundle is associated to a principle $GL(V)$-bundle where the cocycles g_{ji} now act by left multiplication on elements of $GL(V)$.

Example 25.3.15 (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 25.3.16 (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 $(t_{ji}^T)^{-1}$, is called the cotangent bundle on M and is denoted by T^*M. Elements of the fibres are called covariant vectors or covectors.

[20] We could instead use any other kind of topological space. The point is that a vector bundle is a fibre bundle 25.1.3 for which the typical fibres are vector spaces.

Notation 25.3.17. A combination of the cocycle t_{ji} and its dual $(t_{ji}^T)^{-1}$ 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 25.3.18 (Pseudovectors). If we consider the representation

$$\rho : A \mapsto \operatorname{sgn}\det(A)A \tag{25.26}$$

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 25.3.19 (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 from quantum mechanics, for which the sections correspond to the physical "wave functions".

Definition 25.3.20 (Whitney sum). Consider two vector bundles W, W' with fibres E, E' respectively. Then we can construct a new vector bundle $W \oplus W'$ by defining the new typical fibre to be the direct sum $E \oplus E'$, i.e. the fibre above b is given by $E_b \oplus E'_b$. This operation is called the Whitney sum or direct sum of vector bundles.

25.3.3 Sections

Definition 25.3.21 (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 25.3.22. A vector bundle is trivial if and only if there exists a frame of global sections.

Theorem 25.3.23 (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.*

25.4 Vector fields

Definition 25.4.1 (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 25.4.2. The set of all vector fields on a manifold M is often denoted by $\mathfrak{X}(M)$.

Theorem 25.4.3 (Hairy ball theorem). *There exists no nowhere vanishing vector field on an even-dimensional sphere S^{2n}.*

Definition 25.4.4 (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:

$$(\varphi^* X)_p = T\varphi^{-1}(X_{\varphi(p)}) \tag{25.27}$$

Definition 25.4.5 (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:

$$(\varphi_* X)_{\varphi(p)} = T\varphi(X_p) \tag{25.28}$$

which we can rewrite using the pullback as:

$$\varphi_* X = \varphi^{-1*} X \tag{25.29}$$

Or equivalently we can define a vector field on N by:

$$(\varphi_* X)_q(f) = X_{\varphi^{-1}(q)}(f \circ \varphi) \tag{25.30}$$

for all smooth functions $f : N \to \mathbb{R}$ and points $q \in N$.

25.4.1 Integral curves

Definition 25.4.6 (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:

$$\boxed{\gamma'(t) = X(\gamma(t))} \tag{25.31}$$

for all $t \in]a, b[$ where we defined $\gamma'(t) := T\gamma(t, 1)$ using the functorial derivative 25.20.

This equation can be seen as a system of ordinary differential equations in the second argument. Using Picard's existence theorem[21] together with the initial value condition $\gamma(0) = p$ we can find a unique curve on $]a, b[$ satisfying the defining equation 25.31. 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 25.4.7 (Flow). Let $X \in \mathfrak{X}(M)$. The function σ_t:

$$\sigma_t(p) = \gamma_p(t) \tag{25.32}$$

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 25.4.8. Suppose that $D(X) = \mathbb{R} \times M$. The flow σ_t has following properties for all $s, t \in \mathbb{R}$:

- $\sigma_0 = \mathbb{1}_M$
- $\sigma_{s+t} = \sigma_s \circ \sigma_t$
- $\sigma_{-t} = (\sigma_t)^{-1}$

These three properties[22] 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.

[21] Also Picard-Lindelöf theorem.
[22] The third property follows from the other two.

Definition 25.4.9 (Complete vector field). A vector field X is called complete if the flow domain for every flow is whole $\mathbb{R}$.

Property 25.4.10. 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.

25.4.2 Lie derivative

Formula 25.4.11 (Lie derivative for smooth functions). Let $X \in \mathfrak{X}(M)$ and let $f : M \to \mathbb{R}$ be a smooth function. The Lie derivative of f with respect to X at $p \in M$ is defined as:

$$\boxed{(\mathcal{L}_X f)(p) = \lim_{t\to 0} \frac{f(\gamma_p(t)) - f(p)}{t}} \tag{25.33}$$

which closely resembles the standard derivative in Euclidean space.

Formula 25.4.12 (†). Working out previous formula and rewriting it as an operator equality gives:

$$\boxed{\mathcal{L}_X = \sum_k X_k \frac{\partial}{\partial x^k}} \tag{25.34}$$

It is clear that this is just the vector field X expanded in the basis 22.2.3. We also recover the behaviour of a tangent vector as a derivation. So for smooth functions $f : M \to \mathbb{R}$ we obtain:

$$\mathcal{L}_X f(p) = X_p(f) \tag{25.35}$$

Formula 25.4.13 (Lie derivative for vector fields†). Let $X, Y \in \mathfrak{X}(M)$

$$\boxed{\mathcal{L}_X Y = \frac{d}{dt}(\sigma_t^* X)(\gamma_p(t))\bigg|_{t=0}} \tag{25.36}$$

Property 25.4.14. 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**:
$$\mathcal{L}_X Y = [X, Y] \tag{25.37}$$
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.
- The Lie derivative is antisymmetric:
$$\mathcal{L}_X Y = -\mathcal{L}_Y X \tag{25.38}$$
This follows from the previous property.

25.4.3 Grassmann bundle

Looking at property 15.2.11 and noting that $\mathrm{GL}_n(\mathbb{R})$ is a Lie group, we can endow the Grassmannian $\mathrm{Gr}(k, \mathbb{R}^n)$ 15.2.10 with a differentiable structure, turning it into a smooth manifold. With this we can construct a new bundle[23] by applying the usual construction theorem 25.1.9:

Construction 25.4.15 (Grassmann bundle). First define the transition functions:

$$\psi_{ji} : \varphi_i(U_i \cap U_j) \times \mathrm{Gr}(k, \mathbb{R}^n) \rightarrow \varphi_j(U_i \cap U_j) \times \mathrm{Gr}(k, \mathbb{R}^n) : (\varphi_i(x), V) \mapsto (\varphi_j(x), t_{ji}(x) \cdot V) \tag{25.39}$$

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 $\mathbb{R}^n$. This set of transition functions is then used to create a new fibre bundle where every fibre is diffeomorphic to $\mathrm{Gr}(k, \mathbb{R}^n)$, namely it is the Grassmannian $\mathrm{Gr}(k, T_pM)$ associated to the tangent space in every point $p \in M$.

Notation 25.4.16. The Grassman k-plane bundle is denoted by $\mathrm{Gr}(k, TM)$.

25.4.4 Frobenius theorem

Definition 25.4.17 (Distribution). A smooth section of the Grassman k-plane bundle is called a distribution of k-planes.

Definition 25.4.18 (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 25.4.19 (Frobenius integrability condition). A distribution of k-planes 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 25.4.20 (Frobenius' integrability theorem). *Let W be a distribution of k-planes over a smooth manifold M. Then W is integrable if and only if W satisfies the Frobenius integrability condition.*

25.5 Differential k-forms

Definition 25.5.1 (Differential form). A differential k-form is a map

$$\boxed{\omega : T^{\diamond k}M \rightarrow \mathbb{R}} \tag{25.40}$$

[23] Due to the fact that the Grassmannian is not a vector space, we construct a general fibre bundle and not a vector bundle.

such that the restriction of ω to each fibre of the fibre product[24] $T^{\diamond k}M$ 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 $C^\infty : M \to \mathbb{R}$.

Alternative Definition 25.5.2. An alternative definition goes as follows. Consider the representation

$$\rho_k : GL(R^{m*}) \to GL(\Lambda^k(\mathbb{R}^{m*})) : T \mapsto T \wedge ... \wedge T$$

where T is a linear map. This representation induces an associated vector bundle[25] $\rho_k(\tau_M^*)$ 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 25.5.3. We can construct a Grassmann algebra[26] by equipping the graded vector space

$$\Omega(M) = \bigoplus_{k\geq 0} \Omega^k(M) \tag{25.41}$$

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 25.5.4 (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:

$$\boxed{f^*(\omega) = \omega \circ f_* : TM \to \mathbb{R}} \tag{25.42}$$

So f^* can be seen as a map pulling elements from T^*N back to T^*M.

Definition 25.5.5 (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:

$$f_*(\omega) : \omega \circ (f^{-1})_* : TN \to \mathbb{R} \tag{25.43}$$

Remark. Note that the pushforward of differential k-form is only defined for diffeomorphisms, in constrast to pullbacks which only require smooth functions. Furthermore 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.

[24]See definition 25.8.
[25]See definition 25.3.13.
[26]As in definition 19.4.14.

Formula 25.5.6 (Dual basis). Consider the basis $\{\left.\frac{\partial}{\partial x_i}\right|_p\}_{i\leq n}$ from definition 22.2.3 for the tangent space T_pM. From this set we can construct[27] a natural dual basis for the cotangent space T_p^*M using the natural pairing of these spaces:

$$\left\langle \frac{\partial}{\partial x^i}, dx^j \right\rangle = \delta_i^j \tag{25.44}$$

25.5.1 Exterior derivative

Definition 25.5.7 (Exterior derivative). The exterior derivative d_k is a map defined on the graded algebra of differential k-forms:

$$d_k : \Omega^k(M) \rightarrow \Omega^{k+1}(M) \tag{25.45}$$

For $k = 0$ it is given by[28]:

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i \tag{25.46}$$

where we remark that the 'infinitesimals' are in fact unit vectors with norm 1. This formula can be generalized to higher dimensions as follows:

$$\boxed{d(f dx_{i_1} \wedge ... \wedge dx_{i_k}) = df \wedge dx_{i_1} \wedge ... \wedge dx_{i_k}} \tag{25.47}$$

Corollary 25.5.8. It follows immediately from 25.47 that

$$d(dx_i) = 0 \tag{25.48}$$

for all $i \leq n$.

Property 25.5.9. 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 $\mathrm{im}(d_k) \subseteq \ker(d_{k+1})$.
- The exterior derivative is an $\mathbb{R}$-linear map.
- Graded Leibniz rule:

$$d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^j \omega_1 \wedge d\omega_2 \tag{25.49}$$

  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 25.42.

[27] 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 25.5.7 to the coordinate maps x^i.

[28] For $f \in \Omega^0(M)$, we call df the **differential** of f.

Remark 25.5.10 (†). The gradient, rotor (curl) and divergence from standard vector calculus[29] 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 $\sim$.

$$\boxed{\begin{aligned} \sim df &= \nabla f \\ \sim (*d\alpha) &= \nabla \times \vec{\boldsymbol{f}} \\ *d\omega &= \nabla \cdot \vec{\boldsymbol{f}} \end{aligned}} \qquad \begin{aligned} (25.50) \\ (25.51) \\ (25.52) \end{aligned}$$

The properties in section 16.1.2 then follow from the identity $d^2 = 0$.

Example 25.5.11. Let $f \in C^\infty(M, \mathbb{R})$. Let γ be a curve on M. From the definition 25.44 of the basis $\{dx_k\}_{k\leq n}$ we obtain following result:

$$\langle df(x), \gamma'(t)\rangle = \sum_k \frac{\partial f}{\partial x_k}(x)\gamma'_k(t) = (f \circ \gamma')(t) \tag{25.53}$$

Example 25.5.12. An explicit formula for the exterior derivative of a k-form Φ is:

$$\begin{aligned} d\Phi(X_1, ...nX_{k+1}) = &\sum_{i=0}^{k+1}(-1)^i X_i(\Phi(X_1, ..., \hat{X}_i, ..., X_{k+1})) \\ &- \sum_{i<j} \Phi([X_i, X_j], X_1, ..., \hat{X}_i, ..., \hat{X}_j, ..., X_{k+1}) \end{aligned} \tag{25.54}$$

where $\hat{X}$ means that this argument is omitted.

25.5.2 Lie derivative

Formula 25.5.13 (Lie derivative of differential forms).

$$\boxed{\mathcal{L}_X\omega(p) = \lim_{t\to 0} \frac{\sigma_t^*\omega - \omega}{t}(p)} \tag{25.55}$$

Formula 25.5.14 (Lie derivative of smooth functions). Using the definition of the exterior derivative of smooth functions 25.46 and the definition of the dual (cotangent) basis 25.44 we can rewrite the Lie derivative 25.34 as:

$$Xf(p) = df_p(X(p)) \tag{25.56}$$

Property 25.5.15. The Lie derivative also has following Leibniz-type property with respect to differential forms:

$$\mathcal{L}_X(\omega(Y)) = (\mathcal{L}_X\omega)(Y) + \omega(\mathcal{L}_X Y) \tag{25.57}$$

where X, Y are two vector fields and ω is a 1-form.

[29] See section 16.1.

Formula 25.5.16 (Lie derivative of tensor fields). By comparing the definitions of the Lie derivatives of vector fields 25.36 and differential forms 25.55 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 $\mathcal{T} \in \Gamma(T^{(k,l)}M)$:

$$\boxed{\mathcal{L}_X\mathcal{T}(p) = \frac{d}{dt}\sigma_t^*\mathcal{T}(\gamma_p(t))\bigg|_{t=0}} \tag{25.58}$$

25.5.3 Interior product

Definition 25.5.17 (Interior product). Aside from the differential (exterior derivative) we can also define another operation on the algebra of differential forms:

$$\iota_X : (\iota_X\omega)(v_1, ..., v_{k-1}) \mapsto \omega(X, v_1, ..., v_{k-1}) \tag{25.59}$$

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 19.7.

Formula 25.5.18 (Cartan[30]). 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:

$$\mathcal{L}_X\omega = \iota_X(d\omega) + d(\iota_X\omega) \tag{25.60}$$

25.5.4 de Rham Cohomology

Definition 25.5.19 (Exact form). Let $\omega \in \Omega^k(M)$. If ω can be written as $\omega = d\chi + 0$ for some $\chi \in \Omega^{k-1}(M)$ and $0 \in \Omega^0(M)$ the zero function then ω is said to be exact. It follows that

$$\mathrm{im}(d_k) = \{\omega \in \Omega^{k+1}(M) : \omega \text{ is exact}\} \tag{25.61}$$

Definition 25.5.20 (Closed form). Let $\omega \in \Omega^k(M)$. If $d\omega = 0$ then ω i said to be closed. It follows that

$$\{\omega \in \Omega^k(M) : \omega \text{ is closed}\} \subseteq \ker(d_k) \tag{25.62}$$

Remark 25.5.21. From the first item of property 25.5.9 it follows that every exact form is closed. The converse however is not true[31].

Definition 25.5.22 (Cochain complex). Let $(A_k)_{k\in\mathbb{N}}$ be a sequence of Abelian groups or modules together with a sequence $(\partial_k)_{k\in\mathbb{N}}$ of homomorphisms, called the **boundary operators** or **differentials**, such that for all k:

$$\partial_k : A_k \to A_{k+1} \tag{25.63}$$

[30] Sometimes called *Cartan's magic formula* or *Cartan's (infinitesimal) homotopy formula*.
[31] See result 25.5.26 for more information.

Furthermore let ${\partial_k}^2 = 0$ for every $k \in \mathbb{N}$. This structure is called a cochain complex[32]. Elements in $\mathrm{im}(\partial_k)$ are called **coboundaries** and elements in $\ker(\partial_k)$ are called **cocycles**.

Definition 25.5.23 (de Rham complex). The structure given by the chain

$$0 \to \Omega^0(M) \xrightarrow{d} \Omega^1(M) \xrightarrow{d} ... \tag{25.64}$$

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 25.5.24 (de Rham cohomology). The k^{th} de Rham cohomology group on M is defined as the following quotient space:

$$\boxed{H^k_{\mathrm{dr}}(M) = \frac{\ker(d_{k+1})}{\mathrm{im}(d_k)}} \tag{25.65}$$

This quotient space is a vector space. Two elements of the same equivalence class in $H^k_{dr}(M)$ are said to be **cohomologous**.

One can construct a graded ring 3.2.15 from these cohomology groups, called the cohomology ring H^*. The product is called the **cup product** $\smile$ and it is a graded-commutative product (see 3.20).

Definition 25.5.25 (Cup product). Let $[\nu] \in H^k_{\mathrm{dr}}$ and $[\omega] \in H^l_{\mathrm{dr}}$, where we used $[\cdot]$ to show that the elements are in fact equivalence relations belonging to differential forms ν and ω. The cup product is defined as follows: $[\nu] \smile [\omega] = [\nu \wedge \omega]$.

Theorem 25.5.26 (Poincaré's lemma[33]). *For every point $p \in M$ there exists a neighbourhood on which the de Rham cohomology is trivial:*

$$\forall p \in M : \exists U \subseteq M : H^k_{dr}(U) = 0 \tag{25.66}$$

This implies that every closed form is locally exact.

25.5.5 Vector-valued differential forms

Definition 25.5.27 (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:

$$\Omega^k(M, E) = \Gamma(E \otimes \Lambda^k T^*M) \tag{25.67}$$

[32] A chain complex is constructed similarly. For this structure we consider a descending order, i.e.: $\partial_k : A_k \to A_{k-1}$.

[33] The original theorem states that on a contractible space (see definition 4.6.5) every closed form is exact.

Formula 25.5.28 (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:

$$\omega \wedge \nu(v_1, ..., v_{k+p}) = \frac{1}{(k+p)!} \sum_{\sigma \in S_{k+p}} \mathrm{sgn}(\sigma) \omega(v_{\sigma(1)}, ..., v_{\sigma(k)}) \otimes \nu(v_{\sigma(k+1)}, ..., v_{\sigma(p)}) \tag{25.68}$$

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 general tensor algebra). It should be noted that 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 25.5.29 (Lie-algebra-valued differential form). A vector-valued differential form where the vector space V is equipped with a Lie algebra structure.

Formula 25.5.30 (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:

$$[\omega \wedge \nu](v_1, ..., v_{k+p}) = \frac{1}{(k+p)!} \sum_{\sigma \in S_{k+p}} \mathrm{sgn}(\sigma) [\omega(v_{\sigma(1)}, ..., v_{\sigma(k)}), \nu(v_{\sigma(k+1)}, ..., v_{\sigma(p)})] \tag{25.69}$$

where $[\cdot, \cdot]$ is the Lie bracket in $\mathfrak{g}$.

25.6 Connections

25.6.1 Vertical bundle

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 this to construct the notions of horizontal and vertical bundles:

Definition 25.6.1 (Vertical vector). Let $\pi : E \to B$ be a smooth fibre bundle. The subbundle $\ker(T\pi)$ 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 25.6.2. Consider a smooth principal G-bundle $G \hookrightarrow P \xrightarrow{\pi} M$. We first construct a map ι_p for every element $p \in P$:

$$\iota_p : G \to P : g \mapsto p \cdot g \tag{25.70}$$

We then define a tangent vector $v \in T_pP$ to be vertical if it lies in the image of $T_e\iota_p$, i.e. $\mathrm{Vert}(T_pP) = \mathrm{im}(T_e\iota_p)$. This construction is supported by the exactness of following short sequence:

$$0 \to \mathfrak{g} \xrightarrow{T_e\iota_p} T_pP \xrightarrow{T_p\pi} T_xM \to 0 \tag{25.71}$$

Property 25.6.3 (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 $T_e\iota_p : \mathfrak{g} \to \mathrm{Vert}(T_pP)$ is an isomorphism of vector spaces. In particular, it implies that

$$\dim \mathrm{Vert}(T_pP) = \dim \mathfrak{g} = \dim G \tag{25.72}$$

Definition 25.6.4 (Fundamental vector field). Consider a principle 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

$$A^\#(p) = \iota_{p,*}(A) \in \mathrm{Vert}(T_pP) \tag{25.73}$$

is called the fundamental vector field associated to A.

Property 25.6.5. The map $(\cdot)^\# : \mathfrak{g} \to \Gamma(TP)$ is a Lie algebra morphism:

$$[A, B]^\# = [A^\#, B^\#] \tag{25.74}$$

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 25.37.

Property 25.6.6. The vertical bundle satisfies the following G-equivariance condition:

$$R_{g,*}(\mathrm{Vert}(T_pP)) = \mathrm{Vert}(T_{pg}P) \tag{25.75}$$

By differentiating the equality

$$R_g \circ \iota_p = \iota_{pg} \circ \mathrm{ad}_{g^{-1}}$$

and using 25.73 and 24.3.1 we obtain the following algebraic formulation of the G-equivariance condition:

$$R_{g,*}\left(A^\#(p)\right) = \left(\mathrm{Ad}_{g^{-1}}A\right)^\#(pg) \tag{25.76}$$

25.6.2 Horizontal bundle

Definition 25.6.7 (Connection). Consider a principal bundle P with structure group G. A 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.[34]
- The subspace $\mathrm{Hor}(T_pP)$ is G-equivariant:

$$R_{g,*}(\mathrm{Hor}(T_pP)) = \mathrm{Hor}(T_{pg}P) \tag{25.77}$$

The elements of $\mathrm{Hor}(T_pP)$ are said to be **horizontal vectors** with respect to the connection.

[34] See the definiton of a (smooth) distribution 25.4.17.

Remark 25.6.8. Remark that the G-invariance condition for vertical bundles is an intrinsic property while we have to require it by definition for the horizontal bundle.

Definition 25.6.9 (Horizontal bundle). The horizontal bundle $\mathrm{Hor}(TP)$ is defined as $\bigcup_{p\in P}\mathrm{Hor}(T_pP)$. It is a subbundle of TP. The G-invariance condition then implies that this subbundle is invariant under (the pushforward of) the right action of G.

Property 25.6.10 (Dimension). Properties 25.17, 25.72 and the direct sum decomposition of T_pP imply the following relation:

$$\dim \mathrm{Hor}(T_pP) = \dim M \tag{25.78}$$

Property 25.6.11. We take some time to summarize all dimensional relations between the components of a principal G-bundle over a base manifold M:

$$\boxed{\begin{aligned} \dim P &= \dim M + \dim G & (25.79)\\ \dim M &= \dim \mathrm{Hor}(T_pP) & (25.80)\\ \dim G &= \dim \mathrm{Vert}(T_pP) & (25.81)\end{aligned}}$$

for all $p \in P$.

Definition 25.6.12 (Dual connection). First we define the dual of the horizontal bundle:

$$\mathrm{Hor}(T_p^*P) = \{h^* \in T_p^*P | h^*(v) = 0, v \in \mathrm{Vert}(T_pP)\} \tag{25.82}$$

Equivalently, the horizontal covector bundle is defined as the set of linear functionals that annihilate vertical vectors. Just as with the vertical bundle this structure is independent of any connection on P.

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 25.6.7 where Vert and Hor should be interchanged.

25.6.3 Ehresmann connection

Definition 25.6.13 (Ehresmann connection). Let (P, M, π, G) be a principal bundle. An Ehresmann connection is a $\mathfrak{g}$-valued 1-form $\omega : TP \to \mathfrak{g}$ that satisfies following 2 conditions:

1. $$\omega \circ R_{g,*} = \mathrm{Ad}_{g^{-1}} \circ \omega \tag{25.83}$$

2. $$\omega(A^\#) = A \tag{25.84}$$

The horizontal subspaces are then defined as $\mathrm{Hor}(T_pP) = \ker \omega|_p$.

Property 25.6.14. Consider two principal G-bundles ξ_1 and ξ_2. Let ω be an Ehresmann connection on ξ_1 and let $F; \xi_1 \to \xi_2$ be a bundle map covering a smooth map f. The map $F^*\omega$ defines an Ehresmann connection on ξ_2.

Example 25.6.15. Consider a principal G-bundle. An Ehresmann connection on this bundle is given by the following map:

$$\omega = (T_e\iota_p)^{-1} \circ \mathrm{pr}_V \tag{25.85}$$

where pr_V is the projection $TP \to \mathrm{Vert}(TP)$ associated to the decomposition from definition 25.6.7.

Definition 25.6.16 (Horizontal and vertical forms). Let ω be an Ehresmann connection on a principal bundle $G \hookrightarrow P \to M$. Let $\theta \in \Omega^k(P)$ be a k-form. We define following notions:

- θ is said to be horizontal if
$$\theta(v_1, ..., v_k) = 0 \tag{25.86}$$
whenever at least 1 of the v_i lies in $\mathrm{Vert}(T_pP)$.
- θ is said to be vertical if
$$\theta(v_1, ..., v_k) = 0 \tag{25.87}$$
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.

25.6.4 Maurer-Cartan form

Definition 25.6.17 (Maurer-Cartan form). For every $g \in G$ we have that the tangent space T_gG is isomorphic to $T_eG \cong \mathfrak{g}$. The isomorphism $T_gG \to \mathfrak{g}$ is given by the Maurer-Cartan form:

$$\boxed{\Omega := L_{g^{-1},*}} \tag{25.88}$$

Definition 25.6.18. 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 25.6.11 we see that the horizontal spaces are null-spaces (which defines a smooth distribution and thus a connection according to 25.6.7) and that the vertical spaces are equal to the tangent spaces, i.e. $\mathrm{Vert}(T_gG) = T_gG$ (where we already made 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 25.6.19. The Maurer-Cartan form is the unique Ehresmann connection on the bundle $G \hookrightarrow G \to \{x\}$.

25.6.5 Horizontal lifts and parallel transport

Property 25.6.20. 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 25.6.21 (Horizontal curve). Curves satisfying the last condition are said to be horizontal.

Definition 25.6.22 (Parallel transport on principal bundles). The parallel transport map with respect to the curve γ is defined as follows:

$$\mathrm{Par}_t^\gamma : \pi^{-1}(\gamma(0)) \to \pi^{-1}(\gamma(t)) : p_0 \mapsto \widetilde{\gamma}_{p_0}(t) \tag{25.89}$$

This map is G-equivariant and it is a diffeomorphism of fibres.

Formula 25.6.23. 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 this bundle. The horizontal lift of $\gamma(t)$ can locally be parametrized as $(\gamma(t), g(t))$ where $g(t)$ is some unique curve in G. To determine $\widetilde{\gamma}(t)$ it is thus sufficient to find $g(t)$. The following parametrization uniquely characterizes $g(t)$:

$$g'(t) = -\omega(\gamma(t), \mathbb{1}_G, \gamma'(t), 0)g(t) \tag{25.90}$$

Using the trivial section $s : U \to U \times G : x \mapsto (x, \mathbb{1}_G)$ where U is an open subset of M we can rewrite this formula as follows: First we consider the action of the pullback $s^*\omega$ on the derivative $\gamma_* : \mathbb{R} \times \mathbb{R} \to TM : (t,1) \mapsto (\gamma(t), \gamma'(t))$. 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

$$s^*\omega(\gamma(t), \gamma'(t)) = \omega\Big(s_*(\gamma(t), \gamma'(t))\Big) = \omega(\gamma(t), \mathbb{1}_G, \gamma'(t), 0)$$

Combining these relations with the ODE for $g(t)$ gives

$$\left(\frac{d}{dt} + A(\gamma(t))\gamma'(t)\right) g(t) = 0 \tag{25.91}$$

where $\frac{d}{dt}$ is a matrix given by the scalar multiplication of the derivative $\frac{d}{dt}$ and the unit matrix I.

Definition 25.6.24 (Holonomy group). Consider a principal bundle $G \hookrightarrow P \to M$. Let $\Omega_m^{ps}M \subset \Omega M$ be the loop space with basepoint $m \in M$ 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:

$$\mathrm{Hol}_p(\omega) = \{g \in G : p \text{ and } p \cdot g \text{ can be connected by a } \widetilde{\gamma}, \gamma \in \Omega_m^{ps}M\} \tag{25.92}$$

Definition 25.6.25 (Reduced holoomy group). The reduced holonomy group $\mathrm{Hol}_p^0(\omega)$ is defined as the subset of $\mathrm{Hol}_p(\omega)$ using only contractible loops.

25.6.6 Koszul connections and covariant derivatives

Definition 25.6.26 (Parallel transport on vector bundles). Consider a principal bundle $G \hookrightarrow P \to M$ where we explicitly require P to be trivial, i.e. $P = M \times G$. Suppose that the Lie group G acts on a vector space V by a representation $\rho : G \to GL_m$. We can then construct an associated vector bundle $\pi_1 : M \times V \to M$.

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 $x_1 = \gamma(1)$. Furthemore, let the horizontal lift $\widetilde{\gamma}(t)$ have $\widetilde{\gamma}(0) = (x_0, h)$ as initial condition. The parallel transport of the point $(x_0, v_0) \in M \times V$ along γ is given by the following map:

$$\mathrm{Par}_t^\gamma : \pi_1^{-1}(x_0) \to \pi_1^{-1}(\gamma(t)) : (x_0, v_0) \mapsto \big(\gamma(t), \rho\big(g(t)h^{-1}\big)v_0\big) \tag{25.93}$$

It should be noted that this map is independent of the initial element $h \in G$. 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 25.6.27. Two remarks have to be made. First of all, although the previous construction explicitly used trivial bundles, it is also valid for general non-trivial vector bundles. Secondly, following remark 25.3.14 we can construct a principal bundle for any vector bundle and use the parallel transport on this bundle to define parallel transport of vectors. The previous construction is thus possible for every vector bundle.

Definition 25.6.28 (Covariant derivative). Consider a vector bundle with model fibre space V and its associated principal $GL(V)$-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:

$$\nabla_X\sigma(x_0) = \lim_{t\to+\infty} \frac{(\mathrm{Par}_t^\gamma)^{-1}\sigma(\gamma(t)) - \sigma(x_0)}{t} \tag{25.94}$$

where $\gamma(t)$ is any curve such that $\gamma(0) = x_0$ and $\gamma'(0) = X(x_0)$.

Property 25.6.29. Let $\pi : E \to M$ be a vector bundle. Let σ, X and f be respectively a section on E, a vector field on M and a C^∞ function on M. The covariant derivative along X satisfies following properties:

- $\nabla_X\sigma$ is a smooth section on E.
- The map $(X,\sigma)\mapsto\nabla_X\sigma$ is bilinear over $\mathbb{R}$.
- $\nabla_{(fX)}\sigma = f\nabla_X\sigma$
- $\nabla_X(f\sigma) = f\nabla_X\sigma + X(f)\sigma$

Remark 25.6.30. The last two properties show the major difference between the Lie derivative and the covariant derivative when σ is a section of the tangent bundle, i.e. a vector field. 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.

Definition 25.6.31 (Koszul connection). The map

$$\Gamma(TM)\times\Gamma(E)\to\Gamma(E):(X,\sigma)\mapsto\nabla_X\sigma \tag{25.95}$$

is called a Koszul connection if the above properties hold. From the above constructions it also follows that every Ehresmann connection on a principal bundle induces a Koszul connection on all of its associated vector bundles.

Definition 25.6.32 (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:

$$D\theta(v_0,...,v_k) = d\theta(v_0^H,...,v_k^H) \tag{25.96}$$

where d is the exterior derivative 25.5.7 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 covariant derivative $D\theta$ is a horizontal form 25.6.16.

Remark 25.6.33. 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.

Formula 25.6.34. Using the Koszul connection on the tangent bundle TP we can rewrite the action of the exterior covariant derivative as follows:

$$D\theta(v_0,...,v_k) = \sum_i^k(-1)^i\nabla_{v_i}\theta(v_0,...,\hat{v}_i,...,v_k) + \sum_{i<j}^k(-1)^{i+j}\theta([v_i,v_j],v_0,...,\hat{v}_i,...,\hat{v}_j,...,v_k) \tag{25.97}$$

where $\hat{v}_i$ means that this vector is omitted. As an example we explicitly give the formula for a 1-form Φ:

$$D\Phi(X,Y) = \nabla_X(\Phi(Y)) - \nabla_Y(\Phi(X)) - \Phi([X,Y]) \tag{25.98}$$

which should remind the reader of the analogous formula for the ordinary exterior derivative 25.54.

25.6.7 Curvature of a connection

Definition 25.6.35 (Curvature). Let ω be an Ehresmann connection on a principal bundle $G \hookrightarrow P \to M$. The curvature of ω is defined as the exterior covariant derivative $D\omega$.

Definition 25.6.36 (Flat connection). An Ehresmann connection ω is said to be flat if its curvature $D\omega$ vanishes everywhere.

The following property is an immediate consequence of the Frobenius integrability theorem 25.4.20 and the fact that an Ehresmann connection vanishes on the horizontal subbundle.

Property 25.6.37. Let ω be an Ehresmann connection. The associated horizontal distribution[35]

$$p \mapsto \mathrm{Hor}(T_pP)$$

is integrable if and only if the connection ω is flat. The vertical distribution is always integrable.

25.7 Complex bundles

25.7.1 Almost complex structure

Definition 25.7.1 (Almost complex structure). Let M be a real manifold. An almost complex strcuture on M is a smooth $(1,1)$-tensor field $J : TM \to TM$ such that $J|_p : T_pM \to T_pM$ satisifes $J|_p^2 = -1$ for all $p \in M$. An **almost complex manifold** is a real manifold equipped with an almost complex structure.

Property 25.7.2. An almost complex manifold is even-dimensional and orientable.

Property 25.7.3. 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 $GL(\mathbb{C}^n)$.

[35] See 25.4.17 for the definition of a distribution of vector spaces.

Chapter 26

Integration on manifolds

26.1 Orientation

Definition 26.1.1 (Orientation). Similar to definition 19.4.18 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 form $\mathrm{Vol} \in \Omega^n(M)$ where $n = \dim(M)$. The definition of an orientation is then equivalent to that in 19.4.18.

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**.

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 26.1.2. Let $M = \mathbb{R}^n$. The canonical Euclidean volume form is given by the determinant map

$$\det : (u_1, ..., u_n) \mapsto \det(u_1, ..., u_n) \tag{26.1}$$

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 26.1.3. 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.

[1]This is in fact an equivalent definition.

26.2 Integration of top-dimensional forms

Formula 26.2.1. Let θ be a top-dimensional form on M with compact support. Let $\{\varphi_i\}_i$ be a partition of unity[2] subordinate to an atlas $\{(U_i, \varphi_i)\}_i$.

$$\int_M \theta = \sum_i \int_{U_i} \varphi_i \theta dx_1 ... dx_n \tag{26.2}$$

26.3 Stokes' theorem

Theorem 26.3.1 (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:*

$$\boxed{\int_{\partial\Sigma} \omega = \int_\Sigma d\omega} \tag{26.3}$$

Corollary 26.3.2. The Kelvin-Stokes theorem 16.22, the divergence theorem 16.23 and Green's identity 16.24 are immediate results of this (generalized) Stokes' theorem.

26.4 de Rham Cohomology

Now we can also give a little side note about why the de Rham cohomology groups 25.65 really form a cohomology theory. For this we need some concepts from homology which can be found in section 4.7. 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 26.2 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]:

$$\langle \cdot, C \rangle : \Omega^n(M) \to \mathbb{R} : \omega \mapsto \int_C \omega = \sum_{i=0}^k a_i \int_{\Delta_k} \lambda_i^* \omega \tag{26.4}$$

where λ_i^* pulls back ω to Δ^k which is a subset of $\mathbb{R}^k$ as required. Now Stokes' theorem 26.3 tells us that

$$\int_C d\omega = \int_{\partial C} \omega \tag{26.5}$$

[2] See definition 4.5.15.

[3] This requires the chain group to have real coefficients instead of integer coefficients as is mostly used in homology.

Using the paring $\langle \cdot, \cdot \rangle$ this becomes

$$\langle d\omega, C \rangle = \langle \omega, \partial C \rangle \tag{26.6}$$

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 25.65.

[4] Suppose that $A, B \in [C]$ and $\phi, \chi \in [\omega]$ then $\langle \phi, A \rangle = \langle \chi, B \rangle$.

Chapter 27

Riemannian Geometry

27.1 Riemannian manifolds

27.1.1 Metric

Definition 27.1.1 (Bundle metric). Consider the bundle of second order covariant vectors. Following from 19.1.1 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 27.1.2 (Musical isomorphisms). Let $g : TM \times TM \to \mathbb{R}$ be the Riemannian metric on M. The **flat** isomorphism is defined as:

$$\flat : v \mapsto g(v, \cdot) \tag{27.1}$$

The **sharp** isomorphism is defined as the inverse map:

$$\sharp : p \mapsto v \tag{27.2}$$

such that $p(\cdot) = g(v, \cdot)$. These 'musical' isomorphisms can be used to lower and raise tensor indices.

[1]See also the section about Hermitian forms and metric forms 15.4.

[2]See definition 25.8.

27.1.2 Riemannian manifold

Definition 27.1.3 (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.

Definition 27.1.4 (Riemannian isometry). Let (M, g_M) and (N, g_N) be two Riemannian manifolds. An isometry 21.2.10 $f : M \rightarrow N$ is said to be Riemannian if $F^*g_N = g_M$.

Property 27.1.5. 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 pseudometriv 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 27.1.6 (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.

Theorem 27.1.7 (Whitney's embedding theorem). *Every smooth paracompact[3] manifold M can be embedded in $\mathbb{R}^{2\dim M}$.*

Theorem 27.1.8 (Whitney's immersion theorem). *Every smooth paracompact manifold M can be immersed in $\mathbb{R}^{2\dim M-1}$.*

Theorem 27.1.9 (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.*

27.2 Sphere bundle

Definition 27.2.1 (Unit sphere bundle). Let V be a normed vector space. Consider a vector bundle $V \hookrightarrow E \rightarrow B$. 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 27.2.2 (Unit disk bundle). A similar construction can be made by replacing the unit sphere by the unit disk $\{v \in V : ||v|| \leq 1\}$.

[3] See definition 4.5.13.

27.3 Hilbert bundles

Definition 27.3.1 (Hilbert bundle). A vector bundle for which the typical fibre is a Hilbert space is called a Hilbert bundle.

Definition 27.3.2 (Compatible Hilbert bundle). Consider the isomorphisms

$$l_x : F_x \to \mathcal{H} : h \mapsto \varphi_i(x,h) \in \pi(x) \tag{27.3}$$

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:

$$\langle v|w\rangle_x = \langle l_x(v)|l_x(w)\rangle_{\mathcal{H}} \tag{27.4}$$

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.

Chapter 28

Symplectic Topology

28.1 Symplectic manifolds

Definition 28.1.1 (Symplectic form). Let $\omega \in \Omega^2(M)$ be a differential 2-form. ω is said to be a symplectic form if it satisfies following properties:

- Closed: $d\omega = 0$
- Non-degeneracy: if $\omega(u, v) = 0, \forall u \in TM$ then $v = 0$

Definition 28.1.2 (Symplectic manifold). A manifold M equipped with a symplectic 2-form ω is called a symplectic manifold. This structure is often denoted as a pair (M, ω).

Property 28.1.3. From the antisymmetry (valid for all differential k-forms) and the non-degeneracy of the symplectic form, it follows that M is even dimensional.

Theorem 28.1.4 (Darboux). *Let (M, ω) be a symplectic manifold. For every neighbourhood Ω in T^*M there exists a fibered chart (x^i, y^i) such that*

$$\omega|_\Omega = \sum_i dx^i \wedge dy^i \tag{28.1}$$

The charts in Darboux's theorem are called **Darboux charts** and they form a cover for M.

28.2 Lagrangian submanifolds

Definition 28.2.1 (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:

$$T_p^\perp S = \{v \in T_pM : \omega(v, \iota_* w) = 0, \forall w \in T_pS\} \tag{28.2}$$

Definition 28.2.2 (Isotropic submanifold). Let (M, ω) be a symplectic manifold. An embedded submanifold $\iota : S \hookrightarrow M$ is called isotropic if $T_pS \subset T_p^{\perp}S$.

Definition 28.2.3 (Isotropic submanifold). Let (M, ω) be a symplectic manifold. An embedded submanifold $\iota : S \hookrightarrow M$ is called co-isotropic if $T_p^{\perp}S \subset T_pS$.

Definition 28.2.4 (Larangian submanifold). Let (M, ω) be a symplectic manifold. An embedded submanifold $\iota : S \hookrightarrow M$ is called Lagrangian if $T_pS = T_p^{\perp}S$. Therefore they are sometimes called maximal isotropic submanifolds.

Part VI

Probability Theory & Statistics

Chapter 29

Probability

29.1 Probability

Definition 29.1.1 (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 29.1.2. 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 9.1.1.

Definition 29.1.3 (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 29.1.4 (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 29.1.5 (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 2.1.1) of all events. Intuitively this seems to mean that some possible outcomes are not treated as events. However

[1] See definition 9.1.2.

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 9.

Formula 29.1.6. Let A, B be two events.

$$P(A \cup B) = P(A) + P(B) + P(A \cap B) \tag{29.1}$$

Definition 29.1.7 (Disjoint events). Two events A and B are said to be disjoint if they cannot happen at the same time:

$$P(A \cap B) = 0 \tag{29.2}$$

Corollary 29.1.8. If A and B are disjoint, the probability that both A and B are true is just the sum of their individual propabilities.

Formula 29.1.9 (Complement). Let A be an event. The probability of A being false is denoted as $P\left(\overline{A}\right)$ and is given by:

$$P\left(\overline{A}\right) = 1 - P(A) \tag{29.3}$$

Corollary 29.1.10. From the previous equation and de Morgan's laws (equations 2.5 and 2.6) we derive the following formula[2]:

$$P\left(\overline{A} \cap \overline{B}\right) = 1 - P(A \cup B) \tag{29.4}$$

29.2 Conditional probability

Definition 29.2.1 (Conditional probability). Let A, B be two events. The probability of A given that B is true is denoted as $P(A|B)$.

$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)}} \tag{29.5}$$

Corollary 29.2.2. 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:

$$P(A|B)P(B) = P(B|A)P(A) \tag{29.6}$$

Theorem 29.2.3 (Bayes' theorem). *Let A, B be two events. From the conditional probability 29.5 it is possible to derive following important theorem:*

$$\boxed{P(A|B) = \frac{P(B|A)P(A)}{P(B)}} \tag{29.7}$$

[2] Switching the union and intersection has no impact on the validity of the formula.

Definition 29.2.4 (Independent events). Let A, B be two events. A and B are said to be independent if they satisfy the following relation:

$$P(A \cap B) = P(A)P(B) \tag{29.8}$$

Corollary 29.2.5. If A and B are two independent events, then equation 29.7 simplifies to:

$$P(A|B) = P(A) \tag{29.9}$$

Property 29.2.6. 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 29.2.7. 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:

$$P(F_{i_1} \cap ... \cap F_{i_k}) = P(F_{i_1})...P(F_{i_k}) \tag{29.10}$$

Formula 29.2.8. 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:

$$P(A) = \sum_{i=1}^{+\infty} P(A|B_i)P(B_i) \tag{29.11}$$

29.3 Random variables

Definition 29.3.1 (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 29.3.2 (σ-algebra generated by 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:

$$X^{-1}(\mathcal{B}) = \{S \in \Sigma : S = X^{-1}(B \in \mathcal{B})\} \tag{29.12}$$

This measure is called the probability distribution of X.

Notation 29.3.3. The σ-algebra generated by the random variable X is often denoted by $\mathcal{F}_X$, analogous to notation 2.4.7.

Theorem 29.3.4. *Let X, Y be two random variables. X and Y are independent if the σ-algebras generated by them are independent*[4].

[3] $X^{-1}\big([a, +\infty[\big) = \{\omega \in \Omega : X(\omega) \geq a\}$.

[4] See equation 29.10.

29.4 Probability distribution

Definition 29.4.1 (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:

$$\boxed{P_X(B) = P(X^{-1}(B))} \tag{29.13}$$

Formula 29.4.2 (Change of variable). Let X be a random variable defined on a probability space (Ω, Σ, P).

$$\int_\Omega g(X(\omega))dP(\omega) = \int_\mathbb{R} g(x)dP_X(x) \tag{29.14}$$

Definition 29.4.3 (Density). Let f be a non-negative integrable function and recall theorem 9.2.20. 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:

$$\int f dm = 1 \tag{29.15}$$

Definition 29.4.4 (Cumulative distribution function). Let f be a density. The c.d.f. corresponding to f is given by:

$$F(y) = \int_{-\infty}^y f(x)dx \tag{29.16}$$

Theorem 29.4.5 (Skorokhod's representation theorem). *Let $F : \mathbb{R} \rightarrow [0, 1]$ be a function that satisfies following 3 properties:*

- *$F(x)$ is non-decreasing.*
- *$\lim_{x\rightarrow-\infty} F(x) = 0$ and $\lim_{x\rightarrow+\infty} F(x) = 1$*
- *$F(x)$ is right-continuous: $y \geq y_0, y \rightarrow y_0 \implies F(y) \rightarrow F(y_0)$.*

There exists a random variable $X : [0, 1] \rightarrow \mathbb{R}$ defined on the probability space $([0, 1], \mathcal{B}, m_{[0,1]})$ such that $F = F_X$.

Formula 29.4.6. 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 $g : \mathbb{R}^n \rightarrow \mathbb{R}$ be integrable with respect to P_X.

$$\int_{\mathbb{R}^\ltimes} g(x)dP_X(x) = \int_{\mathbb{R}^n} f_X(x)g(x)dx \tag{29.17}$$

Corollary 29.4.7. Previous formula together with formula 29.14 gives rise to:

$$\int_\Omega g(X)dP = \int_{\mathbb{R}^\ltimes} f_X(x)g(x)dx \tag{29.18}$$

29.5 Moments

29.5.1 Expectation value

Definition 29.5.1 (Expectation value). Let X be random variable defined on a probability space (Ω, Σ, P).

$$\boxed{E(X) = \int_\Omega X dP} \tag{29.19}$$

Notation 29.5.2. Other often used notations are $\langle X \rangle$ and μ.

Definition 29.5.3 (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 29.18 this becomes:

$$E(X^r) = \int x^r f_X(x) dx \tag{29.20}$$

Definition 29.5.4 (Central moment of order r).

$$E((X-\mu)^r) = \int (x-\mu)^r f_X(x) dx \tag{29.21}$$

Definition 29.5.5 (Variance). The central moment of order 2 is called the variance: $V(X) = E((X-\mu)^2)$.

Property 29.5.6. 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 29.5.7. 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 29.5.8 (Moment generating function).

$$M_X(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tX} P(X) dX \tag{29.22}$$

Theorem 29.5.9. *If the above function exists we can derive the following useful result[5] by using the series expansion 6.15:*

$$E[X^n] = \left.\frac{d^n M_X(t)}{dt^n}\right|_{t=0} \tag{29.23}$$

Definition 29.5.10 (Characteristic function). Let X be a random variable. The characteristic function of X is defined as follows:

$$\varphi_X(t) = E(e^{itX}) \tag{29.24}$$

[5] This property is the reason why 29.22 is called the moment generating function.

Property 29.5.11. 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 29.5.12. If $\varphi_X(t)$ is k times continuously differentiable then X has finite k^{th} moment and

$$E(X^k) = \frac{1}{i^k}\frac{d^k}{dt^k}\varphi_X(0) \tag{29.25}$$

Conversely, if X has finite k^{th} moment then $\varphi_X(t)$ is k times continuously differentiable and the above formula holds.

Formula 29.5.13 (Inversion formula). Let X be a random varibale. If the c.d.f. of X is continuous at $a, b \in \mathbb{R}$ then

$$F_X(b) - F_X(a) = \lim_{c\rightarrow+\infty}\frac{1}{2\pi}\int_{-c}^{c}\frac{e^{-ita} - e^{-itb}}{it}\varphi_X(t)dt \tag{29.26}$$

Formula 29.5.14. If $\varphi_X(t)$ is integrable then the c.d.f. is given by:

$$f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-itx}\varphi_X(t)dt \tag{29.27}$$

Remark 29.5.15. From previous formula it is clear that the density function and the characteristic function are Fourier transformed quantities.

29.5.2 Correlation

Theorem 29.5.16. *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 9.36. It follows that independence of random variables implies orthogonality. To generalize this concept, we introduce following notions.

Definition 29.5.17 (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 29.5.18 (Covariance). Let X, Y be two random variables. The covariance of X and Y is defined as follows:

$$\text{cov}(X, Y) = \langle X_c|Y_c\rangle = E\big((X - E(X))(Y - E(Y))\big) \tag{29.28}$$

Some basic math gives:

$$\text{cov}(X, Y) = E(XY) - E(X)E(Y) \tag{29.29}$$

[6]See definition 9.1.33.

Definition 29.5.19 (Correlation). Let X, Y be two random variables. The correlation is defined as the cosine of the angle between X_c and Y_c:

$$\rho_{XY} = \frac{\text{cov}(X,Y)}{||X||_2||Y||_2} \tag{29.30}$$

Corollary 29.5.20. From theorem 29.5.16 it follows that independent random variables are also uncorrelated.

Corollary 29.5.21. Uncorrelated X and Y satisfy the following equality: $E(XY) = E(X)E(Y)$.

Property 29.5.22. Let $(X_i)_{i\in\mathbb{N}}$ be a sequence of independent random variables. Their variances satisfy the following equation:

$$V\left(\sum_{i=1}^{+\infty} X_i\right) = \sum_{i=1}^{+\infty} V(X_i) \tag{29.31}$$

29.5.3 Conditional expectation

Let (Ω, Σ, P) be a probability space. Let the random variable $X \in L^2(\Omega, \Sigma, P)$[7]. Consider the sub-$\sigma$-algebra $\mathcal{G} \subset \Sigma$, The spaces $L^2(\Sigma)$ and $L^2(\mathcal{G})$ are complete (see property 9.4.3). The projection theorem 17.2.22 can thus be applied, i.e. there exists for every X 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:

$$\forall Z \in L^2(\mathcal{G}) : \langle X - Y|Z\rangle = \int_\Omega (X - Y)ZdP = 0 \tag{29.32}$$

And since $\mathbb{1}_G \in L^2(\mathcal{G})$ for every $G \in \mathcal{G}$ we find by applying 9.25:

$$\int_G XdP = \int_G YdP \tag{29.33}$$

This leads us to introducing the following notion of conditional expectations:

Definition 29.5.23 (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 29.33 for every $G \in \mathcal{G}$. This Y is called the conditional expectation of X given $\mathcal{G}$ and it is denoted by $Y = E(X|\mathcal{G})$:

$$\boxed{\int_G E(X|\mathcal{G})dP = \int_G XdP} \tag{29.34}$$

Remark 29.5.24. Allthough 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 9.2.23.

[7]This vector space has the same interpretation as the one in section 9.4.2. The difference is that all sets are elements of Σ instead of $\mathcal{M}$, that the functions are Σ-measurable instead of $\mathcal{M}$-measurable and that the integral is calculated with respect to the measure P instead of the Lebesgue measure m.

29.6 Joint distributions

Definition 29.6.1 (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:

$$P_{(X,Y)}(B) = P((X, Y) \in B) \tag{29.35}$$

Definition 29.6.2 (Joint density). If the probability measure from previous definition can be written as

$$P_{(X,Y)}(B) = \int_B f_{(X,Y)}(x, y) dm_2(x, y) \tag{29.36}$$

for some integrable $f_{(X,Y)}$ it is said that X and Y have a joint density.

Definition 29.6.3 (Marginal distribution). The distributions of one-dimensional random variables is determined by the joint distribution:

$$P_X(A) = P_{(X,Y)}(A \times \mathbb{R}) \tag{29.37}$$

$$P_Y(A) = P_{(X,Y)}(\mathbb{R} \times A) \tag{29.38}$$

where $A \subset \mathcal{B}$.

Corollary 29.6.4. If the joint density exists then the marginal distributions are absolutely continuous and given by

$$f_X(x) = \int_{\mathbb{R}} f_{(X,Y)}(x, y) dy \tag{29.39}$$

$$f_Y(y) = \int_{\mathbb{R}} f_{(X,Y)}(x, y) dx \tag{29.40}$$

The converse however is not always true. The one-dimensional densities can be absolutely continuous without the existence of the joint density.

29.6.1 Independence

Theorem 29.6.5. *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, i.e.:*

$$P_{(X,Y)} = P_X \times P_Y$$

Remark 29.6.6. If X and Y are absolutely continuous then the previous theorem also applies with the densities instead of the distributions.

29.6.2 Conditional probability

Formula 29.6.7 (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:

$$h(y|X \in A) = \frac{\displaystyle\int_A f_{(X,Y)}(x,y)dx}{\displaystyle\int_A f_X(x)dx} \tag{29.41}$$

For $X = \{a\}$ this equation fails as the denominator would become 0. However it is possible to avoid this problem by formally putting

$$h(y|A = a) = \frac{f_{(X,Y)}(a,y)}{f_X(a)} \tag{29.42}$$

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 $X = a$:

$$P(Y \in B|X = a) = \int_B h(y|X = a)dy \tag{29.43}$$

Formula 29.6.8 (Conditional expectation).

$$E(Y|X)(\omega) = \int_{\mathbb{R}} yh(y|X(\omega))dy \tag{29.44}$$

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 29.34, holds:

$$\int_A E(Y|X)dP = \int_A YdP \tag{29.45}$$

Remark 29.6.9. 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 29.6.10. As mentioned above, applying Fubini's theorem gives:

$$\boxed{E(E(Y|X)) = E(Y)} \tag{29.46}$$

Chapter 30

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 29.

30.1 Data samples

30.1.1 Moment

Formula 30.1.1 (r^{th} sample moment).

$$\overline{x^r} = \frac{1}{N}\sum_{i=1}^{N} x_i^r \tag{30.1}$$

Formula 30.1.2 (r^{th} central sample moment).

$$m_r = \frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})^r \tag{30.2}$$

30.1.2 Mean

Definition 30.1.3 (Arithmetic mean). The arithmetic mean is used to average out differences between measurements. It is equal to the 1^{st} sample moment:

$$\boxed{\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i} \tag{30.3}$$

Formula 30.1.4 (Weighted mean). Let $f : \mathbb{R} \to \mathbb{R}^+$ be a weight function. The weighted mean is given by:

$$\overline{x} = \frac{\sum_i f(x_i)x_i}{\sum_i f(x_i)} \tag{30.4}$$

Corollary 30.1.5. 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:

$$\overline{x} = \frac{1}{N}\sum_{i=1} n_i x_i \tag{30.5}$$

Remark 30.1.6. 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 30.1.7. It is also important to notice that $\overline{f}(x) \neq f(\overline{x})$. The equality only holds for linear functions.

Definition 30.1.8 (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.

$$g = \left(\prod_{i=1}^{N} x_i\right)^{1/N} \tag{30.6}$$

The following relation exists between the arithmetic and geometic mean:

$$\ln g = \overline{\ln x} \tag{30.7}$$

Definition 30.1.9 (Harmonic mean).

$$h = \left(\frac{1}{N}\sum_{i=1}^{N} x_i^{-1}\right)^{-1} \tag{30.8}$$

The following relation exists between the arithmetic and harmonic mean:

$$\frac{1}{h} = \overline{x^{-1}} \tag{30.9}$$

Property 30.1.10. Let $\{x_i\}$ be a positive data set.

$$h \leq g \leq \overline{x} \tag{30.10}$$

where the equalities only hold when all x_i are equal.

Definition 30.1.11 (Mode). The most occurring value in a dataset.

Definition 30.1.12 (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.

[1] A negative data set is also allowed. The real condition is that all values should have the same sign.

30.1.3 Dispersion

Definition 30.1.13 (Range). The simplest indicator for statistical dispersion. It is however very sensitive for extreme values.

$$R = x_{max} - x_{min} \tag{30.11}$$

Definition 30.1.14 (Mean absolute difference).

$$\mathrm{MD} = \frac{1}{N}\sum_{i=1}^{N}|x_i - \overline{x}| \tag{30.12}$$

Definition 30.1.15 (Sample variance).

$$V(X) = \frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})^2 \tag{30.13}$$

Formula 30.1.16. The variance can also be written in the following way:

$$\boxed{V(X) = \overline{x^2} - \overline{x}^2} \tag{30.14}$$

Remark 30.1.17. A better estimator for the variance of a sample is the following formula:

$$\hat{s} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2 \tag{30.15}$$

Equation 30.13 gives a good estimation when the sample mean $\overline{x}$ is replaced by the "true" mean μ. Otherwise one should use the estimator 30.15.

Definition 30.1.18 (Standard deviation).

$$\sigma(X) = \sqrt{V(x)} \tag{30.16}$$

Definition 30.1.19 (Skewness). The skewness γ describes the asymmetry of a distribution. It is defined in relation to the third central moment m_3;

$$m_3 = \gamma\sigma^3 \tag{30.17}$$

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 30.1.20 (Pearson's mode skewness).

$$\gamma_P = \frac{\overline{x} - \text{mode}}{\sigma} \tag{30.18}$$

Definition 30.1.21 (Kurtosis). The kurtosis c is an indicator for the "tailedness". It is defined in relation to the fourth central moment m_4:

$$m_4 = c\sigma^4 \tag{30.19}$$

Definition 30.1.22 (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 30.1.23 (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 30.1.24 (Interquartile range). The interquartile range is the difference between the upper and lower quartile (75^{th} and 25^{th} percentile respectively).

Definition 30.1.25 (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 30.1.26. For Gaussian distributions the following relation exists between the FWHM and the standard deviation σ:

$$\text{FWHM} = 2.35\sigma \tag{30.20}$$

30.1.4 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 30.1.27 (Covariance). Let X, Y be two random variables. The covariance of X and Y is defined as follows:

$$\text{cov}(X, Y) = \frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})(y_i - \overline{y}) = \overline{xy} - \overline{x}\ \overline{y} \tag{30.21}$$

The covariance of X and Y is often denoted by σ_{XY}.

Formula 30.1.28. The covariance and standard deviation are related by the following equality:

$$\sigma_X^2 = \sigma_{XX} \tag{30.22}$$

Definition 30.1.29 (Correlation coefficent).

$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} \tag{30.23}$$

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 30.1.30. For multivariate distributions the above definitions can be generalized using matrices:

$$V_{ij} = \text{cov}(x_{(i)}, x_{(j)}) \tag{30.24}$$

$$\rho_{ij} = \rho_{(i)(j)} \tag{30.25}$$

where $\text{cov}(x_{(i)}, x_{(j)})$ and $\rho_{(i)(j)}$ are defined using equations 30.21 and 30.23. The following general equality exists:

$$V_{ij} = \rho_{ij}\sigma_i\sigma_j \tag{30.26}$$

30.2 Law of large numbers

Theorem 30.2.1 (Law of large numbers). *If the size N of a sample tends towards infinity, then the observed frequencies tend towards the theoretical propabilities.*

Corollary 30.2.2 (Frequentist probability[2]).

$$P(X) = \lim_{n\rightarrow\infty} \frac{f_n(X)}{n} \tag{30.27}$$

30.3 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.

Definition 30.3.1 (probability density functions p.d.f). Let X be a random variable and $P(X)$ the associated probability distribution. The p.d.f. $f(X)$ is defined as follows:

$$\boxed{P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f(X)dX} \tag{30.28}$$

An alternative definition[3] is the following:

$$f(X) = \lim_{\delta x\rightarrow 0} \frac{P(x \leq X \leq x + \delta x)}{\delta x} \tag{30.29}$$

Definition 30.3.2 (Cumulative distribution function c.d.f.). Let X be a random variable and $f(X)$ the associated p.d.f. The cumulative distribution function $F(X)$ is defined as follows:

$$F(x) = \int_{-\infty}^{x} f(X)dX \tag{30.30}$$

[2] Also called the **empirical probability**.

[3] A more formal definition uses *measure theory* and the *Radon-Nikodym derivative*.

Theorem 30.3.3. *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:*

$$P(x_1 \leq X \leq x_2) = F(x_2) - F(x_1) \tag{30.31}$$

Theorem 30.3.4. *$F(X)$ is continuous if and only if $P_X(\{y\}) = 0$ for every $y \in \mathbb{R}$.*

Remark 30.3.5 (Normalization).

$$F(\infty) = 1 \tag{30.32}$$

Formula 30.3.6. The p^{th} percentile c_p can be computed as follows[4]:

$$c_p = F^{-1}(p) \tag{30.33}$$

Definition 30.3.7 (Parametric family). A parametric family of probability densities $f(X; \vec{\theta})$ is a set of densities described by one or more parameters $\vec{\theta}$.

30.3.1 Function of a random variable

Formula 30.3.8. 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:

$$\boxed{g(a) = f(x(a)) \left|\frac{dx}{da}\right|} \tag{30.34}$$

30.3.2 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.

Definition 30.3.9 (Joint density). Let X, Y be two random variables. The joint p.d.f. $f_{XY}(x, y)$ is defined as follows:

$$f_{XY}(x, y)dxdy = \begin{cases} f_x(x \in [x, x + dx]) \\ f_y(y \in [y, y + dy]) \end{cases} \tag{30.35}$$

Remark 30.3.10. As f_{XY} is a probability density, the normalization condition 30.32 should be fulfilled.

Definition 30.3.11 (Conditional density). The conditional p.d.f. of X when Y has the value y is given by the following formula:

$$g(x|y) = \frac{f_{XY}(x, y)}{f_Y(y)} \tag{30.36}$$

where we should pay attention to the remark made when we defined 29.42.

[4]This is clear from the definition of a percentile, as this implies that $F(c_p) = p$.

Corollary 30.3.12. If X and Y are independent, then by remark 29.6.6 the marginal p.d.f is equal to the conditional p.d.f.

Theorem 30.3.13 (Bayes' theorem). *The conditional p.d.f. can be computed without prior knowledge of the joint p.d.f:*

$$\boxed{g(x|y) = \frac{h(y|x)f_X(x)}{f_Y(y)}} \tag{30.37}$$

Remark. This theorem is the statistical (random variable) analogon of theorem 29.7.

Formula 30.3.14. Let $Z = XY$ with X, Y two independent random variables. The distribution $f(z)$ is given by

$$f(z) = \int_{-\infty}^{+\infty} g(x)h(z/x)\frac{dx}{|x|} = \int_{-\infty}^{+\infty} g(z/y)h(y)\frac{dy}{|y|} \tag{30.38}$$

Corollary 30.3.15. Taking the Mellin transform 10.17 of both the positive and negative part of the above integrand (to be able to handle the absolute value) gives following relation

$$\mathcal{M}\{f\} = \mathcal{M}\{g\}\mathcal{M}\{h\} \tag{30.39}$$

Formula 30.3.16. Let $Z = X + Y$ with X, Y two independent random variables. The distribution $f(z)$ is given by the convolution of $g(x)$ and $h(y)$:

$$f(z) = \int_{-\infty}^{+\infty} g(x)h(z-x)dx = \int_{-\infty}^{+\infty} g(z-y)h(y)dy \tag{30.40}$$

30.3.3 Important distributions

Formula 30.3.17 (Uniform distribution).

$$\boxed{P(x; a, b) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{elsewhere} \end{cases}} \tag{30.41}$$

$$E(x) = \frac{a+b}{2} \tag{30.42}$$

$$V(x) = \frac{(b-a)^2}{12} \tag{30.43}$$

Formula 30.3.18 (Normal distribution). Also called the Gaussian distribution.

$$\boxed{\mathcal{G}(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}} \tag{30.44}$$

Formula 30.3.19 (Standard normal distribution).

$$\boxed{\mathcal{N}(z) = \frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}} \tag{30.45}$$

Remark 30.3.20. Every Gaussian distribution can be rewritten as a standard normal distribution by setting $Z = \frac{X-\mu}{\sigma}$.

Remark 30.3.21. The c.d.f. of the standard normal distribution is given by the error function: $F(z) = \mathrm{Erf}(z)$.

Formula 30.3.22 (Exponential distribution).

$$\boxed{P(x;\tau) = \frac{1}{\tau}e^{-\frac{x}{\tau}}} \tag{30.46}$$

$$E(x) = \tau \tag{30.47}$$

$$V(x) = \tau^2 \tag{30.48}$$

Theorem 30.3.23. *The exponential distribution is memoryless:*

$$P(X > x_1 + x_2 | X > x_2) = P(X > x_1) \tag{30.49}$$

Formula 30.3.24 (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:

$$\boxed{P(x;p) = p^x(1-p)^{1-x}} \tag{30.50}$$

$$E(x) = p \tag{30.51}$$

$$V(x) = p(1-p) \tag{30.52}$$

Formula 30.3.25 (Binomial distribution). A process with n identical independent trials, all Bernoulli processes $P(x;p)$, is described by a binomial distribution:

$$\boxed{P(r;p,n) = p^r(1-p)^{n-r}\frac{n!}{r!(n-r)!}} \tag{30.53}$$

$$E(r) = np \tag{30.54}$$

$$V(r) = np(1-p) \tag{30.55}$$

Formula 30.3.26 (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.

$$\boxed{P(r;\lambda) = \frac{e^{-\lambda}\lambda^r}{r!}} \tag{30.56}$$

$$E(r) = \lambda \tag{30.57}$$

$$V(r) = \lambda \tag{30.58}$$

Theorem 30.3.27. *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 30.3.28. The number of events coming from A is given by a binomial distribution $P(r_a;\Lambda_a,r)$ where $\Lambda_a = \dfrac{\lambda_a}{\lambda_a+\lambda_b}$.

Remark 30.3.29. 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 30.3.30 (χ^2 distribution). The sum of k squared independent (standard) normally distributed random variables Y_i defines the random variable:

$$\chi_k^2 = \sum_{i=1}^k Y_i^2 \tag{30.59}$$

where k is said to be the number of **degrees of freedom**.

$$\boxed{P(\chi^2;n) = \frac{\chi^{n-2}e^{-\frac{\chi^2}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}} \tag{30.60}$$

Remark 30.3.31. 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 30.3.32 (Student-t distribution). The student-t distribution describes a sample with estimated standard deviation $\hat{\sigma}$.

$$\boxed{P(t;n) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)\left(1+\frac{t^2}{n}\right)^{\frac{n+1}{2}}}} \tag{30.61}$$

where

$$t = \frac{(x-\mu)/\sigma}{\hat{\sigma}/\sigma} = \frac{z}{\sqrt{\chi^2/n}} \tag{30.62}$$

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 30.3.33 (Cauchy distribution[5]). The general form $f(x; x_0, \gamma)$ is given by:

$$f(x; x_0, \gamma) = \frac{1}{\pi} \frac{\gamma}{(x - x_0)^2 + \gamma^2} \tag{30.63}$$

The characteristic function 29.24 is given by:

$$E\left(e^{itx}\right) = e^{ix_0 t - \gamma|t|} \tag{30.64}$$

Property 30.3.34. Both the mean and variance of the Cauchy distribution are undefined.

30.4 Central limit theorem (CLT)

Theorem 30.4.1 (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 30.4.2. If the random variables are not independent, property 2 will not be fulfilled.

Remark 30.4.3. The sum of Gaussians will be Gaussian to.

30.4.1 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 30.4.4.

$$< \overline{x} >= \mu \tag{30.65}$$

Property 30.4.5.

$$V(\overline{x}) = \frac{\sigma^2}{N} \tag{30.66}$$

[5] Also known (especially in particle physics) as the **Breit-Wigner** distribution.

30.5 Errors

30.5.1 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 30.3 in a meaningful way because the measurements are not of the same type. Therefore it is also impossible to apply the CLT 30.4.1.

Definition 30.5.1 (Weighted mean). The appropriate alternative is the *weighted mean*:

$$\boxed{\overline{x} = \frac{\sum_i \frac{x_i}{\sigma_i^2}}{\sum_i \frac{1}{\sigma_i^2}}} \tag{30.67}$$

The resolution of the weighted mean is given by:

$$V(\overline{x}) = \frac{1}{\sum_i \sigma_i^{-2}} \tag{30.68}$$

30.5.2 Propagation of errors

Formula 30.5.2. Let X be random variable with variance $V(x)$. The variance of a linear function $f(X) = aX + b$ is given by:

$$V(f) = a^2 V(x) \tag{30.69}$$

Formula 30.5.3. Let X be random variable with **small** (!!) variance $V(x)$. The variance of a general function $f(X)$ is given by:

$$V(f) \approx \left(\frac{df}{dx}\right)^2 V(x) \tag{30.70}$$

Corollary 30.5.4. The correlation coefficient ρ (30.23) of a random variable X and a **linear** function of X is independent of σ_x and is always equal to ± 1.

Formula 30.5.5 (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:

$$\boxed{V(f) = \sum_p \left(\frac{\partial f}{\partial X_{(p)}}\right)^2 V(X_{(p)}) + \sum_p \sum_{q\neq p} \left(\frac{\partial f}{\partial X_{(p)}}\right)\left(\frac{\partial f}{\partial X_{(q)}}\right) \mathrm{cov}(X_{(p)}, X_{(q)})} \tag{30.71}$$

Definition 30.5.6 (Fractional error). Let X, Y be two **independent** random variables. The standard deviation of $f(x, y) = xy$ is given by the fractional error:

$$\left(\frac{\sigma_f}{f}\right)^2 = \left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2 \tag{30.72}$$

Remark 30.5.7. The fractional error of quantity is equal to the fractional error of the reciprocal of that quantity.

Property 30.5.8. Let X be a random variable. The error of the logarithm of X is equal to the fractional error of X.

Definition 30.5.9 (Covariance of functions).

$$\text{cov}(f_1, f_2) = \sum_p \sum_q \left(\frac{\partial f_1}{\partial X_{(p)}}\right) \left(\frac{\partial f_2}{\partial X_{(q)}}\right) \text{cov}(X_{(p)}, X_{(q)}) \tag{30.73}$$

Corollary 30.5.10. Let $\vec{f} = \{f_1, ..., f_k\}$. The covariance matrix V_f of the k functions is given by:

$$V_f = G\ V_X\ G^T \tag{30.74}$$

where G is the Jacobian matrix of $\vec{f}$.

30.5.3 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.

30.6 Estimators

Definition 30.6.1 (Estimator). An estimator is a procedure that, given a sample, produces a numerical value for a property of the parent population.

30.6.1 General properties

Property 30.6.2 (Consistency).

$$\boxed{\lim_{N\rightarrow\infty} \hat{a} = a} \tag{30.75}$$

Property 30.6.3 (Unbiased estimator).

$$\boxed{< \hat{a} >= a} \tag{30.76}$$

Definition 30.6.4 (Bias).

$$B(\hat{a}) = |< \hat{a} > -a| \tag{30.77}$$

Property 30.6.5 (Efficiency). An estimator $\hat{a}$ is said to be efficient if its variance $V(\hat{a})$ is equal to the minimum variance bound 30.86.

Definition 30.6.6 (Mean squared error).

$$\Upsilon(\hat{a}) = B(\hat{a})^2 + V(\hat{a}) \tag{30.78}$$

Remark 30.6.7. If an estimator is unbiased, the MSE is equal to the variance of the estimator.

30.6.2 Fundamental estimators

Property 30.6.8 (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 30.66.

Property 30.6.9 (Variance estimator for known mean). If the true mean μ is known then a consistent and unbiased estimator for the variance is given by:

$$\widehat{V(x)} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2 \tag{30.79}$$

Property 30.6.10 (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]:

$$s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2 \tag{30.80}$$

30.6.3 Estimation error

Formula 30.6.11 (Variance of estimator of variance).

$$V(\widehat{V(x)}) = \frac{(N-1)^2}{N^3} < (x- < x >)^4 > - \frac{(N-1)(N-3)}{N^3} < (x- < x >)^2 >^2 \tag{30.81}$$

Formula 30.6.12 (Variance of estimator of standard deviation).

$$V(\hat{\sigma}) = \frac{1}{4\sigma^2}V(\widehat{V(x)}) \tag{30.82}$$

[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 30.13. The consistency however is guaranteed by the CLT.

Remark 30.6.13. 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.

30.6.4 Likelihood function

Definition 30.6.14 (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 $P(X;a)$:

$$\mathcal{L}(a;\vec{x}) = \prod_{i=1}^{N} P(x_i;a) \tag{30.83}$$

Definition 30.6.15 (Log-likelihood).

$$\log\mathcal{L}(a;\vec{x}) = \sum_i \ln P(x_i;a) \tag{30.84}$$

Property 30.6.16. The expectation value of an estimator ˆ is given by:

$$< \hat{a} >= \int \hat{a}\mathcal{L}(\hat{a};X)dX \tag{30.85}$$

Theorem 30.6.17 (Minimum variance bound). *The variance of an* ***unbiased*** *estimator has a lower bound: the minimum variance bound*[7] *(MVB).*

$$\boxed{V(\hat{a}) \geq \frac{1}{\left\langle\left(\frac{d\ln\mathcal{L}}{da}\right)^2\right\rangle}} \tag{30.86}$$

For a biased estimator with bias b the MVB takes on the following form:

$$V(\hat{a}) \geq \frac{\left(1+\frac{db}{da}\right)^2}{\left\langle\left(\frac{d\ln\mathcal{L}}{da}\right)^2\right\rangle} \tag{30.87}$$

Remark 30.6.18.

$$\left\langle\left(\frac{d\ln\mathcal{L}}{da}\right)^2\right\rangle = -\left\langle\frac{d^2\ln\mathcal{L}}{da^2}\right\rangle \tag{30.88}$$

Definition 30.6.19 (Fisher information).

$$I_X(a) = \left\langle\left(\frac{d\ln\mathcal{L}}{da}\right)^2\right\rangle = N\int\left(\frac{d\ln P}{da}\right)^2 PdX \tag{30.89}$$

[7] It is also known as the Cramer-Rao bound.

30.6.5 Maximum likelihood estimator

Following from definition 30.6.14 it follows that the best estimator $\hat{a}$ is the value for which the likelihood function is maximal. It is the value that makes the measurements the most propable, but it is therefore not the most propable estimator.

Method 30.6.20 (Maximum likelihood estimator). The maximum likelihood estimator $\hat{a}$ is obtained by solving following equation:

$$\left.\frac{d\ln\mathcal{L}}{da}\right|_{a=\hat{a}} = 0 \tag{30.90}$$

Remark 30.6.21. MLH estimators are mostly consistent but often biased.

Property 30.6.22. MLH estimators are invariant under parameter transformations.

Corollary 30.6.23. The invariance implies that the two estimators $\hat{a}$ and $\widehat{f(a)}$ cannot both be unbiased at the same time.

Property 30.6.24. Asymptotically ($N \to \infty$) every **consistent** estimator becomes unbiased and efficient.

30.6.6 Least squares

Method 30.6.25 (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: $\chi^2 = \sum_i \left[\frac{d_i}{\sigma_i}\right]^2$
4. Find the most propably value of $\hat{a}$ by solving the equation $\frac{d\chi^2}{da} = 0$.

Property 30.6.26. The optimalized (minimized) χ^2 is distributed according to a χ^2 dsitribution 30.60 $P(\chi^2; n)$. The number of degrees of freedom n is equal to the number of events N minus the number of fitted parameters k.

Formula 30.6.27 (Linear fit). When all uncertainties σ_i are equal, the slope $\hat{m}$ and intercept $\hat{c}$ are given by following formulas:

$$\hat{m} = \frac{\overline{xy} - \overline{x}\,\overline{y}}{\overline{x^2} - \overline{x}^2} = \frac{\text{cov}(x,y)}{V(x)} \tag{30.91}$$

$$\hat{c} = \overline{y} - \hat{m}\overline{x} = \frac{\overline{x^2} - \overline{x}\,\overline{y}}{\overline{x^2} - \overline{x}^2} \tag{30.92}$$

Remark 30.6.28. 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 30.6.29 (Errors of linear fit).

$$V(\hat{m}) = \frac{1}{N(\overline{x^2} - \overline{x}^2)}\sigma^2 \tag{30.93}$$

$$V(\hat{c}) = \frac{\overline{x^2}}{N(\overline{x^2} - \overline{x}^2)}\sigma^2 \tag{30.94}$$

$$\mathrm{cov}(\hat{m}, \hat{c}) = \frac{-\overline{x}}{N(\overline{x^2} - \overline{x}^2)}\sigma^2 \tag{30.95}$$

Remark 30.6.30. When there are different uncertainties σ_i, the arithmetic means have to be replaced with weighted means, but the expressions remain the same. The quantity σ^2 has to be replaced by its weighted variant:

$$\overline{\sigma^2} = \frac{\sum \sigma_i^2/\sigma_i^2}{\sum \sigma_i^{-2}} = \frac{N}{\sum \sigma_i^{-2}}$$

30.6.7 Binned least squares

The least squares method is very useful to fit data which has been grouped in bins (histograms).

Method 30.6.31 (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: $\chi^2 = \sum_i^{N_B} \frac{(n_i - f_i)^2}{f_i^2}$

30.7 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 30.7.1. 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.

30.7.1 Interval types

Definition 30.7.2 (Two-sided confidence interval).

$$\boxed{P(x_-\leq X\leq x_+)=\int_{x_-}^{x_+}P(x)dx=C} \tag{30.96}$$

There are three possible (often used) two-sided intervals:

- Symmetric interval: $x_+-\mu=\mu-x_-$
- Shortest interval: $|x_+-x_-|$ is minimal
- Central interval: $\int_{-\infty}^{x_-}P(x)dx=\int_{x_+}^{\infty}P(x)dx=\frac{1-C}{2}$

Remark 30.7.3. For Gaussian distributions these three definitions are equivalent.

Remark. The central interval is the (best and) most widely used confidence interval.

Definition 30.7.4 (One-sided confidence interval).

$$\boxed{P(x\geq x_-)=\int_{x_-}^{+\infty}P(x)dx=C} \tag{30.97}$$

$$\boxed{P(x\leq x_+)=\int_{-\infty}^{x_+}P(x)dx=C} \tag{30.98}$$

Remark 30.7.5. 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 30.7.6 (Discrete central confidence interval).

$$x_-=\max_\theta\left[\sum_{x=0}^{\theta-1}P(x;X)\right]\leq\frac{1-C}{2} \tag{30.99}$$

$$x_+=\min_\theta\left[\sum_{x=\theta+1}^{+\infty}P(x;X)\right]\leq\frac{1-C}{2} \tag{30.100}$$

30.7.2 General construction

For every value of the true parameter X it is possible to construct a confidence interval. This leads to the construction of the 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.

Theorem 30.7.7. *Let x_0 be the measured value of a parameter X. From the confidence region, it is possible to infere a confidence interval $[X_-(x); X_+(x)]$. 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 it is 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.*

30.7.3 Extra conditions

Method 30.7.8 (Bayesian statistics).

$$p(\text{theory}|\text{result}) = p(\text{result}|\text{theory})\frac{p(\text{theory})}{p(\text{result})} \tag{30.101}$$

or more mathematically:

$$\boxed{p(X|x) = p(x|X)\frac{p(X)}{p(x)}} \tag{30.102}$$

- The denominator $p(\text{result})$ does not play a real role, it is a normalization constant.
- The probability $p(x|X)$ to have measurement x when the true parameter is X is a Gaussian distribution $\mathcal{G}(x; X, \sigma)$

Remark 30.7.9. If nothing is known about the theory, $p(X)$ is (exaggerated assumption) a uniform probability 30.41.

30.7.4 Interval for a sample mean

Formula 30.7.10 (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 $Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{N}}$ is given by:

$$\left[\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{N}}; \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{N}}\right] \tag{30.103}$$

Remark 30.7.11. If the sample size is not sufficiently large, the measured quantity must follow a normal distribution.

Formula 30.7.12 (Interval with unknown variance). To account for the uncertainty of the estimated standard deviation $\hat{\sigma}$, the student-t distribution 30.61 is used instead of a Gaussian distribution to describe the sample mean $\overline{X}$. The α-level confidence interval is given by:

$$\left[\overline{X} - t_{\alpha/2;(n-1)}\frac{s}{\sqrt{N}}; \overline{X} + t_{\alpha/2;(n-1)}\frac{s}{\sqrt{N}}\right] \tag{30.104}$$

where s is the estimated standard deviation 30.80.

Formula 30.7.13 (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:

$$\left[\frac{(2N\hat{P} + z_{\alpha/2}^2) - z_{\alpha/2}\sqrt{z_{\alpha/2}^2 + 4N\hat{P}(1-\hat{P})}}{2(N + z_{\alpha/2}^2)}; \frac{(2N\hat{P} + z_{\alpha/2}^2) + z_{\alpha/2}\sqrt{z_{\alpha/2}^2 + 4N\hat{P}(1-\hat{P})}}{2(N + z_{\alpha/2}^2)}\right] \tag{30.105}$$

Remark. The expectation value and variance are these of a binomial distribution 30.53 with $r = X/N$.

30.7.5 Confidence region

30.8 Hypotheses and testing

30.8.1 Hypothesis

Definition 30.8.1 (Simple hypothesis). A hypothesis is called simple if the distribution is fully specified.

Definition 30.8.2 (Composite hypothesis). A hypothesis is called composite if the distribution is given relative to some parameter values.

Definition 30.8.3 (Null hypothesis H_0).

Definition 30.8.4 (Alternative hypothesis H_1).

30.8.2 Testing

Definition 30.8.5 (Type I error). Rejecting a true null hypothesis.

Definition 30.8.6 (Type II error). Accepting/retaining a false null hypothesis.

Definition 30.8.7 (Significance). The probability of making a type I error:

$$\boxed{\alpha = \int P_{H_0}(x)dx} \tag{30.106}$$

Property 30.8.8. Let $a_1 > a_2$. An a_2-level test is also significant at the a_1-level.

Remark 30.8.9. 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 30.8.10 (Power). The probability of not making a type II error:

$$\boxed{\beta = \int P_{H_1}(x)dx \qquad \rightarrow \qquad \text{power: } 1-\beta} \tag{30.107}$$

Theorem 30.8.11. *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 30.8.12 (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:

$$\Lambda(x) = \frac{L(x|H_0)}{L(x|H_1)} \leq \eta \tag{30.108}$$

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.

30.8.3 Confindence intervals and decisions

30.9 Goodness of fit

Let $f(x|\vec{\theta})$ be the fitted function with N measurements.

30.9.1 χ^2-test

Formula 30.9.1.

$$\chi^2 = \sum_{i=1}^{N} \frac{[y_i - f(x_i)]^2}{\sigma_i^2} \tag{30.109}$$

Property 30.9.2. If there are $N - n$ fitted parameters we have:

$$\int_{\chi^2}^{\infty} f(\chi^2|n)d\chi^2 \approx 1 \implies \begin{cases} \circ \text{ good fit} \\ \circ \text{ errors were overestimated} \\ \circ \text{ selected measurements} \\ \circ \text{ lucky shot} \end{cases} \tag{30.110}$$

Property 30.9.3 (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.

30.9.2 Runs test

A good χ^2-test does not mean that the fit is good. As mentioned in property 30.110 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 30.111/30.112.

Remark 30.9.4. 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 30.9.5 (Runs distribution).

$$\boxed{P(r_{even}) = 2\frac{C^{N_B-1}_{\frac{r}{2}-1}C^{N_O-1}_{\frac{r}{2}-1}}{C^N_{N_B}}} \tag{30.111}$$

$$\boxed{P(r_{odd}) = \frac{C^{N_B-1}_{\frac{r-3}{2}}C^{N_O-1}_{\frac{r-1}{2}} + C^{N_O-1}_{\frac{r-3}{2}}C^{N_B-1}_{\frac{r-1}{2}}}{C^N_{N_B}}} \tag{30.112}$$

$$E(r) = 1 + 2\frac{N_B N_O}{N} \tag{30.113}$$

$$V(r) = 2\frac{N_B N_O}{N}\frac{2N_B N_O - 1}{N(N-1)} \tag{30.114}$$

Remark 30.9.6. For $r > 10 - 15$ the runs distribution approximates a Gaussian distribution.

30.9.3 Kolmogorov test

Definition 30.9.7 (Empirical distribution function).

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{1}_{]-\infty,x]}(x_i) \tag{30.115}$$

where $\mathbb{1}_A(x)$ is the indicator function 9.19.

Definition 30.9.8 (Kolmogorov-Smirnov statistic). Let $F(x)$ be a given cumulative distribution function. The n^{th} Kolmogorov-Smirnov statistic is defined as:

$$D_n = \sup_x |F_n(x) - F(x)| \tag{30.116}$$

Definition 30.9.9 (Kolmogorov distribution).

$$P(K \leq x) = 1 - 2\sum_{i=1}^{+\infty}(-1)^{i-1}e^{-2i^2x^2} = \frac{\sqrt{2\pi}}{x}\sum_{i=1}^{+\infty}e^{-(2i-1)^2\pi^2/(8x^2)} \tag{30.117}$$

Property 30.9.10 (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:

$$D_n\sqrt{n} > K_\alpha \tag{30.118}$$

where K_α is defined by using the Kolmogorov distribution: $P(K \leq K_\alpha) = 1 - \alpha$

Chapter 31

Stochastic calculus

Part VII

Classical Mechanics

Chapter 32

Equations of motion

32.1 General quantities

32.1.1 Linear quantities

Formula 32.1.1 (Force).

$$\boxed{\vec{\boldsymbol{F}} = \frac{d\vec{\boldsymbol{p}}}{dt}} \tag{32.1}$$

Remark. In classical mechanics, this formula is given by Newton's second law.

Formula 32.1.2 (Work).

$$W = \int \vec{\boldsymbol{F}} \cdot d\vec{\boldsymbol{l}} \tag{32.2}$$

Definition 32.1.3 (Conservative force). If the work done by a force is independent of the path taken, the force is said to be **conservative**.

$$\oint_C \vec{\boldsymbol{F}} \cdot d\vec{\boldsymbol{l}} = 0 \tag{32.3}$$

Stokes' theorem 16.22 together with relation 16.15 lets us rewrite the conservative force as the gradient of a scalar field:

$$\boxed{\vec{\boldsymbol{F}} = -\nabla V} \tag{32.4}$$

Formula 32.1.4 (Kinetic energy).

$$E_{kin} = \frac{p^2}{2m} \tag{32.5}$$

32.1.2 Angular quantities

Formula 32.1.5 (Angular velocity).

$$\omega = \frac{v}{r} \tag{32.6}$$

Formula 32.1.6 (Angular frequency).

$$\nu = \frac{\omega}{2\pi} \tag{32.7}$$

Formula 32.1.7 (Moment of inertia). For a symmetric object the moment of inertia is given by:

$$I = \int_V r^2 \rho(r) dV \tag{32.8}$$

For a general body we can define the moment of inertia tensor:

$$\boxed{\mathcal{I} = \int_V \rho(\vec{\boldsymbol{r}}) \left(r^2 \mathbb{1} - \vec{\boldsymbol{r}} \otimes \vec{\boldsymbol{r}}\right) dV} \tag{32.9}$$

Definition 32.1.8 (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:

$$[I] = [Q][\Lambda][Q]^T \tag{32.10}$$

The columns of $[Q]$ are called the principal axes of inertia. The eigenvalues are called the **principal moments of inertia**.

Example 32.1.9 (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 32.1.10 (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:*

$$\boxed{I_A = I_{CM} + M||\vec{\boldsymbol{r}}_A - \vec{\boldsymbol{r}}_{CM}||^2} \tag{32.11}$$

where M is the mass of the rotating body.

[1] The matrix associated with the inertia tensor 32.9.

[2] See 15.6.15.

Formula 32.1.11 (Angular momentum).

$$\boxed{\vec{\boldsymbol{L}} = \vec{\boldsymbol{r}} \times \vec{\boldsymbol{p}}} \tag{32.12}$$

Given the angular velocity vector we can compute the angular momentum as follows:

$$\vec{\boldsymbol{L}} = \mathcal{I}(\vec{\boldsymbol{\omega}}) \tag{32.13}$$

where $\mathcal{I}$ is the moment of inertia tensor. If $\vec{\boldsymbol{\omega}}$ is parallel to a principal axis, then the formula reduces to:

$$\vec{\boldsymbol{L}} = I\vec{\boldsymbol{\omega}} \tag{32.14}$$

Formula 32.1.12 (Torque).

$$\vec{\boldsymbol{\tau}} = \frac{d\vec{\boldsymbol{L}}}{dt} \tag{32.15}$$

For constant bodies, this formula can be rewritten as follows:

$$\vec{\boldsymbol{\tau}} = I\vec{\boldsymbol{\alpha}} = \vec{\boldsymbol{r}} \times \vec{\boldsymbol{F}} \tag{32.16}$$

Remark 32.1.13. 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 32.1.14 (Rotational energy).

$$E_{\text{rot}} = \frac{1}{2}I\omega^2 \tag{32.17}$$

32.2 Central force

Definition 32.2.1 (Central force). A central force is a force that only depends on the relative position of two objects:

$$\vec{\boldsymbol{F}}_c \equiv F\left(||\vec{\boldsymbol{r}}_2 - \vec{\boldsymbol{r}}_1||\right)\hat{\boldsymbol{e}}_{\boldsymbol{r}} \tag{32.18}$$

32.3 Kepler problem

Formula 32.3.1 (Potential for a point mass).

$$\boxed{V = -G\frac{M}{r}} \tag{32.19}$$

where $G = 6.67 \times 10^{-11}\frac{Nm^2}{\text{kg}^2}$ is the **gravitational constant**.

32.4 Harmonic oscillator

Formula 32.4.1 (Harmonic potential).

$$\boxed{V = \frac{1}{2}kx^2} \tag{32.20}$$

or

$$V = \frac{1}{2}m\omega^2x^2 \tag{32.21}$$

where we have set $\omega = \sqrt{\frac{k}{m}}$.

Formula 32.4.2 (Solution).

$$\begin{aligned} x(t) &= A\sin\omega t + B\cos\omega t & (32.22)\\ &= Ce^{i\omega t} + De^{-i\omega t} & (32.23) \end{aligned}$$

Chapter 33

Lagrangian and Hamiltonian formalism

Definition 33.0.1 (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 33.0.2 (Generalized velocities). The generalized velocities $\dot{q}_k$ are the derivatives of the generalized coordinates with respect to time.

Notation 33.0.3.

$$L\left(\vec{\boldsymbol{q}}(t), \dot{\vec{\boldsymbol{q}}}(t), t\right) \equiv L\left(q_1(t), ..., q_n(t), \dot{q}_1(t), ..., \dot{q}_n(t), t\right) \tag{33.1}$$

Definition 33.0.4 (Action).

$$\boxed{S = \int_{t_1}^{t_2} L\left(\vec{\boldsymbol{q}}(t), \dot{\vec{\boldsymbol{q}}}(t), t\right) dt} \tag{33.2}$$

33.1 Euler-Lagrange equations†

Formula 33.1.1 (Euler-Lagrange equation of the first kind).

$$\boxed{\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}^k}\right) - \frac{\partial T}{\partial q^k} = Q_k} \tag{33.3}$$

where T is the total kinetic energy.

Formula 33.1.2 (Euler-Lagrange equation of the second kind).

$$\boxed{\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\right) - \frac{\partial L}{\partial q^k} = 0} \tag{33.4}$$

33.2 Conservation laws and symmetry properties

Definition 33.2.1 (Conjugate momentum). Also called the **canonically conjugate momentum**.

$$p_k = \frac{\partial L}{\partial \dot{q}^k} \tag{33.5}$$

Definition 33.2.2 (Cyclic coordinate). If the lagrangian L does not explicitly depend on a coordinate q_k, the coordinate is called a cyclic coordinate.

Property 33.2.3. The conjugate momentum of a cyclic coordinate is a conserved quantity.

$$\dot{p}_k \overset{33.5}{=} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\right) \overset{33.4}{=} \frac{\partial L}{\partial q^k} \overset{\substack{\text{cyclic}\\ \text{coord.}}}{=} 0 \tag{33.6}$$

33.3 Noether's theorem

Theorem 33.3.1 (Noether's theorem†). *Consider a field transformation*

$$\phi(x) \rightarrow \phi(x) + \alpha\delta\phi(x) \tag{33.7}$$

where α is an infinitesimal quantity and $\delta\phi$ is a small deformation. In case of a symmetry we obtain the following conservation law:

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi - \mathcal{J}^\mu\right) = 0 \tag{33.8}$$

The factor between parentheses can be interpreted as a conserved current $j^\mu(x)$. Noether's theorem states that every symmetry of the form 33.7 leads to such a current.

The conservation can also be expressed in terms of a charge[1]*:*

$$\frac{dQ}{dt} = \frac{d}{dt}\int j^0 d^3x = 0 \tag{33.9}$$

Definition 33.3.2 (Stress-energy tensor). Consider a field transformation

$$\phi(x) \rightarrow \phi(x+a) = \phi(x) + a^\mu\partial_\mu\phi(x)$$

[1] The conserved current and its associated charge are called the **Noether current** and **Noether charge**.

Because the Lagrangian is a scalar it transforms similarly:

$$\mathcal{L} \rightarrow \mathcal{L} + a^\mu \partial_\mu \mathcal{L} = \mathcal{L} + a^\nu \partial_\mu (\delta^\mu{}_\nu \mathcal{L}) \tag{33.10}$$

This leads to the existence of 4 conserved currents. These can be used to define the stress-energy tensor:

$$\boxed{T^\mu{}_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial_\nu \phi - \mathcal{L} \delta^\mu{}_\nu} \tag{33.11}$$

33.4 Hamilton's equations

Definition 33.4.1 (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 33.5) and leaving the coordinates q^i and t invariant.

Definition 33.4.2 (Hamiltonian function). The (classical) Hamiltonian function is defined as follows:

$$H(q, p, t) = \sum_i p_i \dot{q}^i - L(q, p, t) \tag{33.12}$$

Formula 33.4.3 (Hamilton's equations[2]).

$$\dot{q}^i = \frac{\partial H}{\partial p_i} \tag{33.13}$$

$$-\dot{p}_i = \frac{\partial H}{\partial q^i} \tag{33.14}$$

$$-\frac{\partial L}{\partial t} = \frac{\partial H}{\partial t} \tag{33.15}$$

The formula to obtain the Hamiltonian from the Lagrangian is an application of the following more general Legendre transformation:

Definition 33.4.4 (Legendre transformation). Consider an equation of the following form:

$$df = udx + vdy \tag{33.16}$$

where $u = \frac{\partial f}{\partial x}$ and $v = \frac{\partial f}{\partial y}$.

Suppose we want to perform a coordinate transformation $(x, y) \rightarrow (u, y)$ while preserving the general form of 33.16 for differential quantities. To do this we consider the function

$$g = f - ux \tag{33.17}$$

[2] Also known as the *canonical equations of Hamilton*.

Differentiating gives

$$\begin{aligned} dg &= df - udx - xdu \\ &= (udx + vdy) - udx - xdu \\ &= vdy - xdu \end{aligned}$$

which has the form of 33.16 as desired. The quantities v and x are now given by

$$x = -\frac{\partial g}{\partial u} \qquad \text{and} \qquad v = \frac{\partial g}{\partial y} \tag{33.18}$$

The transition $f \to g$ defined by equations 33.16 and 33.17 is called a Legendre transformation.

Remark 33.4.5. Although the previous derivation used only 2 coordinates, the definition of Legendre transformations can easily be generalized to more coordinates.

33.5 Hamilton-Jacobi equation

33.5.1 Canonical transformations

Definition 33.5.1 (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:

$$\sum_i \dot{q}^i p_i - H(q,p,t) = \sum_i \dot{Q}^i P_i - K(Q,P,t) - \frac{dG}{dt}(Q,P,t) \tag{33.19}$$

The function G is called the generating function of the canonical transformation. The choice of G uniquely determines the transformation.

Formula 33.5.2 (Hamilton-Jacobi equation). Sufficient conditions for the generating function S are given by:

$$\begin{aligned} P_i &= \frac{\partial S}{\partial Q^i} \\ Q^i &= \frac{\partial S}{\partial P_i} \end{aligned}$$

and

$$K = H + \frac{\partial S}{\partial t}$$

Choosing the new Hamiltonian function K to be 0 gives the Hamilton-Jacobi equation:

$$\boxed{H\left(q, \frac{\partial S}{\partial q}\right) + \frac{\partial S}{\partial t} = 0} \tag{33.20}$$

The function S is called **Hamilton's principal function**.

Property 33.5.3. The new coordinates P_i and Q^i are all constants of motion. This follows immediately from the choice $K = 0$.

Definition 33.5.4 (Hamilton's characteristic function). If the system is time-independent it follows from the HJE that the principal function is of the form

$$S(q, p, t) = W(q, p) - Et \tag{33.21}$$

where E is a constant. The time-independent function W is called Hamilton's characteristic function.

Substituting this result in the HJE results in

$$H\left(q, \frac{\partial S}{\partial q}\right) = E \tag{33.22}$$

In time-independent systems the Hamiltonian function is thus a constant of motion and we call it the **energy** of the system.

33.5.2 Stäckel potential

Remark 33.5.5. If the principal function can be separated into n equations, the HJE splits up into n equations of the form

$$h_i\left(q^i, \frac{dS}{dq^i}, \alpha_i\right) = 0 \tag{33.23}$$

The partial differential equation for S can thus be rewritten as a system of n ordinary differential equations.

Theorem 33.5.6 (Stäckel condition). *Using an orthogonal coordinate system, the Hamilton-Jacobi equation is separable if and only if the potential is of the following form:*

$$V(q) = \sum_{i=1}^{n} \frac{1}{G_i^2(q)} W_i(q^i) \tag{33.24}$$

whenever the Hamiltonian function can be written as

$$H(q, p) = \frac{1}{2} \sum_i \frac{p_i^2}{G_i^2(q)} + V(q) \tag{33.25}$$

*These potentials are called **Stäckel potentials**.*

Chapter 34

Phase space

34.1 Phase space

Definition 34.1.1 (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 34.1.2 (Rotation). A rotation is the change of a coordinate for which every possible value is allowed.

Definition 34.1.3 (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.

34.2 Material derivative

Definition 34.2.1 (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:

$$\begin{aligned}\frac{Da}{Dt} &= \lim_{\Delta t\to 0}\frac{a(\vec{\boldsymbol{r}}+\Delta\vec{\boldsymbol{r}}, \vec{\boldsymbol{v}}+\Delta\vec{\boldsymbol{v}}, t+\Delta t) - a(\vec{\boldsymbol{r}}, \vec{\boldsymbol{v}}, t)}{\Delta t}\\ &= \frac{\partial a}{\partial t} + \frac{d\vec{\boldsymbol{r}}}{dt}\cdot\frac{\partial a}{\partial\vec{\boldsymbol{r}}} + \frac{d\vec{\boldsymbol{v}}}{dt}\cdot\frac{\partial a}{\partial\vec{\boldsymbol{v}}}\\ &= \frac{\partial a}{\partial t} + \vec{\boldsymbol{v}}\cdot\nabla a + \frac{d\vec{\boldsymbol{v}}}{dt}\cdot\frac{\partial a}{\partial\vec{\boldsymbol{v}}}\end{aligned}\tag{34.1}$$

The second term $\vec{\boldsymbol{v}}\cdot\nabla a$ in this equation is called the **advective** term.

Remark 34.2.2. 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.

[1] Not only those as given by the equations of motion.
[2] Also known as the **material derivative**, especially when applied to fluidum mechanics.

Corollary 34.2.3. If we take $a(\vec{\boldsymbol{r}}, \vec{\boldsymbol{v}}, t) = \vec{\boldsymbol{r}}$ we obtain:

$$\frac{D\vec{\boldsymbol{r}}}{Dt} = \vec{\boldsymbol{v}} \tag{34.2}$$

34.3 Liouville's theorem

Formula 34.3.1 (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:

$$J(\vec{\boldsymbol{x}}, t) = \frac{dV}{dV_0} = \det\left(\frac{\partial\vec{\boldsymbol{x}}}{\partial\vec{\boldsymbol{x}}_0}\right) = \sum_{ijklmn} \varepsilon_{ijklmn} \frac{\partial x^1}{\partial x_0^i}\frac{\partial x^2}{\partial x_0^j}\frac{\partial x^3}{\partial x_0^k}\frac{\partial x^4}{\partial x_0^l}\frac{\partial x^5}{\partial x_0^m}\frac{\partial x^6}{\partial x_0^n} \tag{34.3}$$

The Lagrangian derivative of this Jacobian then becomes:

$$\frac{DJ}{Dt} = (\nabla \cdot \vec{\boldsymbol{x}})J \tag{34.4}$$

Furthermore using the Hamiltonian equations 33.13 it is easy to prove that

$$\nabla \cdot \vec{\boldsymbol{x}} = 0 \tag{34.5}$$

Theorem 34.3.2 (Liouville's theorem). *Let $V(t)$ be a phase space volume containing a fixed set of particles. Application of Liouville's lemma gives:*

$$\frac{DV}{Dt} = \frac{D}{Dt}\int_{\Omega(t)} d^6x = \frac{D}{Dt}\int_{\Omega_0} J(\vec{\boldsymbol{x}}, t)d^6x_0 = 0 \tag{34.6}$$

It follows that the phase space volume of a Hamiltonian system[3] *is invariant with respect to time-evolution.*

Formula 34.3.3 (Boltzmann's transport equation). Let $F(\vec{\boldsymbol{r}}, \vec{\boldsymbol{v}}, t)$ be the mass distribution function:

$$M_{tot} = \int_{\Omega(t)} F(\vec{\boldsymbol{x}}, t)d^6x \tag{34.7}$$

From the conservation of mass we can derive the following formula:

$$\frac{DF}{Dt} = \frac{\partial F}{\partial t} + \frac{d\vec{\boldsymbol{r}}}{dt}\cdot\frac{\partial F}{\partial\vec{\boldsymbol{r}}} - \nabla V \cdot \frac{\partial F}{\partial\vec{\boldsymbol{v}}} = 0 \tag{34.8}$$

This formula is a partial differential equation in 7 variables which can be solved to obtain $F(\vec{\boldsymbol{x}}, t)$.

Theorem 34.3.4 (Poincaré recurrence theorem). *Consider a Hamiltonian system with a finite phase space $\mathcal{V}$ (for example when the system is trapped in a potential well). By Liouville's theorem, the phase flow generated by the equations of motion is a volume preserving map $g : \mathcal{V} \to \mathcal{V}$. 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 $n \in \mathbb{N}$.*

[3] A system that satisfies Hamilton's equations of motion.

34.4 Continuity equation

Formula 34.4.1 (Reynolds transport theorem[4]). Consider a quantity

$$F = \int_{V(t)} f(\vec{\boldsymbol{r}}, \vec{\boldsymbol{v}}, t)dV$$

Using equation 34.4 and the divergence theorem 16.23 we can obtain:

$$\boxed{\frac{DF}{Dt} = \int_V \frac{\partial f}{\partial t}dV + \oint_S f\vec{\boldsymbol{v}} \cdot d\vec{\boldsymbol{S}}} \tag{34.9}$$

Formula 34.4.2 (Continuity equations). For a conserved quantity the equation above becomes:

$$\frac{Df}{Dt} + (\nabla \cdot \vec{\boldsymbol{v}})f = 0 \tag{34.10}$$

$$\frac{\partial f}{\partial t} + \nabla \cdot (f\vec{\boldsymbol{v}}) = 0 \tag{34.11}$$

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 34.9.

Corollary 34.4.3. Combining the Reynolds transport theorem with the Lagrangian continuity equation gives the following identity for an arbitrary function f:

$$\frac{D}{Dt}\int_V \rho f dV = \int_V \rho \frac{Df}{Dt}dV \tag{34.12}$$

[4]This is a 3D extension of the *Leibniz integral rule*.

Chapter 35

Fluid mechanics

35.1 Cauchy stress tensor

Theorem 35.1.1 (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 35.1.2 (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:

$$\vec{\boldsymbol{t}}_{(\vec{\boldsymbol{n}})} = \mathbf{T}(\vec{\boldsymbol{n}}) \tag{35.1}$$

Example 35.1.3. For identical particles, the stress tensor is given by:

$$\mathbf{T} = -\rho\langle\vec{\boldsymbol{w}} \otimes \vec{\boldsymbol{w}}\rangle \tag{35.2}$$

where $\vec{\boldsymbol{w}}$ is the random component of the velocity vector and $\langle\cdot\rangle$ denotes the expectation value (see 29.19).

Theorem 35.1.4 (Cauchy's lemma). *The stress vectors acting on opposite planes are equal in magnitude but opposite in direction:*

$$\vec{\boldsymbol{t}}_{(-\vec{\boldsymbol{n}})} = -\vec{\boldsymbol{t}}_{(\vec{\boldsymbol{n}})} \tag{35.3}$$

Formula 35.1.5 (Cauchy momentum equation). From Newton's second law 32.1 it follows that:

$$\frac{D\vec{\boldsymbol{P}}}{Dt} = \int_V \vec{\boldsymbol{f}}(\vec{\boldsymbol{x}},t)dV + \oint_S \vec{\boldsymbol{t}}(\vec{\boldsymbol{x}},t)dS \tag{35.4}$$

[1] Also known as **Cauchy's fundamental theorem**.

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 16.23 we get

$$\frac{D\vec{\boldsymbol{P}}}{Dt} = \int_V \left[\vec{\boldsymbol{f}}(\vec{\boldsymbol{x}}, t) + \nabla \cdot \mathbf{T}(\vec{\boldsymbol{x}}, t)\right] dV \tag{35.5}$$

The left-hand side can be rewritten using 34.12 as

$$\int_V \rho \frac{D\vec{\boldsymbol{v}}}{Dt} dV = \int_V \left[\vec{\boldsymbol{f}}(\vec{\boldsymbol{x}}, t) + \nabla \cdot \mathbf{T}(\vec{\boldsymbol{x}}, t)\right] dV \tag{35.6}$$

Chapter 36

Optics

36.1 General

36.1.1 Conservation of energy

From the law of conservation of energy we can derive the following formula:

$$\boxed{T + R + A = 1} \tag{36.1}$$

where

T : Transmission coefficient
R : Reflection coefficient
A : Absorption coefficient

36.1.2 Photon

Formula 36.1.1 (Energy).

$$E = h\nu = \hbar\omega = \frac{hc}{\lambda} \tag{36.2}$$

Formula 36.1.2 (Momentum).

$$p = \frac{h}{\lambda} = \hbar k \tag{36.3}$$

where formula 36.4 was used in the last step.

Remark. These formulas can also be (approximately) used for particles for which the rest mass (energy) is negligible.

36.2 Plane wave

Formula 36.2.1 (Wave number).

$$k = \frac{2\pi}{\lambda} \tag{36.4}$$

Formula 36.2.2 (Plane wave). Following equations represent a plane wave moving in the x-direction and polarized in the xy-plane:

$$\vec{\boldsymbol{E}}(x,t) = \mathrm{Re}\left\{A \exp\left[i\left(kx - \omega t + \phi\right)\right]\right\}\vec{\boldsymbol{e}}_y \tag{36.5}$$

$$\vec{\boldsymbol{E}}(x,t) = \mathrm{Re}\left\{A \exp\left[2\pi i\left(\frac{x}{\lambda} - \frac{t}{T} + \frac{\phi}{2\pi}\right)\right]\right\}\vec{\boldsymbol{e}}_y \tag{36.6}$$

36.3 Refraction

Formula 36.3.1 (Refraction).

$$v_2 = \frac{v_1}{n} \tag{36.7}$$

Formula 36.3.2 (Diëlectric function). In the case of non-magnetic materials ($\mu_r \approx 1$), we can write the diëlectric function as following:

$$\epsilon = \epsilon_r + i\epsilon_i = \widetilde{n}^2 = (n + ik)^2 \tag{36.8}$$

Where $\widetilde{n}$ is the **complex refractive index** and k is the **extinction coefficient**.

36.4 Absorption

Theorem 36.4.1 (Law of Lambert-Beer†).

$$\frac{I(x)}{I(0)} = exp\left(-\frac{4\pi\nu k}{c}x\right) \tag{36.9}$$

Definition 36.4.2 (Absorption coefficient). The constant factor in the Lambert-Beer law is called the absorption coefficient.

$$\alpha = \frac{4\pi\nu k}{c} \tag{36.10}$$

36.5 Diffraction

Chapter 37

Astronomy

37.1 Ellipsoidal coordinates

We start from folowing parametrized equation:

$$f(\tau) = \frac{x^2}{\tau+\alpha} + \frac{y^2}{\tau+\beta} + \frac{z^2}{\tau+\gamma} - 1 \tag{37.1}$$

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

$$-(\tau-\lambda)(\tau-\mu)(\tau-\nu) \tag{37.2}$$

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:

$$\begin{cases} x^2 & = & \dfrac{(\lambda+\alpha)(\mu+\alpha)(\nu+\alpha)}{(\beta-\alpha)(\gamma-\alpha)} \\ y^2 & = & \dfrac{(\lambda+\beta)(\mu+\beta)(\nu+\beta)}{(\beta-\alpha)(\beta-\gamma)} \\ z^2 & = & \dfrac{(\lambda+\gamma)(\mu+\gamma)(\nu+\gamma)}{(\alpha-\gamma)(\beta-\gamma)} \end{cases} \tag{37.3}$$

For these solutions multiple cases can be considered. We can define different surfaces by fixing τ at different values.

37.1.1 Ellipsoid: $\tau = \lambda$

First we look at the surfaces defined by fixing $\tau = \lambda$ in equation 37.1. 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**.

37.1.2 One-sheet hyperboloid: $\tau = \mu$

By fixing $\tau = \mu$ in 37.1 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

$$\frac{x^2}{\alpha - \beta} + \frac{z^2}{\gamma - \beta} = 1 \tag{37.4}$$

This hyperbola intersects the z-plane in the foci of the focal ellipse.

37.1.3 Two-sheet hyperboloid: $\tau = \nu$

By fixing $\tau = \nu$ in 37.1 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 37.4. If $\nu \to -\gamma$ the two sheets coincide in the xy-plane.

37.1.4 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 37.3 it is clear that the Hamiltonian function can be spearated as follows:

$$H = \frac{1}{2}\left(\frac{p_\lambda^2}{Q_\lambda^2} + \frac{p_\mu^2}{Q_\mu^2} + \frac{p_\nu^2}{Q_\nu^2}\right) + V \tag{37.5}$$

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

$$Q_\lambda^2 = \frac{1}{4}\frac{(\lambda-\mu)(\lambda-\nu)}{(\lambda+\alpha)(\lambda+\beta)(\lambda+\gamma)} \tag{37.6}$$

which is also valid for μ and ν by applying cyclic permutation to the coordinates.

Following from the Stäckel conditions 33.24 the potential must be of the form

$$V = \sum_i \frac{W_i(\lambda^i)}{Q_i^2} \tag{37.7}$$

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:

$$G(\tau) = -4(\tau+\beta)W_\tau(\tau) \tag{37.8}$$

The 3D potential is thus completely determined by a 1D function $G(\tau)$.

37.1.5 Hamilton-Jacobi equation

If we consider a time-independent system we can use 33.22 as our starting point. If we multiply this equation by $(\lambda-\mu)(\lambda-\nu)(\mu-\nu)$ we obtain

$$\begin{aligned}(\mu-\nu)\Bigg[2(\lambda+\alpha)(\lambda+\beta)(\lambda+\gamma)\left(\frac{dS^\lambda(\lambda)^2}{d\lambda}\right)\\ -(\lambda+\alpha)(\lambda+\gamma)G(\lambda)-\lambda^2E\Bigg] + \text{cyclic permutations} = 0\end{aligned} \tag{37.9}$$

where we rewrote the multiplication factor in the form $a\lambda^2 + b\mu^2 + c\nu^2$ before multiplying the RHS of 33.22. This equation can be elegantly rewritten as

$$(\mu-\nu)U(\lambda) + (\lambda-\mu)U(\nu) + (\nu-\lambda)U(\mu) = 0 \tag{37.10}$$

Differentiating twice with respect to any λ^i gives $U''(\tau) = 0$ or equivalently

$$U(\tau) = I_3 - I_2\tau \tag{37.11}$$

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

$$p_\tau^2 = \frac{1}{2(\tau+\beta)}\left[E - V_{\text{eff}}(\tau)\right] \tag{37.12}$$

where the effective potential is given by

$$\boxed{V_{\text{eff}} = \frac{J}{\tau+\alpha} + \frac{K}{\tau+\gamma} - G(\tau)} \tag{37.13}$$

where J and K are two conserved quantities given by

$$J = \frac{\alpha^2 E + \alpha I_2 + I_3}{\alpha - \gamma} \quad \text{and} \quad K = \frac{\gamma^2 E + \gamma I_2 + I_3}{\gamma - \alpha}$$

To be physically acceptable, p_τ^2 should be positive. This leads to following conditions on the energy:

$$\begin{cases} E \geq V_{\text{eff}}(\lambda) \\ E \geq V_{\text{eff}}(\mu) \\ E \leq V_{\text{eff}}(\nu) \end{cases} \tag{37.14}$$

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

$$V = -\frac{1}{\lambda - \nu}\left(\frac{F(\lambda) - F(\mu)}{\lambda - \mu} - \frac{F(\mu) - F(\nu)}{\mu - \nu}\right) \leq 0 \tag{37.15}$$

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 37.15 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

$$\boxed{G(\tau) = \frac{GM}{\sqrt{\gamma_0 + \tau}}} \tag{37.16}$$

where G is the gravitational constant and M is the galactic mass.

Theorem 37.1.1 (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 37.1.2. For triaxial mass models in ellipsoidal coordinates the axial ratios are inversely proportional to the axial ratios of the coordinate system.

Part VIII

Electromagnetism

Chapter 38

Electricity

38.1 Resistance R

38.1.1 Conductivity

Definition 38.1.1 (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 electricity[1].

Formula 38.1.2 (Conductivity).

$$\sigma = nq\mu \tag{38.1}$$

Formula 38.1.3 (Resistivity).

$$\rho = \frac{1}{\sigma} \tag{38.2}$$

Formula 38.1.4 (Mobility).

$$\mu = \frac{v_d}{E} \tag{38.3}$$

38.1.2 Current density

Formula 38.1.5. Let A be the cross section of a conductor. Let $\vec{J}$ be the current density though A. The current trough A is then given by:

$$I = \iint_A \vec{J} \cdot \hat{n} dS \tag{38.4}$$

Formula 38.1.6 (Free current). The current density generated by free charges is given by:

$$\vec{J} = nq\vec{v}_d \tag{38.5}$$

[1] It is several orders of magnitude smaller.

38.1.3 Pouillet's law

$$R = \rho\,\frac{l}{A} \tag{38.6}$$

where:

ρ : resistivity of the material
l : length of the resistor
A : cross-sectional area

38.2 Ohm's law

Formula 38.2.1 (Ohm's law).

$$\boxed{\vec{\boldsymbol{J}} = \overset{\leftrightarrow}{\sigma} \cdot \vec{\boldsymbol{E}}} \tag{38.7}$$

where $\overset{\leftrightarrow}{\sigma}$ is the conductivity tensor.

Formula 38.2.2 (Ohm's law in wires). The following formula can be found by combining equations 38.1, 38.2,38.4 and 38.7 and by assuming that the conductivity tensor can be simplified to a scalar (follows from the isotropic behaviour of normal resistors):

$$U = RI \tag{38.8}$$

38.3 Capacitance C

Definition 38.3.1 (Capacitance). The capacitance is a (geometrical) value that reflects the amount of charge a certain body can store.

$$C = \frac{q}{V} \tag{38.9}$$

38.4 Electric dipole

Formula 38.4.1 (Electric dipole).

$$\vec{\boldsymbol{p}} = q\vec{\boldsymbol{a}} \tag{38.10}$$

Where:

q : charge of the positive particle
$\vec{\boldsymbol{a}}$: vector pointing from the negative to the positive particle

Formula 38.4.2 (Energy). If an electric dipole is placed in an electric field, its potential energy is:

$$U = -\vec{\boldsymbol{p}} \cdot \vec{\boldsymbol{E}} \tag{38.11}$$

Formula 38.4.3 (Torque). If an electric dipole is placed in an electric field, a torque is generated:

$$\vec{\boldsymbol{\tau}} = \vec{\boldsymbol{p}} \times \vec{\boldsymbol{E}} \tag{38.12}$$

Chapter 39

Magnetism

39.1 Magnetic field

39.1.1 Magnetizing field $\vec{H}$

The magnetizing field $\vec{H}$ is the field resulting from all exterior sources.

39.1.2 Magnetization $\vec{M}$

$$\vec{\boldsymbol{M}} = \chi \vec{\boldsymbol{H}} \tag{39.1}$$

where χ is the magnetic susceptibility.

39.1.3 Magnetic induction $\vec{B}$

The magnetic induction $\vec{B}$ is the field resulting from exterior sources and interior magnetization. (It is the 'real', detectable field.) In vacuum we have the following relation between the magnetic induction B, the magnetizing field H and the magnetization M:

$$\vec{\boldsymbol{B}} = \mu_0 \left(\vec{\boldsymbol{H}} + \vec{\boldsymbol{M}} \right) \tag{39.2}$$

By combining this formula with formula 39.1 we get[1]:

$$\vec{\boldsymbol{B}} = \mu_0 \left(1 + \chi\right) \vec{\boldsymbol{H}} \tag{39.3}$$

[1]This equation is only valid in linear media.

Definition 39.1.1 (Magnetic permeability). The proportionality constant in formula 39.3 is called the magnetic permeability:

$$\mu = \mu_0(1+\chi) \tag{39.4}$$

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.

39.1.4 Tensorial formulation

In anistropic materials we have to use the tensorial formulation.

$$B_i = \sum_j \mu_{ij} H_j \tag{39.5}$$

$$M_i = \sum_j \chi_{ij} H_j \tag{39.6}$$

Both μ and χ are tensors of rank 2.

39.2 Magnetic multipoles

39.2.1 Dipole

$$\vec{\boldsymbol{m}} = IS\vec{\boldsymbol{u}}_n \tag{39.7}$$

39.3 Electric charges in a magnetic field

39.3.1 Cyclotron

Formula 39.3.1 (Gyroradius).

$$r = \frac{mv_\perp}{|q|B} \tag{39.8}$$

Formula 39.3.2 (Gyrofrequency[2]).

$$\omega = \frac{|q|B}{m} \tag{39.9}$$

[2] Also called the Larmor frequency.

Chapter 40

Maxwell equations

40.1 Lorentz force

Formula 40.1.1 (Lorentz force).

$$\vec{\boldsymbol{F}} = q\left(\vec{\boldsymbol{E}} + \vec{\boldsymbol{v}} \times \vec{\boldsymbol{B}}\right) \tag{40.1}$$

Formula 40.1.2 (Lorentz force density).

$$\vec{\boldsymbol{f}} = \rho\vec{\boldsymbol{E}} + \vec{\boldsymbol{J}} \times \vec{\boldsymbol{B}} \tag{40.2}$$

40.2 Differential Maxwell equations

Formula 40.2.1 (Gauss' law for electricity).

$$\nabla \cdot \vec{\boldsymbol{E}} = \frac{\rho}{\varepsilon} \tag{40.3}$$

Formula 40.2.2 (Gauss' law for magnetism).

$$\nabla \cdot \vec{\boldsymbol{B}} = 0 \tag{40.4}$$

Formula 40.2.3 (Faraday's law).

$$\nabla \times \vec{\boldsymbol{E}} = -\frac{\partial \vec{\boldsymbol{B}}}{\partial t} \tag{40.5}$$

Formula 40.2.4 (Maxwell's law[1]).

$$\nabla \times \vec{\boldsymbol{H}} = \frac{\partial \vec{\boldsymbol{D}}}{\partial t} + \vec{\boldsymbol{J}} \tag{40.6}$$

[1] Also called the law of Maxwell-Ampère.

40.3 Potentials

40.3.1 Decomposition in potentials

Remembering the Helmholtz decomposition (equation 16.18) we can derive the following general form for $\vec{\boldsymbol{B}}$ starting from Gauss' law 40.4:

$$\boxed{\vec{\boldsymbol{B}} = \nabla \times \vec{\boldsymbol{A}}} \tag{40.7}$$

where $\vec{\boldsymbol{A}}$ is the magnetic potential.

Combining equation 40.7 with Faraday's law 40.5 and rewriting it a bit, gives the following general form for $\vec{\boldsymbol{E}}$:

$$\boxed{\vec{\boldsymbol{E}} = -\nabla V - \frac{\partial \vec{\boldsymbol{A}}}{\partial t}} \tag{40.8}$$

where V is the electrostatic potential.

40.3.2 Conditions

Substituting the expressions 40.7 and 40.8 into Gauss' law 40.3 and Maxwell's law 40.6 gives the following two (coupled) conditions for the electromagnetic potentials:

$$\triangle\, \vec{\boldsymbol{A}} - \varepsilon\mu \frac{\partial^2 \vec{\boldsymbol{A}}}{\partial t^2} = \nabla \left(\nabla \cdot \vec{\boldsymbol{A}} + \varepsilon\mu \frac{\partial V}{\partial t} \right) - \mu \vec{\boldsymbol{J}} \tag{40.9}$$

$$\triangle\, V - \varepsilon\mu \frac{\partial^2 V}{\partial t^2} = -\frac{\partial}{\partial t} \left(\nabla \cdot \vec{\boldsymbol{A}} + \varepsilon\mu \frac{\partial V}{\partial t} \right) - \frac{\rho}{\varepsilon} \tag{40.10}$$

40.3.3 Gauge transformations

Looking at equation 40.7, it is clear that a transformation $\vec{\boldsymbol{A}} \to \vec{\boldsymbol{A}} + \nabla\psi$ has no effect on $\vec{\boldsymbol{B}}$ due to property 16.15. To compensate this in equation 40.8, we also have to perform the transformation $V \to V - \frac{\partial\psi}{\partial t}$.

The (scalar) function $\psi(\vec{\boldsymbol{r}}, t)$ is called a **gauge function**. The transformations are called **gauge transformations**.

Definition 40.3.1 (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.

40.3.4 Lorenz gauge

A first example of a gauge fixing condition is the Lorenz gauge[2] :

$$\boxed{\nabla \cdot \vec{\boldsymbol{A}} + \varepsilon\mu\frac{\partial V}{\partial t} = 0} \tag{40.11}$$

When using this gauge fixing condition, equations 40.9 and 40.10 become uncoupled and can be rewritten as:

$$\Box\vec{\boldsymbol{A}} = -\mu\vec{\boldsymbol{J}} \tag{40.12}$$

$$\Box V = -\frac{\rho}{\varepsilon} \tag{40.13}$$

To see which gauge functions ψ are valid in this case we perform a transformation as explained above:

$$\vec{\boldsymbol{A}}' = \vec{\boldsymbol{A}} + \nabla\psi \qquad \text{and} \qquad V' = V - \frac{\partial\psi}{\partial t}$$

Substituting these transformations in equation 40.11 and using the fact that both sets of potentials $(\vec{\boldsymbol{A}}, V)$ and $(\vec{\boldsymbol{A}}', V)$ satisfiy the Lorenz gauge 40.11 gives the following condition for the gauge function ψ:

$$\Box\psi = 0 \tag{40.14}$$

Example 40.3.2 (Alternative gauges). Apart from the Lorenz gauge 40.11, there is also the Coulomb gauge:

$$\nabla \cdot \vec{\boldsymbol{A}} = 0 \tag{40.15}$$

40.4 Energy and momentum

Definition 40.4.1 (Poynting vector).

$$\boxed{\vec{\boldsymbol{S}} = \vec{\boldsymbol{E}} \times \vec{\boldsymbol{H}}} \tag{40.16}$$

Definition 40.4.2 (Energy density).

$$\boxed{W = \frac{1}{2}\left(\vec{\boldsymbol{E}} \cdot \vec{\boldsymbol{D}} + \vec{\boldsymbol{B}} \cdot \vec{\boldsymbol{H}}\right)} \tag{40.17}$$

[2]Named after Ludvig Lorenz. Not to be confused with Hendrik Lorentz.

Part IX

Relativity Theory

Chapter 41

Special relativity

41.1 Lorentz transformations

Formula 41.1.1.

$$\beta = \frac{v}{c} \tag{41.1}$$

Formula 41.1.2 (Lorentz factor).

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} \tag{41.2}$$

Formula 41.1.3 (Lorentz transformations). Let $\mathbf{V}$ be a 4-vector. A Lorentz boost along the x^1-axis is given by the following transformation:

$$\boxed{\begin{array}{rcl} V'^0 & = & \gamma\left(V^0 - \beta V^1\right) \\ V'^1 & = & \gamma\left(V^1 - \beta V^0\right) \\ V'^2 & = & V^2 \\ V'^3 & = & V^3 \end{array}} \tag{41.3}$$

Remark 41.1.4. Putting $c = +\infty$ in the previous transformation formulas gives the Galilean transformations from classical mechanics.

41.2 Energy and momentum

Formula 41.2.1 (4-velocity).

$$U^\mu = \left(\frac{dx^0}{d\tau}, \frac{dx^1}{d\tau}, \frac{dx^2}{d\tau}, \frac{dx^3}{d\tau}\right) \tag{41.4}$$

or by applying the formulas for proper time and time dilatation we obtain:

$$U^\mu = (\gamma c, \gamma \vec{\boldsymbol{u}}) \tag{41.5}$$

Formula 41.2.2 (4-momentum).

$$p^\mu = m_0 U^\mu \tag{41.6}$$

or by setting $E = cp^0$:

$$p^\mu = \left(\frac{E}{c}, \gamma m_0 \vec{\boldsymbol{u}}\right) \tag{41.7}$$

Definition 41.2.3 (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 41.2.4 (Relativistic energy relation).

$$\boxed{E^2 = p^2c^2 + m^2c^4} \tag{41.8}$$

This formula is often called the **Einstein relation**.

Chapter 42

General Relativity

42.1 Einstein field equations

Formula 42.1.1 (Einstein field equations). The Einstein field equations without the cosmological constant Λ read:

$$\boxed{G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}} \tag{42.1}$$

where $G_{\mu\nu}$ is the Einstein tensor 22.11 and $T_{\mu\nu}$ is the stress-energy tensor 33.11.

42.2 Schwarzschild metric

Formula 42.2.1 (Schwarzschild metric).

$$ds^2 = \left(1 - \frac{R_s}{r}\right) c^2 dt^2 - \left(1 - \frac{R_s}{r}\right)^{-1} dr^2 - r^2 d\Omega^2 \tag{42.2}$$

where R_s is the Schwarzschild radius given by

$$R_s = \frac{2GM}{c^2} \tag{42.3}$$

Theorem 42.2.2 (Birkhoff). *The Schwarzschild metric is the unique solution of the vacuum field equation under the additional constraints of asymptotic flatness and staticity.*

Part X

Quantum Mechanics

Chapter 43

Schrödinger equation

43.1 One dimension

43.1.1 Time independent Schrödinger equation (TISE)

Formula 43.1.1 (TISE).

$$\boxed{\hat{H}\psi(x) = E\psi(x)} \tag{43.1}$$

where $\hat{H}$ is the Hamiltonian of the system.

Property 43.1.2 (Orthonormality). Let $\{\psi_i\}$ be a set of eigenfunctions of the TISE. These functions obey the following orthogonality relations:

$$\int \psi_i^*(x)\psi_j(x)dx = \delta_{ij} \tag{43.2}$$

43.1.2 Time dependent Schrödinger equation (TDSE)

Formula 43.1.3 (TDSE).

$$\boxed{i\hbar\frac{\partial\psi}{\partial t} = \hat{H}\psi} \tag{43.3}$$

where $\hat{H}$ is the Hamiltonian of the system.

Formula 43.1.4 (Massive particle in a time-independent potential).

$$\boxed{i\hbar\frac{\partial}{\partial t}\psi(x,t) = \left(\frac{\hat{p}\cdot\hat{p}}{2m} + \hat{V}(x)\right)\psi(x,t)} \tag{43.4}$$

Formula 43.1.5 (General solution).

$$\boxed{\psi(x,t) = \sum_E c_E \psi_E(x) e^{-\frac{i}{\hbar}Et}} \tag{43.5}$$

where the functions $\psi_E(x)$ are the eigenfunctions of the TISE 43.1. The coefficients c_E can be found using the orthogonality relations:

$$c_E = \left(\int \psi_E^*(x')\psi(x',t_0)dx'\right) e^{\frac{i}{\hbar}Et_0} \tag{43.6}$$

Chapter 44

Mathematical formalism

44.1 Postulates

44.1.1 Postulate 6: eigenfunction expansion

Definition 44.1.1 (Observable). An operator $\hat{A}$ which possesses a complete set of eigenfunctions is called an observable.

Formula 44.1.2. 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:

$$\boxed{|\Psi\rangle = \sum_n c_n|\psi_n\rangle + \int c_a|\psi_a\rangle da} \tag{44.1}$$

where the summation ranges over the discrete spectrum and the integral over the continuous spectrum.

Formula 44.1.3 (Closure relation). For a complete set of discrete eigenfunctions the closure relation[1] reads:

$$\sum_n |\psi_n\rangle\langle\psi_n| = \mathbb{1} \tag{44.2}$$

For a complete set of continuous eigenfunctions we have the following counterpart:

$$\int |i\rangle\langle i|di = \mathbb{1} \tag{44.3}$$

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 44.2 but implicitly integrate over the continuous spectrum.

[1] This relation is also called the **resolution of the identity**.

44.2 Uncertainty relations

Definition 44.2.1 (Commutator). Let $\hat{A}, \hat{B}$ be two operators. We define the commutator of $\hat{A}$ and $\hat{B}$ as follows:

$$\boxed{\left[\hat{A}, \hat{B}\right] = \hat{A}\hat{B} - \hat{B}\hat{A}} \tag{44.4}$$

Formula 44.2.2.

$$\left[\hat{A}\hat{B}, \hat{C}\right] = \hat{A}\left[\hat{B}, \hat{C}\right] + \left[\hat{A}, \hat{C}\right]\hat{B} \tag{44.5}$$

Definition 44.2.3 (Anticommutator). Let $\hat{A}, \hat{B}$ be two operators. We define the anti-commutator of $\hat{A}$ and $\hat{B}$ as follows:

$$\boxed{\left\{\hat{A}, \hat{B}\right\}_+ = \hat{A}\hat{B} + \hat{B}\hat{A}} \tag{44.6}$$

Definition 44.2.4 (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 44.2.5 (Heisenberg uncertainty relation). Let $\hat{A}, \hat{B}$ be two observables. Let $\Delta A, \Delta B$ be the corresponding uncertainties.

$$\boxed{\Delta A \Delta B = \frac{1}{4}\left|\left\langle\left[\hat{A}, \hat{B}\right]\right\rangle\right|^2} \tag{44.7}$$

44.3 Matrix representation

Formula 44.3.1. The following formula gives the (m, n)-th element of the matrix representation of $\hat{A}$ with respect to the orthonormal basis $\{\psi_n\}$:

$$\boxed{A_{mn} = \langle\psi_m|\hat{A}|\psi_n\rangle} \tag{44.8}$$

Remark 44.3.2. The basis $\{\psi_n\}$ need not consist out of eigenfunctions of $\hat{A}$.

44.4 Slater determinants

Theorem 44.4.1 (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 44.4.2. 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 44.4.3. Let $\{\sigma\}$ be the set of all permutations of the sequence $(1, ..., n)$. Let $|\psi\rangle$ be the single-particle wave function. Fermionic systems are described by a wave function of the form

$$|\Psi_F\rangle = \sum_\sigma |\psi_{\sigma(1)}\rangle \cdots |\psi_{\sigma(n)}\rangle \tag{44.9}$$

Bosonic systems are described by a wave function of the form

$$|\Psi_B\rangle = \sum_\sigma \mathrm{sgn}(\sigma)|\psi_{\sigma(1)}\rangle \cdots |\psi_{\sigma(n)}\rangle \tag{44.10}$$

Definition 44.4.4 (Slater determinant). Let $\{\phi_i(\vec{\boldsymbol{q}})\}_{i\leq N}$ be a set of wave functions (spin orbitals) describing a system of N identical particles. The (totally antisymmetric) wave function of the system is given by:

$$\boxed{\psi(\vec{\boldsymbol{q}}_1, ..., \vec{\boldsymbol{q}}_N) = \frac{1}{\sqrt{N!}} \det \begin{pmatrix} \phi_1(\vec{\boldsymbol{q}}_1) & \cdots & \phi_N(\vec{\boldsymbol{q}}_1) \\ \vdots & & \vdots \\ \phi_1(\vec{\boldsymbol{q}}_N) & \cdots & \phi_N(\vec{\boldsymbol{q}}_N) \end{pmatrix}} \tag{44.11}$$

44.5 Interaction picture

Let $\hat{H} = \hat{H}_0 + \hat{V}(t)$ be the total Hamiltonian of a system where $\hat{V}(t)$ is the interaction Hamiltonian. Let $|\psi(t)\rangle$ and $\hat{O}$ be the state vector and operator in the Schrödinger picture.

Formula 44.5.1. In the interaction picture we define the state vector as follows:

$$|\psi(t)\rangle_I = e^{\frac{i}{\hbar}\hat{H}_0 t}|\psi(t)\rangle \tag{44.12}$$

From this it follows that the operators in the interaction picture are given by:

$$\hat{O}_I(t) = e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{O}e^{-\frac{i}{\hbar}\hat{H}_0 t} \tag{44.13}$$

Formula 44.5.2 (Schrödinger equation). Using the previous definition the Schrödinger equation can be rewritten as follows:

$$i\hbar\frac{d}{dt}|\psi(t)\rangle_I = \hat{V}_I(t)|\psi(t)\rangle_I \tag{44.14}$$

The time-evolution of operators is given by:

$$\frac{d}{dt}\hat{O}_I(t) = \frac{i}{\hbar}\left[\hat{H}_0, \hat{O}_I(t)\right] \tag{44.15}$$

Chapter 45

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 $\{|j, m\rangle\}$.

45.1 General operator

Property 45.1.1. The mutual eigenbasis of $\hat{J}^2$ and $\hat{J}_z$ is defined by the following two eigenvalue equations:

$$\hat{J}^2|j, m\rangle = j(j+1)\hbar^2|j, m\rangle \tag{45.1}$$

$$\hat{J}_z|j, m\rangle = m\hbar|j, m\rangle \tag{45.2}$$

Property 45.1.2. The angular momentum operators generate a Lie algebra 23.2.1. The Lie bracket is defined by following commutation relation:

$$\boxed{\left[\hat{J}_i, \hat{J}_j\right] = i\hbar\varepsilon_{ijk}\hat{J}_k} \tag{45.3}$$

Definition 45.1.3 (Ladder operators[1]). The raising and lowering operators[2] $\hat{J}_+$ and $\hat{J}_-$ are defined as:

$$\hat{J}_+ = \hat{J}_x + i\hat{J}_y \qquad \text{and} \qquad \hat{J}_- = \hat{J}_x - i\hat{J}_y \tag{45.4}$$

Corollary 45.1.4. From the commutation relations of the angular momentum operators we can derive the commutation relations of the ladder operators:

$$\left[\hat{J}_+, \hat{J}_-\right] = 2\hbar\hat{J}_z \tag{45.5}$$

[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.

Formula 45.1.5. 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 45.3:

$$\hat{J}^2 = \hat{J}_+\hat{J}_- + \hat{J}_z^2 - \hbar\hat{J}_z \tag{45.6}$$

Remark 45.1.6 (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 $\hat{J}_i$.

45.2 Rotations

45.2.1 Infinitesimal rotation

Formula 45.2.1. An infinitesimal rotation $\hat{R}(\delta\vec{\boldsymbol{\varphi}})$ is given by the following formula:

$$\boxed{\hat{R}(\delta\vec{\boldsymbol{\varphi}}) = \mathbb{1} - \frac{i}{\hbar}\vec{\boldsymbol{J}}\cdot\delta\vec{\boldsymbol{\varphi}}} \tag{45.7}$$

A finite rotation can then be produced by applying this infinitesimal rotation repeatedly, which gives:

$$\hat{R}(\vec{\boldsymbol{\varphi}}) = \left(\mathbb{1} - \frac{i}{\hbar}\vec{\boldsymbol{J}}\cdot\frac{\vec{\boldsymbol{\varphi}}}{n}\right)^n = \exp\left(-\frac{i}{\hbar}\vec{\boldsymbol{J}}\cdot\vec{\boldsymbol{\varphi}}\right) \tag{45.8}$$

Formula 45.2.2 (Matrix elements). Applying a rotation over an angle φ around the z-axis to a state $|j,m\rangle$ gives:

$$\hat{R}(\varphi\vec{\boldsymbol{e}}_z)|j,m\rangle = \exp\left(-\frac{i}{\hbar}\hat{J}_z\varphi\right)|j,m\rangle = \exp\left(-\frac{i}{\hbar}m\varphi\right)|j,m\rangle \tag{45.9}$$

Multiplying these states with a bra $\langle j',m'|$ and using the orthonormality of the eigenstates gives the matrix elements of the rotation operator:

$$\boxed{\hat{R}_{ij}(\varphi\vec{\boldsymbol{e}}_z) = \exp\left(-\frac{i}{\hbar}m\varphi\right)\delta_{jj'}\delta_{mm'}} \tag{45.10}$$

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:

$$\boxed{\langle j,m'|\hat{R}(\varphi\vec{\boldsymbol{n}})|j,m\rangle = \mathcal{D}^{(j)}_{m,m'}(\hat{R})} \tag{45.11}$$

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

[3]See definition 23.2.25.

45.2.2 Spinor representation

Definition 45.2.3 (Pauli matrices).

$$\boxed{\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}} \tag{45.12}$$

From this definition it is clear that the Pauli matrices are Hermitian and unitary. Together with the 2×2 identity matrix they form a basis for the space of 2×2 Hermitian matrices.

Formula 45.2.4. In the spinor representation ($\mathrm{J} = \frac{1}{2}$) the Wigner-D matrix reads:

$$\mathcal{D}^{(1/2)}(\varphi\vec{\boldsymbol{e}}_z) = \begin{pmatrix} e^{-i/2\varphi} & 0 \\ 0 & e^{i/2\varphi} \end{pmatrix} \tag{45.13}$$

45.3 Coupling of angular momenta

45.3.1 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**.

45.3.2 Clebsch-Gordan series

Let $\vec{\boldsymbol{J}}$ denote the total angular momentum defined as:

$$\vec{\boldsymbol{J}} = \hat{J}_1 + \hat{J}_2 \tag{45.14}$$

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 45.3.1 (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:

$$\boxed{|\mathbf{J}, \mathbf{M}\rangle = \sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} |j_1, j_2, m_1, m_2\rangle\langle j_1, j_2, m_1, m_2|\mathbf{J}, \mathbf{M}\rangle} \tag{45.15}$$

These coefficients are called the Clebsch-Gordan coefficients.

Property 45.3.2. By acting with the operator $\hat{J}_z$ on both sides of equation 45.15 it is possible to proof that the CG coefficient are non-zero if and only if $\mathbf{M} = m_1 + m_2$.

Chapter 46

Dirac equation

46.1 Dirac equation

Definition 46.1.1 (Dirac matrices). The time-like Dirac matrix γ^0 is defined as:

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \tag{46.1}$$

where $\mathbb{1}$ is the 2-dimensional identity matrix. The space-like Dirac matrices γ^k, $k = 1, 2, 3$ are defined using the Pauli matrices[1] σ^k:

$$\gamma^k = \begin{pmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{pmatrix} \tag{46.2}$$

This form of the Dirac matrices fixes a basis called the **Dirac basis**. The **Weyl** or **chiral** basis is fixed by replacing the time-like matrix γ^0 by

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \tag{46.3}$$

We can also define a fifth matrix γ^5:

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{46.4}$$

Property 46.1.2. The Dirac matrices satisfy

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu}\mathbb{1} \tag{46.5}$$

This has the form of equation 20.4. The Dirac matrices can thus be used as the generating set of a Clifford algebra[2].

[1] See definition 45.12.
[2] See defiinition 20.1.1.

Notation 46.1.3 (Feynman slash notation). Let $a = a_\mu x^\mu \in V$ be a general 4-vector. The Feynman slash $\not{a}$ is defined as follows:

$$\not{a} = \gamma^\mu a_\mu \tag{46.6}$$

A more formal treatment of the Feynman slash notation shows that it gives us a canonical map:

$$/ : V \to C\ell_V : a_\mu x^\mu \mapsto a_\mu \gamma^\mu \tag{46.7}$$

Formula 46.1.4 (Dirac equation). In covariant form the Dirac equation reads:

$$\boxed{(i\hbar\not{\partial} - mc)\psi = 0} \tag{46.8}$$

Chapter 47

Perturbation theory

47.1 Rayleigh-Schrödinger perturbation 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 47.1.1. The perturbed eigenfunctions and eigenvalues can be expanded in the following way, where we assume that λ is a small perturbation parameter:

$$|\psi_n\rangle = \sum_{i=0}^{+\infty} \lambda^i |\psi_n^{(i)}\rangle \tag{47.1}$$

$$E_n = \sum_{i=0}^{+\infty} \lambda^i E_n^{(i)} \tag{47.2}$$

where i denotes the order of the perturbation.

47.2 Time-dependent perturbation theory

In this section we consider perturbed Hamiltoninians of the following form:

$$\hat{H}(t) = \hat{H}_0 + \lambda \hat{V}(t) \tag{47.3}$$

47.2.1 Dyson series

Formula 47.2.1 (Tomonaga-Schwinger equation). The evolution operator $\hat{U}(t)$ satisfies the following Schrödinger-type equation in the interaction image[1]:

$$i\hbar\frac{d}{dt}\hat{U}_I|\psi(0)\rangle_I = \hat{V}_I(t)\hat{U}_I|\psi(0)\rangle_I \tag{47.4}$$

Formula 47.2.2 (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:

$$\hat{U}_I(t) = \mathbb{1} - \frac{i}{\hbar}\int_0^t \hat{V}_I(t')\hat{U}_I(t')dt' \tag{47.5}$$

This solution can be iterated to obtain a series expansion of the evolution operator:

$$\hat{U}(t) = 1 - \frac{i}{\hbar}\int_0^t \hat{V}(t_1)dt_1 + \left(-\frac{i}{\hbar}\right)^2\int_0^t dt_1\int_0^{t_1} dt_2\hat{V}(t_1)\hat{V}(t_2) + ... \tag{47.6}$$

It is clear that the integrands obey a time-ordering. By introducing the **time-ordering operator** $\mathcal{T}$:

$$\mathcal{T}\left(\hat{V}(t_1)\hat{V}(t_2)\right) = \begin{cases} \hat{V}(t_1)\hat{V}(t_2) & , \quad t_1 \geq t_2 \\ \hat{V}(t_2)\hat{V}(t_1) & , \quad t_2 > t_1 \end{cases} \tag{47.7}$$

the integrals can be rewritten in a more symmetric form:

$$\hat{U}(t) = 1 - \frac{i}{\hbar}\int_0^t \hat{V}(t_1)dt_1 + \frac{1}{2!}\left(-\frac{i}{\hbar}\right)\int_0^t dt_1\int_0^{t} dt_2\mathcal{T}\left(\hat{V}(t_1)\hat{V}(t_2)\right) + ... \tag{47.8}$$

or by comparing with the series expansion for exponential functions:

$$\boxed{\hat{U}(t) = \mathcal{T}\left(e^{-\frac{i}{\hbar}\int_0^t\hat{V}(t')dt'}\right)} \tag{47.9}$$

This expansion is called the **Dyson series**.

47.3 Variational method

Definition 47.3.1 (Energy functional).

$$E(\psi) = \frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle} \tag{47.10}$$

Property 47.3.2. The energy functional 47.10 satisfies following inequality:

$$E(\psi) \geq E_0 \tag{47.11}$$

where E_0 is the ground state energy.

[1] See section 44.5.

Method 47.3.3. Assume that the trial function $|\psi\rangle$ depends on a set of parameters $\{c_i\}_{i\in I}$. The 'optimal' wave function is found by solving the following system of equations:

$$\frac{\partial\psi}{\partial c_i} = 0 \qquad \forall i \in I \tag{47.12}$$

47.4 Adiabatic approximation

47.4.1 Berry phase

Consider a system for which the adiabatic approximation is valid. We then have a wavefunction of the form

$$\psi(t) = C_a(t)\psi_a(t)\exp\left[-\frac{i}{\hbar}\int_{t_0}^{t} E_a(t')dt'\right] \tag{47.13}$$

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

$$C_a(t) = e^{i\gamma_a(t)} \tag{47.14}$$

Substituting this ansatz in the wavefunction and the Schödinger equation gives a differential equation for the phase factor $\gamma_a(t)$. It can be readily integrated to obtain:

$$\gamma_a(t) = i\int_{t_0}^{t}\left\langle \psi_a(t') \middle| \frac{\partial\psi_a(t')}{\partial t'}\right\rangle dt' \tag{47.15}$$

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:

$$\psi'_a(t) = \psi_a(t)e^{i\eta(t)} \tag{47.16}$$

Entering this in equation 47.15 gives

$$\bar{\gamma}'_a(t) = \bar{\gamma}_a(t) - \eta(t_f) + \eta(t_0) \tag{47.17}$$

where the overhead bar denotes the integration between t_0 and t_f in equation 47.15. If the system is cyclic then $\psi_a(t_0) = \psi_a(t_f)$. Combining this with equation 47.16 gives us:

$$\eta(t_f) - \eta(t_0) = 2k\pi \qquad k \in \mathbb{N} \tag{47.18}$$

which implies that the Berry phase cannot be eliminated through a basis transformation and is thus an observable property of the system.

Definition 47.4.1 (Berry connection). The quantity

$$\mathbf{A}(\vec{\boldsymbol{c}}) = i\langle\psi_a(\vec{\boldsymbol{c}})|\nabla_{\vec{\boldsymbol{c}}}\psi_a(\vec{\boldsymbol{c}})\rangle \tag{47.19}$$

where $\nabla_{\vec{\boldsymbol{c}}}$ denotes the gradient in phase space, is called the Berry connection (or Berry gauge potential). Applying Stokes' theorem to 47.15 gives us:

$$\bar{\gamma}_a = \int \boldsymbol{\mathcal{B}} \cdot d\vec{\boldsymbol{S}} \tag{47.20}$$

where $\boldsymbol{\mathcal{B}} = \nabla_{\vec{\boldsymbol{c}}} \times \mathbf{A}(\vec{\boldsymbol{c}})$ is called the **Berry curvature**. Although the Berry connection is gauge dependent, the Berry curvature is gauge invariant!

Chapter 48

Scattering theory

48.1 Cross section

Formula 48.1.1 (Differential cross section).

$$\frac{d\sigma}{d\Omega} = \frac{N(\theta,\varphi)}{F} \tag{48.1}$$

where is the incoming particle flux and N the detected flow rate[1].

48.1.1 Fermi's golden rule

Formula 48.1.2 (Fermi's golden rule). The transition probability from state i to state f is given by:

$$\boxed{\Gamma_{i\rightarrow f} = \frac{2\pi}{\hbar}|\langle f|\hat{H}|i\rangle|^2\frac{dn}{dE_f}} \tag{48.2}$$

48.2 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 48.2.1 (Lippman-Schwinger equation).

$$|\psi^{(\pm)}\rangle = |\varphi\rangle + \frac{1}{E - \hat{H}_0 \pm i\varepsilon}\hat{V}|\psi^{(\pm)}\rangle \tag{48.3}$$

where $|\varphi\rangle$ is an eigenstate of the free Hamiltonian with the same energy as $|\psi\rangle$.

[1] As N is not defined as a rate per unit area (flux), the differential cross section has the dimension of area.

Remark 48.2.2. The term $\pm i\varepsilon$ is added to the denominator as otherwise it would be singular. It has no real physical meaning.

Formula 48.2.3 (Born series). If we rewrite the Lippman-Schwinger equation as $|\psi\rangle = |\varphi\rangle + \hat{G}_0\hat{V}|\psi\rangle$, were $\hat{G}_0$ is the Green's operator, then we can derive the following series expansion by iterating the equation:

$$|\psi\rangle = |\varphi\rangle + \hat{G}_0\hat{V}|\varphi\rangle + \left(\hat{G}_0\hat{V}\right)^2|\varphi\rangle + ... \tag{48.4}$$

Formula 48.2.4 (Born approximation). If we cut off the Born series at the first order term in $\hat{V}$ then we obtain the Born approximation:

$$|\psi\rangle = |\varphi\rangle + \hat{G}_0\hat{V}|\varphi\rangle \tag{48.5}$$

Chapter 49

Entanglement & Quantum computing

49.1 Bipartite systems

49.1.1 Marginal density operators

Definition 49.1.1 (Marginal density operator). Let $|\Psi\rangle_{AB}$ be the state of a bipartite system. The marginal density operator $\hat{\rho}_A$ of system A is defined as follows:

$$\hat{\rho}_A = \mathrm{Tr}_B |\Psi\rangle_{AB AB}\langle\Psi| \tag{49.1}$$

Definition 49.1.2 (Purification). Let $\hat{\rho}_A$ be the desnity operator of a system A. A purification of $\hat{\rho}_A$ is a pure state $|\Psi\rangle_{AB}$ of a composite system AB such that:

$$\hat{\rho}_A = \mathrm{Tr}_B |\Psi\rangle_{AB AB}\langle\Psi| \tag{49.2}$$

Property 49.1.3. 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 $\mathcal{H}_B$.

Chapter 50

Quantum Field Theory

In this chapter we adopt the standard Minkowskian signature $(+,-,-,-)$ unless otherwise stated. This follows the introductory literature and courses such as [4]. Furthermore we also work in natural units unless stated otherwise, i.e. $\hbar = c = 1$.

50.1 Klein-Gordon Field

50.1.1 Lagrangian and Hamiltonian

The simplest Lagrangian (density) is given by:

$$\boxed{\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}m^2\phi^2} \tag{50.1}$$

Using the principle of least action we obtain the following Euler-Lagrange equations[1]:

$$\left(\partial^\mu\partial_\mu + m^2\right)\phi = 0 \tag{50.2}$$

or by introducing the **d'Alembertian** $\Box = \partial^\mu\partial_\mu$:

$$\boxed{(\Box + m^2)\phi = 0} \tag{50.3}$$

This equation is called the Klein-Gordon equation. In the limit $m \rightarrow 0$ it reduces to the well-known wave equation.

From the Lagrangian 50.1 we can also derive a Hamiltonian function using relation 33.12:

$$\boxed{H = \int d^3x\frac{1}{2}\left[\pi^2(x) + (\nabla\phi(x))^2 + m^2\phi^2(x)\right]} \tag{50.4}$$

[1] See formula 33.4.

50.1.2 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:

$$\left(\partial_t^2 + p^2 + m^2\right)\phi(\vec{\boldsymbol{p}}, t) = 0 \tag{50.5}$$

This is the equation for a simple harmonic oscillator with frequency $\omega_{\vec{\boldsymbol{p}}} = \sqrt{p^2 + m^2}$.

Analogous to ordinary quantum mechanics we define the raising and lowering operators $a^\dagger_{\vec{\boldsymbol{p}}}$ and $a_{\vec{\boldsymbol{p}}}$ such that:

$$\phi(\vec{\boldsymbol{x}}) = \iiint \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{\boldsymbol{p}}}}}\left(a_{\vec{\boldsymbol{p}}}e^{i\vec{\boldsymbol{p}}\cdot\vec{\boldsymbol{x}}} + a^\dagger_{\vec{\boldsymbol{p}}}e^{-i\vec{\boldsymbol{p}}\cdot\vec{\boldsymbol{x}}}\right) \tag{50.6}$$

$$\pi(\vec{\boldsymbol{x}}) = \iiint \frac{d^3p}{(2\pi)^3}(-i)\sqrt{\frac{\omega_{\vec{\boldsymbol{p}}}}{2}}\left(a_{\vec{\boldsymbol{p}}}e^{i\vec{\boldsymbol{p}}\cdot\vec{\boldsymbol{x}}} - a^\dagger_{\vec{\boldsymbol{p}}}e^{-i\vec{\boldsymbol{p}}\cdot\vec{\boldsymbol{x}}}\right) \tag{50.7}$$

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 $\pi(\vec{\boldsymbol{x}})$:

$$\boxed{\phi(\vec{\boldsymbol{x}}) = \iiint \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{\boldsymbol{p}}}}}\left(a_{\vec{\boldsymbol{p}}} + a^\dagger_{-\vec{\boldsymbol{p}}}\right)e^{i\vec{\boldsymbol{p}}\cdot\vec{\boldsymbol{x}}}} \tag{50.8}$$

$$\boxed{\pi(\vec{\boldsymbol{x}}) = \iiint \frac{d^3p}{(2\pi)^3}(-i)\sqrt{\frac{\omega_{\vec{\boldsymbol{p}}}}{2}}\left(a_{\vec{\boldsymbol{p}}} - a^\dagger_{-\vec{\boldsymbol{p}}}\right)e^{i\vec{\boldsymbol{p}}\cdot\vec{\boldsymbol{x}}}} \tag{50.9}$$

When we impose the commutation relation

$$[a_{\vec{\boldsymbol{p}}}, a^\dagger_{\vec{\boldsymbol{q}}}] = (2\pi)^3\delta(\vec{\boldsymbol{p}} - \vec{\boldsymbol{q}}) \tag{50.10}$$

we obtain the following commutation relation for the scalar field and its conjugate momentum:

$$[\phi(\vec{\boldsymbol{x}}), \pi(\vec{\boldsymbol{y}})] = i\delta(\vec{\boldsymbol{x}} - \vec{\boldsymbol{y}}) \tag{50.11}$$

Combining the previous formulas gives us the following important commutation relations:

$$\boxed{[H, a^\dagger_{\vec{\boldsymbol{p}}}] = \omega_p a^\dagger_{\vec{\boldsymbol{p}}}} \tag{50.12}$$

$$\boxed{[H, a_{\vec{\boldsymbol{p}}}] = -\omega_p a_{\vec{\boldsymbol{p}}}} \tag{50.13}$$

The Hamiltonian can also be explicitly calculated:

$$H = \int \frac{d^3p}{(2\pi)^3}\omega_{\vec{\boldsymbol{p}}}\left(a^\dagger_{\vec{\boldsymbol{p}}}a_{\vec{\boldsymbol{p}}} + \frac{1}{2}[a_{\vec{\boldsymbol{p}}}, a^\dagger_{\vec{\boldsymbol{p}}}]\right) \tag{50.14}$$

It is however immediately clear from 50.10 that the second term in this integral diverges. This is a consequence of both the fact that space is infinite, i.e. the d^3x integral diverges, and the "large p" limit in the the d^3p integral. The first divergence can be resolved by applying some kind of boundary and considering the energy density instead of the energy itself. The second divergence follows from the fact that 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. A more practical solution is to note that only energy differences are physical and so we can drop the second term altogether as it is merely a "constant".

50.1.3 Complex scalar fields

Formula 50.1.1 (Pauli-Jordan function).

$$[\phi(x), \phi(y)] = i \underbrace{\int \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_p} \left(e^{-ip\cdot(x-y)} - e^{ip\cdot(x-y)}\right)}_{\Delta(x-y)} \tag{50.15}$$

In the case that $x^0 = y^0$ (ETCR) or $(x-y)^2 < 0$ the Pauli-Jordan function is identically 0.[2]

50.2 Lorentz invariant integrals

When applying a Lorentz boost Λ the delta function $\delta(\vec{\boldsymbol{p}} - \vec{\boldsymbol{q}})$ transforms[3] as $\delta(\Lambda\vec{\boldsymbol{p}} - \Lambda\vec{\boldsymbol{q}})\frac{\Lambda E}{E}$. This is clearly not a Lorentz invariant quantity and cannot be used for normalisation. It is however also clear that the quantity $2E\delta(\vec{\boldsymbol{p}} - \vec{\boldsymbol{q}})$ is Lorentz invariant. (The constant 2 is merely introduced for future convenience.) The correct normalisation for the momentum representation thus becomes:

$$\langle p|q\rangle = 2E_p(2\pi)^3\delta(\vec{\boldsymbol{p}} - \vec{\boldsymbol{q}}) \tag{50.16}$$

where the factors 2 and $(2\pi)^3$ are again a matter of convention.

The factor $2E_p$ does not only occur in the normalisation conditions. To define a Lorentz invariant measure for evaluating integrals in spacetime we define define the following integral:

$$\int \frac{d^3p}{2E_p} = \int d^4p\ \delta(p^2 - m^2)\Big|_{p^0>0} \tag{50.17}$$

This means that the integral of any Lorentz invariant function $f(p)$ using the measure $\frac{d^3p}{2E_p}$ will be Lorentz invariant.

Computing the quantity $\langle 0|\phi(\vec{\boldsymbol{x}})|\vec{\boldsymbol{p}}\rangle$ gives $e^{i\vec{\boldsymbol{x}}\cdot\vec{\boldsymbol{p}}}$. This coincides with the position representation from quantum mechanics of the state $|\vec{\boldsymbol{p}}\rangle$ and so we will also interpret it in QFT as the position representation of the single particle state $|\vec{\boldsymbol{p}}\rangle$.

[2] See also the axiom of microcausality 50.4.1

[3] This follows from property 11.9.

50.3 Wick's theorem

50.3.1 Bosonic fields

Definition 50.3.1 (Contraction for neutral bosonic fields).

$$\contraction{}{\phi}{(x)}{\phi}\phi(x)\phi(y) = \begin{cases} [\phi(x)^{(+)}, \phi(y)^{(-)}] & x^0 > y^0 \\ [\phi(y)^{(+)}, \phi(x)^{(-)}] & y^0 > x^0 \end{cases} \tag{50.18}$$

Formula 50.3.2 (Feynman propagator).

$$\contraction{}{\phi}{(x)}{\phi}\phi(x)\phi(y) = i \underbrace{\lim_{\varepsilon \to 0^+} i \int \frac{d^4k}{(2\pi)^4} \frac{e^{-ik\cdot(x-y)}}{k^2 - m^2 + i\varepsilon}}_{\Delta_F(x-y)} \tag{50.19}$$

Definition 50.3.3 (Contraction for charged bosonic fields).

$$\contraction{}{\phi}{(x)}{\overline{\phi}}\phi(x)\overline{\phi}(y) = \begin{cases} [\phi(x)^{(+)}, \overline{\phi}(y)^{(-)}] & x^0 > y^0 \\ [\phi(y)^{(+)}, \overline{\phi}(x)^{(-)}] & y^0 > x^0 \end{cases} \tag{50.20}$$

Remark 50.3.4. In the case of charged bosons, only contractions of the form $\contraction{}{\phi}{(x)}{\overline{\phi}}\phi(x)\overline{\phi}(y)$ will remain because $[a, b^+] = 0$ for charged bosons.

50.3.2 Fermionic fields

Definition 50.3.5 (Contraction).

$$\contraction{}{\psi}{(x)}{\overline{\psi}}\psi(x)\overline{\psi}(y) = \begin{cases} \{\psi(x)^{(+)}, \overline{\psi}(y)^{(-)}\}_+ & x^0 > y^0 \\ -\{\psi(y)^{(+)}, \overline{\psi}(x)^{(-)}\}_+ & y^0 > x^0 \end{cases} \tag{50.21}$$

Remark 50.3.6. Only contractions of the form $\contraction{}{\psi}{(x)}{\overline{\psi}}\psi(x)\overline{\psi}(y)$ will remain because $\{a, b^+\}_+ = 0$.

Formula 50.3.7 (Feynman propagator).

$$\contraction{}{\psi}{(x)}{\overline{\psi}}\psi(x)\overline{\psi}(y) = i \underbrace{\lim_{\varepsilon \to 0^+} \int \frac{d^4p}{(2\pi)^4} \frac{\not{p} + m}{p^2 - m^2 + i\varepsilon} e^{-ip\cdot(x-y)}}_{S_F(x-y)} \tag{50.22}$$

50.4 Axiomatic approach

Theorem 50.4.1 (Axiom of microcausality). *Let $\hat{O}$ be an observable and let x, y be two spacetime points. If $x - y$ is a space-like vector then $[\hat{O}(x), \hat{O}(y)] = 0$.*

50.4.1 Wightman axioms

50.5 Quantum Chromodynamics

Property 50.5.1 (OZI rule). Decay processes for which the corresponding Feynman diagrams become disconnected (initial states and final states are disconnected) when removing internal gluon lines are suppressed with respect to other processes.

Part XI

Thermal Physics & Statistical Mechanics

Chapter 51

Thermodynamics

51.1 General definitions

Definition 51.1.1 (System). The part of space that we are examining.

Definition 51.1.2 (Surroundings). Everything outside the system.

Definition 51.1.3 (Immediate surrounding). The part of the surroundings that 'lies' immediately next to the system.

Definition 51.1.4 (Environment). Everything outside the immediate surroundings.

Definition 51.1.5 (Thermodynamic coordinates). Macroscopical quantities that describe the system.

Definition 51.1.6 (Intensive coordinate). Coordinate that does not depend on the total amount of material (or system size).

Definition 51.1.7 (Extensive coordinate). Coordinate that does depend on the amount of material.

Definition 51.1.8 (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 51.1.9. *During thermodynamic equilibrium, all intensive coordinates are uniform throughout the system.*

Definition 51.1.10 (Isolated system). An isolated system can't interact with its surroundings (due to the presence of impenetrable walls).

Definition 51.1.11 (Diathermic wall). A diathermic wall is a wall that allows heat transfer.

Definition 51.1.12 (Adiabatic wall). An adiabatic wall is a wall that does not allow heat transfer.

Definition 51.1.13 (Open system). An open system is a system that allows matter exchange.

Definition 51.1.14 (Closed system). A closed system is a system that does not allow matter exchange.

Definition 51.1.15 (Quasistatic process). A quasistatic process is a sequence of equilibrium states separated by infinitesimal changes.

Definition 51.1.16 (Path). The sequence of equilibrium states in a process is called the path.

51.2 Postulates

Theorem 51.2.1 (Zeroth law). *If two object are in thermal equilibrium with a third object then they are also in thermal equilibrium with eachother.*

Theorem 51.2.2 (First law).

$$U_f - U_i = W + Q \tag{51.1}$$

$$dU = \delta W + \delta Q \tag{51.2}$$

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 51.2.3 (Kelvin-Planck formulation of the second law). *No machine can absorb an amount of heat and completely transform it into work.*

Theorem 51.2.4 (Clausius formulation of the second law). *Heat cannot be passed from a cooler object to a warmer object without performing work.*

Formula 51.2.5 (Clausius' inequality). In differential form, the inequality reads:

$$\frac{\delta Q}{T} \geq 0 \tag{51.3}$$

Theorem 51.2.6 (Third law). *No process can reach absolute zero in a finite sequence of operations.*

51.3 Gases

51.3.1 Ideal gases

Theorem 51.3.1 (Ideal gas law).

$$\boxed{PV = nRT} \tag{51.4}$$

Chapter 52

Statistical mechanics

52.1 Axioms

Theorem 52.1.1 (Ergodic principle). *All microstates corresponding to the same macroscopic state are equally propable.*

Theorem 52.1.2 (Boltzmann formula). *The central axiom of statistical mechanics gives following formula for the entropy:*

$$\boxed{S = k \ln \Omega(E, V, N, \alpha)} \tag{52.1}$$

where Ω denotes the number of microstates corresponding to the system at energy E, volume V, and so on.

52.2 Temperature

The temperature of a system in contact with a heat bath is defined as:

$$\boxed{T = \left(\frac{\partial E}{\partial S}\right)_V} \tag{52.2}$$

52.3 Maxwell-Boltzmann statistics

Consider a system of N indistinguishable non-interacting particles. Let ε_i be the energy associated with the i-th energy level with degeneracy g_i. The probability p_i of finding a particle in the i-th energy level is given by:

$$\boxed{p_i = \frac{g_i e^{-\beta\varepsilon_i}}{Z}} \tag{52.3}$$

where Z is the single particle **partition function** defined as:

$$\boxed{Z = \sum_i g_i e^{-\beta\varepsilon_i}} \tag{52.4}$$

52.4 Grand canonical system

Formula 52.4.1 (Grand canonical partition function).

$$\mathcal{Z}_i = \sum_{\varepsilon_i} e^{\beta n_i(\mu-\varepsilon_i)} \tag{52.5}$$

Corollary 52.4.2. In the case that $n_i \in \{0,1\}$ this formula reduces to $\mathcal{Z}_i = e^{\beta\mu}Z_i$.

Definition 52.4.3 (Fugacity).

$$z = e^{\mu N} \tag{52.6}$$

52.5 Energy

Theorem 52.5.1 (Virial theorem).

$$\boxed{\langle T\rangle = -\frac{1}{2}\sum_k \langle \vec{\boldsymbol{r}}_k \cdot \vec{\boldsymbol{F}}_k\rangle} \tag{52.7}$$

Corollary 52.5.2. For potentials of the form $V = ar^{-n}$ this becomes:

$$2\langle T\rangle = -n\langle V\rangle \tag{52.8}$$

Theorem 52.5.3 (Equipartition theorem). *Let x be any generalized coordinate (both position or momentum).*

$$\boxed{\left\langle x^k \frac{\partial H}{\partial x^l}\right\rangle = \delta_{kl}k_bT} \tag{52.9}$$

Corollary 52.5.4. For quadratic Hamiltonians this can be rewritten using Euler's theorem for homogeneous functions 6.10 as:

$$\langle T\rangle = \frac{1}{2}k_bT \tag{52.10}$$

Chapter 53

Photon gas

53.1 Black-body radiation

Formula 53.1.1 (Planck's law).

$$\boxed{B_\nu(\nu,T) = \frac{2h\nu^3}{c^2}\frac{1}{e^{\frac{h\nu}{kt}}-1}} \tag{53.1}$$

Formula 53.1.2 (Wien's displacement law).

$$\boxed{\lambda_{max}T = b} \tag{53.2}$$

where $b = 2.897\,772\,9(17) \times 10^{-3}\ \text{Km}$ is **Wien's displacement constant**.

Part XII

Solid State Physics

Chapter 54

Material physics

54.1 Crystals

Theorem 54.1.1 (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 54.1.2 (Zone). The collection of faces parallel to a given axis, is called a zone. The axis itself is called the zone axis.

54.1.1 Analytic representation

Definition 54.1.3 (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 54.1.4. Negative numbers are written as $\overline{a}$ instead of $-a$.

Formula 54.1.5 (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

$$u = \begin{vmatrix} k_1 & l_1 \\ k_2 & l_2 \end{vmatrix} \qquad v = \begin{vmatrix} l_1 & h_1 \\ l_2 & h_2 \end{vmatrix} \qquad w = \begin{vmatrix} h_1 & k_1 \\ h_2 & k_2 \end{vmatrix} \tag{54.1}$$

Theorem 54.1.6 (Hauy's law of rational indices). *The Miller indices of every natural face of a crystal will always have rational proportions.*

54.2 Symmetries

Definition 54.2.1 (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 54.2.2 (Rotational symmetry). Only $1, 2, 3, 4$ and 6-fold rotational symmetries can occur.

54.3 Crystal lattice

Formula 54.3.1. For an orthogonal crystal lattice, the distance between planes of the family $(h\ k\ l)$ is given by:

$$d_{hkl} = \frac{1}{\sqrt{\left(\frac{h}{a}\right)^2 + \left(\frac{k}{b}\right)^2 + \left(\frac{l}{c}\right)^2}} \tag{54.2}$$

54.3.1 Bravais lattice

Definition 54.3.2 (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 54.3.3 (Wigner-Seitz cell). The part of space consisting of all points closer to a given lattice point than to any other.

54.3.2 Reciprocal lattice

Formula 54.3.4 (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

$$\vec{\boldsymbol{a}}^* = 2\pi \frac{\vec{\boldsymbol{b}} \times \vec{\boldsymbol{c}}}{\vec{\boldsymbol{a}} \cdot (\vec{\boldsymbol{b}} \times \vec{\boldsymbol{c}})} \tag{54.3}$$

The vectors $\vec{\boldsymbol{b}}^*$ and $\vec{\boldsymbol{c}}^*$ are obtained by permutation of (a, b, c). These vectors satisfy the relations

$$\begin{aligned} \vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{a}}^* &= 2\pi \\ \vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{b}}^* &= 2\pi \\ \vec{\boldsymbol{c}} \cdot \vec{\boldsymbol{c}}^* &= 2\pi \end{aligned} \tag{54.4}$$

$$\tag{54.5}$$

Notation 54.3.5 (Reciprocal lattice vector). The reciprocal lattice vector $\vec{\boldsymbol{r}}^*_{hkl}$ is defined as follows:

$$\vec{\boldsymbol{r}}^*_{hkl} = h\vec{\boldsymbol{a}}^* + k\vec{\boldsymbol{b}}^* + l\vec{\boldsymbol{c}}^* \tag{54.6}$$

Property 54.3.6. 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.
- $||\vec{\boldsymbol{r}}^*_{hkl}|| = \frac{2\pi n}{d_{hkl}}$

54.4 Diffraction

54.4.1 Constructive interference

Formula 54.4.1 (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 54.4.2. 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 54.4.3 (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:

$$\boxed{\vec{k} - \vec{k}_0 = \vec{r}^*_{hkl}} \tag{54.7}$$

Formula 54.4.4 (Bragg's law). Another equivalent formulation of the Laue conditions is given by following formula:

$$\boxed{2d_{hkl}\sin\theta = n\lambda} \tag{54.8}$$

where

λ : wavelength of the incoming beam

θ : the **Bragg angle**

d_{hkl} : distance between neighbouring planes

Remark 54.4.5. The angle between the incident and diffracted beams is given by 2θ.

Construction 54.4.6 (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 54.7. Therefore Bragg diffraction will occur in the direction of all the intersections of the Ewald sphere and the reciprocal lattice.

54.4.2 Intensity of diffracted beams

Definition 54.4.7 (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 54.4.8 (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 54.4.9 (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:

$$F(hkl) = \sum_j f_j \exp\left[2\pi i(hx_j + ky_j + lz_j)\right] \tag{54.9}$$

where f_j is the atomic scattering factor of the j^{th} atom in the motive.

Example 54.4.10. 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 54.4.11 (Laue indices). Higher order diffractions can be rewritten as a first order diffraction in the following way:

$$2d_{nhnknl}\sin\theta = \lambda \qquad \text{with} \qquad d_{nhnknl} = \frac{d_{hkl}}{n} \tag{54.10}$$

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.

54.5 Alloys

Theorem 54.5.1 (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.

54.6 Lattice defects

Definition 54.6.1 (Vacancy). A lattice point where an atom is missing. Also called a **Schottky defect**.

Formula 54.6.2 (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:

$$\frac{n}{n+N} = e^{-E_v/kT} \tag{54.11}$$

where T is the temperature and E_v the energy needed to create a vacancy.

Remark. A similar relation holds for interstitials.

Definition 54.6.3 (Interstitial). An atom placed at a position which is not a lattice point.

Definition 54.6.4 (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 54.6.5 (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:

$$\frac{n_i}{\sqrt{NN_i}} = e^{-E_{fr}/2kT} \tag{54.12}$$

where E_{fr} denotes the energy needed to create a Frenkel pair.

Remark 54.6.6. In compounds the number of vacancies can be much higher than in mono-atomic lattices.

Remark 54.6.7. The existence of these defects creates the possibility of diffusion.

54.7 Electrical properties

54.7.1 Charge carriers

Formula 54.7.1 (Conductivity). Definition 38.1 can be modified to account for both positive and negative charge carriers:

$$\sigma = n_n q_n \mu_n + n_p q_p \mu_p \tag{54.13}$$

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.

54.7.2 Band structure

Definition 54.7.2 (Valence band). The energy band corresponding to the outermost (partially) filled atomic orbital.

Definition 54.7.3 (Conduction band). The first unfilled energy band.

Definition 54.7.4 (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 54.7.5 (Fermi level). The energy level having a 50% chance of being occupied at thermodynamic equilibrium.

Formula 54.7.6 (Fermi function). The following distribution gives the probability of a state with energy E_i being occupied by an electron:

$$\boxed{f(E_i) = \frac{1}{e^{(E_i - E_f)/kT} + 1}} \tag{54.14}$$

where E_f is the Fermi level as defined above.

54.7.3 Intrinsic semiconductors

Formula 54.7.7. Let n denote the charge carrier density as before. We find the following temperature dependence:

$$n \propto e^{-E_g/2kt} \tag{54.15}$$

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 $E_g \gg kT$.

54.7.4 Extrinsic semiconductors

Definition 54.7.8 (Doping). Intentionally introducing impurities to modify the (electrical) properties.

Definition 54.7.9 (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 54.7.10 (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**.

[1] For a basic derivation see [14].

54.7.5 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 54.7.11. 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 54.7.12 (Saturation polarization). The maximum polarization obtained by a ferroelectric material. It it obtained when the domain formation also reaches a maximum.

Definition 54.7.13 (Remanent polarization). The residual polarization of the material when the external electric field is turned off.

Definition 54.7.14 (Coercive field). The electric field needed to cancel out the remanent polarization.

Definition 54.7.15 (Piezoelectricity). Materials that obtain a polarization when exposed to mechanical stress are called piezoelectric materials.

Remark 54.7.16. 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 54.7.17 (Transducer). A device that converts electrical to mechanical energy (and vice versa).

54.8 Magnetic properties

Definition 54.8.1 (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 54.8.2. All materials exhibit a diamagnetic character.

Definition 54.8.3 (Paramagnetism). The susceptibility is small, positive and inversely proportional to the temperature.

Definition 54.8.4 (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.

54.8.1 Paramagnetism

Formula 54.8.5 (Curie's law). If the interactions between the particles can be neglected, we obtain the following law:

$$\chi = \frac{C}{T} \tag{54.16}$$

Materials that satisfy this law are called **ideal paramagnetics**.

Formula 54.8.6 (Curie-Weiss law). If the interactions between particles cannot be neglected, we obtain the following law:

$$\chi = \frac{C}{T-\theta} \tag{54.17}$$

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 54.8.7 (Brillouin function B_J).

$$\boxed{B_J(y) = \frac{2J+1}{2J}\coth\left(\frac{2J+1}{2J}y\right) - \frac{1}{2J}\coth\left(\frac{y}{2J}\right)} \tag{54.18}$$

where $y = \frac{g\mu_B JB}{kT}$

Remark 54.8.8. Because $\coth(y\rightarrow\infty) \approx 1$ we have:

$$\text{if } T \rightarrow 0 \quad \text{then} \quad M = Ng\mu_B J B_J(y\rightarrow\infty) = Ng\mu_B J \tag{54.19}$$

This value is called the **absolute saturation magnetization**.

54.8.2 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 54.8.9 (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 54.8.10 (Bloch walls). A wall between two magnetic domains.

Definition 54.8.11 (Ferromagnetic Curie temperature). Above this this temperature the material loses its ferromagnetic properties and it becomes a paramagnetic material following the Curie-Weiss law.

Remark 54.8.12. 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.

54.8.3 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 54.8.13 (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:

$$\chi = \frac{C}{T+\theta} \tag{54.20}$$

This resembles a generalization of the Curie-Weiss law with a negative and therefore virtual critical temperature.

54.8.4 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 54.8.14 (Néel hyperbola). Above the Néel temperature it is possible to define a susceptibility given by:

$$\frac{1}{\chi} = \frac{T}{C} - \frac{1}{\chi_0} - \frac{\sigma}{T-\theta'} \tag{54.21}$$

[2]This can be seen from equation 39.1: $M = \chi H$. The susceptibility should be infinite.

[3]This will occur if it is energetically more favourable.

54.9 Mathematical description

Theorem 54.9.1 (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.*

Part XIII

Appendices

Appendix A

Derivations: Mathematics

A.1 Group theory

A.1.1 Explanation for property 3.1.40

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. $\square$

A.2 Calculus

A.2.1 Proof of method 7.2.2

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 6.17 on line 3 and the relation between the factorial function and the Gamma function 6.18 on line 4. □

A.3 Linear algebra

A.3.1 Proof for the equality of definitions 19.4.14 and 19.4.15

$$(u+v) \otimes (u+v) - u \otimes u - v \otimes v = u \otimes v + v \otimes u \tag{A.3}$$

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 19.16 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 19.4.15 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 tangent space constructions. (Definitions 22.2.3 and 22.2.7)

Let (U, φ) be a chart around the point $p \in M$. Using the first definition of a tangent vector (22.2.3), 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 22.7:

$$v_p(f) = \frac{\partial(f \circ \varphi^{-1})}{\partial q^i}(\varphi(p))\frac{dq^i}{dt}(0)$$

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 22.2.3 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

$$\chi(t) \equiv \varphi^{-1} \circ \Big(\varphi \circ \gamma(t) + \varphi \circ \delta(t) - \varphi(p)\Big)$$

[1] More precisely: representatives of equivalence classes of vectors tangent to curves.

where φ is again the coordinate map in some chart (U, φ) around $p \in M$. Using equation 22.7 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 25.4.12

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 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 25.31 for integral curves on line 2. If we now rewrite this equation as an operator equality, we obtain:

$$\boxed{\mathcal{L}_X = \sum_k X_k \frac{\partial}{\partial x^k}} \tag{A.6}$$

[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.

A.4.3 Explanation for formula 25.4.13

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 25.32:

$$\mathcal{L}_X Y = \lim_{t \to 0} \frac{(T\sigma_t)^{-1}[X(\gamma_p(t))] - X(p)}{t} \tag{A.7}$$

where the $T\sigma_t$ is the differential 25.3.6 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 25.32.

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 25.28 the Lie derivative can be rewritten as:

$$\begin{aligned}\mathcal{L}_X Y &= \lim_{t \to 0} \frac{\sigma_{-t*}X(\gamma_p(t)) - \sigma_{0*}X(\gamma_p(0))}{t}\\ &= \frac{d}{dt}(\sigma_{-t*}X)(\gamma_p(t))\Big|_{t=0}\end{aligned}$$

Or finally by using the relation between pushforward and pullback 25.29 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 Connection between vector calculus and differential geometry (Remark 25.5.10)

Looking at formula 25.46 for the exterior derivative of a smooth function and remembering the definition of the gradient 16.2 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 16.9 and divergence 16.7, 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 19.26 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:

$$\begin{aligned} d\alpha &= \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3}\right) dx_2 \wedge dx_3 + \left(\frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1}\right) dx_3 \wedge dx_1 + \left(\frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2}\right) dx_1 \wedge dx_2 \\ d\omega &= \left(\frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} + \frac{\partial f_3}{\partial x_3}\right) dx_1 \wedge dx_2 \wedge dx_3 \end{aligned}$$

Using result 19.4.22 and the isomorphism $\sim : \mathbb{R}^{3*} \to \mathbb{R}^3$ we can rewrite this as:

$$\boxed{\begin{aligned} \sim df &= \nabla f \\ \sim (*d\alpha) &= \nabla \times \vec{\boldsymbol{f}} \\ *d\omega &= \nabla \cdot \vec{\boldsymbol{f}} \end{aligned}} \qquad \begin{aligned} &(A.9) \\ &(A.10) \\ &(A.11) \end{aligned}$$

Appendix B

Derivations: Lagrangian formalism

B.1 d'Alembert's principle

In the following derivation we assume a constant mass.

$$\begin{aligned}
&\sum_k \left(\vec{\boldsymbol{F}}_k - \dot{\vec{\boldsymbol{p}}}_k\right) \dot{\vec{\boldsymbol{r}}}_k = 0 \\
\iff &\sum_k \left(\vec{\boldsymbol{F}}_k - \dot{\vec{\boldsymbol{p}}}_k\right) \cdot \left(\sum_l \frac{\partial \vec{\boldsymbol{r}}}{\partial q_l}\dot{q}_l\right) = 0 \\
\iff &\sum_l \left(\sum_k \vec{\boldsymbol{F}}_k \cdot \frac{\partial \vec{\boldsymbol{r}}}{\partial q_l} - \sum_k m\ddot{\vec{\boldsymbol{r}}} \cdot \frac{\partial \vec{\boldsymbol{r}}}{\partial q_l}\right) \dot{q}_l = 0 \\
\iff &\sum_l \left(Q_l - \sum_k m\ddot{\vec{\boldsymbol{r}}} \cdot \frac{\partial \vec{\boldsymbol{r}}}{\partial q_l}\right) \dot{q}_l = 0
\end{aligned} \tag{B.1}$$

Now we look at the following derivative:

$$\begin{aligned}
&\frac{d}{dt}\left(\dot{\vec{\boldsymbol{r}}} \cdot \frac{\partial \vec{\boldsymbol{r}}}{\partial q_l}\right) = \ddot{\vec{\boldsymbol{r}}} \cdot \frac{\partial \vec{\boldsymbol{r}}}{\partial q_l} + \dot{\vec{\boldsymbol{r}}} \cdot \frac{d}{dt}\left(\frac{\partial \vec{\boldsymbol{r}}}{\partial q_l}\right) \\
\iff &\ddot{\vec{\boldsymbol{r}}} \cdot \frac{\partial \vec{\boldsymbol{r}}}{\partial q_l} = \frac{d}{dt}\left(\dot{\vec{\boldsymbol{r}}} \cdot \frac{\partial \vec{\boldsymbol{r}}}{\partial q_l}\right) - \dot{\vec{\boldsymbol{r}}} \cdot \frac{d}{dt}\left(\frac{\partial \vec{\boldsymbol{r}}}{\partial q_l}\right) \\
\iff &\ddot{\vec{\boldsymbol{r}}} \cdot \frac{\partial \vec{\boldsymbol{r}}}{\partial q_l} = \frac{d}{dt}\left(\dot{\vec{\boldsymbol{r}}} \cdot \underbrace{\frac{\partial \vec{\boldsymbol{r}}}{\partial q_l}}_{A}\right) - \dot{\vec{\boldsymbol{r}}} \cdot \left(\frac{\partial \dot{\vec{\boldsymbol{r}}}}{\partial q_l}\right)
\end{aligned} \tag{B.2}$$

To evaluate A we can take a look at another derivative:

$$\begin{aligned}\frac{\partial \dot{\vec{r}}}{\partial \dot{q}_l} &= \frac{\partial}{\partial \dot{q}_l}\left(\sum_k \frac{\partial r}{\partial q_k}\dot{q}_k\right)\\ &= \sum_k \frac{\partial r}{\partial q_k}\delta_{kl}\\ &= \frac{\partial \vec{r}}{\partial q_l}\\ &= A\end{aligned}$$

Substituting this in formula B.2 gives:

$$\begin{aligned}\ddot{\vec{r}}\cdot\frac{\partial \vec{r}}{\partial q_l} &= \frac{d}{dt}\left(\dot{\vec{r}}\cdot\frac{\partial \dot{\vec{r}}}{\partial \dot{q}_l}\right) - \dot{\vec{r}}\cdot\left(\frac{\partial \dot{\vec{r}}}{\partial q_l}\right)\\ &= \frac{d}{dt}\left(\frac{1}{2}\frac{\partial \dot{\vec{r}}^2}{\partial \dot{q}_l}\right) - \frac{1}{2}\frac{\partial \dot{\vec{r}}^2}{\partial q_l}\end{aligned} \tag{B.3}$$

If we multiply this by the mass m and sum over all particles we get :

$$\begin{aligned}\sum_k m_k \ddot{\vec{r}}_k\cdot\frac{\partial \vec{r}_k}{\partial q_l} &= \frac{d}{dt}\frac{\partial}{\partial \dot{q}_l}\left(\sum_k \frac{1}{2}m\dot{\vec{r}}_k^2\right) - \frac{\partial}{\partial q_l}\left(\sum_k \frac{1}{2}m\dot{\vec{r}}_k^2\right)\\ &= \frac{d}{dt}\frac{\partial T}{\partial \dot{q}_l} - \frac{\partial T}{\partial q_l}\end{aligned} \tag{B.4}$$

Where we have denoted the total kinetic energy in the last line as T.
Plugging this result into formula B.1 gives us:

$$\sum_l \left(Q_l - \frac{d}{dt}\frac{\partial T}{\partial \dot{q}_l} - \frac{\partial T}{\partial q_l}\right)\dot{q}_l = 0 \tag{B.5}$$

As all the q_l are independent the following relation should hold for all l:

$$\begin{aligned}Q_l - \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_l}\right) - \frac{\partial T}{\partial q_l} = 0\\ \iff \boxed{\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_l}\right) - \frac{\partial T}{\partial q_l} = Q_l}\end{aligned} \tag{B.6}$$

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:

$$F_i = -\nabla_i V \tag{B.7}$$

With this in mind, lets take a look at the derivative of the potential V with respect to the l-th generalized coordinate:

$$\begin{aligned}\frac{\partial V}{\partial q_l} &= \sum_i (\nabla_i V)\cdot\frac{\partial \vec{r}_i}{\partial q_l}\\ &= -Q_l\end{aligned} \tag{B.8}$$

The differentiation of V with respect to any generalized velocity $\dot{q}_l$ is trivially zero. This combined with the last formula B.8 and with formula B.6 gives:

$$\begin{aligned} & \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_l}\right) - \frac{\partial T}{\partial q_l} = Q_l \\ \iff & \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_l}\right) - \frac{\partial T}{\partial q_l} = -\frac{\partial V}{\partial q_l} + \frac{\partial V}{\partial \dot{q}_l} \\ \iff & \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_l} - \frac{\partial V}{\partial \dot{q}_l}\right) - \frac{\partial}{\partial q_l}(T - V) = 0 \end{aligned} \tag{B.9}$$

If we introduce a new variable L, called the **Lagrangian**, we get the **Lagrangian equation of the second kind**:

$$\boxed{\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_l}\right) - \frac{\partial L}{\partial q_l} = 0} \tag{B.10}$$

B.2 Hamilton's principle

In this part we start from the principle of least action. First we the define the **action** as following:

$$\boxed{I = \int_{t_1}^{t_2} L\left(y(t), \dot{y}(t), t\right) dt} \tag{B.11}$$

Then we require that this action is minimal for the physically acceptable path. To do this we define a family of paths:

$$y(t, \alpha) = y(t) + \alpha\eta(t) \tag{B.12}$$

Where $\eta(t)$ is an arbitrary function with the following boundary conditions:

$$\begin{cases} \eta(t_1) = 0 \\ \eta(t_2) = 0 \end{cases} \tag{B.13}$$

If we define the action integral over this family of paths, the integral B.11 becomes a function of α:

$$I(\alpha) = \int_{t_1}^{t_2} L\left(y(t, \alpha), \dot{y}(t, \alpha), t\right) dt \tag{B.14}$$

Requiring that the action integral is stationary for $y(t)$ (thus $\alpha = 0$) is equivalent to:

$$\left(\frac{dI}{d\alpha}\right)_{\alpha=0} = 0 \tag{B.15}$$

This condition combined with formula B.14 gives us:

$$\frac{dI}{d\alpha} = \int_{t_1}^{t_2} \frac{d}{d\alpha} L\left(y(t, \alpha), \dot{y}(t, \alpha), t\right) dt \tag{B.16}$$

As we evaluate this derivative in $\alpha = 0$ we can replace $y(t, \alpha)$ by $y(t)$ due to definition B.12.

$$\begin{aligned}\frac{dI}{d\alpha} &= \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial L}{\partial \dot{y}}\frac{\partial \dot{y}}{\partial \alpha}\right] dt \\ &= \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial y}\eta(t) + \frac{\partial L}{\partial \dot{y}}\dot{\eta}(t)\right] dt \end{aligned} \tag{B.17}$$

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:

$$\begin{aligned}\frac{dI}{d\alpha} &= \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial y}\eta(x) + h(t)\dot{\eta}(t)\right] dt \\ &= \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial y}\eta(t) + h(t)\frac{d\eta}{dt}\right] dt \\ &= \int_{t_1}^{t_2} \frac{\partial L}{\partial y}\eta(t)dt + \eta(t_2)h(t_2) - \eta(t_1)h(t_1) - \int_{t_1}^{t_2} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}}\right)\eta(t)dt \end{aligned} \tag{B.18}$$

Due to the initial conditions B.15 for the function $\eta(t)$, the two terms in the middle vanish and we obtain:

$$\frac{dI}{d\alpha} = \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial y} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}}\right)\right] \eta(t)dt \tag{B.19}$$

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:

$$\boxed{\frac{\partial L}{\partial y} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}}\right) = 0} \tag{B.20}$$

If we compare this result with formula B.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 B.2.1. Differential equations of the form

$$\boxed{\frac{\partial f}{\partial y}(y, \dot{y}, x) = \frac{d}{dx}\left(\frac{\partial f}{\partial \dot{y}}(y, \dot{y}, x)\right)} \tag{B.21}$$

are known as **Euler-Lagrange equations**.

B.3 Explanation for Noether's theorem 33.3.1

The general transformation rule for the Lagrangian is:

$$\mathcal{L}(x) \to \mathcal{L}(x) + \alpha\delta\mathcal{L}(x) \tag{B.22}$$

To have a symmetry, i.e. keep the action invariant, the deformation factor has to be a 4-divergence:

$$\mathcal{L}(x) \to \mathcal{L}(x) + \alpha \partial_\mu \mathcal{J}^\mu(x) \tag{B.23}$$

To obtain formula 33.8 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 B.20. Combining these formulas gives us:

$$\partial_\mu\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\delta\phi\right) - \partial_\mu \mathcal{J}^\mu(x) = 0 \tag{B.24}$$

From this equation we can conclude that the current

$$\boxed{j^\mu(x) = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\delta\phi - \mathcal{J}^\mu(x)} \tag{B.25}$$

is conserved.

Appendix C

Derivations: Optics and material physics

C.1 Optics

C.1.1 Law of Lambert-Beer 36.9

From formula 36.8 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 36.10. $\square$

Appendix D

Derivations: Classical and Statistical Mechanics

D.1 Moments of inertia

In this section we will always use formula 32.8 to calculate the moment of inertia.

D.1.1 Disk

The volume of a (solid) disk is given by:

$$V_{disk} = \pi R^2 d \tag{D.1}$$

where R is the radius and d is the thickness. The mass density is then given by:

$$\rho = \frac{M}{\pi R^2 d} \tag{D.2}$$

Using cylindrical coordinates the moment of inertia then becomes:

$$I = \frac{M}{\pi R^2 d}\int_0^{2\pi} d\varphi \int_0^d dz \int_0^R r^3 dr \tag{D.3}$$

$$= \frac{M}{\pi R^2 d} 2\pi d \frac{R^4}{4} \tag{D.4}$$

$$= \frac{1}{2} M R^2 \tag{D.5}$$

D.1.2 Solid sphere

The volume of a solid sphere is given by:

$$V_{sphere} = \frac{4}{3}\pi R^3 \tag{D.6}$$

where R is the radius. The mass density is then given by:

$$\rho = \frac{M}{\frac{4}{3}\pi R^3} \tag{D.7}$$

We will use spherical coordinates to derive the moment of inertia, but we have to be carefull. The r in formula 32.8 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:

$$r = r' \sin\theta \tag{D.8}$$

where r' is the spherical coordinate. Now we can calculate the moment of inertia as follows:

$$\begin{align} I &= \frac{M}{\frac{4}{3}\pi R^3}\int_0^{2\pi} d\varphi \int_0^R r'^4 dr' \int_0^\pi \sin^3\theta d\theta \tag{D.9}\\ &= \frac{M}{\frac{4}{3}\pi R^3} 2\pi \frac{R^5}{5}\frac{4}{3} \tag{D.10}\\ &= \frac{2}{5}MR^2 \tag{D.11} \end{align}$$

D.2 Schottky defects

Let E_v be the energy needed to remove a particle from its lattice point and move it to the surface. Furthermore we will neglect any surface effects and assume that the the energy E_v is indepedent 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

$$\Omega = \frac{(N+n)!}{n!N!} \tag{D.12}$$

where we used the fact that the removal of n particles creates n more lattice points at the surface. Using Boltzmann's entropy formula 52.1 and Stirling's formula we obtain

$$S(N,n) = k\ln\Omega = k\big[(N+n)\ln(N+n) - n\ln n - N\ln N\big] \tag{D.13}$$

Using 52.2 we can find the temperature:

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{N,V} = \frac{dS}{dn}\frac{dn}{dE} = \frac{k}{E_v}\ln\frac{N+n}{n} \tag{D.14}$$

which can be rewritten as

$$\boxed{\frac{n}{N+n} = \exp\left(-\frac{E_v}{kT}\right)} \tag{D.15}$$

The density of Frenkel defects can be derived analogously.

Appendix E

Units and symbols

pk	745.7 W

Table E.1: Units

List of Symbols

The following symbols are used throughout the summary:

Operations

$\approx$	is approximately equal to
$\cong$	is isomorphic to
$\hookrightarrow$	is included in
$\mapsto$	mapsto
$\mathbb{1}_X$	The identity map on the set X.
Im	Imaginary part of a complex number.
Re	Real part of a complex number.
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.
Par_t^γ	Parallel transport map with respect to the curve γ.
Res	Residue of a complex function.
e	The identity map on a group.
$X + Y$	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.
$X \times Y$	Cartesian product of the sets X and Y.

Sets

$[a, b]$	Closed interval
$\emptyset$	Empty set
$\mathrm{Hom}(V, W)$	The set of (homo)morphisms from a set V to a set W.
$\Lambda^n(V)$	Space of antisymmetric rank n tensors over a vector space V.
$\mathcal{H}$	Hilbert space
$\mathcal{B}(V, W)$	Space of bounded continuous maps from the space X to the space Y.
$\mathfrak{X}(M)$	$C^\infty(M)$-module of vector fields on the manifold M.

$\Omega^k(M)$	$C^\infty(M)$-module of differential k-forms on the manifold M.
$\pi_n(X, x_0)$	The n^{th} homotopy space on X based at $x_0 \in X$.
$\mathrm{Aut}(V)$	The set of automorphisms (invertible endomorphisms) on a set V.
$\mathrm{End}(V)$	The set of endomorphisms on a set V.
$\mathrm{Hol}_p(\omega)$	Holonomy group at p with respect to the connection ω.
$]a, b[$	Open interval
$C_p^\infty(M)$	Ring of all smooth functions $f : M \to \mathbb{R}$ defined on a neighbourhood of $p \in M$.
D^n	Standard n-disk
$GL(V)$	General linear group: group of all automorphisms on a vector space V.
$GL_n(K)$	General linear group: group of all invertible n-dimensional matrices over the field K.
S^n	Standard n-sphere
$S^n(V)$	Space of symmetric rank n tensors over a vector space V.
T^n	Standard n-torus. Cartesian product of n times S^1.
ΩX	The loop space on X.

Bibliography

[1] Analysis, Manifolds and Physics, Part 1: Basics (Y. Choquet-Bruhat, C. DeWitt-Morette & M. Dillard-Bleick)

[2] Analysis, Manifolds and Physics, Part 2 (Y. Choquet-Bruhat, C. DeWitt-Morette)

[3] Berry phase and quantum structure (Holger Lyre, arXiv:1408.6867)

[4] An Introduction to Quantum Field Theory (Michael E. Peskin & Daniel V. Schroeder, Westview Press)

[5] An Introduction to Tensors and Group Theory for Physicists (Nadir Jeevanjee, Birkhäuser)

[6] Branch cuts (`http://www.math.odu.edu/~jhh/ch45.PDF`)

[7] Classical Mechanics (Goldstein, Safko & Poole, Pearson New International Edition)

[8] Elementary Symplectic Topology and Mechanics (Franco Cardin, Springer)

[9] Field Quantization (Walter Greiner & Joachim Reinhardt, Springer)

[10] Gauge Theory (Florin Belgun, `http://www.math.uni-hamburg.de/home/belgun/Gauge4.pdf`)

[11] Invariant Subspaces (H. Radjavi & P. Rosenthal)

[12] A Theorem Of Looman-Menchoff (Donald Carvel Ferguson, `http://digitool.library.mcgill.ca/thesisfile111406.pdf`)

[13] Principal Bundles: The Classical Case (Stephen Bruce Sontz, Springer)

[14] Quantum Mechanics (B.H. Bransden & C.J. Joachain, Prentice Hall)

[15] Quantum Mechanics (Walter Greiner, Springer)

[16] Representation Theory: A First Course (William Fulton & Joe Harris, Springer)

[17] Topology and Geometry for Physicists (Charles Nash & Siddharta Sen, Dover Publications)

[18] Lots and lots of wikipedia.

Index