Compendium of Mathematics & Physics

Nicolas Dewolf

July 22, 2022

Contents

Chapter 1

Calculus

1.1 General definitions

Definition 1.1.1 (Domain). A connected, open subset of $\mathbb{R}^n$. (Not to be confused with the domain of a function **??**.)

Definition 1.1.2 (Factorial).

$$n! := n(n-1)\cdots 1\,, \tag{1.1}$$

where $n \in \mathbb{N}$. The convention is that $0! = 1$. (This for example agrees with the combinatorial result that there is only one way to order zero objects.)

Definition 1.1.3 (Envelope). Consider a set $\mathcal{F}$ of real-valued functions with common domain X. An envelope (function) for $\mathcal{F}$ is any function $F : X \to \mathbb{R}$ such that

$$\forall f \in \mathcal{F}, x \in X : |f(x)| \leq F(x). \tag{1.2}$$

1.2 Sequences

Definition 1.2.1 (Limit superior). Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of real numbers. The limit superior is defined as follows:

$$\limsup_{n\to\infty} x_n := \inf_{n\geq 1} \sup_{k\geq n} x_k. \tag{1.3}$$

Definition 1.2.2 (Limit inferior). Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of real numbers. The limit superior is defined as follows:

$$\liminf_{n\to\infty} x_n := \sup_{n\geq 1} \inf_{k\geq n} x_k. \tag{1.4}$$

Property 1.2.3. A sequence $(x_n)_{n\in\mathbb{N}}$ converges pointwise if and only if

$$\limsup_{n\to\infty} x_n = \liminf_{n\to\infty} x_n. \tag{1.5}$$

1.3 Continuity

Definition 1.3.1 (Darboux function). A function that has the intermediate value property **??**.

Theorem 1.3.2 (Darboux). *Every differentiable function defined on a closed interval has the intermediate value property.*

Corollary 1.3.3 (Bolzano). If $f(a) < 0$ and $f(b) > 0$ (or vice versa), there exists at least one point x_0 for which $f(x_0) = 0$.

Theorem 1.3.4 (Weierstrass's extreme value theorem). *Let $I = [a, b]$ be a closed interval and let $f : I \to \mathbb{R}$ be a continuous function. Then f attains a minimum and maximum at least once on I.*

Definition 1.3.5 (Absolute continuity). A function $f : \mathbb{R} \to \mathbb{R}$ is said to be absolutely continuous if for every $\varepsilon > 0$ there exists a $\delta_\varepsilon > 0$ such that for every finite collection of disjoint intervals $]x_i, y_i[$ satisfying

$$\sum_i (y_i - x_i) < \delta_\varepsilon\,, \tag{1.6}$$

the function f satisfies

$$\sum_i |f(y_i) - f(x_i)| < \varepsilon. \tag{1.7}$$

Property 1.3.6. The different types of continuity form the following hierarchy:

Lipschitz-continuous $\subset$ absolutely continuous $\subset$ uniformly continuous $\subset$ continuous.

Definition 1.3.7 (Function of bounded variation). A function f is said to be of bounded variation on the interval $[a, b]$ if the following quantity is finite:

$$V_{a,b}(f) := \sup_{P\in\mathcal{P}} \sum_{i=0}^{|P|-1} |f(x_{i+1}) - f(x_i)|\,, \tag{1.8}$$

where the supremum is taken over all partitions of $[a, b]$.

Property 1.3.8. Every function of bounded variation can be decomposed as the difference of two monotonically increasing functions.

Example 1.3.9. Every absolutely continuous function is of bounded variation.

1.4 Convergence

Definition 1.4.1 (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 if

$$\forall x \in \mathrm{dom}(f_n) : \lim_{n\to\infty} f_n(x) = f(x). \tag{1.9}$$

Definition 1.4.2 (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 if

$$\lim_{n\to\infty} \sup_{x\in\mathrm{dom}(f_n)} |f_n(x) - f(x)| = 0. \tag{1.10}$$

1.5 Series

1.5.1 Convergence tests

Property 1.5.1. A necessary condition for the convergence of a series $\sum_{i=1}^{\infty} a_i$ is that

$$\lim_{n\to\infty} a_n = 0. \tag{1.11}$$

Property 1.5.2 (Absolute/conditional convergence). If $S' = \sum_{i=1}^{\infty} |a_i|$ converges, so does $S = \sum_{i=1}^{\infty} a_i$. In this case, S is said to be absolutely convergent. If S converges but S' does not, S is said to be conditionally convergent.

Definition 1.5.3 (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 sums satisfy $S_{a,k} \geq S_{b,k}$, i.e.

$$\sum_{i=1}^{k} a_i \geq \sum_{i=1}^{k} b_i \tag{1.12}$$

for all $k \in \mathbb{N}$.

Method 1.5.4 (Comparison test). Let S_a, S_b be two series such that S_a majorizes S_b.

- 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 1.5.5 (Maclaurin-Cauchy integral test). Let f be a nonnegative, continuous and monotonically decreasing function defined on the interval $[n, \infty[$ for some $n \in \mathbb{N}$. If $\int_n^{\infty} f(x)\,dx$ is convergent, so is $\sum_{k=n}^{\infty} f(k)$. On the other hand, if the integral is divergent, so is the series.

Remark 1.5.6. The function does not have to be nonnegative and decreasing on the complete interval. As long as it does on the interval $[N, \infty[$ for some $N \geq n$, the statement holds. 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 (and similarly for the integral).

Property 1.5.7. If the integral in the previous theorem converges, the series is bounded in the following way:

$$\int_n^{\infty} f(x)dx \leq \sum_{i=n}^{\infty} a_i \leq f(n) + \int_n^{\infty} f(x)dx. \tag{1.13}$$

Method 1.5.8 (d'Alembert's ratio test). Consider the quantity

$$R := \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|. \tag{1.14}$$

The following cases can be distinguished:

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

Method 1.5.9 (Cauchy's root test). Consider the quantity

$$R := \limsup_{n\to\infty} \sqrt[n]{|a_n|}. \tag{1.15}$$

The following cases can be distinguished:

- $R < 1$: 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 1.5.10 (Radius of convergences). The number $1/R$ is called the radius of convergence.

Remark 1.5.11. The root test is stronger than the ratio test. However, if the ratio test can determine the convergence of a series, the radius of convergence of both tests will coincide and, hence, it is a well-defined quantity.

Method 1.5.12 (Gauss's test). If $a_n > 0$ for all $n \in \mathbb{N}$, one can write the ratio of successive terms as follows:

$$\left|\frac{a_n}{a_{n+1}}\right| = 1 + \frac{h}{n} + \frac{B(n)}{n^k}, \tag{1.16}$$

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

Definition 1.5.13 (Asymptotic expansion). Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function. A series $\sum_{i=0}^{\infty} a_n x^n$ is called an asymptotic expansion of f if there exists an $n \in \mathbb{N}$ such that

$$f(x) - \sum_{i=0}^{n} a_i x^i = O(x^{n+1}) \tag{1.17}$$

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

1.6 Differentiation

Formula 1.6.1 (Derivative). Consider a function $f : \mathbb{R} \to \mathbb{R}$. If it exists, the following limit is called the derivative of f at $x \in \mathbb{R}$:

$$\frac{df}{dx} \equiv f'(x) := \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}. \tag{1.18}$$

If the derivative exists at every point of some interval I, then f is said to be differentiable on I. For multivariate functions $f : \mathbb{R}^n \to \mathbb{R}$ one can similarly define the partial derivatives:

$$\frac{\partial f}{\partial x_i} := \frac{f(x + he_i) - f(x)}{h}, \tag{1.19}$$

where e_i is the i^{th} coordinate vector, i.e. the partial derivatives determine the rate of change in the coordinate directions.

Theorem 1.6.2 (Mean value theorem). *Let f be a continuous function defined on the closed interval $[a, b]$ and differentiable on the open interval $]a, b[$. There exists a point $c \in\ ]a, b[$ such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}. \tag{1.20}$$

Definition 1.6.3 (Differentiablity class). A function $f : \mathbb{R} \to \mathbb{R}$ is said to be of class C^n if it is n times continuously differentiable, i.e. $f^{(i)}$ exists and is continuous for $i = 1, \ldots, n$. Multivariate functions are said to be of class C^n if all of their partial derivatives are.

Definition 1.6.4 (Smooth function). A function f is said to be smooth if it is of class C^∞.

Theorem 1.6.5 (Boman). *Consider a function $f : \mathbb{R}^d \to \mathbb{R}$. If for every smooth function $g : \mathbb{R} \to \mathbb{R}^d$ the composition $f \circ g$ is smooth, the function f is also smooth.*

Property 1.6.6 (Taylor expansion). Let $f : \mathbb{R} \to \mathbb{R}$ be a smooth function. Around every point $x \in \mathbb{R}$ one can express f as the following series:

$$f(y) = f(x) + f'(x)(y-x) + \frac{f''(x)}{2}(y-x)^2 + \cdots = \sum_{n=0}^{\infty} \frac{f^{(n)}(x)}{n!}(y-x)^n. \tag{1.21}$$

For the special case $x = 0$ the name **Maclaurin series** is sometimes used. A similar expression exists for multivariate functions, where derivatives are replaced by partial derivatives.

Definition 1.6.7 (Analytic function). A function f is said to be analytic if it is smooth and if its Taylor series expansion around any point x converges to f in some neighbourhood of x. The set of analytic functions defined on V is denoted by $C^\omega(V)$.

Theorem 1.6.8 (Hadamard lemma). *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a smooth function defined on an open, star-convex set U. One can expand the function as follows:*

$$f(x) = f(x_0) + \sum_{i=1}^{n} (x^i - x_0^i) g_i(x_0)\,, \tag{1.22}$$

where all functions g_i are also smooth on U.

From this expression one can also see that the functions g_i, evaluated at 0, give the partial derivatives of f. These functions are sometimes called the **Hadamard quotients**.

Remark 1.6.9. This lemma gives a finite order approximation of the Taylor expansion of f.

Theorem 1.6.10 (Schwarz[1]). *Consider a twice differentibale function $f \in C^2(\mathbb{R}^n, \mathbb{R})$. The mixed partial derivatives of f coincide for all indices $i, j \leq n$:*

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right) = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right). \tag{1.23}$$

Formula 1.6.11 (Derivative of $f(x)^{g(x)}$). Consider a function of the form

$$u(x) = f(x)^{g(x)}\,,$$

with $f, g : \mathbb{R} \to \mathbb{R}$ differentiable. After taking the logarithm and applying the standard rules of differentiation, one can obtain the following expression:

$$\frac{d}{dx}\left(f(x)^{g(x)}\right) = f(x)^{g(x)}\left(\frac{dg}{dx}(x)\ln[f(x)] + \frac{g(x)}{f(x)}\frac{df}{dx}(x)\right). \tag{1.24}$$

Definition 1.6.12 (Euler operator). On the space $C^{n>1}(\mathbb{R}^n, \mathbb{R})$, the Euler operator $\mathbb{E}$ is defined as follows:

$$\mathbb{E} := \sum_{i=1}^{n} x_i \frac{\partial}{\partial x^i}. \tag{1.25}$$

Theorem 1.6.13 (Euler). *Let f be a homogeneous function, i.e.*

$$f(\lambda x_1, \ldots, \lambda x_n) = \lambda^n f(x_1, \ldots, x_n). \tag{1.26}$$

This function satisfies the following equality:

$$\mathbb{E}(f) = nf(x_1, \ldots, x_n). \tag{1.27}$$

[1] Also called **Clairaut's theorem**.

1.7 Integration theory

Definition 1.7.1 (Improper Riemann integral).

$$\int_{-\infty}^{+\infty} f(x)\,dx := \lim_{\substack{a\to-\infty\\ b\to+\infty}} \int_a^b f(x)\,dx \tag{1.28}$$

One-sided improper integrals are defined in a similar fashion.

Theorem 1.7.2 (First fundamental theorem of calculus). *Let f be a continuous function defined on an open interval I and consider any number $c \in I$. The following expression establishes the relation between integration and differentiation:*

$$\exists F \in C^1(I) : \forall x \in I : F'(x) = f(x). \tag{1.29}$$

Furthermore, the function F is uniformly continuous on I and is given by the following integral:

$$F(x) = \int_c^x f(x')\,dx'. \tag{1.30}$$

Remark 1.7.3. The function F in the previous theorem is called a **primitive (function)** of f. Remark that F is just "a" primitive function, since adding a constant to F does not change anything because the derivative of a constant is zero (the number $c \in \mathbb{R}$ was arbitrary).

Theorem 1.7.4 (Second fundamental theorem of calculus). *Let f be a C^1-function on the interval $[a, b]$.*

$$\int_a^b f'(x)\,dx = f(b) - f(a) \tag{1.31}$$

Formula 1.7.5 (Differentiation under the integral sign[2]).

$$\frac{d}{dx}\int_{a(x)}^{b(x)} f(x,y)\,dy = f\big(x,b(x)\big)b'(x) - f\big(x,a(x)\big)a'(x) + \int_{a(x)}^{b(x)} \frac{\partial f(x,y)}{\partial x}\,dy \tag{1.32}$$

Definition 1.7.6 (Borel transform). Consider the following function:

$$F(x) := \sum_{n=0}^{\infty} \frac{a_n}{n!} x^n. \tag{1.33}$$

If

$$\int_0^{+\infty} e^{-t} F(xt)\,dt < +\infty \tag{1.34}$$

for all $x \in \mathbb{R}$, then F is called the Borel transform of

$$f(x) = \sum_{n=0}^{\infty} a_n x^n. \tag{1.35}$$

Furthermore, the integral gives a convergent expression for f.

[2] This is a more general version of the *Leibniz integral rule*.

Proof. The function F is defined as follows:

$$F(x) := \sum_{n=0}^{+\infty} \frac{a_n}{n!} x^n. \tag{1.36}$$

The Borel transform gives:

$$\begin{aligned}
\int_0^{+\infty} F(xt)e^{-t}\,dt &= \sum_{n=0}^{N} \int_0^{+\infty} \frac{a_n}{n!} x^n t^n e^{-t}\,dt \\
&= \sum_{n=0}^{N} \frac{a_n}{n!} x^n \int_0^{+\infty} t^n e^{-t}\,dt \\
&= \sum_{n=0}^{N} \frac{a_n}{n!} x^n\, \Gamma(n+1) \\
&= \sum_{n=0}^{N} a_n x^n,
\end{aligned}$$

where the definition of the Gamma function 1.7.9 was used on line 3 and the relation (1.41) between the factorial function and the Gamma function was used on line 4. □

Theorem 1.7.7 (Watson). *The Borel transform F is unique if the function f is holomorphic (Definition* **??***) on the domain $\{z \in \mathbb{C} \mid |\arg(z)| < \frac{\pi}{2} + \varepsilon\}$.*

1.7.1 Euler integrals

Formula 1.7.8 (Beta function). The beta function (also known as the **Euler integral of the first kind**) is defined as follows:

$$B(x, y) := \int_0^1 t^{x-1}(1-t)^{y-1}\,dt. \tag{1.37}$$

Formula 1.7.9 (Gamma function). The gamma function (also known as the **Euler integral of the second kind**) is defined as follows:

$$\Gamma(x) := \int_0^{+\infty} t^{x-1} e^{-t}\,dt. \tag{1.38}$$

Formula 1.7.10. The following formula relates the beta and gamma functions:

$$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}. \tag{1.39}$$

Property 1.7.11 (Recursion). The gamma function satisfies the following recursion relation for all points x in its domain:

$$\Gamma(x+1) = z\Gamma(x). \tag{1.40}$$

Formula 1.7.12 (Factorial). For integers $n \in \mathbb{N}$ the gamma function can be expressed in terms of the factorial 1.1.2:

$$\Gamma(n) = (n-1)!\,. \tag{1.41}$$

Formula 1.7.13 (Stirling). This formula (originally stated for the factorial of natural numbers) gives an asymptotic expansion of the gamma function:

$$\ln\Gamma(z) \approx z\ln z - z + \frac{1}{2}\ln\left(\frac{2\pi}{z}\right). \tag{1.42}$$

1.7.2 Gaussian integrals

Formula 1.7.14 (n-dimensional Gaussian integral). An integral of the form

$$I\left(A,\vec{\boldsymbol{b}}\right) := \int_{\mathbb{R}^n} \exp\left(-\frac{1}{2}\vec{\boldsymbol{x}}\cdot A\vec{\boldsymbol{x}} + \vec{\boldsymbol{b}}\cdot\vec{\boldsymbol{x}}\right) d^n x\,, \tag{1.43}$$

where A is a real symmetric matrix. By performing the transformation $\vec{\boldsymbol{x}} \to A^{-1}\vec{\boldsymbol{b}} - \vec{\boldsymbol{x}}$ and diagonalizing A, one can obtain the following expression:

$$I\left(A,\vec{\boldsymbol{b}}\right) = \sqrt{\frac{(2\pi)^n}{\det(A)}}\exp\left(\frac{1}{2}\vec{\boldsymbol{b}}\cdot A^{-1}\vec{\boldsymbol{b}}\right). \tag{1.44}$$

More generally one has the following result:

$$\int_{\mathbb{R}^n} \exp\left(-\frac{1}{2}\vec{\boldsymbol{x}}\cdot A\vec{\boldsymbol{x}}\right) f(\vec{\boldsymbol{x}})\, d^n x = \sqrt{\frac{(2\pi)^n}{\det(A)}}\exp\left(\frac{1}{2}\sum_{i,j=1}^{n} A^{-1}_{ij}\partial_i\partial_j\right) f(\vec{\boldsymbol{x}})\Bigg|_{\vec{\boldsymbol{x}}=0}. \tag{1.45}$$

This result is sometimes called **Wick's lemma**.

Corollary 1.7.15. A functional generalization is given by:

$$\begin{aligned} I(iA,iJ) &= \int \exp\left(-i\int_{\mathbb{R}^n\times\mathbb{R}^n} \varphi(x)A(x,y)\varphi(y)\, d^n x\, d^n y + i\int_{\mathbb{R}^n}\varphi(x)J(x)\, d^n x\right)[d\varphi] \\ &= C\det(A)^{-1/2}\exp\left(\frac{i}{2}\int_{\mathbb{R}^n\times\mathbb{R}^n} J(x)A^{-1}(x,y)J(y)\, d^n x\, d^n y\right), \end{aligned} \tag{1.46}$$

where the analytic continuation $I(iA,iJ)$ of Equation (1.44) was used. One should pay attention to the normalization factor C which is infinite in general.

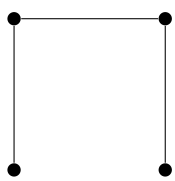
Method 1.7.16 (Feynman diagrams). The expansion of the exponential in the general expression for Gaussian integrals admits a diagrammatic expression. Let $f(\vec{\boldsymbol{x}})$ be a polynomial function of the coordinates.

If the number of factors in a monomial is odd, the resulting integral will vanish (since the integral of an odd function over an even domain is zero). For an even number of factors one gets the following expression:

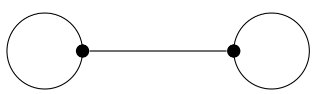
$$\int_{\mathbb{R}^n} \exp\left(-\frac{1}{2}\vec{\boldsymbol{x}}\cdot A\vec{\boldsymbol{x}}\right) x^{i_1}\cdots x^{i_k}\, d^n x = \sqrt{\frac{(2\pi)^n}{\det(A)}}\sum_{\sigma\in S_k} A^{-1}_{\sigma(i_1)\sigma(i_2)}\cdots A^{-1}_{\sigma(i_{k-1})\sigma(i_k)}. \tag{1.47}$$

To every coordinate dimension one can assign a vertex in the place, i.e. the object x_i can be interpreted as a real-valued function on the set of k elements. The sum on the right-hand side above can then be expressed as a "sum" over all possible diagrams, where a factor A^{-1}_{ij} is represented by a line connecting the vertices i and j.

Example 1.7.17 (Feynman diagrams). Some simple examples are given:

$A^{-1}_{13}A^{-1}_{12}A^{-1}_{24} =$ 

Higher powers of a given coordinate would then for example give rise to diagrams with loops at a given vertex:

$A^{-1}_{11}A^{-1}_{12}A^{-1}_{22} =$ 

Remark 1.7.18 (Normalization). In practice one often divides all Gaussian integrals by the quantity $I(A,0)$ to cancel the normalization factor. In the functional setting this even imperative since, as mentioned above, the normalization factor diverges for infinite-dimensional spaces.

1.8 Convexity

Definition 1.8.1 (Convex set). A subset of X of a vector space V (Definition 2.1.1) is said to be convex if $x, y \in X$ implies that $\{\lambda x + (1-\lambda)y \mid \lambda \in [0,1]\} \subset X$, i.e. if all straight lines connecting elements of the set are completely contained in that set. The **convex hull** of a subset X is defined as the smallest convex subset containing X.

Definition 1.8.2 (Convex function). Let X be a convex set. A function $f : X \to \mathbb{R}$ is said to be convex if for all $x, y \in X$ and $\lambda \in [0,1]$:

$$f\big(\lambda x + (1-\lambda)y\big) \leq t\lambda(x) + (1-\lambda)f(y). \tag{1.48}$$

For the definition of a **concave** function the inequality has to be turned around.

Definition 1.8.3 (Linear map). A function $f : X \to \mathbb{R}$ is linear if and only if it is both convex and concave.

Theorem 1.8.4 (Karamata's inequality). *Consider an interval $I \subset \mathbb{R}$ and let $f : I \to \mathbb{R}$ be a convex function. If $(x_1, \ldots, x_n)$ is a tuple that majorizes $(y_1, \ldots, y_n)$, i.e.*

$$\sum_{i=1}^n x_i = \sum_{i=1}^n y_i \tag{1.49}$$

and

$$x_{(1)} + \cdots + x_{(k)} \geq y_{(1)} + \cdots + y_{(k)} \tag{1.50}$$

for all $k \leq n$, where $x_{(i)}$ denotes the i^{th} largest element of $(x_1, \ldots, x_n)$, then

$$\sum_{i=1}^n f(x_i) \geq \sum_{i=1}^n f(y_i). \tag{1.51}$$

The following inequality can be derived directly from the definition of convexity by induction:

Theorem 1.8.5 (Jensen's inequality). *Let f be a convex function and consider a point $\{a_i\}_{i\leq n}$ in the probability simplex Δ^n (Definition **??**).*

$$f\left(\sum_{i=1}^n a_i x_i\right) \leq \sum_{i=1}^n a_i f(x_i). \tag{1.52}$$

Definition 1.8.6 (Legendre transformation). Consider a function $f : \mathbb{R} \to \mathbb{R}$. In certain cases (especially in physics) it is sometimes useful to replace the argument x by the slope of f at x, i.e. to perform the transformation

$$x \longrightarrow f'(x). \tag{1.53}$$

However, it should be clear that this transformation is not always well-defined and, even if it is, it does not always preserve all the information contained in f.

These conditions are satisfied exactly if f is convex (or concave). In this case the Legendre transform of f is defined as

$$f^*(x^*) := \sup_x \big(x^*x - f(x)\big). \tag{1.54}$$

Now, consider the case where f is differentiable. The above supremum can then be obtained by differentiating the right-hand side and equating it to zero. This results in $x^* = f'(x)$, which is exactly the transformation that was required. By expressing everything in terms of the Legendre tranformed quantity x^*, one can also find the derivative of f^*:

$$\frac{df^*}{dx^*}(x^*) = x(x^*). \tag{1.55}$$

Property 1.8.7 (Alternative characterization). In fact, up to an additive constant, the condition

$$(f^*)' = (f')^{-1} \tag{1.56}$$

uniquely determines the Legendre transformation.

Remark 1.8.8. The above definitions can easily be extended to higher dimensions ($n \geq 2$).

Chapter 2

Linear Algebra

2.1 Vector spaces

Definition 2.1.1 (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 axioms:

1. V forms an Abelian group under vector addition.
2. Scalar multiplication is associative: $\lambda(\mu v) = (\lambda\mu)v$ for all $\lambda, \mu \in K$ and $v \in V$.
3. The identity of the field K acts as a neutral element for scalar multiplication: $1_K v = v$ for all $v \in V$.
4. Scalar multiplication is distributive with respect to vector addition: $\lambda(v + w) = \lambda v + \lambda w$ for all $\lambda \in K$ and $v, w \in V$.
5. Vector addition is distributive with respect to scalar multiplication: $(\lambda + \kappa)v = \lambda v + \kappa w$ for all $\lambda, \kappa \in K$ and $v \in V$.

From here on the underlying field K will be left implicit unless the results depend on it.

Remark 2.1.2. The above definition can be restated in abstract algebraic terms. A K-vector space is a module **??** over K.

2.1.1 Linear independence

Definition 2.1.3 (Linear combination). The vector w is a linear combination of elements in the set $\{v_i\}_{i\leq n} \subset V$ if it can be written as

$$w = \sum_{i=1}^{n} \lambda_i v_i \tag{2.1}$$

for some $\{\lambda_i\}_{i\leq n} \subset K$. One can generalize this to general subsets $S \subseteq V$, but the number of nonzero elements λ_i is always required to be finite.[1] (See the remark about *Hamel bases* in next section.)

[1] Generalizations are possible in the context of topological vector spaces (see Chapters **??** and **??**), where one can define the notion of convergence.

Definition 2.1.4 (Linear independence). A finite set $\{v_i\}_{i\leq n}$ is said to be linearly independent if the following relation holds:

$$\sum_{i=1}^{n} \lambda_i v_i = 0 \iff \forall i \leq n : \lambda_i = 0. \tag{2.2}$$

A general set $S \subset V$ is said to be linearly independent if every finite subset of it is linearly independent.

Definition 2.1.5 (Span). A set of vectors $S \subseteq V$ is said to span V if every vector $v \in V$ can be written as a linear combination of elements in S.

Definition 2.1.6 (Frame). A k-frame is an ordered set of k linearly independent vectors.

2.1.2 Bases

Definition 2.1.7 (Basis). A subset $\mathcal{B} \subset V$ that is linearly independent and spans V.

Property 2.1.8. Every spanning set contains a basis.

Remark 2.1.9 (Hamel basis). In the previous definition the concept of a Hamel basis was implicitly used. This concept is based on two conditions:

1. The basis is linearly independent.
2. Every element in the vector space can be written as a linear combination of a finite subset of the basis.

For bases consisting of a finite number of vectors, one does not have to worry. However, for infinite bases one has to keep this in mind. An alternative construction that allows for combinations of a countably infinite number of elements, is given by that of a *Schauder basis*.

Nonetheless, it can be shown that every vector space admits a Hamel basis:

Construction 2.1.10 (Hamel basis ♣). Let V be a vector space and consider the set of all linearly independent subsets of V. Under the relation of inclusion this set becomes a partially ordered set **??**. Zorn's lemma **??** then says that there exists at least one maximal linearly independent set.

Now, one can show that this maximal subset S is also a spanning set of V. Choose a vector $v \in V$ that is not already in S. From the maximality of S it follows that $S \cup v$ is linearly dependent and, hence, there exists a finite sequence of scalars $(a^1, \ldots, a^n, b)$ and a finite sequence of elements $(e_1, \ldots, e_n)$ in S such that:

$$\sum_{i=0}^{n} a^i e_i + bv = 0, \tag{2.3}$$

where not all scalars are zero. This implies that $b \neq 0$, because otherwise the set $\{e_i\}_{i\leq n}$ and, hence, also S would be linearly dependent. It follows that v can be written as[2]

$$v = -\frac{1}{b}\sum_{i=0}^{n} a^i e_i. \tag{2.4}$$

Because v was randomly chosen, one can conclude that S is a spanning set for V.

[2]It is this step that requires R to be a division ring in Property **??** because otherwise one would in general not be able to divide by $b \in R$.

Remark. This construction assumes the axiom of choice in 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.

Property 2.1.11. Every basis of a vector space has the same number of elements. For infinite-dimensional spaces this means that all bases have the same *cardinality*.

Definition 2.1.12 (Dimension). Let V be a finite-dimensional vector space and let $\mathcal{B}$ be a basis for V with n elements. With the previous property in mind, the dimension of V is defined as follows:

$$\dim(V) := n. \tag{2.5}$$

Definition 2.1.13 (Subspace). Let V be a vector space. A subset W of V is called a subspace if W is itself a vector space under (the restriction of) the operations of V:

$$W \leq V \iff \forall w_1, w_2 \in W, \forall \lambda \in K : \lambda w_1 + w_2 \in W. \tag{2.6}$$

2.1.3 Sum and direct sum

Definition 2.1.14 (Sum). Let V be a vector space and consider a finite collection of subspaces $\{W_1, \dots, W_k\}$. The sum of these subspaces is defined as follows:

$$W_1 + \cdots + W_k := \left\{ \sum_{i=1}^k w_i \,\middle|\, w_i \in W_i \right\}. \tag{2.7}$$

For an infinite collection of subspaces the linear combinations have to be finite.

Definition 2.1.15 (Direct sum). If every element v of the sum can be written as a unique linear combination, the sum is called a direct sum.

Notation 2.1.16 (Direct sum). The direct sum of vector spaces is denoted by

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

Formula 2.1.17. Let V be a finite-dimensional vector space and consider two subspaces $W_1, W_2 \leq V$. The dimensions of these spaces can be related in the following way:

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2). \tag{2.8}$$

Property 2.1.18. Let V be a vector space and assume that V can be decomposed as $W = W_1 \oplus W_2$. If $\mathcal{B}_1$ is a basis of W_1 and if $\mathcal{B}_2$ is a basis of W_2, then $\mathcal{B}_1 \cup \mathcal{B}_2$ is a basis of W.

Definition 2.1.19 (Complement). Let V be a vector space and let W be a subspace of V. A subspace W' of V is called a complement of W if $V = W \oplus W'$.

Property 2.1.20 (Existence of complements). Let V be a vector space and let U, W be two subspaces of V. If $V = U + W$, there exists a subspace $Y \leq U$ such that $V = Y \oplus W$. In particular, every subspace of V has a complement in V.

2.2 Linear maps

Definition 2.2.1 (Linear map). Let V, W be K-vector spaces. A function $f : V \to W$ is said to be linear if $f(\lambda v + w) = \lambda f(v) + f(w)$ for all $\lambda \in K$ and $v, w \in V$.

Remark 2.2.2. Linear maps are also called **linear transformations** or **linear mappings**.

2.2.1 Homomorphisms

Definition 2.2.3 (Homomorphism space). Let V, W be two vector spaces. The set of all linear maps between V and W is called the homomorphism space from V to W:

$$\mathrm{Hom}_K(V, W) := \big\{f : V \to W \,\big|\, f \text{ is linear}\big\}. \tag{2.9}$$

The collection of K-vector spaces and linear maps between them form a category $\mathbf{Vect}_K$.

Formula 2.2.4. Let V, W be two finite-dimensional vector spaces.

$$\dim\big(\mathrm{Hom}_K(V, W)\big) = \dim(V)\dim(W) \tag{2.10}$$

Definition 2.2.5 (Endomorphism ring). The space $\mathrm{Hom}_K(V, V)$ with composition of maps as multiplication forms a ring, the endomorphism ring. It is denoted by $\mathrm{End}_K(V)$ or $\mathrm{End}(V)$ when the underlying field is clear.

Property 2.2.6 (Commutator). The endomorphism ring $\mathrm{End}(V)$ can also be endowed with the structure of a Lie algebra (see Property **??**) by equipping it with the commutator

$$[A, B] := A \circ B - B \circ A. \tag{2.11}$$

Property 2.2.7. Let V be finite-dimensional vector space and let $f : V \to V$ be an endomorphism. The following statements are equivalent:

- f is injective.
- f is surjective.
- f is bijective.

Definition 2.2.8 (Automorphism). An isomorphism from V to V is called an automorphism. The set of all automorphisms on V is denoted by $\mathrm{Aut}(V)$. It forms a group under composition. Often this group is called the general linear group[3] $\mathrm{GL}_K(V)$ or $\mathrm{GL}(V)$ when the underlying field is clear.

Remark 2.2.9. Sometimes automorphisms are also called **linear operators**. However, this terminology is also used for a general linear map in operator theory (Chapter **??**) and so this terminology is not adopted in this text.

Definition 2.2.10 (Kernel). Consider a linear map $f : V \to W$. The kernel of f is defined as the following subspace of V:

$$\ker(f) := \{v \in V \mid f(v) = 0\}. \tag{2.12}$$

Property 2.2.11. A linear map $f : V \to W$ is injective if and only if $\ker(f) = 0$.

Definition 2.2.12 (Rank). The dimension of the image of a linear map.

Definition 2.2.13 (Nullity). The dimension of the kernel of a linear map.

Theorem 2.2.14 (Dimension theorem[4]). *Let $f : V \to W$ be a linear map.*

$$\dim(\mathrm{im}(f)) + \dim(\ker(f)) = \dim(V) \tag{2.13}$$

Corollary 2.2.15. Two finite-dimensional vector spaces are isomorphic if and only if they have the same dimension.

[3]It is isomorphic to the general linear group of invertible matrices 2.4.3 (hence the similar name and notation).
[4]Also called the **rank-nullity theorem**.

Definition 2.2.16 (Minimal polynomial). Let $f \in \mathrm{End}(V)$ with V a finite-dimensional vector space. The monic polynomial μ_f of lowest order such that $\mu_f(f) = 0$ is called the minimal polynomial of f.

Property 2.2.17. Let $f \in \mathrm{End}(V)$ with minimal polynomial μ_f. If $\varphi(f) = 0$ for some polynomial φ, the minimal polynomial μ_f divides φ.

Property 2.2.18 (Jordan-Chevalley decomposition). Every endomorphism A can be decomposed as follows:

$$A = A_{ss} + A_n, \tag{2.14}$$

where

- A_{ss} is **semisimple**: for every invariant subspace of A_{ss} there exists an invariant complementary subspace.
- A_n is **nilpotent**: $\exists k \in \mathbb{N} : A_n^k = 0$.

Furthermore, this decomposition is unique and the endomorphisms A_{ss}, A_n can be written as polynomials in A.

2.2.2 Dual maps

Definition 2.2.19 (Dual space). Let V be a vector space. The (algebraic) dual V^* of V is defined as the following vector space:

$$V^* := \mathrm{Hom}_K(V, K) = \big\{f : V \to K \,\big|\, f \text{ is linear}\big\}. \tag{2.15}$$

The elements of V^* are called **linear forms** or (linear) **functionals**.

Property 2.2.20 (Dimension). From Theorem 2.2.4 it follows that $\dim(V^*) = \dim(V)$ whenever V is finite-dimensional. If V is infinite-dimensional, this property is never valid. In the infinite-dimensional case $\mathrm{card}(V^*) > \mathrm{card}(V)$ always holds.

Definition 2.2.21 (Dual basis). Let $\mathcal{B} = \{e_1, e_2, \dots, e_n\}$ be a basis for a finite-dimensional vector space V. One can construct a basis $\mathcal{B}^* = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$ for V^*, called the dual basis of $\mathcal{B}$, as follows:

$$\varepsilon_i : \sum_{j=1}^{n} a_i e_i \mapsto a_i. \tag{2.16}$$

The relation between a basis and its associated dual basis can also be expressed as

$$\varepsilon^i(e_j) = \delta^i_j. \tag{2.17}$$

Definition 2.2.22 (Natural pairing). The definition of the dual basis extends to a natural pairing of V and its dual V^* in terms of the following bilinear map:

$$\langle v, v^* \rangle := v^*(v). \tag{2.18}$$

(See Definition **??** for a generalization of this map.)

Definition 2.2.23 (Dual map). Let $f : V \to W$ be a linear map. The linear map

$$f^* : W^* \to V^* : \varphi \to \varphi \circ f \tag{2.19}$$

is called the dual map or **transpose** of f. It is also often denoted by f^T.

2.3 Inner product

In this section all vector spaces V will be defined over $\mathbb{R}$ or $\mathbb{C}$.

2.3.1 Inner product space

Definition 2.3.1 (Inner product). A function $\langle\cdot|\cdot\rangle : V \times V \to \mathbb{C}$ is called an inner product on V if it satisfies the following properties for all $u, v, w \in V$ and $\lambda \in \mathbb{C}$:

1. **Conjugate symmetry**: $\langle v|w\rangle = \langle w|v\rangle^*$,
2. **Linearity in the second argument**: $\langle u|\lambda v + w\rangle = \lambda\langle u|v\rangle + \langle u|w\rangle$,
3. **Nondegeneracy**: $\langle v|v\rangle = 0 \iff v = 0$, and
4. **Positive-definiteness**: $\langle v|v\rangle \geq 0$.

Remark 2.3.2. Inner products are special cases of **nondegenerate Hermitian forms** which do not satisfy the positive-definiteness property.

Corollary 2.3.3. The first two properties have the result of conjugate linearity in the first argument:

$$\langle \lambda f + \mu g|h\rangle = \overline{\lambda}\langle f|h\rangle + \overline{\mu}\langle g|h\rangle \tag{2.20}$$

Therefore, these two properties together are often combined into a **sesquilinearity** axiom. When the underlying field is restricted to $\mathbb{R}$, such that the conjugate symmetry property is replaced by proper symmetry, the inner product becomes a bilinear form.

Definition 2.3.4 (Inner product space). A vector space equipped with an inner product $\langle\cdot|\cdot\rangle$. This is sometimes called a **pre-Hilbert space**.

Definition 2.3.5 (Metric dual). Using the inner product (or any other nondegenerate Hermitian form) one can define the metric dual of a vector by the following map:

$$L : V \to V^* : v \mapsto \langle v|\cdot\rangle. \tag{2.21}$$

(See Equation (**??**) for a generalization.) If the sesquilinearity condition would have been stated in the reversed convention, i.e. conjugate linearity in the second argument, metric duals would be conjugate linear and, hence, would not be proper elements of the dual space.

Definition 2.3.6 (Adjoint map). Let A be a linear map on V. The (**Hermitian**) adjoint of A is defined as the linear map $A^\dagger$ that satisfies

$$\langle A^\dagger v|w\rangle = \langle v|Aw\rangle \tag{2.22}$$

for all $v, w \in V$. Alternatively, one can define the adjoint using the transpose and metric dual as follows:

$$A^\dagger = L^{-1} \circ A^* \circ L. \tag{2.23}$$

If $A = A^\dagger$, A is said to be **Hermitian** or **self-adjoint**. (In Chapter **??** a distinction will be made between these two notions.)

2.3.2 Orthogonality

Definition 2.3.7 (Orthogonal). Consider two vectors $v, w \in V$ in an inner product space. These vectors are said to be orthogonal, denoted by $v \perp w$, if they obey the following relation:

$$\langle v|w\rangle = 0. \tag{2.24}$$

An **orthogonal system** is a collection of vectors, none of them equal to 0, that are mutually orthogonal.

Property 2.3.8. Orthogonal systems are linearly independent.

Definition 2.3.9 (Orthonormal). A set of vectors S is said to be orthonormal if it forms an orthogonal system and if all the elements $v \in S$ obey the following relation:

$$\langle v|v\rangle = 1. \tag{2.25}$$

Definition 2.3.10 (Orthogonal complement). Let W be a subspace of an inner product space V. The orthogonal complement of W is defined as the following subspace:

$$W^\perp := \{v \in V \mid \forall w \in W : \langle v|w\rangle = 0\}. \tag{2.26}$$

Remark. $W^\perp$ is pronounced as "W-perp".

Property 2.3.11 (Complements). Let V be a finite-dimensional inner product space. The orthogonal complement $W^\perp$ is a complementary subspace to W, i.e. $W \oplus W^\perp = V$.

Corollary 2.3.12. Let $W \leq V$ with V a finite-dimensional inner product vector space. Forming orthogonal complements defines an involution:

$$(W^\perp)^\perp = W. \tag{2.27}$$

Definition 2.3.13 (Orthogonal projection). Let V be a finite-dimensional inner product vector space and consider a subspace $W \leq V$. Consider a vector $w \in W$ and let $\{w_1, \ldots, w_k\}$ be an orthonormal basis of W. The projections of $v \in V$ on W and $w \in W$ are defined as follows:

$$\mathrm{proj}_W(v) := \sum_{i=1}^{k} \langle v|w_i\rangle w_i \tag{2.28}$$

$$\mathrm{proj}_w(v) := \frac{\langle v|w\rangle}{\langle w|w\rangle} w. \tag{2.29}$$

Property 2.3.14. Orthogonal projections satisfy the following conditions:

$$\forall w \in W : \mathrm{proj}_W(w) = w \qquad \text{and} \qquad \forall u \in W^\perp : \mathrm{proj}_W(u) = 0. \tag{2.30}$$

Method 2.3.15 (Gram-Schmidt orthonormalization). Let $\{u_i\}_{i\leq n}$ be a set of linearly independent vectors. An orthonormal set $\{e_i\}_{i\leq n}$ can be constructed out of $\{u_i\}_{i\leq n}$ using the following procedure:

1. Orthogonalization:

$$\begin{aligned}
w_1 &= u_1 \\
w_2 &= u_2 - \frac{\langle u_2|w_1\rangle}{\|u_2\|^2} w_1 \\
&\vdots \\
w_n &= u_n - \sum_{i=1}^{n-1} \frac{\langle u_n|w_i\rangle}{\|u_n\|^2} w_i
\end{aligned} \tag{2.31}$$

2. Normalization:

$$\begin{aligned} e_1 &= \frac{w_1}{\|w_1\|} \\ e_2 &= \frac{w_2}{\|w_2\|} \\ &\vdots \\ e_n &= \frac{w_n}{\|w_n\|} \end{aligned} \tag{2.32}$$

Definition 2.3.16 (Householder transformation). Let v be an element of an inner product space V. The Householder transformation generated by v is defined as the linear map

$$\sigma_v : V \to V : w \mapsto w - 2\frac{\langle w|v\rangle}{\langle v|v\rangle}v. \tag{2.33}$$

This transformation amounts to a reflection in the hyperplane orthogonal to v.

Definition 2.3.17 (Angle). Let v, w be elements of an inner product space V. The angle θ between v and w is defined by the following formula:

$$\cos\theta := \frac{\langle v|w\rangle}{\|v\|\|w\|}. \tag{2.34}$$

The angle between two vectors v, w is sometimes denoted by $\sphericalangle(v, w)$.

2.4 Matrices

Notation 2.4.1. The vector space of all $m \times n$-matrices defined over the field K is denoted by $M_{m,n}(K)$. If $m = n$, the space is denoted by $M_n(K)$ or $M(n, K)$.

Property 2.4.2 (Dimension). The dimension of $M_{m,n}(K)$ is mn.

Definition 2.4.3 (General linear group). The set of invertible matrices is called the general linear group and is denoted by $\mathrm{GL}_n(K)$ or $\mathrm{GL}(n, K)$.

Property 2.4.4. For all $A \in \mathrm{GL}_n(K)$ one has:

- $A^T \in \mathrm{GL}_n(K)$, and
- $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$.

Definition 2.4.5 (Trace). Let $A \equiv (a_{ij}) \in M_n(K)$. The trace of A is defined as follows:

$$\mathrm{tr}(A) := \sum_{i=1}^n a_{ii}. \tag{2.35}$$

Property 2.4.6. Let $A, B \in M_n(K)$. The trace satisfies the following properties:

- $\mathrm{tr} : M_n(K) \to K$ is a linear map,
- $\mathrm{tr}(AB) = \mathrm{tr}(BA)$, and
- $\mathrm{tr}(A^T) = \mathrm{tr}(A)$.

Formula 2.4.7 (Hilbert-Schmidt norm). The Hilbert-Schmidt (or **Frobenius**) norm is defined by the following formula:

$$\|A\|^2_{HS} := \sum_{i,j} |A_{ij}|^2 = \mathrm{tr}\Big(A^\dagger A\Big). \tag{2.36}$$

If one identifies $M_n(\mathbb{C})$ with $\mathbb{C}^{2n}$, this norm equals the standard Hermitian norm.

Formula 2.4.8 (Hadamard product). The Hadamard product of two matrices is defined as the entry-wise product:

$$(A \circ B)_{ij} := A_{ij}B_{ij}. \tag{2.37}$$

Property 2.4.9. Let $A \in M_{m,n}(K)$. Denote the set of columns as $\{A_1, A_2, \ldots, A_n\}$ and the set of rows as $\{R_1, R_2, \ldots, R_m\}$. The set of columns is a subspace of K^m and the set of rows is a subspace of K^n. Their spans satisfy the following property:

$$\dim(\mathrm{span}(A_1, \ldots, A_n)) = \dim(\mathrm{span}(R_1, \ldots, R_m)). \tag{2.38}$$

Definition 2.4.10 (Rank). Using the invariance relation from the previous property, one can define the rank of a matrix $A \in M_{m,n}(K)$ as follows:

$$\mathrm{rk}(A) := \dim(\mathrm{span}(A_1, \ldots, A_n)) = \dim(\mathrm{span}(R_1, \ldots, R_m)). \tag{2.39}$$

Property 2.4.11. Let $A \in M_{m,n}(K), B \in \mathrm{GL}_n(K), C \in M_{n,r}(K)$ and $D \in M_{r,n}(K)$. The ranks of these matrices satisfy the following properties:

- $\mathrm{rk}(AC) \leq \mathrm{rk}(A)$,
- $\mathrm{rk}(AC) \leq \mathrm{rk}(C)$,
- $\mathrm{rk}(BC) = \mathrm{rk}(C)$, and
- $\mathrm{rk}(DB) = \mathrm{rk}(D)$.

Property 2.4.12. Let $A \in M_{m,n}(K)$. The linear map

$$L_A : K^n \to K^m : v \mapsto Av \tag{2.40}$$

satisfies $\mathrm{im}(L_A) = \mathrm{span}(A_1, \ldots, A_n)$.

2.4.1 System of equations

Property 2.4.13. Let $Ax = b$ with $A \in M_{m,n}(K), x \in K^n$ and $b \in K^m$ be a system of m equations in n variables and let L_A be the linear map as defined in Property (2.4.12). The following properties hold:

- The system is inconsistent if and only if $b \notin \mathrm{im}(L_A)$.
- If the system is not inconsistent, the solution set is an affine space. If $x_0 \in K^n$ is a solution, the solution set is given by: $x_0 + \ker(L_A)$.
- If the system is homogeneous, i.e. $b = 0$, the solution set is equal to $\ker(L_A)$.

Property 2.4.14 (Uniqueness). Let $Ax = b$ with $A \in M_n(K)$ be a system of n equations in n variables. If $\mathrm{rk}(A) = n$, the system has a unique solution.

Formula 2.4.15 (Cramer's rule). Let $Ax = b$ be a system of linear equations where the matrix A has a nonzero determinant. There exists a unique solution:

$$x_i = \frac{\det(A_i)}{\det(A)}, \tag{2.41}$$

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

2.4.2 Coordinates and matrix representations

Definition 2.4.16 (Coordinate vector). Let $\mathcal{B} = \{b_1, \dots, b_n\}$ be a basis of V and consider the vector $v = \sum_{i=1}^n \lambda_i b_i$. The coordinate vector of v with respect to $\mathcal{B}$ is defined as the column vector $(\lambda_1, \dots, \lambda_n)^T$. The scalars λ_i are called the **coordinates** of v with respect to $\mathcal{B}$.

Definition 2.4.17 (Coordinate isomorphism). With the previous definition in mind one can define the coordinate isomorphism induced by $\mathcal{B}$ as follows:

$$\beta : V \to K^n : \sum_{i=1}^n \lambda_i b_i \mapsto (\lambda_1, \dots, \lambda_n)^T. \tag{2.42}$$

Construction 2.4.18 (Matrix representation). Let V, W be m- and n-dimensional vector spaces with bases $\mathcal{B} = \{b_1, \dots, b_m\}, \mathcal{C} = \{c_1, \dots, c_n\}$ and consider a linear map $f : V \to W$. The matrix representation of f with respect to $\mathcal{B}$ and $\mathcal{C}$ is defined as the matrix $A_{f,\mathcal{B},\mathcal{C}}$ that satisfies the following condition for all vectors $v \in V$. Let $(\lambda_1, \dots, \lambda_n)^T$ be the coordinate vector of v with respect to $\mathcal{B}$ and let $(\mu_1, \dots, \mu_m)^T$ be the coordinate vector of $f(v)$ with respect to $\mathcal{C}$, then

$$\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \end{pmatrix} = A_{f,\mathcal{B},\mathcal{C}} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}. \tag{2.43}$$

This matrix can be constructed as follows. For every $j \in \{1, \dots, m\}$, write $f(b_j) = \sum_{i=1}^n a_{ij} c_i$. The matrix $A_{f,\mathcal{B},\mathcal{C}} \equiv (a_{ij}) \in M_{n,m}(K)$ is called the matrix representation of f. The j^{th} column of $A_{f,\mathcal{B},\mathcal{C}}$ coincides with the coordinate vector of $f(b_j)$ with respect to $\mathcal{C}$.

The following property shows that the matrix algebra $M_{m,n}(K)$ is isomorphic to the algebra[5] of linear maps $\mathcal{L}(K^n, K^m)$, thereby explaining why the same notation for the space of invertible matrices 2.4.3 and the space of automorphisms 2.2.8 was used:

Property 2.4.19 (Matrices and linear maps). For every matrix $A \in M_{m,n}(K)$ there exists a linear map $f : K^n \to K^m$ such that $A_{f,\mathcal{B},\mathcal{C}} = A$. Conversely, for every linear map $f : K^m \to K^n$ there exists a matrix $A \in M_{n,m}(K)$ such that $f = L_A$ (given by the previous construction).

Corollary 2.4.20. Let $f \in \mathrm{End}(V)$ and 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,

$$(A_f)^{-1} = A_{f^{-1}}.$$

In other words, the linear isomorphism $\mathrm{End}(V) \to M_n(K)$ descends to a group isomorphism

$$\mathrm{GL}_K(V) \to \mathrm{GL}_n(K) : f \mapsto A_f, \tag{2.44}$$

where $n = \dim(V)$.

Formula 2.4.21 (Linear forms). Let $V \cong K^n$ and consider a linear form $f \in V^*$. Equation (2.43) can be rewritten as

$$f\big((\lambda_1, \dots, \lambda_n)^T\big) = (f(e_1), \dots, f(e_n))(\lambda_1, \dots, \lambda_n)^T = \sum_{i=1}^n f(e_i)\lambda_i, \tag{2.45}$$

where $\{e_i\}_{i \in I}$ is the standard basis of K^n. In terms of the standard dual basis $\{\varepsilon_1, \dots, \varepsilon_n\}$ this becomes:

$$f = \sum_{i=1}^n f(e_i)\varepsilon_i. \tag{2.46}$$

[5] The multiplication is given by the composition of linear maps.

Property 2.4.22 (Transpose). Let $f : V \to W$ be a linear map and 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 $\mathcal{C}$, 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 $\mathcal{B}$.

Corollary 2.4.23. The Hermitian adjoint of a linear map 2.3.6 induces the (Hermitian) adjoint of matrices $A \in \mathbb{C}^{m\times n}$. It is given by

$$A^\dagger = \overline{A}^T, \tag{2.47}$$

where $\overline{A}$ denotes the complex conjugate of A.

2.4.3 Coordinate transformations

Definition 2.4.24 (Transition matrix). Let $\mathcal{B} = \{b_1, \ldots, b_n\}$ and $\mathcal{B}' = \{b'_1, \ldots, b'_n\}$ be two bases of V. By definition, every element of $\mathcal{B}'$ can be written as a linear combination of elements in $\mathcal{B}$:

$$b'_j = q_{1j}b_1 + \cdots + q_{nj}b_n. \tag{2.48}$$

The matrix $Q \equiv (q_{ij}) \in M_n(K)$ is called the transition matrix from $\mathcal{B}$ to $\mathcal{B}'$.

Property 2.4.25. Let $\mathcal{B}, \mathcal{B}'$ be two bases of V and let Q be the transition matrix from $\mathcal{B}$ to $\mathcal{B}'$. The following statements hold:

- $Q \in \mathrm{GL}_n(K)$ and Q^{-1} is the transition matrix from $\mathcal{B}'$ to $\mathcal{B}$.
- Let $\mathcal{C}$ be an arbitrary basis of V with γ the corresponding coordinate isomorphism and define the following matrices:

  $$B := (\gamma(b_1), \ldots, \gamma(b_n)) \quad \text{and} \quad B' := (\gamma(b'_1), \ldots, \gamma(b'_n)).$$

  In terms of these matrices one finds that $BQ = B'$.
- Consider $v \in V$. Let $(\lambda_1, \ldots, \lambda_n)^T$ be the coordinate vector with respect to $\mathcal{B}$ and let $(\lambda'_1, \ldots, \lambda'_n)^T$ be the coordinate vector with respect to $\mathcal{B}'$, then

  $$Q\begin{pmatrix}\lambda'_1\\ \vdots\\ \lambda'_n\end{pmatrix} = \begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_n\end{pmatrix} \quad \text{and} \quad \begin{pmatrix}\lambda'_1\\ \vdots\\ \lambda'_n\end{pmatrix} = Q^{-1}\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_n\end{pmatrix}. \tag{2.49}$$

Corollary 2.4.26 (Basis change). Let V, W be two finite-dimensional vector spaces. Consider two bases $\mathcal{B}, \mathcal{B}'$ of V and two bases $\mathcal{C}, \mathcal{C}'$ of W. Let Q, P be the transition matrices from $\mathcal{B}$ to $\mathcal{B}'$ and from $\mathcal{C}$ to $\mathcal{C}'$, respectively. The matrix representations $A = A_{f,\mathcal{B},\mathcal{C}}$ and $A' = A_{f,\mathcal{B}',\mathcal{C}'}$ of a linear map $f : V \to W$ are related in the following way:

$$A' = P^{-1}AQ. \tag{2.50}$$

Remark 2.4.27. From the definition of the transition matrix and the above property it follows that the basis vectors and coordinate representations transform by Q and Q^{-1} respectively. That they transform in an inverse manner makes sense, since a vector should be independent from its coordinate representation:

$$v' = \sum_{i=1}^n \lambda'_i e'_i = \sum_{i,j,k=1}^n Q^{-1}_{ji}\lambda_j Q_{ik} e_k = \sum_{i,j,k=1}^n \delta_{jk}\lambda_j e_k = v.$$

This remark gives a new way to define a vector $v \in V$:

Alternative Definition 2.4.28 (Vector). Consider an n-dimensional vector space V. One can define an equivalence relation on the set $K^n \times FV$, where FV denotes the set of all bases of V, by saying that the pairs $(c, \mathfrak{b})$ and $(c', \mathfrak{b}')$ are equivalent if and only if there exists a matrix $A \in \mathrm{GL}_n(K)$ such that $c' = Ac$ and $\mathfrak{b} = A\mathfrak{b}'$. A vector $v \in V$ is then defined as an equivalence class of such pairs.

Definition 2.4.29 (Matrix conjugation). Let $A \in M_n(K)$. The set

$$\left\{Q^{-1}AQ \,\middle|\, Q \in \mathrm{GL}_n(K)\right\} \tag{2.51}$$

is called the conjugacy class of A in accordance with group theory (Definition **??**). Another term for conjugation is **similarity transformation**.

Remark 2.4.30. If A is a matrix representation of a linear operator f, the conjugacy class of A consists of all matrix representations of f.

Property 2.4.31 (Trace). Property 2.4.6 implies that the trace of a matrix is invariant under conjugation:

$$\mathrm{tr}(Q^{-1}AQ) = \mathrm{tr}(A). \tag{2.52}$$

Definition 2.4.32 (Matrix congruence). Let $A, B \in M_n(K)$. The matrices are said to be congruent if there exists a matrix P such that

$$A = P^T BP. \tag{2.53}$$

Property 2.4.33. Every matrix congruent to a symmetric matrix is also symmetric.

Property 2.4.34 (Orthogonality of basis changes). Let V be an inner product space and let $\mathcal{B}, \mathcal{B}'$ be two orthonormal bases of V with transition matrix Q. Q is *orthogonal* (Definition 2.4.56):

$$Q^T Q = \mathbb{1}_n. \tag{2.54}$$

2.4.4 Determinant

Definition 2.4.35 (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 2.4.36 (Cofactor). The cofactor α_{ij} of the matrix element a_{ij} is defined as $(-1)^{i+j}\det(A_{ij})$.

Definition 2.4.37 (Adjugate matrix). The adjugate matrix of $A \in M_n(K)$ is defined as follows:

$$\mathrm{adj}(A) := \begin{pmatrix} \alpha_{11} & \alpha_{21} & \cdots & \alpha_{n1} \\ \alpha_{12} & \alpha_{22} & \cdots & \alpha_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{1n} & \alpha_{2n} & \cdots & \alpha_{nn} \end{pmatrix}, \tag{2.55}$$

or in terms of the cofactors: $\mathrm{adj}(A) = (\alpha_{ij})^T$, where the transpose is taken after the elements have been replaced by their cofactor.

Formula 2.4.38 (Laplace). The determinant of a matrix $A \equiv (a_{ij}) \in M_n(K)$ can be evaluated as follows:

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+k} a_{ik} \det(A_{ik}). \tag{2.56}$$

Property 2.4.39. Let $A, B \in M_n(K)$ and denote the columns of A by $A_1, \ldots, A_n$. The determinant has the following properties:

- $\det(AB) = \det(A)\det(B)$,
- $\det(A^T) = \det(A)$,
- $\det(A_1, \ldots, A_i + \lambda A'_i, \ldots, A_n) = \det(A_1, \ldots, A_i, \ldots, A_n) + \lambda \det(A_1, \ldots, A'_i, \ldots, A_n)$ for all $A_i, A'_i \in M_{n,1}(K)$, and
- $\det(A_{\sigma(1)}, \ldots, A_{\sigma(n)}) = \mathrm{sgn}(\sigma)\det(A_1, \ldots, A_n)$

Items 2, 3 and 4 further imply that a matrix with two identical rows or columns has a vanishing determinant.

Property 2.4.40. Let $A \in M_n(K)$. The following statements are equivalent:

- $\det(A) \neq 0$,
- $\mathrm{rk}(A) = n$, or
- $A \in \mathrm{GL}_n(K)$.

Property 2.4.41. For all $A \in M_n(K)$ one finds that $A\,\mathrm{adj}(A) = \mathrm{adj}(A)A = \det(A)I_n$.

Corollary 2.4.42. For all $A \in \mathrm{GL}_n(K)$ one finds

$$A^{-1} = \det(A)^{-1}\,\mathrm{adj}(A). \tag{2.57}$$

Alternative Definition 2.4.43 (Minor). Let $A \in M_{m,n}(K)$ and choose $k \leq \min(m, n)$. 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.

Property 2.4.44. Let $A \in M_{m,n}(K)$ and choose $k \leq \min(m, n)$. Then $\mathrm{rk}(A) \geq k$ if and only if A contains a nonzero $k \times k$-minor.

Property 2.4.45 (Invariance of determinant). Let $f \in \mathrm{End}(V)$. The determinant of the matrix representation of f is invariant under basis transformations.

Definition 2.4.46 (Determinant of a linear map). The previous property allows for an unambiguous definition of the determinant of $f \in \mathrm{End}(V)$:

$$\det(f) := \det(A) \tag{2.58}$$

for any matrix representation A of f.

2.4.5 Characteristic polynomial

Definition 2.4.47 (Characteristic polynomial). Consider a linear map $f \in \mathrm{End}(V)$ and denote its matrix representation by A_f. The function

$$\chi_f(x) := \det(x\mathbb{1}_n - A_f) \in K[x] \tag{2.59}$$

is a monic polynomial of degree n in the variable x. The following equation is called the **characteristic equation** or **secular equation** of f:

$$\chi_f(x) = 0. \tag{2.60}$$

Formula 2.4.48. Consider a matrix $A \equiv (a_{ij}) \in M_n(K)$ with characteristic polynomial

$$\chi_A(x) = x^n + c_{n-1}x^{n-1} + \ldots + c_1 x + c_0.$$

The first and last of the coefficients c_i have a simple expression:

$$\begin{cases} c_0 = (-1)^n \det(A), \\ c_{n-1} = -\operatorname{tr}(A) \end{cases} \tag{2.61}$$

Theorem 2.4.49 (Cayley-Hamilton). *Consider a linear map $f \in \mathrm{End}(V)$ with characteristic polynomial χ_f.*

$$\chi_f(f) = f^n + \sum_{i=1}^{n-1} c_i f^i = 0. \tag{2.62}$$

Corollary 2.4.50. From Property 2.2.17 and the Cayley-Hamilton theorem it follows that the minimal polynomial μ_f is a divisor of the characteristic polynomial χ_f.

2.4.6 Matrix groups

Definition 2.4.51 (Elementary matrix). An elementary matrix is a matrix of the form

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & a & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & b & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}, \ldots$$

i.e. it is equal to the sum of an identity matrix and a multiple of a matrix unit U_{ij}. The elementary matrix with the scalar c at position (i,j) is denoted by $E_{ij}(c)$.

A second type of elementary matrix is one of the form

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & a & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & a & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

These matrices are sometimes denoted by $T_{i,j}$.

Property 2.4.52 (Invertibility). Elementary matrices have determinant 1 and, accordingly, are elements of $\mathrm{GL}_n(K)$.

Property 2.4.53. Multiplication by an elementary matrix has the following properties:

- 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.
- 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.
- Left multiplication by an elementary matrix $T_{i,j}$ interchanges the i^{th} and j^{th} rows.

Property 2.4.54. Every invertible matrix can be written as a product of elementary matrices.

Definition 2.4.55 (Special linear group). The subgroup of $\mathrm{GL}_n(K)$ consisting of all matrices with determinant 1:

$$\mathrm{SL}_n(K) := \{A \in \mathrm{GL}_n(K) \mid \det(A) = 1\}. \tag{2.63}$$

Definition 2.4.56 (Orthogonal group). The orthogonal and special orthogonal group are defined as follows:

$$\begin{aligned} \mathrm{O}(n, K) &:= \{A \in \mathrm{GL}_n(K) \mid AA^T = A^T A = \mathbb{1}_n\} \\ \mathrm{SO}(n, K) &:= \mathrm{O}_n(K) \cap \mathrm{SL}_n(K). \end{aligned}$$

Definition 2.4.57 (Unitary group). Consider a field K equipped with an involution $\sigma : \lambda \mapsto \overline{\lambda}$. The unitary and special unitary group are defined as follows:

$$\begin{aligned} \mathrm{U}_n(K, \sigma) &:= \{A \in \mathrm{GL}_n(K) \mid A\sigma(A)^T = \sigma(A)^T A = \mathbb{1}_n\} \\ \mathrm{SU}_n(K, \sigma) &:= \mathrm{U}_n(K) \cap \mathrm{SL}_n(K). \end{aligned}$$

In practice K is often $\mathbb{C}$ with complex conjugation as the involution. For this reason the notation $A^\dagger := \sigma(A)^T$ is common. Moreover, in the case $K = \mathbb{C}$ the notation is further simplified to $\mathrm{U}(n)$ and $\mathrm{SU}(n)$.

Definition 2.4.58 (Unitary equivalence). Let $A, B \in M_n(K)$ over a field K with an involution. The matrices are said to be unitarily equivalent if there exists a unitary matrix U such that

$$A = U^\dagger B U.$$

Property 2.4.59. For orthogonal matrices, conjugacy 2.4.29 and congruency 2.4.32 coincide. More generally, for unitary matrices conjugacy and unitary equivalence coincide.

Definition 2.4.60 (Symplectic group). Consider a vector space V with an antisymmetric nonsingular matrix Ω. The symplectic group $\mathrm{Sp}(V, \Omega)$ is defined as follows:

$$\mathrm{Sp}(V, \Omega) := \{A \in \mathrm{GL}(V) \mid A^T \Omega A = \Omega\}. \tag{2.64}$$

Over the real or complex numbers one can define the canonical **symplectic** matrix

$$\Omega_{st} := \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}. \tag{2.65}$$

The groups of matrices that preserve this matrix are denoted by $\mathrm{Sp}(n, \mathbb{R})$ and $\mathrm{Sp}(n, \mathbb{C})$.

Remark 2.4.61. Symplectic groups can only be defined on even-dimensional spaces because antisymmetric matrices can only be nonsingular if the dimension n is even.

Definition 2.4.62 (Compact symplectic group). The compact symplectic group is defined as follows (although the notation is confusing, it is standard):

$$\mathrm{Sp}(n) := \mathrm{Sp}(2n, \mathbb{C}) \cap \mathrm{U}(2n). \tag{2.66}$$

This group is in fact isomorphic to the *quaternionic* unitary group in n quaternionic dimensions.

Property 2.4.63.

$$\mathrm{Sp}(1) \cong \mathrm{SU}(2)$$

2.4.7 Matrix decompositions

Method 2.4.64 (QR Decomposition). Every square complex matrix M can be decomposed as

$$M = QR \tag{2.67}$$

with Q unitary and R upper-triangular. The easiest way to achieve this decomposition is by applying the Gram-Schmidt orthonormalization process:

> Let $\{v_i\}_{i\leq n}$ be a basis for the column space of M. By applying the Gram-Schmidt process to this basis one obtains a new orthonormal basis $\{e_i\}_{i\leq n}$. The matrix M can then be written as a product QR of the following matrices:
>
> - an upper-triangular matrix R with entries $R_{ij} = \langle e_i|\text{col}_j(M)\rangle$, where $\text{col}_j(M)$ denotes the j^{th} column of M.
> - a unitary matrix $Q = (e_1, \ldots, e_n)$ constructed by setting the i^{th} column equal to the i^{th} basis vector e_i.

Property 2.4.65. If M is invertible and if the diagonal elements of R are required to have positive norm, the QR-decomposition is unique.

?? COMPLETE (Cholesky, polar, ...) ??

2.5 Eigenvectors

Definition 2.5.1 (Eigenvector). A vector $v \in V \setminus \{0\}$ is called an **eigenvector** of the linear map $f : V \to V$ if it satisfies

$$f(v) = \lambda v \tag{2.68}$$

for some $\lambda \in K$. The scalar λ is called the **eigenvalue** associated to v.

Definition 2.5.2 (Eigenspace). The subspace of V spanned by the eigenvectors of a linear map is called the eigenspace of that linear map. It is given by

$$\ker(\lambda \mathbb{1}_V - f). \tag{2.69}$$

It follows that the eigenvalues are exactly those scalars for which the linear map $\lambda\mathbb{1}_V - f$ is not injective. (This is generalized in Section **??**.)

Property 2.5.3 (Characteristic equation). Consider a linear map $f \in \text{End}(V)$. A scalar $\lambda \in K$ is an eigenvalue of f if and only if it satisfies the characteristic equation (2.60).

Property 2.5.4. A linear map $f \in \text{End}(V)$ defined over an n-dimensional vector space V has at most n different eigenvalues.

These property lead to the following method for finding eigenvectors:

Method 2.5.5 (Finding the eigenvectors of a matrix). To calculate the eigenvectors of a matrix one should perform the following steps:

1. Find the eigenvalues λ_i of A by solving the characteristic equation (2.60).
2. Find the eigenvector v_i associated to the eigenvalue λ_i by solving

$$(A - \lambda_i \mathbb{1}_V)v_i = 0. \tag{2.70}$$

2.5.1 Diagonalization

Definition 2.5.6 (Diagonalizable map). Let V be a finite-dimensional vector space. A linear map $f \in \mathrm{End}(V)$ is said to be diagonalizable if it admits a diagonal matrix representation.

Property 2.5.7. Every diagonalizable map is semisimple 2.2.18. Conversely, in finite dimensions (and over an algebraically closed field), a semisimple map is diagonalizable.

Theorem 2.5.8. *A matrix $A \in M_n(K)$ is diagonalizable if and only if there exists a matrix $P \in \mathrm{GL}_n(K)$ such that $P^{-1}AP$ is diagonal.*

Corollary 2.5.9 (Trace). Using the Property 2.4.31 that the trace of a linear map is invariant under similarity transformations, the following useful formula can be proven:

$$\mathrm{tr}(f) = \sum_{i=0}^{n} \lambda_i, \tag{2.71}$$

where $\{\lambda_i\}_{i\leq n}$ are the eigenvalues of f.

Property 2.5.10. Let V be an n-dimensional vector space and let $f \in \mathrm{End}(V)$ be a linear map. The eigenvalues and eigenvectors of f satisfy the following properties:

- The eigenvectors of f belonging to different eigenvalues are linearly independent.
- If f has exactly n eigenvalues, f is diagonalizable.
- If f is diagonalizable, then V is the direct sum of the eigenspaces of f belonging to the different eigenvalues of f.

Theorem 2.5.11. *A linear map defined on a finite-dimensional vector space is diagonalizable if and only if its set of eigenvectors forms a basis of the vector space.*

2.5.2 Multiplicity

Definition 2.5.12 (Multiplicity). Let V be a vector space and let $f \in \mathrm{End}(V)$ have characteristic polynomial

$$\chi_f(x) = \prod_{i=1}^{n} (x - \lambda_i)^{n_i}. \tag{2.72}$$

The multiplicities are defined as follows:

- The **algebraic multiplicity** of an eigenvalue λ_i is equal to n_i.
- The **geometric multiplicity** of an eigenvalue λ_i is equal to the dimension of the eigenspace belonging to that eigenvalue.

Remark 2.5.13 (Splitting field). In the previous definition it was assumed that the characteristic polynomial can be completely factorized. However, this depends on the possibility to completely factorize the polynomial over K (i.e. if it has "enough' roots" in K). If not, f cannot even be diagonalized. In general there always exists a field f containing K, called a *splitting field*, over which the polynomial can be completely factorized. Note that in general this field is strictly smaller than the algebraic closure of K, which is the *splitting field* of the collection of all polynomials over K.

Property 2.5.14. The algebraic multiplicity is always greater than or equal to the geometric multiplicity.

Theorem 2.5.15. *A linear map $f \in \mathrm{End}(V)$ is diagonalizable if and only if for every eigenvalue the algebraic multiplicity is equal to the geometric multiplicity.*

Property 2.5.16. Every Hermitian linear map $f \in \mathrm{End}(\mathbb{C}^n)$ has the following properties:

- All the eigenvalues of f are real.
- Eigenvectors belonging to different eigenvalues are orthogonal.
- f is diagonalizable and there always exists an orthonormal basis of eigenvectors of f, in particular, the diagonalizing matrix P is unitary, i.e. $P^{-1} = P^\dagger$.

Property 2.5.17 (Commutator). Let $f, g \in \mathrm{End}(V)$ be two diagonalizable maps. If the commutator $[f, g]$ is zero, the two maps have a common eigenbasis.

Theorem 2.5.18 (Sylvester's law of inertia). *The number of positive and negative eigenvalues of a Hermitian matrix is invariant with respect to $\dagger$-congruence (or conjugation due to Property 2.4.59).*

2.6 Euclidean space

A finite-dimensional $\mathbb{R}$-vector space is sometimes called a **Euclidean** or **Cartesian space**.

Notation 2.6.1. When working in a Euclidean space, the inner product $\langle v|w\rangle$ is often written as $v \cdot w$.

Definition 2.6.2 (Orientation). Let $\mathcal{B}, \mathcal{B}'$ be two ordered bases of $\mathbb{R}^n$ and let Q be the transition matrix from $\mathcal{B}$ to $\mathcal{B}'$. If $\det(Q) > 0$, the bases are said to have the same orientation (or to be **consistently oriented**). If $\det(Q) < 0$, the bases are said to have an opposite orientation.

Corollary 2.6.3 (Positive orientation). The previous definition imposes an equivalence relation on the set of bases of $\mathbb{R}^n$ with exactly two equivalence classes. The bases in one of these classes are said to be **positively** (or **directly**) oriented. The bases in the other class are then said to be **negatively** (or **indirectly**) oriented.

Remark 2.6.4. It is convenient to take the standard basis $(e_1, \ldots, e_n)$ to be positively oriented.

2.7 Algebras

Definition 2.7.1 (Algebra). Let V be a vector space equipped with a binary operation $\star : V \times V \to V$. The pair $(V, \star)$ is called an algebra over K if it satisfies the following conditions:

1. **Right distributivity**: $(x + y) \star z = x \star z + y \star z$,
2. **Left distributivity**: $x \star (y + z) = x \star y + x \star z$, and
3. **Compatibility with scalars**: $(\lambda x) \star (\mu y) = \lambda\mu(x \star y)$.

These conditions say that the binary operation is bilinear. An algebra V is said to be unital if it contains an identity element with respect to the bilinear map $\star$.

Remark 2.7.2 (Over rings). More generally one can define an algebra over a commutative unital ring R. The defining conditions remain the same, except that one requires V to be an R-module instead of a vector space.

Definition 2.7.3 (Division algebra). A unital algebra in which every nonzero element has both a left and right multiplicative inverse. If the algebra is associative, these inverses coincide. A normed division algebra is a division algebra equipped with a multiplicative quadratic form q such that $\langle a|b\rangle := \frac{1}{2}[q(a + b) - q(a) - q(b)]$ is a nondegenerate inner product (2.3.1).

Theorem 2.7.4 (Frobenius). *There exist three inequivalent finite-dimensional real associative division algebras:* $\mathbb{R}, \mathbb{C}$ *and* $\mathbb{H}$.

Theorem 2.7.5 (Hurwitz). *There exist four inequivalent finite-dimensional real normed division algebras:* $\mathbb{R}, \mathbb{C}, \mathbb{H}$ *and* $\mathbb{O}$.

Example 2.7.6 (Frobenius algebra). An associative algebra A equipped with a nondegenerate bilinear form $\eta : A \times A \to K$ satisfying the following condition for all $a, b, c \in A$:

$$\eta(ab, c) = \eta(a, bc). \tag{2.73}$$

Equivalently, an associative algebra (A, μ) equipped with a linear form $\varepsilon : A \to K$ such that $\varepsilon \circ \mu$ is nondegenerate.[6]

A Frobenius algebra is said to be symmetric if η is symmetric.

Example 2.7.7 (Temperley-Lieb algebra). Let R be a commutative unital ring and fix an element $\delta \in R$. The Temperley-Lieb algebra $\mathrm{TL}_n(\delta)$ is the unital R-algebra with generators $\{U_i\}_{i<n}$ that satisfy the **Jones relations**:

1. $U_i^2 = \delta U_i$,
2. $U_i U_j = U_j U_i$ if $|i-j| \neq 1$, and
3. $U_i U_j U_i = U_i$ if $|i-j| = 1$.

One can represent the elements of a Temperley-Lieb algebra diagrammatically. All elements of $\mathrm{TL}_n(\delta)$ are represented as diagrams with n inputs and n outputs:

> The unit is given by the diagram where all inputs are connected to the outputs directly across the diagram. The generators $\{U_i\}_{i<n}$ are constructed by connecting the i^{th} input (resp. output) to the $i+1^{th}$ input (resp. output) and all other inputs are connected to the output directly across the diagram. Multiplication in $\mathrm{TL}_n(\delta)$ is performed diagrammatically by placing two diagrams side by side. Closed loops are replaced by a factor δ.

(a) Unit in $\mathrm{TL}_4(\delta)$. (b) Generator U_2 in $\mathrm{TL}_4(\delta)$.

Figure 2.1: Temperley-Lieb algebra.

Definition 2.7.8 (Jordan algebra). A nonassociative, commutative algebra A such that

$$(xy)(xx) = x(y(xx)) \tag{2.74}$$

for all $x, y \in A$.

Property 2.7.9 (Power associativity). It can be shown that the Jordan condition implies that powers of elements are well-defined:

$$(xx)x = x(xx) =: x^3 \tag{2.75}$$

for all $x \in A$ and likeiwse for higher-order powers.

[6]A third equivalent definition is given in **??**.

The original definition of a Jordan algebra does not admit a lot of intuition. However, by the power-associativity property one also has expressions of the form

$$(x^m y)x^n = x^m(yx^n). \tag{2.76}$$

By commutativity one obtains that the multiplication maps $L_{x^m} : y \mapsto x^m y$ associated to powers commute:

$$L_{x^m}L_{x^n} = L_{x^n}L_{x^m}. \tag{2.77}$$

This leads to the following equivalent definition:

Alternative Definition 2.7.10. A Jordan algebra is a commutative, power-associative algebra A such that Equation (2.77) holds for all $x \in A$.

Property 2.7.11. Every associative algebra over a field of characteristic not 2 (or over a ring in which 2 is a unit) the multiplication induces a Jordan structure as follows:

$$x \circ y := \frac{1}{2}(xy + yx), \tag{2.78}$$

i.e. the Jordan product is given by the anticommutator. Jordan algebras of this form are said to be **special**, while all other Jordan algebras are said to be **exceptional**.

2.8 Grassmanians

Definition 2.8.1 (Grassmannian). Let V be a vector space. The set of all subspaces of V of dimension k is called the Grassmannian $\mathrm{Gr}(k, V)$.

Property 2.8.2. $\mathrm{GL}(V)$ acts transitively **??** on the k-dimensional subspaces of V. Property **??** implies that the coset space $\mathrm{GL}(V)/H_W$ for the stabilizer H_W of any $W \in \mathrm{Gr}(k, V)$ is isomorphic (as a set) to $\mathrm{Gr}(k, V)$. When V is an n-dimensional real vector space one can show that this quotient is isomorphic to $\mathrm{O}(n)/(\mathrm{O}(k) \times \mathrm{O}(n-k))$. For complex vector spaces the orthogonal groups should be replaced by unitary groups.

Example 2.8.3 (Projective space). Recall Definition **??**. The Grassmannian $\mathrm{Gr}(1, V)$ is given by the projective space $K\mathbb{P}^{\dim(V)-1}$.

Definition 2.8.4 (Flag). Let V be a finite-dimensional vector space. A sequence of proper subspaces $V_1 < \cdots < V_n = V$ is called a flag of V. The sequence $(\dim(V_1), \ldots, \dim(V_n) = \dim(V))$ is called the **signature** of the flag. If $\forall i \leq \dim(V) : \dim(V_i) = i$, the flag is said to be **complete**.

Grassmannians are a specific instance of the following object:

Definition 2.8.5 (Flag variety). The set of all flags of a given signature is called the (generalized) flag variety (of that signature). If the underlying field is the field of real (or complex) numbers, the flag variety is a smooth (or complex) manifold (Chapter **??**), called the **flag manifold**.

Property 2.8.2 generalizes as follows:

Property 2.8.6 (Parabolic subgroups). Every flag variety has the structure of a homogeneous space: $\mathrm{Fl}_{n,\underline{d}} = \mathrm{GL}(V)/P_{n,\underline{d}}$, where $\underline{d}$ denotes the signature of the flags. The subgroups $P_{n,\underline{d}}$ are called **parabolic subgroups**. The maximal parabolic subgroups are those that define the Grassmannian varieties. The flag variety of all complete flags defines the **Borel subgroup** B_n. It can be shown that every parabolic subgroup contains the Borel subgroup.

Chapter 3

Vector & Tensor Calculus

References for this chapter are [52, 53]. For a more geometric approach to some of the concepts and results in this chapter, see the content of Chapters **??** and **??** and Section **??**.

Remark. In this chapter a *vector field* will mean a vector-valued function with smooth projections.

3.1 Nabla-operator

Remark. The geometric approach to this section is summarized in Remark **??**.

Definition 3.1.1 (Gradient). Let $\varphi : \mathbb{R}^3 \to \mathbb{R}$ be a smooth function.

$$\nabla\varphi := \left(\frac{\partial\varphi}{\partial x}, \frac{\partial\varphi}{\partial y}, \frac{\partial\varphi}{\partial z}\right) \tag{3.1}$$

This definition can easily be generalized to arbitrary dimensions.

Property 3.1.2. The gradient of a smooth real-valued function is perpendicular to its level sets **??**.

Definition 3.1.3 (Directional derivative). Consider a smooth function $\varphi : \mathbb{R}^3 \to \mathbb{R}$ and let $\hat{a}$ be a unit vector. The directional derivative $\nabla_{\hat{a}}\varphi$ is defined as the change of the function φ in the direction of $\hat{a}$:

$$\nabla_{\hat{a}}\varphi := (\hat{a} \cdot \nabla)\varphi. \tag{3.2}$$

Example 3.1.4. Let $\varphi : \mathbb{R}^3 \to \mathbb{R}$ be a smooth function and let $\frac{d\vec{\boldsymbol{r}}}{ds}$ denote the tangent vector to a curve $\vec{\boldsymbol{r}}(s)$ with natural parameter[1] s. The variation of φ along $\vec{\boldsymbol{r}}(s)$ is given by

$$\frac{\partial\varphi}{\partial s} = \frac{d\vec{\boldsymbol{r}}}{ds} \cdot \nabla\varphi. \tag{3.3}$$

Definition 3.1.5 (Conservative vector field). A vector field that can be expressed as the gradient of a scalar function.

Definition 3.1.6 (Gradient of a tensor). Let T be a tensor field on $\mathbb{R}^3$ and let $\vec{\boldsymbol{e}}_i$ be the coordinate basis. The gradient of T is defined as follows:

$$\nabla T := \sum_{i=1}^{3} \frac{\partial T}{\partial x^i} \otimes \vec{\boldsymbol{e}}_i. \tag{3.4}$$

[1] See Definition **??** for a formal definition.

Definition 3.1.7 (Divergence). Let $\vec{A}$ be a vector field on $\mathbb{R}^3$.

$$\nabla\cdot\vec{A} := \frac{\partial A_x}{\partial x}+\frac{\partial A_y}{\partial y}+\frac{\partial A_z}{\partial z} \tag{3.5}$$

Definition 3.1.8 (Solenoidal vector field). A vector field $\vec{A}$ that satisfies

$$\nabla\cdot\vec{A} = 0. \tag{3.6}$$

Such a vector field is also said to be **divergence-free** due to Equation (3.13) below.

Definition 3.1.9 (Rotor / curl). Let $\vec{A}$ be a vector field on $\mathbb{R}^3$.

$$\nabla\times\vec{A} := \left(\frac{\partial A_z}{\partial y}-\frac{\partial A_y}{\partial z},\frac{\partial A_x}{\partial z}-\frac{\partial A_z}{\partial x},\frac{\partial A_y}{\partial x}-\frac{\partial A_x}{\partial y}\right) \tag{3.7}$$

Definition 3.1.10 (Irrotational vector field). A vector field $\vec{A}$ that satisfies

$$\nabla\times\vec{A} = 0. \tag{3.8}$$

Definition 3.1.11 (Laplacian). Let φ and $\vec{A}$ be respectively a smooth function and smooth vector field on $\mathbb{R}^3$.

$$\Delta\varphi := \nabla^2\varphi = \frac{\partial^2\varphi}{\partial x^2}+\frac{\partial^2\varphi}{\partial y^2}+\frac{\partial^2\varphi}{\partial z^2} \tag{3.9}$$

$$\Delta\vec{A} := \nabla^2\vec{A} = \nabla\left(\nabla\cdot\vec{A}\right)-\nabla\times\left(\nabla\times\vec{A}\right) \tag{3.10}$$

$$= (\Delta A_x, \Delta A_y, \Delta A_z) \tag{3.11}$$

The latter is sometimes called the **vector Laplacian**.

Property 3.1.12 (Mixed properties). The differential operators introduced above satisfy the following identities:

$$\nabla\times(\nabla\varphi) = 0\,, \tag{3.12}$$

$$\nabla\cdot\left(\nabla\times\vec{A}\right) = 0. \tag{3.13}$$

Corollary 3.1.13. All conservative vector fields are irrotational. However, the converse is only true if the domain is simply-connected **??**. (All of this is formalized in the Poincaré lemma **??**.)

Formula 3.1.14 (Helmholtz decomposition). If $\vec{A}$ is a vector field that decays faster than $1/r$ when $r\longrightarrow\infty$, it can be written as

$$\vec{A} = \nabla\times\vec{B}+\nabla\varphi \tag{3.14}$$

for some smooth vector field $\vec{B}$ and smooth function φ.

The differential operators introduced above can also be generalized to curvilinear coordinates. To do this one needs the scale factors as formally defined in Definition **??**. For the remainder of this section the Einstein summation convention will not be used to make everything as explicit as possible.

Formula 3.1.15 (Unit vectors).

$$\frac{\partial\vec{r}}{\partial q^i} = h_i\hat{e}_i \tag{3.15}$$

Formula 3.1.16 (Gradient).

$$\nabla\varphi = \sum_{i=1}^{3}\frac{1}{h_i}\frac{\partial\varphi}{\partial q^i}\hat{e}_i \tag{3.16}$$

Formula 3.1.17 (Divergence).

$$\nabla\cdot\vec{\boldsymbol{A}} = \frac{1}{h_1h_2h_3}\left(\frac{\partial}{\partial q^1}(A_1h_2h_3) + \frac{\partial}{\partial q^2}(A_2h_3h_1) + \frac{\partial}{\partial q^3}(A_3h_1h_2)\right) \tag{3.17}$$

Formula 3.1.18 (Rotor).

$$\left(\nabla\times\vec{\boldsymbol{A}}\right)_i = \sum_{j,k=1}^{3}\frac{\varepsilon_{ijk}}{h_jh_k}\left(\frac{\partial}{\partial q^j}(A_kh_k) - \frac{\partial}{\partial q^k}(A_jh_j)\right), \tag{3.18}$$

where ε_{ijk} is the 3-dimensional Levi-Civita symbol 3.4.8.

Formula 3.1.19 (Laplacian in different coordinate systems). In general the Laplace operator is defined as

$$\Delta f := \nabla\cdot\nabla f. \tag{3.19}$$

The Laplacian can also be expressed in different coordinate systems:

- Cylindrical coordinates (ρ, ϕ, z):

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial}{\partial\rho}\right) + \frac{1}{\rho^2}\frac{\partial^2}{\partial\phi^2} + \frac{\partial^2}{\partial z^2}. \tag{3.20}$$

- Spherical coordinates (r, ϕ, θ):

$$\frac{1}{r^2}\left[\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)\right]. \tag{3.21}$$

3.2 Integration

3.2.1 Line integrals

Formula 3.2.1 (Line integral of a continuous function). Let $f:\mathbb{R}^3\rightarrow\mathbb{R}$ be a continuous function and let Γ be a piecewise smooth curve $\vec{\boldsymbol{\varphi}}:[a,b]\rightarrow\mathbb{R}^3$. The line integral of f along Γ is defined as follows:

$$\int_\Gamma f\,ds := \int_a^b f(\vec{\boldsymbol{\varphi}}(t))\,\|\vec{\boldsymbol{\varphi}}'(t)\|\,dt. \tag{3.22}$$

Formula 3.2.2 (Line integral of a continuous vector field). Let $\vec{\boldsymbol{F}}$ be a continuous vector field on $\mathbb{R}^3$ and let Γ be a piecewise smooth curve $\vec{\boldsymbol{\varphi}}:[a,b]\rightarrow\mathbb{R}^3$. The line integral of $\vec{\boldsymbol{F}}$ along Γ is defined as follows:

$$\int_\Gamma \vec{\boldsymbol{F}}\cdot d\vec{\boldsymbol{s}} := \int_a^b \vec{\boldsymbol{F}}(\vec{\boldsymbol{\varphi}}(t))\cdot\vec{\boldsymbol{\varphi}}'(t)\,dt. \tag{3.23}$$

Property 3.2.3 (Conservative vector fields). A vector field is conservative if and only if its line integral is path-independent, i.e. if it only depends on the values at the end points. (This is a corollary of Stokes's theorem **??**.)

3.2.2 Integral theorems[2]

Theorem 3.2.4 (Fundamental theorem of calculus for line integrals). *Let* $\Gamma : \mathbb{R} \to \mathbb{R}^3$ *be a piecewise smooth curve defined on the interval* $[a, b]$*.*

$$\int_\Gamma \nabla f \cdot d\vec{r} = \varphi(\Gamma(b)) - \varphi(\Gamma(a)) \tag{3.24}$$

Theorem 3.2.5 (Kelvin-Stokes theorem). *Let* $\vec{A}$ *be a vector field defined on* $\mathbb{R}^3$ *and let* S *be a smooth surface with boundary* ∂S*.*

$$\oint_{\partial S} \vec{A} \cdot d\vec{l} = \iint_S \left(\nabla \times \vec{A}\right) dS \tag{3.25}$$

Theorem 3.2.6 (Divergence theorem[3]). *Let* $\vec{A}$ *be a vector field defined on* $\mathbb{R}^3$*.*

$$\oiint_{\partial V} \vec{A} \cdot d\vec{S} = \iiint_V \left(\nabla \cdot \vec{A}\right) dV \tag{3.26}$$

Corollary 3.2.7 (Green's identity). Let ϕ, ψ be smooth real-valued functions defined on $\mathbb{R}^3$.

$$\oiint_{\partial V} (\psi \nabla \phi - \phi \nabla \psi) \cdot d\vec{S} = \iiint_V \left(\psi \nabla^2 \phi - \phi \nabla^2 \psi\right) dV \tag{3.27}$$

3.3 Tensors

3.3.1 Tensor product

There are two possible (equivalent) ways to introduce the concept of a "tensor" on finite-dimensional vector spaces. One is to interpret tensors as multilinear maps, while the other is to work in a local fashion and express work with the expansion coefficients with respect to a chosen basis.

Definition 3.3.1 (Tensor product space). The tensor product of two finite-dimensional vector spaces V and W is defined as[4] the set of bilinear maps on the Cartesian product $V^* \times W^*$. Let v, w be vectors in respectively V and W and let g, h be vectors in the corresponding dual spaces. The tensor product of v and w is then defined as follows:

$$(v \otimes w)(g, h) := v(g)w(h). \tag{3.28}$$

In this incarnation the tensor product is sometimes known as the **outer product**. Outer products are also frequently called **pure** or **simple tensors**.

Definition 3.3.2 (Tensor component). Let $\mathbf{T}$ be a tensor that takes r vectors and s covectors as input and returns a scalar (element of the underlying field). The components of $\mathbf{T}$ with respect to a frame $\{e_i\}_{i \leq n}$ and a coframe $\{e^i\}_{i \leq n}$ are defined as $T_{i \ldots j}{}^{k \ldots l} := \mathbf{T}(e_i, \ldots, e_j, e^k, \ldots, e^l)$.

The above definition can be restated as a universal property (this is also the right way to generalize tensors to infinite-dimensional spaces and avoid the awkward definition involving dual spaces):

Universal Property 3.3.3. Let Z be a vector space. For every bilinear map $T : V \times W \to Z$ there exists a unique linear map $f : V \otimes W \to Z$ such that $T = f \circ \varphi$, where $\varphi : V \times W \to V \otimes W$ is the bilinear map $(v, w) \mapsto v \otimes w$.

[2]These theorems follow from a more general theorem by Stokes **??**.

[3]Also known as **Gauss's theorem** or the **Gauss-Ostrogradsky theorem**.

[4]"*isomorphic to*" would be better terminology. See the universal property 3.3.3 below.

Corollary 3.3.4. The tensor product is unique up to linear isomorphisms. This results in the commutativity of the tensor product:

$$V \otimes W \cong W \otimes V. \tag{3.29}$$

Notation 3.3.5 (Tensor power).

$$V^{\otimes n} := \underbrace{V \otimes \cdots \otimes V}_{n \text{ copies}} \tag{3.30}$$

More generally, the tensor product of r copies of V and s copies of V^* is denoted by

$$\mathcal{T}_s^r(V) = V^{\otimes r} \otimes V^{*\otimes s}. \tag{3.31}$$

Tensors in this space are said to be of **type** (r, s).

Definition 3.3.6 (Scalar). The scalars, i.e. the elements of the underlying field are by definition the $(0,0)$-tensors.

Definition 3.3.7 (Tensor algebra). The tensor algebra over a vector space V is defined as follows:

$$T(V) := \bigoplus_{k\geq 0} V^{\otimes k}. \tag{3.32}$$

The following remark is strongly related to Property 2.2.20:

Remark 3.3.8. For finite-dimensional vector spaces the space $\mathcal{T}_1^1 V$ is isomorphic to $\mathrm{End}(V)$ and the space $\mathcal{T}_0^1 V$ is isomorphic to V itself.

However, when including infinite-dimensional spaces, the space $\mathcal{T}_1^1 V$ is only isomorphic to the endomorphism space $\mathrm{End}(V^*)$ of the dual. This isomorphism is given by the map $\hat{T} : V^* \to V^* : \omega \mapsto \mathbf{T}(-, \omega)$ for every $\mathbf{T} \in \mathcal{T}_1^1 V$. Moreover, in this general setting, the spaces $\mathcal{T}_1^0 V$ and V^* are also isomorphic.

The tensor product space can also be defined as follows:

Alternative Definition 3.3.9 (Tensor product). Consider two vector spaces V, W over a field K. First, construct the free vector space $F(V \times W)$ over K. Then, construct the subspace N of $F(V \times W)$ spanned by elements of the form

- $(v + v', w) - (v, w) - (v', w)$,
- $(v, w + w') - (v, w) - (v, w')$,
- $(\lambda v, w) - \lambda(v, w)$, or
- $(v, \mu w) - \mu(v, w)$,

where $v \in V, w \in W$ and $\lambda, \mu \in K$. The tensor product $V \otimes W$ is defined as the quotient $F(V \times W)/N$. It can be shown that this construction is associative, i.e. $U \otimes (V \otimes W) \cong (U \otimes V) \otimes W$, and as such these brackets will be omitted in all expressions.

Now, consider the case where $W = V^*$. In this case the basis of the tensor product $\mathcal{T}_s^r(V)$ will be denoted by

$$\underbrace{e_i \otimes \cdots \otimes e_j}_{r \text{ basis vector}} \otimes \underbrace{\varepsilon^k \otimes \cdots \otimes \varepsilon^l}_{s \text{ dual basis vectors}}$$

and the expansion coefficients will be denoted by $T_{i\ldots j}{}^{k\ldots l}$.

Property 3.3.10 (Dimension). From the previous construction it follows that the dimension of $\mathcal{T}^r_s(V)$ is equal to rs.

For completeness the proof that the values of the tensor operating on r basis vectors and s basis covectors are equal to the corresponding expansion coefficients is given:

Proof. Consider a general tensor $\mathbf{T} = T_{i\dots j}{}^{k\dots l} e_k \otimes \dots \otimes e_l \otimes \varepsilon^i \otimes \dots \otimes \varepsilon^j$. Combining Definition 3.3.1 and the pairing of dual vectors (2.17) gives

$$\begin{aligned}\mathbf{T}(\varepsilon^m, \dots, \varepsilon^n, e_a, \dots, e_b) &= T_{i\dots j}{}^{k\dots l} e_k(\varepsilon^m)\dots e_l(\varepsilon^n)\varepsilon^i(e_a)\dots\varepsilon^j(e_b) \\ &= T_{i\dots j}{}^{k\dots l}\delta^m_k\dots\delta^n_l\delta^i_a\dots\delta^j_b \\ &= T_{a\dots b}{}^{m\dots n}.\end{aligned}$$

□

3.3.2 Transformation rules

In this section the behaviour of tensors under basis transformations of the form $e'_i = A^i{}_j\, e_j$ is considered.

Definition 3.3.11 (Contravariant). A tensor component that transforms by the following rule is said to be contravariant:

$$v^i = A^i{}_j\, v'^j. \tag{3.33}$$

Definition 3.3.12 (Covariant). A tensor component that transforms by the following rule is said to be covariant:

$$p'_i = A^j{}_i\, p_j. \tag{3.34}$$

Example 3.3.13 (Mixed tensor). This example gives the transformation rule of a mixed third-order tensor $T \in \mathcal{T}^1_2$:

$$T^k_{\ ij} = A^k_{\ w}(A^{-1})^u_{\ i}(A^{-1})^v_{\ j}T'^w_{\ uv}. \tag{3.35}$$

Method 3.3.14 (Quotient rule). Assume that an equation such as $Q_i{}^j A_{jl}{}^k = B_{il}{}^k$ is given, with A and B two known tensors. The quotient rule asserts the following: "*If the equation holds under all transformations, then Q is a tensor of the indicated type.*" Note that this rule does not necessarily hold when $B = 0$ because transformation rules are not well-defined for the zero tensor.

Remark. This rule is a useful substitute for the "illegal" division of tensors.

3.3.3 Tensor operations

Definition 3.3.15 (Contraction). Let A be a tensor of type (m, n). Taking a subscript and superscript to be equal and summing over all possible values of this index gives a new tensor of type $(m-1, n-1)$. This operation is called the contraction of A. It is induced by the evaluation map/pairing 2.2.22.

Definition 3.3.16 (Direct product). Let A and B be two tensors. The tensor constructed by the componentwise multiplication of A and B is called the direct product of A and B. This is a generalization of the Hadamard product 2.4.8.

Example 3.3.17. Let ${A^i}_k$ and ${B^j}_{lm}$ represent two tensors. Their direct product is equal to

$$C^{i\ j}_{\ k\ lm} = {A^i}_k {B^j}_{lm}.$$

Formula 3.3.18 (Operator product). It is also possible to combine operators acting on different vector spaces to make them act on the tensor product space:

$$(A \otimes B)(v \otimes w) := Av \otimes Bw. \tag{3.36}$$

Notation 3.3.19 (Abuse of notation). Consider an operator A acting on a vector space V_1. When working with a tensor product space $V_1 \otimes V_2$, the operator A can be extended to the product as $A \otimes \mathbb{1}$. However, it is often still denoted by A.

Notation 3.3.20 (Symmetric part). Consider a second-order tensor T (here taken to be of covariant type for notational simplicity). The symmetric and antisymmetric part of T are sometimes denoted by

$$T_{(ij)} = \frac{1}{2}(T_{ij} + T_{ji}) \tag{3.37}$$

and

$$T_{[ij]} = \frac{1}{2}(T_{ij} - T_{ji}). \tag{3.38}$$

This notation is easily generalized to other types of tensors.

Property 3.3.21 (Gradient of tensor products). The gradient of an outer product is defined through the Leibniz rule:

$$\nabla \cdot (v \otimes w) := (\nabla \cdot v)w + (v \cdot \nabla)w. \tag{3.39}$$

Definition 3.3.22 (Complexification). Let V be a real vector space. The complexification of V is defined as the following tensor product:

$$V^{\mathbb{C}} := V \otimes \mathbb{C}. \tag{3.40}$$

This space can still be considered a real vector space, but it can also be turned into a complex vector space by generalizing the scalar product as follows for all $\alpha \in \mathbb{C}$:

$$\alpha(v \otimes \beta) := v \otimes (\alpha\beta). \tag{3.41}$$

Property 3.3.23. By (multi)linearity every element $v_{\mathbb{C}} \in V^{\mathbb{C}}$ can be written as

$$v_{\mathbb{C}} = (v_1 \otimes 1) + i(v_2 \otimes 1).$$

Therefore, the complexification can be (formally) decomposed as

$$V^{\mathbb{C}} \cong V \oplus iV. \tag{3.42}$$

3.4 Exterior algebra

3.4.1 Antisymmetric tensors

Definition 3.4.1 (Antisymmetric tensor). A tensor that changes sign under the interchange of any two indices.

Notation 3.4.2 (Symmetric tensors). The space of symmetric $(n, 0)$-tensors is denoted by $S^n(V)$. The space of symmetric $(0, n)$-tensors is denoted by $S^n(V^*)$.

Notation 3.4.3 (Antisymmetric tensors). The space of antisymmetric $(n, 0)$-tensors is denoted by $\Lambda^n(V)$. The space of antisymmetric $(0, n)$-tensors is denoted by $\Lambda^n(V^*)$.

Property 3.4.4. Let $n = \dim(V)$. The space $\Lambda^r(V)$ equals the zero space for all $r \geq n$.

3.4.2 Determinant

Definition 3.4.5 (Form). An n-form is a totally antisymmetric element of $\mathcal{T}^0_n V$.

Definition 3.4.6 (Volume form). A form of rank $\dim(V)$ is also called a **top form** or **volume form**.

Definition 3.4.7 (Determinant). Consider a finite-dimensional vector space V with basis $\{e_i\}_{i\leq n}$. Let φ be a tensor in $\mathcal{T}^1_1 V \cong \mathrm{End}(V)$ and let ω be a volume form on V. The determinant of φ is defined as follows:

$$\det(\varphi) := \frac{\omega(\varphi(e_1), \dots, \varphi(e_n))}{\omega(e_1, \dots, e_n)}. \tag{3.43}$$

This definition is well-defined, i.e. it is independent of the choice of volume form and basis. Furthermore, it coincides with Definition 2.4.46.

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, for these the above formula would not be independent of a choice of basis anymore. A more general concept can be defined using the language principal bundles (see Section **??**).

3.4.3 Levi-Civita symbol

Definition 3.4.8 (Levi-Civita symbol). In n dimensions the Levi-Civita symbol is defined as follows:

$$\varepsilon_{i_1\dots i_n} = \begin{cases} 1 & \text{if } (i_1 \dots i_n) \text{ is an even permutation of } (1\dots n) \\ -1 & \text{if } (i_1 \dots i_n) \text{ is an odd permutation of } (1\dots n) \\ 0 & \text{if any of the indices occurs more than once.} \end{cases} \tag{3.44}$$

Remark 3.4.9 (Pseudotensor). The Levi-Civita symbol is not a tensor, it is a *pseudotensor*. This means that the sign changes under reflections or any transformation with determinant -1. (To turn it into a proper tensor, one should multiply it by a factor $\sqrt{g}$, where g is the determinant of the metric.)

Definition 3.4.10 (Cross product). Using the Levi-Civita symbol, one can define the i^{th} component of the cross product as follows:

$$(v \times w)_i = \sum_{j,k=1}^{3} \varepsilon_{ijk} v_j w_k. \tag{3.45}$$

The previous remark implies that the cross product is in fact not a vector, instead it is a "pseudovector".

Remark 3.4.11 (Generalization and Hurwitz theorem). The cross product actually exists in four cases: $\mathbb{R}^0, \mathbb{R}^1, \mathbb{R}^3$ and $\mathbb{R}^7$. In general it is characterized by the following conditions:

1. **Bilinearity**: $(\lambda v) \times (\kappa w) = \lambda\kappa(v \times w)$.
2. **Orthogonality**: $v \cdot (v \times w) = 0 = w \cdot (v \times w)$.
3. **Magnitude**: $\|v \times w\|^2 = \|v\|^2\|w\|^2 - (v \cdot w)^2$.

These conditions imply that on $\mathbb{R}^1$ the cross product is identically zero. However, one $\mathbb{R}^3$ and $\mathbb{R}^7$ one obtains an anticommutative bilinear operation. (On $\mathbb{R}^3$ it is unique, while on $\mathbb{R}^7$ different choices exist.)

This construction is related to the Hurwitz classification theorem 2.7.5, since one can construct the cross products on $\mathbb{R}^n$ by embedding it as the imaginary part of the (real, normed) division algebra of dimension $n+1$. The cross product is then obtained from the ordinary product after discarding the real component. For example, for $\mathbb{R}^1$ embedded in $\mathbb{C}$, one obtains a product of two purely imaginary numbers, which is real. Discarding this component gives exactly zero, as mentioned above.

Property 3.4.12 (Exceptional Lie group G_2). Using the vector product, one can define an associative 3-form, the **triple product**:

$$u \otimes v \otimes w \mapsto u \cdot (v \times w). \tag{3.46}$$

The group of linear isomorphisms that preserve the triple product on $\mathbb{R}^7$ is denoted by G_2. (By the relation between vector products and division algebras, this group is also the automorphism group of the octonions.)

3.4.4 Wedge product

Definition 3.4.13 (Antisymmetrization). Let $S_k =$ denote the permutation group **??** on k elements. The antisymmetrization operator is defined as follows:

$$\mathrm{Alt}(e_1 \otimes \cdots \otimes e_k) := \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) e_{\sigma(1)} \otimes \cdots \otimes e_{\sigma(k)}. \tag{3.47}$$

Note that many authors introduce a factor $1/k!$. This convention is not adopted here to keep the subsequent constructions clean. If the factor is included, Formula 3.4.16 below should be modified.

Definition 3.4.14 (Wedge product). Let $\{e_i\}_{i \leq \dim(V)}$ be a basis for V. The wedge product of basisvectors is defined as follows:

$$e_1 \wedge \ldots \wedge e_k = \mathrm{Alt}(e_1 \otimes \cdots \otimes e_k) \tag{3.48}$$

From this definition it immediately follows that the wedge product is (totally) antisymmetric.

Construction 3.4.15. Let $\{e_i\}_{i \leq \dim(V)}$ be a basis for V. The above definition implies that a basis for $\Lambda^r(V)$ is given by

$$\{e_{i_1} \wedge \ldots \wedge e_{i_r} \mid \forall k \leq r : 1 \leq i_k \leq \dim(V)\}.$$

Accordingly, the dimension of this space is given by

$$\dim \Lambda^r(V) = \binom{n}{r}. \tag{3.49}$$

For $r = 0$ this construction would be vacuous, so one just defines $\Lambda^0(V) := \mathbb{R}$.

Formula 3.4.16. Let $v \in \Lambda^r(V)$ and $w \in \Lambda^m(V)$. The wedge product 3.4.14 can be generalized as follows:

$$v \wedge w = \frac{1}{r!m!} \mathrm{Alt}(v \otimes w). \tag{3.50}$$

Definition 3.4.17 (Blades). Elements of $\Lambda^k(V)$ that can be written as the wedge product of k vectors are known as k-**blades** or **pure** k-**vectors**.

Formula 3.4.18 (Cross product). In dimension 3 there exists an important isomorphism $J : \Lambda^2(\mathbb{R}^3) \to \mathbb{R}^3$:

$$J(\lambda)^i = \frac{1}{2}\varepsilon^i{}_{jk}\lambda^{jk}, \tag{3.51}$$

where $\lambda \in \Lambda^2(\mathbb{R}^3)$. See also the Hodge $*$-operator 3.4.26 further below.

Looking at the definition of the cross product 3.4.10, one can see that $v \times w$ is actually the same as $J(v \wedge w)$. One can thus use the wedge product to generalize the cross product to arbitrary dimensions.

3.4.5 Exterior algebra

Definition 3.4.19 (Exterior power). In the theory of tensor calculus, the space $\Lambda^k(V)$ is often called the k^{th} exterior power of V. As mentioned before, its elements are called (exterior) k**-forms**.

Definition 3.4.20 (Exterior algebra). One can define a graded vector space **??** as follows:

$$\Lambda^\bullet(V) := \bigoplus_{k\geq 0} \Lambda^k(V). \tag{3.52}$$

This graded vector space can be turned into a graded algebra by taking the wedge product as the multiplication:

$$\wedge : \Lambda^k(V) \times \Lambda^l(V) \to \Lambda^{k+l}(V). \tag{3.53}$$

This algebra is called the exterior algebra or **Grassmann algebra** over V. Elements of the space $\otimes_{k\text{even}}\Lambda^k(V)$ are said to be **Grassmann-even** and elements of $\otimes_{k\text{odd}}\Lambda^k(V)$ are said to be **Grassmann-odd**.

Alternative Definition 3.4.21. Let $T(V)$ be the tensor algebra 3.3.7 over the vector space V, i.e.

$$T(V) = \bigoplus_{k\geq 0} V^{\otimes k}. \tag{3.54}$$

The exterior algebra $\Lambda^\bullet(V)$ over V is defined as the quotient of $T(V)$ by the two-sided ideal I generated by the elements $\{v \otimes v \mid v \in V\}$.

Proof of equivalence. Consider the equality

$$(u+v)\otimes(u+v) - u\otimes u - v\otimes v = u\otimes v + v\otimes u \tag{3.55}$$

The left-hand side is an element of the ideal I generated by $\{v \otimes v \mid v \in V\}$. Using the ideal generated by elements of the form of the right-hand side gives the usual definition of the exterior algebra based on the wedge product as defined in 3.4.14 because it imposes the relation $u \wedge v = -v \wedge u$. □

However, one should pay attention to one little detail. As mentioned in 3.4.21 the general definition uses the ideal I to construct the quotient space. The other construction is only equivalent when working over a field with characteristic different from 2. This follows from the fact that one has to divide by 2 when trying to obtain the ideal I from the right-hand side when setting $u = v$.

Property 3.4.22 (Graded-commutativity). The exterior algebra is both a unital associative algebra (with identity $1 \in K$) and a coalgebra. Furthermore, it is also commutative in the graded sense **??**.

Property 3.4.23 (Nilpotency). Graded-commutativity implies that the wedge product of any odd exterior form with itself is identically 0. The wedge product of an even exterior form with itself vanishes if and only if the form can be decomposed as a product of one-forms, i.e. if it is a pure k-form.

3.4.6 Hodge star

Equation (3.49) says that the spaces $\Lambda^k(V)$ and $\Lambda^{n-k}(V)$ have the same dimension and , hence, that there exists a linear isomorphism between them. Such an isomorphism is given by the Hodge star operator if one restricts to vector spaces equipped with a nondegenerate Hermitian form 2.3.2.

When equipped with an inner product and, hence, an orthonormal basis $\{e_i\}_{i\leq\dim(V)}$, every finite-dimensional vector space admits a canonical volume form given by

$$\mathrm{Vol} = e_1 \wedge \ldots \wedge e_n. \tag{3.56}$$

This convention will also be adopted in the remainder of this section.

Definition 3.4.24 (Orientation). Let $\mathrm{Vol}(V)$ be the standard volume form on the vector space V as defined above. From the definition of a volume form it follows that every other $\dim(V)$-form is a scalar multiple of $\mathrm{Vol}(V)$. Denote this number by r. This also implies that a choice of volume form induces an equivalence relation on top-dimensional forms. An equivalence class under this relation is called an orientation on V. If $r > 0$, the orientation is said to be **positive** and, if $r < 0$, the orientation is said to be **negative**.

Formula 3.4.25 (Inner product). Let V be equipped with an inner product $\langle\cdot|\cdot\rangle$. One can extend this to an inner product on $\Lambda^k(V)$ by first defining it on decomposable forms and extending it by linearity to all of $\Lambda^k(V)$:

$$\langle v_1 \wedge \ldots \wedge v_k | w_1 \wedge \ldots \wedge w_k\rangle_k := \det(\langle v_i|w_j\rangle). \tag{3.57}$$

For an orthogonal basis this formula factorizes as follows:

$$\langle v_1 \wedge \ldots \wedge v_k | w_1 \wedge \ldots \wedge w_k\rangle_k = \langle v_1|w_1\rangle \cdots \langle v_k|w_k\rangle. \tag{3.58}$$

Definition 3.4.26 (Hodge star). The Hodge star $* : \Lambda^k(V) \to \Lambda^{n-k}(V)$ is defined as the unique isomorphism such that for all $\omega \in \Lambda^k(V)$ and $\rho \in \Lambda^{n-k}(V)$ the following equality holds:

$$\omega \wedge \rho = \langle *\omega|\rho\rangle_{n-k}\,\mathrm{Vol}(V)\,, \tag{3.59}$$

where $\langle\cdot|\cdot\rangle_{n-k}$ is the inner product (3.57) on $\Lambda^{n-k}(V)$. The element $*\omega$ is often called the **(Hodge) dual** of ω.

Proof. Fix an element $\omega \in \Lambda^k(V)$. For every element $\rho \in \Lambda^{n-k}(V)$ one can see that $\omega \wedge \rho$ is an element of $\Lambda^n(V)$ and as such it is a scalar multiple of $\mathrm{Vol}(V)$. This implies that it can be written as

$$c_\omega(\rho)\,\mathrm{Vol}(V).$$

The map $c_\omega : \Lambda^{n-k}(V) \to \mathbb{R} : \rho \mapsto c_\omega(\rho)$ is a bounded (and thus continuous) linear map, so Riesz's representation theorem **??** can be applied to identify c_ω with a unique element $*\omega \in \Lambda^{n-k}(V)$ such that

$$c_\omega(\rho) = \langle *\omega|\rho\rangle_{n-k}.$$

□

Formula 3.4.27. Let $\{e_i\}_{i\leq n}$ be a positively oriented orthonormal basis for V. An explicit formula for the Hodge star is given by the following construction. Let $\{i_1,\ldots,i_k\}$ and $\{j_1,\ldots,j_{n-k}\}$ be two ordered, complementary index sets and consider an element $\omega = e_{i_1}\wedge\ldots\wedge e_{i_k}\in\Lambda^k(V)$.

$$*\omega = \mathrm{sgn}(\tau)\prod_{m=1}^{n-k}\langle e_{j_m}|e_{j_m}\rangle e_{j_1}\wedge\ldots\wedge e_{j_{n-k}}\,, \tag{3.60}$$

where τ is the permutation that maps $e_{i_1}\wedge\ldots\wedge e_{i_k}\wedge e_{j_1}\wedge\ldots\wedge e_{j_{n-k}}$ to $\mathrm{Vol}(V)$.

Using this formula one can easily prove the following important property:

Property 3.4.28. Consider an inner product space V. The Hodge dual is involutive up to a factor:

$$**\omega = (-1)^{k(n-k)}\omega. \tag{3.61}$$

Taking the defining relation of the Hodge star operator together with the above property implies the following formula (which is often found in the literature as the defining relation):

Formula 3.4.29. For all $\omega,\rho\in\Lambda^k(V)$ the Hodge star operator satisfies the following formula:

$$\omega\wedge *\rho = \langle\omega|\rho\rangle\,\mathrm{Vol}(V). \tag{3.62}$$

Corollary 3.4.30. Consider three vectors $u,v,w\in\mathbb{R}^3$.

$$*(v\wedge w) = v\times w \tag{3.63}$$

$$*(v\times w) = v\wedge w \tag{3.64}$$

$$*(u\wedge v\wedge w) = u\cdot(v\times w) \tag{3.65}$$

Remark 3.4.31. Formula (3.51) is an explicit evaluation of the first equation (3.63).

Proof. The signs $\mathrm{sgn}(\sigma)$ in the definition of wedge products can be written using the Levi-Civita symbol ε_{ijk} as defined in 3.4.8. The factor $\frac{1}{2}$ is introduced to correct for the double counting due to the contraction over both the indices j and k. □

Definition 3.4.32 (Self-dual form). Let V be a 4-dimensional inner product space and consider an element $\omega\in\Lambda^2(V)$. Then ω is said to be self-dual if

$$*\omega = \omega. \tag{3.66}$$

Furthermore, every element $\rho\in\Lambda^2(V)$ can be uniquely decomposed as the sum of a self-dual and an anti-self-dual two-form:

$$\rho = \frac{1}{2}\big((\rho + *\rho) + (\rho - *\rho)\big). \tag{3.67}$$

Chapter 4

Representation Theory

References for this chapter are [52, 66]. Sections **??** and **??** can be visited for an introduction to groups and group actions.

4.1 Group representations

Group actions on vector spaces are so important that they receive their own name:

Definition 4.1.1 (Representation). A representation of a group G on a vector space V is a group morphism $\rho : G \to \mathrm{GL}(V)$ from G to the automorphism group 2.2.8 of V.

Property 4.1.2 (Freeness). Because every linear map takes the zero vector to itself, a representation can never be free.

Definition 4.1.3 (Subrepresentation). A subrepresentation of a representation on V is a subspace of V invariant under the action of the group G (together with the restricted action).

Example 4.1.4 (Permutation representation). Consider a vector space V with basis $\{e_i\}_{i\leq n}$ and let $G = S_n$ be the symmetric group on n elements. Based on Definition **??**, one can consider the action of G on the index set $\{1, \ldots, n\}$. This induces a representation given by

$$\rho(g) : \sum_{i=1}^{n} v_i e_i \mapsto \sum_{i=1}^{n} v_i e_{g\cdot i}. \tag{4.1}$$

Example 4.1.5 (Contragredient representation). For every representation ρ on V, there exists a natural representation on the dual space V^*:

$$\rho^*(g) := \rho^T(g^{-1}) : V^* \to V^*, \tag{4.2}$$

where ρ^T is the transpose 2.2.23. It is implicitly defined by requiring

$$\Big\langle \rho^*(g)(v^*), \rho(g)(v) \Big\rangle = \langle v^*, v\rangle \tag{4.3}$$

for all $v \in V$ and $v^* \in V^*$, where $\langle\cdot,\cdot\rangle$ is the natural pairing.

Example 4.1.6 (Tensor product representation). A group G that acts on vector spaces V, W also has a representation on the tensor product $V \otimes W$ in the following way:

$$\rho_{V\otimes W}(g)(v \otimes w) := \rho_V(g)(v) \otimes \rho_W(g)(w). \tag{4.4}$$

More generally, consider two representations $\phi : G \to \mathrm{GL}(V)$ and $\psi : H \to \mathrm{GL}(W)$. A representation of the direct product $G \times H$ on $V \otimes W$ is given by the tensor product of representations:

$$\rho(g,h)(v \otimes w) := \phi(g)(v) \otimes \psi(h)(w). \tag{4.5}$$

The former case can be obtained as a subrepresentation induced by the diagonal subgroup inclusion $\Delta_G : G \hookrightarrow G \times G$.

Definition 4.1.7 (Intertwiner). If one views G-representations as G-modules, the natural morphisms are the intertwiners **??**.

4.2 Irreducible representations

Definition 4.2.1 (Irreducibility). A representation is said to be irreducible if there exist no proper nonzero subrepresentations.

Example 4.2.2 (Standard representation). Consider the action of S_n on a vector space V with basis $\{e_i\}_{i\leq n}$. The line generated by $e_1 + e_2 + \ldots + e_n$ is invariant under the permutation action of S_n. It follows that the permutation representation (on finite-dimensional spaces) is never irreducible.

The $(n-1)$-dimensional complementary subspace

$$W = \left\{ \sum_{i=1}^n \lambda_i e_i \,\middle|\, \sum_{i=1}^n \lambda_i = 0 \right\} \tag{4.6}$$

forms an irreducible representation. It is called the standard representation of S_n on V.

Schur's lemma **??** and its corollary are usually found in the following form:

Theorem 4.2.3 (Schur's lemma). *Let V, W be two finite-dimensional irreducible representations of a group G and let $\varphi : V \to W$ be an intertwiner.*

- *Either 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 4.2.4 (Complementary representation). If W is a subrepresentation of V, there exists an invariant complementary subspace W'. This space can be found as follows. Choose an arbitrary complement U such that $V = W \oplus U$ with associated projection map $\pi_0 : V \to W$. Averaging over G gives the intertwiner

$$\pi(v) := \sum_{g\in G} g \circ \pi_0(g^{-1}v). \tag{4.7}$$

On W it is given by multiplication by $|G|$. Its kernel is an invariant subspace of V, complementary to W.

Theorem 4.2.5 (Maschke). *Let G be a finite group with a representation space V such that the characteristic of the underlying field does not divide the order of G. This representation can be uniquely (up to isomorphism) decomposed as*

$$V = V_1^{\oplus a_1} \oplus \cdots \oplus V_k^{\oplus a_k}, \tag{4.8}$$

where all V_k's are distinct irreducible representations.

4.3 Classification by Young tableaux

For an introduction to Young tableaux, see Section **??**.

Definition 4.3.1 (Permutation module). Let λ be a partition. The permutation module M^λ is defined as the vector space generated by the Young tabloids of shape λ.

Definition 4.3.2 (Specht module). Consider a permutation module M^λ for some λ with $|\lambda| = n$. Since the permutation group S_n acts on Young tableaux by permuting the entries, it also has an induced action[1] on M^λ. For every Young tableau y, it has a subrepresentation induced by the subgroup Q_t of permutations that leave the columns invariant, which is spanned by the following elements (called **polytabloids**):

$$e_t := \sum_{\sigma\in Q_t} \operatorname{sgn}(\sigma)\sigma \cdot \{t\}, \tag{4.9}$$

where t ranges over the Young tableaux of shape λ and $\{t\}$ denotes the Young tabloid associated to the tableau t. In fact, one can just take the standard Young tableaux as generators. This is sometimes called the **Specht basis**.

Property 4.3.3. A representation (over $\mathbb{C}$) of S_n is irreducible if and only if it is a Specht module S^λ for some partition λ of n.

One can restate the definition of a Specht module using the following operator:

Definition 4.3.4 (Young symmetrizer). Given a Young tableau t of shape λ, one can decompose the permutation group $S_{|\lambda|}$ as the union of two types of permutations. First one has the permutations that preserve the rows, denote these by P_λ. Then one has also the permutations that preserve the columns, denote these by Q_λ. These two subgroups induce elements in the group algebra $\mathbb{C}[S_{|\lambda|}]$ as follows:

$$a_\lambda := \sum_{p\in P_\lambda} p\,, \tag{4.10}$$

$$b_\lambda := \sum_{q\in Q_\lambda} \operatorname{sgn}(q)q. \tag{4.11}$$

The product $c_\lambda := b_\lambda a_\lambda$ is called the Young symmetrizer of λ.

Comparing the above definition of the Specht module to the Young symmetrizers leads to the following alternative definition.

Alternative Definition 4.3.5 (Specht module). The space $\mathbb{C}[S_{|\lambda|}]c_\lambda$ is called the Specht module S_λ.

Consider a vector space V together with its general linear group $\mathrm{GL}(V)$. For all $n \in \mathbb{N}$ there is an induced (diagonal) representation on $V^{\otimes n}$ by $\mathrm{GL}(V)$. There is also an action by the permutation group S_n that permutes the elements in a monomial $v_1 \otimes \cdots \otimes v_n \in V^{\otimes n}$.

Theorem 4.3.6 (Schur-Weyl). *The above representation of* $\mathrm{GL}(V) \times S_n$ *can be decomposed as follows:*

$$V^{\otimes n} \cong \bigoplus_{|\lambda|=n} V_\lambda \otimes S_\lambda V, \tag{4.12}$$

where

[1] This can be generalized to an action of the group algebra $K[S_n]$.

- *the sum ranges over all partitions of n or, equivalently, all Young diagrams with n boxes,*
- *the V_λ are Specht modules and, hence, irreducible representations of S_n, and*
- *the $S_\lambda V$ are (possibly zero) irreducible representations of* $\mathrm{GL}(V)$ *of the form* $S_\lambda V \equiv \mathrm{Hom}_{S_n}(V_\lambda, V^{\otimes n})$.

The spaces V_λ can be interpreted as multiplicity spaces. The spaces $S_\lambda V$ can be rewritten more explicitly as $V_\lambda \otimes_{S_n} V^{\otimes n}$ by self-duality of S_n-representations or, by using the explicit characterization of Specht modules given above, as $V^{\otimes n}c_\lambda$. Because of the functoriality of the involved operations, one can also see that S_λ is in fact a functor, the **Schur functor**.

Example 4.3.7 (Algebraic curvature tensor). The above definition of the Schur spaces $S_\lambda V$ allows for an a concise description of representations in terms of Young diagrams (and tableaux). Consider for example the Riemann curvature tensor R_{ijkl} from Chapter **??**. This tensor has the following symmetries:

- $R_{ijkl} = -R_{jikl} = -R_{ijlk}$,
- $R_{ijkl} = R_{klij}$, and
- $R_{ijkl} + R_{iljk} + R_{iklj} = 0$.

By looking at the definition of the Young symmetrizer, one can see that these symmetries are exactly those of the irreducible component in $S_\lambda V$ associated to the partition $(2, 2)$.

?? COMPLETE ??

4.4 Tensor operators

Definition 4.4.1 (Representation operators). An intertwiner $\psi : (\rho, V_0) \to \mathrm{End}(V)$ between a G-representation on an auxiliary vector space V_0 to the space of linear maps on a G-vector space V (equipped with the adjoint action).

More explicitly, consider a set of operators $\{\hat{O}_i\}_{i\in I} \subset \mathrm{End}(V)$ acting on a vector space V equipped with a representation ρ of the group G. This collection defines a representation operator with respect to G if there exists a matrix representation R of G such that the following equation holds:

$$\rho(g)\hat{O}_i\rho(g)^{-1} = \sum_{j\in I} R(g)_{ij}\hat{O}_j. \tag{4.13}$$

Example 4.4.2 (Tensor operators). Consider $G = \mathrm{SO}(3)$ and $V_0 = \mathcal{T}^r_s(\mathbb{R}^3)$. This choice gives a set of operators that transform as tensors under rotations. By choosing $V_0 = \mathbb{R}^3$ or $V_0 = \mathcal{H}_l(\mathbb{R}^3)$, the space of spherical harmonics of degree l, one obtains the **vector** and **spherical operators**.

The following property is often used in quantum mechanics to quickly find forbidden transitions in atomic or molecular systems:

Property 4.4.3 (Selection rules). Let G be a semisimple group and let W_1, W_2 be two inequivalent (finite-dimensional), irreducible, unitary subrepresentations of a Hilbert space $\mathcal{H}$. Let $\hat{O}$ be a representation operator indexed by a vector space V. For all $v \in V, w_i \in W_i$ one has

$$\langle w_1|\hat{O}(v)w_2\rangle = 0\,, \tag{4.14}$$

unless $V \otimes W_2$ contains a subrepresentation equivalent to W_1.

Theorem 4.4.4 (Wigner-Eckart). *Consider two irreducible* $\mathrm{SU}(2)$*-subrepresentations* W_j *and* $W_{j'}$ *of some unitary representation* $\mathcal{H}$*, together with two degree-q spherical tensors* $\hat{O}, \hat{O}' : V_0 \to \mathrm{End}(\mathcal{H})$*. If there exists at least one index* $k \leq q$ *and one pair of vectors* $(v, v') \in W_j \times W_{j'}$ *such that*

$$\langle v' | \hat{O}_k v \rangle \neq 0\,,$$

then for all indices $k \geq 0$ *and pairs* $(v, v') \in W_j, W_{j'}$ *the following equality holds*

$$\langle v' | \hat{O}'_k v \rangle = C \langle v' | \hat{O}_k v \rangle \tag{4.15}$$

for some constant C *that only depends on* q, j *and* j'*.*

By noting that the Clebsch-Gordan coefficients are the components of the projection $W_q \otimes W_j \to W_{j'}$, which is itself an intertwiner, one can recast the Wigner-Eckart theorem as a statement about matrix elements:

Corollary 4.4.5. Consider an irreducible tensor operator T_j^m (with respect to the rotation group). The matrix elements of this operator with respect to a symmetry-adapted basis (“angular momentum” basis) decompose as a product of a Clebsch-Gordan coefficient and a factor that only depends on the eigenvalues of the Casimir operator:

$$\langle j', m' | R^{(q)} | j, m \rangle = \langle j' \| R_k^{(q)} \| j \rangle \langle j', m' | q, j; k, m \rangle. \tag{4.16}$$

The factor $\langle j' \| R^{(q)} \| j \rangle$ is sometimes called the **reduced matrix element**.

List of Symbols

The following symbols are used throughout the summary:

Abbreviations

AIC	Akaike information criterion
ARMA	autoregressive moving-average model
BCH	Baker-Campbell-Hausdorff
CCR	canonical commutation relation
CDF	cumulative distribution function
CFT	conformal field theory
CIS	completely integrable system
CP	completely positive
CPTP	completely positive, trace-preserving
CR	Cauchy-Riemann
DGA	differential graded algebra
DGCA	differential graded-commutative algebra
EPR	Einstein-Podolsky-Rosen
ETCS	Elementary Theory of the Category of Sets
FWHM	full width at half maximum
GA	geometric algebra
GHZ	Greenberger-Horne-Zeilinger
GNS	Gel'fand-Naimark-Segal
HoTT	Homotopy Type Theory
KKT	Karush-Kuhn-Tucker
LIVF	left-invariant vector field
MPO	matrix product operator
MPS	matrix product state
MTC	modular tensor category
NDR	neighbourhood deformation retract
OPE	operator product expansion
OZI	Okubo-Zweig-Iizuka
PAC	probably approximately correct
PL manifold	piecewise-linear manifold
PVM	projection-valued measure

RKHS	reproducing kernel Hilbert space
SVM	support-vector machine
TQFT	topological quantum field theory
VIF	variance inflation factor
ZFC	Zermelo-Frenkel set theory with the axiom of choice
TVS	topological vector space
Operations	
Ad_g	adjoint representation of a Lie group G
ad_X	adjoint representation of a Lie algebra $\mathfrak{g}$
arg	argument of a complex number
$\square$	d'Alembert operator
$\deg(f)$	degree of the polynomial f
e	identity element of a group
$\Gamma(E)$	set of global sections of a fibre bundle E
Im	imaginary part of a complex number
$\mathrm{Ind}_f(z)$	index of a point $z \in \mathbb{C}$ with respect to a function f
$\hookrightarrow$	injective function
$\cong$	is isomorphic to
Par_t^γ	parallel transport map with respect to the curve γ
Re	real part of a complex number
Res	residue of a complex function
$\twoheadrightarrow$	surjective function
$\{\cdot,\cdot\}$	Poisson bracket
∂X	boundary of a topological space X
$\overline{X}$	closure of a topological space X
$X^\circ, \mathring{X}$	interior of a topological space X
$\sphericalangle(\cdot,\cdot)$	angle between two vectors
$X \times Y$	cartesian product of the sets X and Y
$X + Y$	sum of the vector spaces X and Y
$X \oplus Y$	direct sum of the vector spaces X and Y
$V \otimes W$	tensor product of the vector spaces V and W
$\mathbb{1}_X$	identity morphism on the object X
$\approx$	is approximately equal to
$\hookrightarrow$	is included in
$\cong$	is isomorphic to
$\mapsto$	mapsto
Collections	
Ab	category of Abelian groups
$\mathrm{Aut}(X)$	automorphism group of an object X
$\mathcal{B}_0(V,W)$	space of compact bounded operators between the Banach spaces V and W

$\mathcal{B}(V, W)$	space of bounded linear maps from the space V to the space W
CartSp	the category of Euclidean spaces and "suitable" homomorphisms (e.g. linear maps, smooth maps, ...)
$C_\bullet$	chain complex
$\mathbf{Ch}(\mathbf{A})$	category of chain complexes with objects in the additive category $\mathbf{A}$
$\mathbf{C}^\infty$	category of smooth spaces
$C_p^\infty(M)$	ring of smooth functions $f : M \to \mathbb{R}$ on a neighbourhood of $p \in M$
$C^\omega(V)$	the set of all analytic functions defined on the set V
$\mathrm{Conf}(M)$	conformal group of (pseudo-)Riemannian manifold M
$C(X, Y)$	set of continuous functions between two topological spaces X and Y
$\mathbf{C}^\infty\mathbf{Ring}, \mathbf{C}^\infty\mathbf{Alg}$	category of smooth algebras
Diff	category of smooth manifolds
DiffSp	category of diffeological spaces and smooth maps
D^n	standard n-disk
$\mathrm{dom}(f)$	domain of a function f
$\mathrm{End}(X)$	endomorphism monoid of a an object X
$\mathcal{E}\mathrm{nd}$	endomorphism operad
$\mathbf{FormalCartSp}_{\mathbf{diff}}$	category of infinitesimally thickened Euclidean spaces
$\mathrm{GL}(V)$	general linear group, the group of automorphisms of a vector space V
$\mathrm{GL}(n, K)$	general linear group: the group of all invertible $n \times n$-matrices over the field K
Grp	category of groups and group homomorphisms
Grpd	category of groupoids
$\mathrm{Hol}_p(\omega)$	holonomy group at the point p with respect to the principal connection ω
$\mathrm{Hom}_{\mathbf{C}}(V, W)$	set of homomorphisms from an object V to an object W in a category $\mathbf{C}$
hTop	homotopy category
$\mathrm{im}(f)$	image of a function f
$K^0(X)$	K-theory over a (compact Hausdorff) space X
Kan	category of Kan complexes
$\mathcal{K}_n(A, v)$	Krylov subspace of dimension n generated by the matrix A and the vector v
L^1	space of integrable functions
Law	category of Lawvere theories
Lie	category of Lie groups
$\mathfrak{Lie}$	category of Lie algebras
$\mathfrak{X}^L$	space of left-invariant vector fields on a Lie group
LX	free loop space on X
$\mathbf{Man}^p$	category of C^p-manifolds
Meas	category of measure spaces and measure-preserving functions
$N\mathbf{C}$	the simplicial nerve of a small category $\mathbf{C}$
$\mathbf{Open}(X)$	category of open subsets of a topological space X
$\mathrm{O}(n, K)$	group of $n \times n$ orthogonal matrices over a field K
$P(S), 2^S$	power set of S

$\mathrm{Pin}(V)$	pin group of the Clifford algebra $C\ell(V, Q)$
$\mathbf{Psh}(\mathbf{C}), \widehat{\mathbf{C}}$	category of presheaves on a (small) category $\mathbf{C}$
$\mathbf{Sh}(X)$	category of sheaves on a topological space X
$\mathbf{Sh}(\mathbf{C}, J)$	category of J-sheaves on a site $(\mathbf{C}, J)$
$\boldsymbol{\Delta}$	simplex category
$\mathrm{SL}_n(K)$	special linear group: group of all invertible n-dimensional matrices with unit determinant over the field K
S^n	standard n-sphere
$S^n(V)$	space of symmetric rank n tensors over a vector space V
$W^{m,p}(U)$	the Sobolov space in L^p of order m
$\mathbf{Span}(\mathbf{C})$	span category over $\mathbf{C}$
$\mathrm{Spec}(R)$	spectrum of a commutative ring R
$\mathrm{supp}(f)$	support of a function f
$\mathrm{Syl}_p(G)$	set of Sylow p-subgroups of a finite group G
S_n	symmetric group of degree n
$\mathrm{Sym}(X)$	symmetric group on the set X
$\mathrm{Sp}(n, K)$	group of matrices preserving a canonical symplectic form over the field K
$\mathrm{Sp}(n)$	compact symplectic group
$\mathrm{TL}_n(\delta)$	Temperley-Lieb algebra with $n-1$ generators and parameter δ.
T^n	standard n-torus (the n-fold Cartesian product of S^1)
Top	category of topological spaces
Topos	the 2-category of (elementary) topoi and geometric morphisms
$U(\mathfrak{g})$	universal enveloping algebra of a Lie algebra $\mathfrak{g}$
$\mathrm{U}(n, K)$	group of $n \times n$ unitary matrices over a field K
$\mathbf{Vect}(X)$	category of vector bundles over a manifold X
$\mathbf{Vect}_K$	category of vector spaces and linear maps over a field K
Y^X	set of functions from a set X to a set Y
$\emptyset$	empty set
$\pi_n(X, x_0)$	n^{th} homotopy space over X with basepoint x_0
$[a, b]$	closed interval
$]a, b[$	open interval
$\Lambda^n(V)$	space of antisymmetric rank n tensors over a vector space V
ΩX	(based) loop space on X
$\Omega^k(M)$	$C^\infty(M)$-module of differential k-forms on the manifold M
$\rho(A)$	resolvent set of a bounded linear operator A
$\mathfrak{X}(M)$	$C^\infty(M)$-module of vector fields on the manifold M
Units	
C	coulomb
T	tesla

Bibliography

[1] Andreas Gathmann. Algebraic geometry. `https://www.mathematik.uni-kl.de/~gathmann/class/alggeom-2019/alggeom-2019.pdf`.

[2] David Mumford. *The Red Book of Varieties and Schemes: Includes the Michigan Lectures (1974) on Curves and Their Jacobians*, volume 1358. Springer Science & Business Media, 1999.

[3] Emily Riehl. Homotopical categories: From model categories to $(\infty, 1)$-categories. 2019. arXiv:1904.00886.

[4] Mark Hovey. *Model Categories*. Number 63. American Mathematical Soc., 2007.

[5] Emily Riehl. Monoidal algebraic model structures. *Journal of Pure and Applied Algebra*, 217(6):1069--1104, 2013.

[6] Charles Rezk. A model for the homotopy theory of homotopy theory. *Transactions of the American Mathematical Society*, 353(3):973--1007, 2001.

[7] Emily Riehl and Dominic Verity. The theory and practice of Reedy categories. *Theory and Applications of Categories*, 29, 2013.

[8] Pascal Lambrechts.

[9] Emily Riehl. Homotopy (limits and) colimits. `http://www.math.jhu.edu/~eriehl/hocolimits.pdf`.

[10] Jade Master. Why is homology so powerful? 2020. arXiv:2001.00314.

[11] Peter T. Johnstone. *Topos Theory*. Dover Publications, 2014.

[12] Olivia Caramello. Lectures on topos theory at the university of Insubria. `https://www.oliviacaramello.com/Teaching/Teaching.htm`.

[13] Angelo Vistoli. Notes on Grothendieck topologies, fibered categories and descent theory. *arXiv:math/0412512*, 2004.

[14] Nima Moshayedi. 4-manifold topology, donaldson-witten theory, floer homology and higher gauge theory methods in the BV-BFV formalism. 2021.

[15] John C. Baez and Urs Schreiber. Higher gauge theory. 2005. arXiv:math/0511710.

[16] Urs Schreiber. *From Loop Space Mechanics to Nonabelian Strings*. PhD thesis, 2005.

[17] Richard Sanders. Commutative spectral triples & the spectral reconstruction theorem.

[18] John Baez and Alexander Hoffnung. Convenient categories of smooth spaces. *Transactions of the American Mathematical Society*, 363(11):5789--5825, 2011.

[19] John C. Baez and Aaron D. Lauda. Higher-dimensional algebra v: 2-groups. 2003. arXiv:math/0307200.

[20] Thomas Augustin, Frank PA Coolen, Gert De Cooman, and Matthias CM Troffaes. *Introduction to imprecise probabilities.* John Wiley & Sons, 2014.

[21] Tetsuji Miwa, Michio Jimbo, Michio Jimbo, and E Date. *Solitons: Differential Equations, Symmetries and Infinite-dimensional Algebras*, volume 135. Cambridge University Press, 2000.

[22] Vladimir I. Arnol'd. *Mathematical Methods of Classical Mechanics*, volume 60. Springer Science & Business Media, 2013.

[23] Edwin J. Beggs and Shahn Majid. *Quantum Riemannian Geometry.* Springer, 2020.

[24] Marc Henneaux and Claudio Teitelboim. *Quantization of Gauge Systems.* Princeton university press, 1992.

[25] Gregory M. Kelly. *Basic Concepts of Enriched Category Theory*, volume 64. CUP Archive, 1982.

[26] Vladimir Vovk, Alex Gammerman, and Glenn Shafer. *Algorithmic Learning in a Random World.* Springer Science & Business Media, 2005.

[27] Mukund Rangamani and Tadashi Takayanagi. *Holographic Entanglement Entropy.* Springer, 2017.

[28] Saunders Mac Lane. *Categories for the Working Mathematician*, volume 5. Springer Science & Business Media, 2013.

[29] Shun-ichi Amari. *Information Geometry and Its Applications.* Springer Publishing Company, Incorporated, 2016.

[30] Charles W. Misner, Kip S. Thorne, and John A. Wheeler. *Gravitation.* Princeton University Press, 2017.

[31] Carlo Rovelli and Francesca Vidotto. *Covariant Loop Quantum Gravity: An Elementary Introduction to Quantum Gravity and Spinfoam Theory.* Cambridge University Press, 2014.

[32] Richard W. Sharpe. *Differential Geometry: Cartan's Generalization of Klein's Erlangen Program*, volume 166. Springer Science & Business Media, 2000.

[33] John C. Baez, Irving E. Segal, and Zhengfang Zhou. *Introduction to Algebraic and Constructive Quantum Field Theory.* Princeton University Press, 2014.

[34] Raoul Bott and Loring W. Tu. *Differential Forms in Algebraic Topology.* Graduate Texts in Mathematics. Springer New York, 1995.

[35] The Univalent Foundations Program. *Homotopy Type Theory: Univalent Foundations of Mathematics.* `https://homotopytypetheory.org/book`, Institute for Advanced Study, 2013.

[36] Bruce Blackadar. *Operator Algebras: Theory of C∗-Algebras and von Neumann Algebras.* Springer, 2013.

[37] Marek Capinski and Peter E. Kopp. *Measure, Integral and Probability.* Springer Science & Business Media, 2013.

[38] Georgiev Svetlin. *Theory of Distributions.* Springer, 2015.

[39] Gerd Rudolph and Matthias Schmidt. *Differential Geometry and Mathematical Physics: Part II. Fibre Bundles, Topology and Gauge Fields*. Springer, 2017.

[40] Martin Schottenloher. *A Mathematical Introduction to Conformal Field Theory*, volume 759. 2008.

[41] Dusa McDuff and Deitmar Salamon. *Introduction to Symplectic Topology*. Oxford Graduate Texts in Mathematics. Oxford University Press, 2017.

[42] John C. Baez and Peter May. *Towards Higher Categories*, volume 152 of *IMA Volumes in Mathematics and its Applications*. Springer, 2009.

[43] Mikhail. M. Kapranov and Vladimir A. Voevodsky. *2-categories and Zamolodchikov Tetrahedra Equations*, volume 56 of *Proc. Sympos. Pure Math.* Amer. Math. Soc., Providence, RI, 1994.

[44] Geoffrey Compère. *Advanced Lectures on General Relativity*, volume 952. Springer, 2019.

[45] Pavel Etingof, Shlomo Gelaki, Dmitri Nikshych, and Victor Ostrik. *Tensor Categories*, volume 205. American Mathematical Soc., 2016.

[46] Charles A. Weibel. *An Introduction to Homological Algebra*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1994.

[47] Peter J. Hilton and Urs Stammbach. *A Course in Homological Algebra*. Springer.

[48] Jean-Luc Brylinski. *Loop Spaces, Characteristic Classes and Geometric Quantization*. Birkhauser.

[49] Antoine Van Proeyen and Daniel Freedman. *Supergravity*. Cambridge University Press.

[50] William S. Massey. *A Basic Course in Algebraic Topology*. Springer.

[51] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory*. Westview Press.

[52] Nadir Jeevanjee. *An Introduction to Tensors and Group Theory for Physicists*. Birkhauser.

[53] Yvonne Choquet-Bruhat, Cecile DeWitt-Morette, and Margaret Dillard-Bleick. *Analysis, Manifolds and Physics, Part 1: Basics*. North-Holland.

[54] Yvonne Choquet-Bruhat and Cecile DeWitt-Morette. *Analysis, Manifolds and Physics, Part 2*. North-Holland.

[55] Herbet Goldstein, John L. Safko, and Charles P. Poole. *Classical Mechanics*. Pearson.

[56] Franco Cardin. *Elementary Symplectic Topology and Mechanics*. Springer.

[57] Walter Greiner and Joachim Reinhardt. *Field Quantization*. Springer.

[58] Walter Greiner. *Quantum Mechanics*. Springer.

[59] B. H. Bransden and Charles J. Joachain. *Quantum Mechanics*. Prentice Hall.

[60] Heydar Radjavi and Peter Rosenthal. *Invariant Subspaces*. Dover Publications.

[61] Max Karoubi. *K-Theory: An Introduction*. Springer.

[62] Damien Calaque and Thomas Strobl. *Mathematical Aspects of Quantum Field Theories*. Springer, 2015.

[63] Ivan Kolar, Peter W. Michor, and Jan Slovak. *Normal Operations in Differential Geometry*. Springer.

[64] Stephen B. Sontz. *Principal Bundles: The Classical Case.* Springer.

[65] Stephen B. Sontz. *Principal Bundles: The Quantum Case.* Springer.

[66] William Fulton and Joe Harris. *Representation Theory: A First Course.* Springer.

[67] Peter Petersen. *Riemannian Geometry.* Springer.

[68] Charles Nash and Siddharta Sen. *Topology and Geometry for Physicists.* Dover Publications.

[69] Ian M. Anderson. *The Variational Bicomplex.*

[70] Joel Robbin and Dietmar Salamon. The maslov index for paths. *Topology*, 32(4):827--844, 1993.

[71] Edward Witten. Global anomalies in string theory. In *Symposium on Anomalies, Geometry, Topology*, 6 1985.

[72] F.A. Berezin and M.S. Marinov.

[73] Paul A. M. Dirac. Generalized Hamiltonian dynamics. *Canadian Journal of Mathematics*, 2:129–148, 1950.

[74] Floris Takens. A global version of the inverse problem of the calculus of variations. *Journal of Differential Geometry*, 14(4):543--562, 1979.

[75] John C. Baez and Alissa S. Crans. Higher-dimensional algebra vi: Lie 2-algebras. 2003. arXiv:math/0307263.

[76] Edward Witten. Supersymmetry and Morse theory. *J. Diff. Geom*, 17(4):661--692, 1982.

[77] Marcus Berg, Cécile DeWitt-Morette, Shangjr Gwo, and Eric Kramer. The Pin groups in physics: C, P and T. *Reviews in Mathematical Physics*, 13(08):953--1034, 2001.

[78] Arindam Banerjee, Srujana Merugu, Inderjit S Dhillon, and Joydeep Ghosh. Clustering with Bregman divergences. *Journal of Machine Learning Research*, 6(Oct):1705--1749, 2005.

[79] Jean-Daniel Boissonnat, Frank Nielsen, and Richard Nock. Bregman Voronoi diagrams. *Discrete & Computational Geometry*, 44(2):281--307, 2010.

[80] Richard Palais. The symmetries of solitons. *Bulletin of the American Mathematical Society*, 34(4):339--403, 1997.

[81] Michael F. Atiyah. Topological quantum field theory. *Publications Mathématiques de l'IHÉS*, 68:175--186, 1988.

[82] Jens Eisert, Christoph Simon, and Martin B Plenio. On the quantification of entanglement in infinite-dimensional quantum systems. *Journal of Physics A: Mathematical and General*, 35(17):3911--3923, 2002.

[83] Benoît Tuybens. Entanglement entropy of gauge theories. 2017.

[84] Lotfi A. Zadeh. Fuzzy sets. *Information and Control*, 8(3):338--353, 1965.

[85] John C. Baez, Alexander E. Hoffnung, and Christopher Rogers. Categorified symplectic geometry and the classical string. *Communications in Mathematical Physics*, 293:701--725, 2010.

[86] Glenn Shafer and Vladimir Vovk. A tutorial on conformal prediction. *J. Mach. Learn. Res.*, 9:371-–421, 2008.

[87] Victor Chernozhukov, Kaspar Wüthrich, and Zhu Yinchu. Exact and robust conformal inference methods for predictive machine learning with dependent data. In *Proceedings of the 31st Conference On Learning Theory*, volume 75 of *Proceedings of Machine Learning Research*, pages 732--749. PMLR, 2018.

[88] Peter May. A note on the splitting principle. *Topology and Its Applications*, 153(4):605--609, 2005.

[89] Irina Markina. Group of diffeomorphisms of the unit circle as a principal $U(1)$-bundle.

[90] Sjoerd E. Crans. Localizations of transfors. 1998.

[91] Tom Leinster. Basic bicategories. 1998. arXiv:math/9810017.

[92] Alexander E. Hoffnung. Spans in 2-categories: A monoidal tricategory. 2011. arXiv:1112.0560.

[93] Eugenia Cheng and Nick Gurski. The periodic table of n-categories for low dimensions ii: Degenerate tricategories. 2007. arXiv:0706.2307.

[94] Mehmet B. Şahinoğlu, Dominic J. Williamson, Nick Bultinck, Michael Mariën, Jutho Haegeman, Norbert Schuch, and Frank Verstraete. Characterizing topological order with matrix product operators. 2014. arXiv:1409.2150.

[95] Dominic J. Williamson, Nick Bultinck, Michael Mariën, Mehmet B. Şahinoğlu, Jutho Haegeman, and Frank Verstraete. Matrix product operators for symmetry-protected topological phases: Gauging and edge theories. *Phys. Rev. B*, 94, 2016.

[96] Guifré Vidal. Efficient classical simulation of slightly entangled quantum computations. *Phys. Rev. Lett.*, 91, 2003.

[97] Aaron D. Lauda and Hendryk Pfeiffer. Open–closed strings: Two-dimensional extended TQFTs and Frobenius algebras. *Topology and its Applications*, 155(7):623--666, 2008.

[98] Domenico Fiorenza. An introduction to the Batalin-Vilkovisky formalism. 2004. arXiv:math/0402057v2.

[99] Stefan Cordes, Gregory Moore, and Sanjaye Ramgoolam. Lectures on 2d Yang-Mills theory, equivariant cohomology and topological field theories. arXiv:hep-th/9411210v2.

[100] Donald C. Ferguson. A theorem of Looman-Menchoff. `http://digitool.library.mcgill.ca/thesisfile111406.pdf`.

[101] Holger Lyre. Berry phase and quantum structure. arXiv:1408.6867.

[102] Florin Belgun. Gauge theory. `http://www.math.uni-hamburg.de/home/belgun/Gauge4.pdf`.

[103] Vladimir Itskov, Peter J. Olver, and Francis Valiquette. Lie completion of pseudogroups. *Transformation Groups*, 16:161--173, 2011.

[104] Richard Borcherds. Lie groups. `https://math.berkeley.edu/~reb/courses/261/`.

[105] Andrei Losev. From Berezin integral to Batalin-Vilkovisky formalism: A mathematical physicist's point of view. 2007.

[106] Edward Witten. Coadjoint orbits of the Virasoro group. *Comm. Math. Phys.*, 114(1):1--53, 1988.

[107] Sidney R. Coleman and Jeffrey E. Mandula. All possible symmetries of the S-matrix. *Phys. Rev.*, 159, 1967.

[108] Valter Moretti. Mathematical foundations of quantum mechanics: An advanced short course. *International Journal of Geometric Methods in Modern Physics*, 13, 2016.

[109] Antonio Michele Miti. Homotopy comomentum maps in multisymplectic geometry, 2021.

[110] Niclas Sandgren and Petre Stoica. On moving average parameter estimation. Technical Report 2006-022, Department of Information Technology, Uppsala University, 2006.

[111] John E. Roberts. Spontaneously broken gauge symmetries and superselection rules. 1974.

[112] Jean Gallier. Clifford algebras, Clifford groups, and a generalization of the quaternions, 2008. arXiv:0805.0311.

[113] Bozhidar Z. Iliev. Normal frames for general connections on differentiable fibre bundles. arXiv:math/0405004.

[114] Piotr Stachura. Short and biased introduction to groupoids. arXiv:1311.3866.

[115] Fosco Loregian. Coend calculus. arXiv:1501.02503.

[116] Frederic Schuller. Lectures on the geometric anatomy of theoretical physics. `https://www.youtube.com/channel/UC6SaWe7xeOp31Vo8cQG1oXw`.

[117] Nima Amini. Infinite-dimensional Lie algebras. `https://people.kth.se/~namini/PartIIIEssay.pdf`.

[118] Peter Selinger. Lecture notes on lambda calculus.

[119] Nigel Hitchin. Lectures on special Lagrangian submanifolds. `https://arxiv.org/abs/math/9907034v1`, 1999.

[120] Derek Sorensen. An introduction to characteristic classes. `http://derekhsorensen.com/docs/sorensen-characteristic-classes.pdf`, 2017.

[121] Arun Debray. Characteristic classes. `https://web.ma.utexas.edu/users/a.debray/lecture_notes/u17_characteristic_classes.pdf`.

[122] Jonathan R. Shewchuk. An introduction to the conjugate gradient method without the agonizing pain. Technical report, 1994.

[123] Chris Tiee. Contravariance, covariance, densities, and all that: An informal discussion on tensor calculus. `https://ccom.ucsd.edu/~ctiee/notes/tensors.pdf`, 2006.

[124] Will J. Merry. Algebraic topology. `https://www.merry.io/algebraic-topology`.

[125] Stacks project. `https://stacks.math.columbia.edu/`.

[126] The nlab. `https://ncatlab.org/nlab`.

[127] Wikipedia. `https://www.wikipedia.org/`.

[128] Joost Nuiten. Cohomological quantization of local prequantum boundary field theory. Master's thesis, 2013.

Index