Mathematical Structures in Physics

Kolleg Studienstiftung
Christoph Schweigert
Hamburg University
Department of Mathematics
Section Algebra and Number Theory
and
Center for Mathematical Physics
(as of 106.01.2023)

Contents

The current version of these notes can be found under
`http://www.math.uni-hamburg.de/home/schweigert/skripten/pskript.pdf`
as a pdf file.

Please send comments and corrections to `christoph.schweigert@uni-hamburg.de`!

These notes are based on lectures delivered at the University of Hamburg in the fall terms 2007/2008, 2009/2010 and 2011/2012 as part of the master programs in physics and mathematical physics. A first version of these notes was taken in German by Linda Sass.

I am grateful to the students in these years for their remarks and comments, in particular to Sion Chan-lang, Marc Lange and Thomas Reichenbächer.

References:

References per Chapter:
1. Classical Mechanics and Lagrangian systems:
Textbooks:

- Vladimir I. Arnold: *Mathematical methods of classical mechanics*, Springer Graduate Text in Mathematics 60, Springer, New York, 1978.
- Herbert Goldstein; Charles Poole and John Safko: *Classical mechanics*, Pearson/Addison Wesley, Upper Saddle River, 2002

Articles and monographs:

- Anderson, Ian M.: Introduction to the variational bicomplex, in M. Gotay, J. Marsden and V. Moncrief (eds), Mathematical Aspects of Classical Field Theory, Contemporary Mathematics 132, Amer. Math. Soc., Providence, 1992, pp. 5173. `http://www.math.usu.edu/~fg_mp/Publications/ComtempMath/IntroVariationalBicomplex.ps`

- You are also encouraged to go back to the classics:
  Emmy Noether: *Invariante Variationsprobleme* Gött. Nachr. 1918 235-257. English translation at `http://arxiv.org/abs/physics/0503066v1`

- Peter Olver: *Applications of Lie Groups to Differential Equations*. Springer Graduate Texts in Mathematics 107, 1986

- Enrique G. Reyes: *On covariant phase space and the variational bicomplex*. Internat. J. Theoret. Phys. 43 (2004), no. 5, 1267–1286.
  `http://www.springerlink.com/content/q151731w71l71q23/`

- Ron Fulp and Jim Stasheff: *Noether's variational theorem II and the BV formalism*
  `http://arxiv.org/abs/math/0204079`

3. Electrodynamics and relativity:
Textbooks:

- Stephen Hawking and George Ellis: *The large scale structure of space-time*. Cambridge monographs on mathematical physics, 21st printing, 2008.

- John David Jackson: *Classical electrodynamics*. 3rd edition, Wiley, 1999.

- Bertel Laurent: *Introduction to Spacetime*. A First Course on Special Relativity. World Scientific Publishing Company, 1994

- Hartmann Römer, M. Forger: *Elementare Feldtheorie*. VCH, Weinheim 1993. `http://www.freidok.uni-freiburg.de/volltexte/405/pdf/efeld1.pdf`

  Articles and monographs:

- Some of our treatment of electrodynamics follows the review:
  Friedrich W. Hehl: *Maxwell's equations in Minkowski's world: their premetric generalization and the electromagnetic energy-momentum tensor*
  `http://arxiv.org/abs/0807.4249v1`

4. Hamiltonian mechanics and symplectic geometry

- Rolf Berndt: *An introduction to symplectic geometry*, American Mathematical Society, Providence, 2001

- Jean-Marie Souriau: *Structure of dynamical systems*. Birkhäuser, Boston, 1997.
  `http://www.jmsouriau.com/structure_des_systemes_dynamiques.htm`

5. Quantum mechanics:
Textbooks:

- Ludwig D. Faddeev, O.A. Yakubovskii: *Lectures on quantum mechanics for mathematics students*. AMS Student Mathematical Library, Volume 47, 2009

- Albert Messiah: *Quantum mechanics*. Dover Publications, 1999.
- Strocchi, F.: *An introduction to the mathematical structure of quantum mechanics*. Advanced Series in mathematical physics 27, World Scientific, Singapur, 2005.
- Jun John Sakurai: *Modern Quantum Mechanics*. Benjamin Cummings, 1985.
- Leon A. Takhtajan: *Quantum Mechanics for Mathematicians*. Graduate Studies in Mathematics, American Mathematical Society, 2008.

Articles and monographs:

- For C^*-algebras: Christian Bär and Christian Becker: *C^*-algebras*, in: Quantum Field Theory on Curved Spacetimes, Lecture Notes in Physics, 786, Springer 2009. Also available at
`http://geometrie.math.uni-potsdam.de/documents/baer/QFT-ProceedingsCh01.pdf`
- For spectral theory: William Arveson: *A Short Course on Spectral Theory*. Springer Graduate Texts in Mathematics, 2002, Volume 209
- For quantization: Domenico Giulini: *That Strange Procedure Called Quantisation*. Lecture Notes in Physics 631, Springer, Berlin, 2003.
`http://de.arxiv.org/abs/quant-ph/0304202`

Appendix A: Manifolds

- Glen E. Bredon: *Topology and geometry*. Springer Graduate Texts in Mathematics 139, 1993.
- Detlef Gromoll, Wilhelm Klingenberg and Wolfgang Meyer: *Riemannsche Geometrie im Großen*, Lecture Notes in Math. 55, Springer (1975).
- Ib Madsen and Jørgen Tornehave: *From Calculus to cohomology: de Rham cohomology and characteristic classes*. Cambridge University Press 1997.
- Raoul Bott and Loring W. Tu: *Differential forms in algebraic topology*. Springer Graduate Texts in Mathematics 82, New York, 1982.

0 Introduction

Eugene Wigner writes in his famous article "*The Unreasonable Effectiveness of Mathematics in the Natural Sciences*"[1]

> The miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics is a wonderful gift which we neither understand nor deserve.

It is one of the goals of this class to show at least few instances of this miracle. Our goal is twofold: the first goal is to show the appropriate mathematics to a student of physics, roughly familiar with all classes of theoretical physics except for quantum field theory. The second goal is to show to a student of mathematics some mature mathematical theories at work: to see how they describe aspects of our reality in the inanimate world - this is after what physics is about.

Both goals in themselves are so ambitious that the course is bound to fail; hence it only matters that the course fails in such a way that students taking the course have a maximum benefit. I am not sure that the combination of these two goals necessarily increases the likelihood to fail. Indeed, the relation between developments in physics and in mathematics is a rather complicated one and there are more interrelation than one naively expects: Let me again quote Wigner:

> It is true, of course, that physics chooses certain mathematical concepts for the formulation of the laws of nature, and surely only a fraction of all mathematical concepts is used in physics. It is true also that the concepts which were chosen were not selected arbitrarily from a listing of mathematical terms but were developed, in many if not most cases, independently by the physicist and recognized then as having been conceived *before* by the mathematician.

The course is based on rather personal choices. It does not at all pretend to show the only possible approach to mathematical physics, not even the most appropriate one. It emphasizes structural aspects and concepts and thereby prefers a general point of view to the example.

The idea is to cover essentially those concepts of classical physics and quantum physics that are needed to understand quantum field theory. One should emphasize that classical mechanics and quantum mechanics cover an enormous range of aspects of our physical reality and that we do meet quite a few of the core achievements of theoretical physics. If we insist on preparing the reader to quantum field theory, then for the reason that quantum field theory is not only connected to many important recent developments in mathematics, but also to what one might consider the two main challenges of physics in our time: a unified description of all forces and particles in nature and an understanding of the collective behaviour of qunautm mechanical particles.

I would like to highlight in the context the following aspects:

- We try to set up a geometric setting for Lagrangian systems that allows to appreciate *both* theorems of Emmy Noether. Here, Einstein's comments on Noether's work might lead the path:

[1] Eugene Wigner, "*The Unreasonable Effectiveness of Mathematics in the Natural Sciences,*" in: Communications in Pure and Applied Mathematics, vol. 13, No. I (February 1960).

> "Gestern erhielt ich von Fr. Nöther eine sehr interessante Arbeit über Invariantenbildung. Es imponiert mir, dass man diese Dinge von so allgemeinem Standpunkt übersehen kann. Es hätte den Göttinger Feldgrauen nicht geschadet, wenn sie zu Frl. Nöther in die Schule geschickt worden wären. Sie scheint ihr Handwerk gut zu verstehen!" [2]

- We consistently use differential forms and a geometric approach. In this spirit, we also see electrodynamics with gauge potentials as an instance of differential cohomology.
- We emphasize the role of observables. For this reason, we treat the Hilbert space as a derived concept in quantum mechanics.

These notes are the (preliminary result) of 25 years of attempts to understand some aspects of physics in a conceptual and mathematically clear way. I am aware of some of their short comings and I am sure these notes have many short comings I am not even aware o. I am just asking the reader for the minimum amount of sympathy without which any understanding is impossible.

1 Newtonian mechanics

1.1 Galilei space, equations of motion

In classical physics, the idea that there exists empty space should be accepted as central. A basic postulate requires empty space to be spacially homogeneous. Also, not direction should be distinguished: space is required to be isotropic. A similar homogeneity requirement is imposed on time.

The mathematical model for these requirements is provided by the notion of an affine space. We formulate it over an arbitrary field k.

Definition 1.1.1
An affine space is a pair $(\mathbb{A}, V)$, consisting of a set $\mathbb{A}$ and a k-vector space V together with an action of the abelian group $(V,+)$ underlying V on the set $\mathbb{A}$ that is transitive and free.

We comment on terms used in the definition:

Remarks 1.1.2.

1. *In more detail, an action of the abelian group $(V,+)$ on the set $\mathbb{A}$ is a map*

$$\rho : V \times \mathbb{A} \to \mathbb{A}$$

such that

$$\rho(v+w,a) = \rho(v,\rho(w,a)) \quad \text{for all } v,w \in V, a \in \mathbb{A}\ .$$

2. *An action is called transitive, if for all $p,q \in \mathbb{A}$ exists $v \in V$ such that $\rho(v,p) = q$; an action is called free, if this $v \in V$ is unique.*

3. *We call $\dim_k V$ the dimension of the affine space $\mathbb{A}$ and write $\dim \mathbb{A} = \dim_k V$. We also say that the affine space $(\mathbb{A}, V)$ is modelled over the vector space V.*

[2] Einstein in a letter to David Hilbert, dated May 24, 1918

4. *We introduce the notation $\rho(v,p) = p+v$. If v is the unique vector in V such that $q = p+v$, we write $q-p = v$. We also say that $\mathbb{A}$ is a $(V,+)$-Torsor or a principal homogenous space for the group $(V,+)$. The group $(V,+)$ is also called the difference space of the affine space.*

5. *While a vector space has the zero vector as a distinguished element, there is no distinguished element in an affine space.*

In the application to classical mechanics, the field k is usually taken to be the field of real numbers, $k = \mathbb{R}$: the difference of three positions in space can be described by three real coordinates. In fact, for all practical purposes, one might restrict to rational coordinates, but mathematically it is convenient to complete the field. The transitive action accounts for spacial homogeneity. The fact that an affine space does not have a distinguished element is a non-trivial feature of Newtonian mechanics: historically, there are many views of the world with a distinguished point in space, including the Garden of Eden, Rome, Jerusalem, the sun of our solar system ....

We also need morphisms of affine spaces:

Definition 1.1.3
Let $(\mathbb{A}_1, V_1)$ and $(\mathbb{A}_2, V_2)$ be affine spaces modelled over vector spaces V_1, V_2 over the same field k. A morphism $(\mathbb{A}_1, V_1) \to (\mathbb{A}_2, V_2)$ or affine map is a map

$$\varphi : \mathbb{A}_1 \to \mathbb{A}_2$$

for which there exists a k-linear map $A_\varphi : V_1 \to V_2$ such that

$$\varphi(p) - \varphi(q) = A_\varphi(p-q) \quad \text{for all } p, q \in \mathbb{A}_1 \ .$$

Remarks 1.1.4.

1. *Note that the $(V,+)$-equivariant morphisms are those morphisms for which $A_\varphi = \mathrm{id}_V$.*

2. *Any two affine spaces of the same dimension over the same vector space are isomorphic, but not canonically isomorphic.*

3. *Recall the definition of a semi-direct product of two groups H, N. Given a group homomorphism $\varphi : H \to \mathrm{Aut}(N)$ the set $N \times H$ can be endowed with the structure of a group $N \rtimes_\varphi H$ by*

$$(n,h) \cdot (n',h') := (n\varphi_h(n), hh') \ .$$

The automorphism group of an affine space is isomorphic to the semi-direct product

$$\mathrm{Aut}(\mathbb{A}) = V \rtimes \mathrm{GL}(V) \ ,$$

where V acts on $\mathbb{A}$ by translations.

4. *The choice of any point $p \in \mathbb{A}$ induces a bijection $\rho(\cdot, p) : V \to \mathbb{A}$ of sets. As a finite-dimensional $\mathbb{R}$-vector space, V has a unique topology as a normed vector space. By considering preimages of open sets in V as open in $\mathbb{A}$, we get a topology on $\mathbb{A}$ that does not depend on the choice of base point. We endow $\mathbb{A}$ with this topology.*

5. *We will see in the appendix that affine space $\mathbb{A}^n$ is an n-dimensional manifold. The choice of a point $p \in \mathbb{A}$ provides a natural global coordinate chart.*

Suppose that we wish to model space. We are used in daily life life to specify points in space by three coordinates, length, breadth and height. Since now point in space is distinguished, one starts to model space by a three-dimensional affine space.

To be able to talk about lengths, we have to model measurements with rods (or rulers). We resort to an idealization and introduce infinitely long rods. Moreover, we assume that on such a rod, a point is marked and one direction is distinguished. One can see that this structure is already enough to introduce a scalar product on the difference vector space of space.

To state a theorem, we formalize the situation:

Definition 1.1.5

1. *A ray L in a real vector space V is a subset of the form $L = \mathbb{R}_{\geq 0}v$ with $v \in V \setminus \{0\}$.*

2. *A halfplane in a real vector space V is a subset $H \subset V$ such that there are two linearly independent vectors $v, w \in V$ with $H = \mathbb{R}v + \mathbb{R}_{\geq 0}w$. The boundary of a halfplane is the only line through the origin contained in it. (For the halfplane just given, this is the line $\mathbb{R}v$.)*

3. *A rotation group for a real three-dimensional vector space V is a subgroup $D \subset \mathrm{GL}(V)$ which acts transitively and freely on the set of pairs consisting of a halfplane and a ray on its boundary.*

For any positive definite non-degenerate scalar product $b : V \times V \to \mathbb{R}$ on V one can consider the corresponding orthogonal group $\mathrm{O}(V, b)$ of linear endomorphisms preserving the scalar product and its subgroup $\mathrm{SO}(V, b)$ of endomorphisms of determinant one which we call "rotations". It is a good exercise in linear algebra to show that for any scalar product b the group $\mathrm{SO}(V, b)$ is a rotation group in the sense of definition 1.1.5.3.

Proposition 1.1.6.
For any three-dimensional vector space, the map

$$\{\textit{Scalar products on } V\}/\mathbb{R}_{>0} \to \textit{Rotation groups}$$

which maps the scalar product b to its special orthogonal group $\mathrm{SO}(V, b)$ is a bijection.

We have to construct an inverse of the map. If one assumes that the rotation group is compact, an invariant scalar product can be obtained by integration. For an elementary (but lengthy) proof without this assumption, we refer to W. Soergel, Herleitung von Skalarprodukten aus Symmetrieprinzipien. Mathematische Semesterberichte 68 (2008) 197-202.

Next we wish to talk about length scales. Such a length scale is e.g. the prototype meter. It consists of a ruler with two marked points. Fixing one point, the other point moves on the orbit of any chosen rotation group acting on $\mathbb{R}^3$.

Definition 1.1.7

1. *Given a rotation group D, we call an orbit $l \subset V \setminus \{0\}$ a unit length.*

2. *Given a unit length, we define a norm on V as follows. We first remark that any ray intersects a given unit length l in precisely one point. If the ray through $w \in V$ intersects l in $v \in l$ and $w = \lambda v$ with $\lambda \geq 0$, we define the norm on W by $|w| := \lambda$.*

We thus find that the three-dimensional space of our intuition should have the following mathematical structure:

Definition 1.1.8

1. *An n-dimensional Euclidean space $\mathbb{E}^n$ is an n-dimensional affine space $\mathbb{A}^n$ together with the structure of a Euclidean vector space on the difference vector space.*
2. *As morphisms of Euclidean spaces, we only admit those affine maps φ for which the linear map A_φ is an isometry, i.e. an orthogonal map.*
3. *The group of automorphisms of a Euclidean space is called a Euclidean group.*

Proposition 1.1.9.
The Euclidean group of a Euclidean vector space $\mathbb{E}^n$ is a semi-direct product of the subgroup V of translations given by the action of V and the rotation group $\mathrm{O}(V, b)$:

$$\mathrm{Aut}(\mathbb{E}) = V \rtimes \mathrm{O}(V, b)$$

It is a non-compact Lie group of dimension $n + \frac{n(n-1)}{2}$.

The Euclidean group is thus a proper subgroup of the affine group.
Including also time, we obtain the model of empty space in classical physics:

Definition 1.1.10

1. *A Galilei space $(\mathbb{A}, V, t, \langle \cdot, \cdot \rangle)$ consists of*
   - *An affine space $\mathbb{A}$ over a real four-dimensional vector space V. The elements of $\mathbb{A}$ are called events or space time points.*
   - *A non-zero linear functional*
     $$t : V \to \mathbb{R}$$
     called absolute time difference function. We could have fixed one one-dimensional affine space $\mathbb{I}$ asked for a morphism $\tilde{t} : \mathbb{A} \to \mathbb{I}$ of affine spaces as well. This would have determined as one additional (uninteresting) piece of information an element $\tilde{t}(p) \in \mathbb{I}$ for one point $p \in \mathbb{A}$.
   - *The structure of a Euclidean vector space with positive definite scalar product $\langle \cdot, \cdot \rangle$ on $\ker t$.*
2. *$t(a - b) \in \mathbb{R}$ is called the time difference between the events $a, b \in \mathbb{A}$.*
3. *Two events $a, b \in \mathbb{A}$ with $t(a - b) = 0$ are called simultaneous. This gives an equivalence relation on Galilei space. The equivalence class*
   $$\mathrm{Cont}(a) = \{b \in \mathbb{A} \mid t(b - a) = 0\}$$
   is the subset of events simultaneous to $a \in \mathbb{A}$.

Remarks 1.1.11.

1. *Notice that on Galilei space, absolute time differences are defined. This is sometimes summarized in the statement that "classically, there is absolute time". In relativistic theories, this does not hold any longer.*

2. $\mathrm{Cont}(a)$ *for any* $a \in \mathbb{A}$ *is a three-dimensional Euclidean space.*

3. *While it makes sense to talk about simultaneous events* a, b *at different places, the phrase: "The two events* a, b *are at different time, but at the same place in three-dimensional space." does not make sense. Consider, for example, two observers. An observer is defined by its worldline which is required to be an affine line in Galilei space, parameterized by an affine parameter, that is not contained in any hypersurface of simultaneous events. (T) sitting in a train moving at uniform velocity on straight rails and (S) at rest on the earth (which we consider at itself rest). Then two events happening at 1pm and 3pm in the dining car of the train happen at the same place for (T), but definitely not for (S).*

Definition 1.1.12
As the morphisms of two Galilei spaces $(\mathbb{A}_1, V_1, t_1, \langle\cdot,\cdot\rangle_1)$ *and* $(\mathbb{A}_2, V_2, t_2, \langle\cdot,\cdot\rangle_2)$ *we consider those affine maps*

$$\varphi : \mathbb{A}_1 \to \mathbb{A}_2$$

which respect time differences

$$t_1(b-a) \;=\; t_2(\varphi(b) - \varphi(a)) \quad \text{for all} \quad a, b \in \mathbb{A}_1,$$

and the Euclidean structure on space in the sense that the restriction $A_\varphi : \ker t_1 \to \ker t_2$ *is an orthogonal linear map.*

Remarks 1.1.13.

1. *The group of automorphisms of a Galilei space is a proper subgroup of the affine group of the underlying affine space.*

2. *An example for a Galilei space is given by the Galileian coordinate space* $\mathbb{G} = (\mathbb{R}^1 \times \mathbb{R}^3, \mathbb{R}^4, pr_1, \langle\cdot,\cdot\rangle)$. *This is the set* $\mathbb{R} \times \mathbb{R}^3$ *seen as an affine space over the vector space* $\mathbb{R}^4 \cong \mathbb{R} \times \mathbb{R}^3$ *with the projection on the first component*

$$\begin{array}{rccc} t \;=\; \mathrm{pr}_1 : \mathbb{R}^4 \cong \mathbb{R}^1 \times \mathbb{R}^3 & \to & \mathbb{R} \\ (x^0, x^1, x^2, x^3) & \mapsto & x^0 \end{array}$$

   as the time difference functional and the standard Euclidean scalar product $\langle\cdot,\cdot\rangle$ *on* $\mathbb{R}^3$.

3. *Two Galilei spaces are isomorphic, but not canonically isomorphic. The automorphism group* $\mathrm{Aut}(\mathbb{G})$ *of the Galileian coordinate space* $\mathbb{G}$ *is called the* Galilei group.

4. *We consider three classes of automorphisms in* $\mathrm{Aut}(\mathbb{G})$*:*

   - *Uniform motions with velocity* $v \in \mathbb{R}^3$*:*

$$g_1(t, x) \;=\; (t, x + vt)$$

- *Spacial translations by $x_0 \in \mathbb{R}^3$ combined with time translations by $t_0 \in \mathbb{R}$:*
$$g_2(t, x) \;=\; (t + t_0, x + x_0)$$
- *Rotations and reflections in space with $R \in \mathrm{O}(3)$:*
$$g_3(t, x) \;=\; (t, Rx)$$

One can show that $\mathrm{Aut}(\mathbb{G})$ is generated by these elements. It is a ten-dimensional non-compact Lie group.

Definition 1.1.14

1. *Let X be any set and*
$$\phi : X \to \mathbb{G}$$
be a bijection of sets. (On the right hand side, it would be formally more correct to write the set underlying $\mathbb{G}$.) We say that ϕ provides a global Galilean coordinate system on X.
2. *We say that two Galilean coordinate systems $\phi_1, \phi_2 : X \to \mathbb{G}$ are in relative uniform motion, if*
$$\phi_1 \circ \phi_2^{-1} \in \mathrm{Aut}(\mathbb{G}) \ .$$

Any global Galilean coordinate system $\phi : X \to \mathbb{G}$ endows the set X with the structure of a Galilean space over the vector space $\mathbb{R}^4$:
we define on X the structure of an affine space over $\mathbb{R}^4$ by requiring ϕ to be an affine map:
$$x + v := \phi^{-1}(\phi(x) + v) \quad \text{for all} \quad x \in X, v \in \mathbb{R}^4 \ .$$
Then we use the standard time difference functional $pr_1 : \mathbb{R}^4 \cong \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}$ and the standard Euclidean structure on $\mathbb{R}^3$ to get a Galilei space $(X, \mathbb{R}^4)_\phi$ that depends on ϕ. For this Galilei space, we have an isomorphism
$$(X, \mathbb{R}^4)_\phi \xrightarrow{\phi} \mathbb{G}$$
of Galilei spaces such that the linear map A_ϕ associated to ϕ is the identity.

Lemma 1.1.15.

1. *Being in relative uniform motion is an equivalence relation on the set of global Galilean coordinate systems on X. (There is, of course, more than one equivalence class.)*
2. *Two Galilean coordinate systems ϕ_1, ϕ_2 in relative uniform motion endow the set X with the structures of Galilei spaces for which the identity is an isomorphism of Galilei spaces.*

This leads us to the following

Definition 1.1.16

1. *A Galilean structure on a set X is an equivalence class of Galilean coordinate systems.*
2. *Given a Galilean structure on a set X, any coordinate system of the defining equivalence class is called an inertial system or inertial frame for this Galilean structure.*

Remarks 1.1.17.

1. *By definition, two different inertial systems for the same Galilean structure are in uniform relative motion.*
2. *There are no distinguished inertial systems.*

1.2 Dynamics of Newtonian systems

To describe how actual particles move in the space-time we have just describe, we first state the fundamental principle:

Principle 1.2.1. *("Galilean principle of relativity")*
All laws of nature are of the same form in all inertial systems.

To make this principle more precise for mechanical system, we have to describe the motions of mass points and to complement the principle of relativity with another important principle.

Definition 1.2.2

1. *A trajectory of a mass point in $\mathbb{R}^3$ is an (at least twice) differentiable map*

$$\varphi : I \to \mathbb{R}^3$$

with $I \subset \mathbb{R}$ an interval. To simplify our exposition, we will from now on restrict to smooth trajectories, i.e. trajectories that are infinite-many times differentiable.

2. *A trajectory of N mass points in $\mathbb{R}^3$ is an N-tuple of (at least twice) differentiable maps*

$$\varphi^{(i)} : I \to \mathbb{R}^3$$

with $I \subset \mathbb{R}$ an interval. Equivalently, we can consider an (at least twice) differentiable map

$$\vec{\varphi} : I \to (\mathbb{R}^3)^N \ .$$

3. *The velocity in $t_0 \in I$ is defined as the derivative:*

$$\dot{\varphi}(t_0) = \frac{\mathrm{d}\varphi}{\mathrm{d}t}\Big|_{t_0} \ .$$

4. *The acceleration in $t_0 \in I$ is defined as the second derivative:*

$$\ddot{\varphi}(t_0) = \frac{\mathrm{d}^2\varphi}{\mathrm{d}t^2}\Big|_{t_0} \ .$$

5. *The graph*

$$\{(t, x(t)) \,|\, t \in I\} \subset \mathbb{R} \times \mathbb{R}^3$$

of a trajectory is called the world line of the mass point. We consider a world line as a subset of Galilean coordinate space $\mathbb{G}$.

The following principle is the basic axiom of Newtonian mechanics. It cannot be derived mathematically but should rather seen as a deep abstraction from many observations in nature.
We first discuss the situation in a fixed coordinate system:

Definition 1.2.3 *[Newtonian determinism]*

1. *A Newtonian trajectory of a point particle*

$$\varphi : I \to \mathbb{R}^3, \qquad with \quad I = (t_0, t_1) \subset \mathbb{R}$$

is completely determined by the initial position $\sigma(\tau_0)$ and the inital velocity

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\varphi\,\Big|_{t=t_0}.$$

2. *In particular, the acceleration at t_0 is determined by the initial position and the initial velocity. As a consequence, there exists a function, called the force field,*

$$F : \mathbb{R}^{3N} \times \mathbb{R}^{3N} \times \mathbb{R} \to \mathbb{R}^{3N},$$

where the first factor are the three spacial coordinates of N particles, the second are their velocities and the third is time, such that for all Newtonian trajectories the Newtonian equation

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\,x^i \;=\; \ddot{x}^i \;=\; F^i(x, \dot{x}, t), \quad i = 1, ..., N.$$

holds. This is a second order ordinary differential equation that completely determines the evolution of the system in time.

Remarks 1.2.4.

1. *The function F has to be measured experimentally and describes the forces that act. It determines the physical system. For the particular case $F = 0$, no forces act and $\ddot{x} = 0$. As a consequence, the trajectory is a uniform motion. (Aristotle had a different idea about the situation: he believed that any body would ultimately come to rest.)*

2. *I find the fact that Newton was able to realize the importance of* second *derivatives so quickly after the discovery of differential calculus a truly amazing strike of genius.*

3. *Standard mathematical propositions assure that a sufficiently smooth function F determines, together with the initial conditions on position and velocity, the trajectory. Implicitly, we will sometimes assume the existence of global solutions or of solutions for sufficiently large times.*

4. *We will see that not all graphs $\{(t, \varphi(t))\}_{t\in I} \subset \mathbb{G}$ describe physical motions.*

So far we have described trajectories using maps $\sigma : I \to \mathbb{R}^3$ giving rise to a graph $\{(t, \varphi(t))\} \subset \mathbb{G}$. Suppose we are given a Galilei space $\mathbb{A}$ with inertial frames $\psi : \mathbb{A} \to \mathbb{G}$. We then get a map

Definition 1.2.5
Let $\mathbb{A}$ be a Galilei space, I an interval of eigentime and $\varphi : I \to \mathbb{A}$ a smooth function. Then φ is called a physical motion if for all inertial frames $\psi : \mathbb{A} \to \mathbb{G}$ the function $\psi \circ \varphi : I \to \mathbb{G}$ is the graph of a Newtonian trajectory.

This is can be seen as a more precise version of the Newtonian principle of relativity.

Remarks 1.2.6.

1. *We immediately have the following consequences of the Newtonian principle of relativity.*

   - *Invariance under time translations: the force F does not depend on time t.*
   - *Invariance under spacial translations: F depends only on the relative coordinates $\varphi^i - \varphi^1$.*
   - *Invariance under relative uniform motion: F only depends on the relative velocities.*
   - *Invariance under rotations: the force F is a vector, i.e. transforms in the same representation of the rotation group of three-dimensional space as φ and $\dot{\varphi}$.For any orthogonal transformation $M : \mathbb{R}^3 \to \mathbb{R}^3$, we have*

   $$F(Mx, M\dot{x}) = MF(x, \dot{x}) \quad .$$

2. *Altogether, this implies that the motion of n points is described by a function*

   $$\ddot{x}^{(k)}(t) = f(x^{(j)} - x^{(l)}, \dot{x}^{(s)} - \dot{x}^{(r)})$$

   with $x^{(i)} : I \to \mathbb{R}^3$.

3. *We deduce Newton's first law: a system consisting of a single point is described in an inertial system by a uniform motion $(t, x_0 + tv_0)$. In particular, the acceleration vanishes, $\ddot{x} = 0$.*

1.3 Examples

Our discussion of Newtonian relativity applied to a system in empty space. In general, it is interesting to consider systems where F is a function violating these principles. For example, interesting functions are not invariant under translation.

Using such a function, we can conveniently summarize the effect of other mass points in the system. For example, one might switch the perspective from considering the two-body system consisting of the sun and the earth by considering just the motion of earth in the background of the force exerted by the mass of the sun.

We investigate some examples in a fixed inertial system with the help of a function $x : \mathbb{R} \to \mathbb{R}^3$. We call $\mathbb{R}^3$ the configuration space of the system.

Examples 1.3.1.

1. *Freely falling particle:*

   $$\ddot{\varphi} = -g\hat{\varphi}_3 \quad \text{with } g \approx 9,81\ ms^{-2}.$$

   One should notice that a direction is distinguished by the unit vector $\hat{x}_3$ in the direction of the x_3-axis. Hence the isotropy of space is violated. We introduce a function, called potential energy

   $$U(x_1, x_2, x_3) : \mathbb{R}^3 \to \mathbb{R} \quad \text{such that } -\operatorname{grad} U = F\ ;$$

   in our case $U(x_1, x_2, x_3) = gx_3$. The function U is not unique, but can be added by any function whose derivative vanishes. The equations of motion then read $\ddot{x} = -\operatorname{grad} U$.

2. *The harmonic oscillator is defined by the potential*

$$V(x,y,z) = \frac{1}{2}Dx \ .$$

The force experience by a particle with x-coordinate x is then in x-direction and equals $-Dx$, i.e. it is proportional to the elongation. The equations of motion in one-dimension read

$$m\ddot{x} = \ -Dx$$

They have the general solution $\varphi(t) = \ A\cos(\omega t - \varphi_0)$ with $\omega = \sqrt{\frac{D}{m}}$ and constants A, φ_0 that have to be determined from the initial conditions.

3. *We just assumed that a force $F = F(x)$ depending only on coordinates is described as the gradient of a potential function U. One can investigate what force fields can be described in this way.*

A force field is called conservative, if for any trajectory $\varphi : \ (t_a, t_b) \to \mathbb{R}^N$ the so-called work integral

$$\int_a^b F \mathrm{d}\varphi = \int_{t_a}^{t_b} F(\varphi(t))\dot{\varphi}(t) \ \mathrm{d}t$$

only depends on the end points $\varphi(t_a)$ and $\varphi(t_b)$ and not on the particular choice of trajectory connecting them. Then, there exists a potential that is unique up to unique isomorphism.

4. *To write down Newton's law of gravity, we consider a potential energy depending only on the distance r from the center of gravity:*

$$U(x_1,x_2,x_3) = -\frac{k}{r} \quad \textit{with} \ \ r := \sqrt{x_1^2 + x_2^2 + x_3^2}$$

$$\ddot{\varphi} = -\operatorname{grad} U = -\frac{k}{r^3}\varphi$$

This potential is central for the description not only of macroscopic systems like a planet turning around the sun; in this case, the potential described the gravitational force exerted by the sun. It also enters in the description of microscopic systems like the hydrogen atom where the potential describes electrostatic force exerted on the electron by the proton that constitutes the nucleus of the atom.

Newton's discovery that the same force is responsible for objects falling down on earth and the trajectory of planets, i.e. the discovery of the universality of gravity is a truly amazing achievement.

5. *One-dimensional harmonic oscillator. To investigate the equation of motion $\ddot{x} = -\alpha^2 x$, we introduce the potential $U = \frac{\alpha^2 x^2}{2}$. This can be realized e.g. using a massive pointlike body connected to a spring.*

Experimentally, one finds that for equal springs with equal initial conditions, but different bodies as mass points, the ratio

$$\frac{\ddot{x}_1}{\ddot{x}_2} = const_{1,2}$$

only depends on the balls, but not on the initial conditions. We put

$$\frac{\ddot{x}_1}{\ddot{x}_2} = \frac{m_2}{m_1}$$

with a quantity m_i called inertial mass which is a property of the i-th body. Again, we introduce a physical unit by comparing to a standard, e.g. the prototype kilogram in Paris.

Observation 1.3.2.
Consider a mechanical system given by a force $F(x, \dot{x}, t)$ which we suppose right away to be given by a potential of the form

$$V : \mathbb{R}^3 \to \mathbb{R}$$

and thus independent of t and $\dot{x}$. We have to study the coupled system of ordinary differential equations

$$\ddot{\varphi} = -\mathrm{grad}\, V(\varphi) \qquad (*)$$

of second order. For any trajectory $\varphi : I \to \mathbb{R}^3$ that is a solution of $(*)$ we consider the real-valued function

$$\begin{aligned} \epsilon :\ & I \to \mathbb{R} \\ & \epsilon(t) = \frac{1}{2}\|\dot{\varphi}(t)\|^2 + V(\varphi(t)) \ . \end{aligned}$$

For its derivative, we find

$$\frac{\mathrm{d}}{\mathrm{d}t}\, \epsilon(t) = \langle \dot{\varphi}, \ddot{\varphi} \rangle \ + \ \langle \mathrm{grad}\, V \ , \dot{\varphi} \rangle = 0 \ ,$$

where in the last step we use the equation of motion $(*)$.

In the theory of ordinary differential equations, it is standard to reduce differential equations of higher order to first order equations by introducing additional variables.

Lemma 1.3.3.
The system $(*)$ is equivalent to the following system of ordinary differential equation of first order for the function $y : I \to \mathbb{R}^6$

$$\begin{aligned} \dot{y}_1 &= y_2, \quad \dot{y}_2 = -\mathrm{grad}_{y_1} V(y_1, y_3, y_5) \\ \dot{y}_3 &= y_4, \quad \dot{y}_4 = -\mathrm{grad}_{y_3} V(y_1, y_3, y_5) \\ \dot{y}_5 &= y_6, \quad \dot{y}_6 = -\mathrm{grad}_{y_5} V(y_1, y_3, y_5) \end{aligned}$$

The solutions of this system are called phase curves.

We thus enlarged the space by considering the first derivatives as independent geometric quantities.

Observation 1.3.4.
This suggests to introduce $\mathbb{R}^6$ with coordinates x, y, z, u_x, u_y, u_z. This is space is sometimes called the phase space $\mathbb{P}$ of the system. (This use of the word "phase space" is somewhat confusing in view of the use of the same word in the context of Hamiltonian mechanics.)

We equip the phase space $\mathbb{P}$ *with the energy function*

$$E(x_1, x_2, x_3, u_{x_1}, u_{x_2}, u_{x_3}) := \frac{1}{2}({u_{x_1}}^2, {u_{x_2}}^2, {u_{x_3}}^2) + V(x_1, x_2, x_3) \ .$$

It should be appreciated that the first term in E *is a quadratic form.*

On phase space, we consider the ordinary differential equations

$$\frac{\mathrm{d}x^i}{\mathrm{d}t} = u_{x_i}, \quad \frac{\mathrm{d}u_{x_i}}{\mathrm{d}t} = \ -\mathrm{g}rad_i V(x_1, x_2, x_3) \quad (**) \ .$$

The solutions of $(**)$ *are called* <u>phase curves</u>*. They are contained in subspaces of* $\mathbb{P}$ *of constant value of* E*.*

Observation 1.3.5.

1. *Assume that for any point* $M \in \mathbb{P}$ *a global solution* $x_M(t)$ *of* $(**)$ *with initial conditions* M *exists. This allows us to define a mapping*

$$\mathrm{g}_t : \mathbb{P} \to \mathbb{P} \qquad by \quad \mathrm{g}_t(M) = x_M(t).$$

2. *Standard theorems about ordinary differential equations imply that* g_t *is a diffeomorphism, i.e. a differential map with differentiable inverse mapping.*

3. *We find* $\mathrm{g}_{t_1} \circ \mathrm{g}_{t_2} = \mathrm{g}_{t_1+t_2}$ *and thus a smooth action*

$$\begin{aligned} \mathrm{g} : \ & \mathbb{R} \times \mathbb{P} \to \mathbb{P} \\ & (t, M) \longmapsto \mathrm{g}_t(M) \end{aligned}$$

called the <u>phase flow</u>*.*

2 Lagrangian mechanics

2.1 Variational calculus

We consider the following mathematical problem:

Observation 2.1.1.

1. *Our objects are trajectories, i.e. smooth functions $\varphi : I \to \mathbb{R}^N$, $I = [t_1, t_2]$ on an interval with values in $\mathbb{R}^N$. We say that the trajectory is parametrize by its eigentime in I. Later on, we will consider more general situations, e.g. smooth functions on intervals with values in smooth manifolds as well. Possibly after choosing additional structure on $\mathbb{R}^N$, we can associate to each such function a real number. For example, if we endow $\mathbb{R}^N$ with the standard Euclidean scalar product $\langle \cdot, \cdot \rangle$ with corresponding norm $\| \cdot \| : \mathbb{R}^N \to \mathbb{R}$, we can define the length of a trajectory by*

$$\mathrm{L}_1(\varphi) = \int_I \|\dot{\varphi}\| \mathrm{d}t,$$

or its energy by

$$\mathrm{L}_2(\varphi) = \int_I \|\dot{\varphi}\|^2 \mathrm{d}t \ .$$

Such scalar valued functions on spaces of trajectories are also called functionals.

Finding trajectories that extremize functionals, e.g. finding curves of extremal length, is obviously a natural and important mathematical problem.

Our goal is to show that extremal trajectories are solutions of ordinary differential equations. For our purposes, the logic will be just the opposite: we are interested in equations of motion which are ordinary differential equations. We will learn more about solutions of these differential equations by deriving them from functionals.

2. *Like the length of a trajectory, the functional we are interested in depend on the trajectory φ and its derivatives, i.e. on positions and velocities. Positions take their values in positions space, in our case*

$$\mathbb{R}^N \ \text{ with Cartesian coordinates } \ (x^1, ..., x^N) \ .$$

We also need coordinates for derivatives of a trajectory φ with respect to its eigentime. To this end, we introduce the large space

$$\mathrm{J}^1(\mathbb{R}^N) := \mathbb{R}^N \oplus \mathbb{R}^N$$

with Cartesian coordinates $(x^1, ..., x^N, x^1_t, ..., x^N_t)$. There is a natural injection

$$\begin{aligned} \mathbb{R}^N &\hookrightarrow \mathrm{J}^1(\mathbb{R}^N) \\ (x^1, ..., x^N) &\mapsto (x^1, ..., x^N, 0, ..., 0). \end{aligned}$$

We now iterate this procedure. This will not only be natural from a mathematical point of view, but also give us a natural place for second derivatives with respect to eigentime, i.e. a place for accelerations.

Since the equations of motion are second order differential equations, we continue by adding a recipient for the second derivatives:

$$\mathrm{J}^2(\mathbb{R}^N) = \mathbb{R}^N \oplus \mathbb{R}^N \oplus \mathbb{R}^N \quad \textit{with coordinates} \quad (x^i, x^i_t, x^i_{tt}) \ .$$

We also use the abbreviation $x^i_2 := x^i_{tt}$. If we continue this way, we get a system of vector spaces for derivatives up to order α

$$\mathrm{J}^\alpha(\mathbb{R}^N) \cong (\mathbb{R}^N)^{\alpha+1}$$

with embeddings

$$\mathrm{J}^\alpha \hookrightarrow \mathrm{J}^{\alpha+1}(\mathbb{R}^N).$$

We call $\mathrm{J}^{\alpha+1}(\mathbb{R}^N)$ the jet space or order α.

3. *Let us now explain in which sense the spaces we have just constructed are recipients for the derivatives of a smooth trajectory*

$$\varphi : \ I \to \mathbb{R}^N \ .$$

For any $n \in \mathbb{N}$ and $i = 1, \dots N$, we can consider the i-th component of the n-th derivative of the trajectory φ with respect to eigentime:

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n}\varphi^i \ ,$$

which is a smooth function on I. For any $\alpha \in \mathbb{N}$, we can combine these functions to a single function

$$\mathrm{j}^\alpha\varphi : \ I \to \mathrm{J}^\alpha\mathbb{R}^N$$

with components given by

$$(\mathrm{j}^\alpha\varphi)^i_n = \frac{\mathrm{d}^n}{\mathrm{d}t^n}\varphi^i.$$

We call this function the prolongation of the trajectory.

4. *Not much has really happened up to this point: instead of a function $\varphi : I \to \mathbb{R}^N$, we consider all its derivatives up to order α at the same time. From the point of view of trajectories, we have just introduced some redundancy. What we have gained, is a natural space $\mathrm{J}^\alpha\mathbb{R}^N$ to discuss velocities, accelerations and higher derivatives of trajectories.*

   Let us discuss the functionals of length and energy in this language. We first introduce the following functions on $\mathrm{J}^1\mathbb{R}^N$:

$$\begin{aligned} l_1, l_2 : \ & \mathrm{J}^1(\mathbb{R}^N) \to \mathbb{R} \\ & l_1(x^i, x^i_t) = \Big(\sum_{i=1}^N (x^i_t)^2\Big)^{\frac{1}{2}} \quad \textit{for the length} \\ & l_2(x^i, x^i_t) = \sum_{i=1}^N (x^i_t)^2 \quad \textit{for the energy.} \end{aligned}$$

   They capture how the length or the energy of a trajectory depends on a trajectory and its derivatives.

Given a trajectory $\varphi:\ I \to \mathbb{R}^N$, we extend it to a function

$$\mathrm{j}^1\varphi:\ I\ \to \mathrm{J}^1(\mathbb{R}^N)$$

by including its first derivative. The functional we are interested in is then described by evaluating the function l_i on $\mathrm{j}^1\varphi$ and integrating the real-valued function thus obtained over I:

$$\mathrm{L}_i(\varphi) = \int_I \mathrm{d}t\ \ l_i\,(\mathrm{j}^1\varphi).$$

5. *We now formulate the variational problem we wish to solve:*

   Given a function like length or energy

   $$l : \mathrm{J}^\alpha(\mathbb{R}^N) \to \mathbb{R}$$

   find all smooth trajectories

   $$\varphi : I \to \mathbb{R}^N$$

   which extremize the function

   $$\mathrm{L}(\varphi) = \int_I \mathrm{d}t\ l(\mathrm{j}^\alpha\varphi).$$

   This problem is not yet correctly posed: if we minimize the length on all trajectories, the constant trajectories are obvious and trivial minima. We therefore fix boundary values $a_i, b_i \in \mathbb{R}^N,\ i = 0, 1, ..., \alpha - 1$ for the trajectory and its derivatives up to order $\alpha - 1$ and restrict ourselves to trajectories with $I = [t_1, t_2]$ in the subset

   $$T^{(\alpha)}_{(a_i)(b_i)} = \Big\{\varphi : I \to \mathbb{R}^N : \frac{\mathrm{d}^i\varphi}{\mathrm{d}t^i}(t_1) = a_i,\ \frac{\mathrm{d}^i\varphi}{\mathrm{d}t^i}(t_2) = b_i\Big\}.$$

6. *Our problem is thus to find extremal trajectories in $T^{(\alpha)}_{(a_i)(b_i)}$ which is an infinite-dimensional space. To reduce the problem to a finite-dimensional problem and to derive necessary conditions, we consider differentiable one-parameter families of trajectories*

   $$\varphi : I \times (-\epsilon_0, \epsilon_0) \to \mathbb{R}^N\ .$$

   We call the real parameter the variational parameter; for fixed variational parameter $\epsilon \in (-\epsilon_0, \epsilon_0)$, we have a trajectory

   $$\begin{aligned}\varphi_\epsilon :\ & I \to \mathbb{R}^N \\ & t \mapsto \varphi(t, \epsilon),\end{aligned}$$

   which we require to have the correct boundary values, $\varphi_\epsilon \in T^{(\alpha)}_{(a_i)(b_i)}$. We call such a family of trajectories a variational family for the given boundary values.

7. *For each value of the variational parameter ϵ, we can evaluate our function $l : \mathrm{J}^\alpha(R^N) \to \mathbb{R}$ on the corresponding trajectory. This gives a real valued function*

   $$\begin{aligned}(-\epsilon_0, \epsilon_0) &\to \mathbb{R} \\ \epsilon &\mapsto \textstyle\int_I l \circ \mathrm{j}^\alpha(\varphi_\epsilon)\ .\end{aligned}$$

Suppose the trajectory φ_0 for $\epsilon = 0$ is an extremum. Then the derivative

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0} \int_I l \circ \mathrm{j}^\alpha(\varphi_\epsilon) = 0$$

at $\epsilon = 0$ has to vanish. We compute this derivative using the chain rule:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0} \int_I l \circ \mathrm{j}^\alpha(\varphi_\epsilon) &= \frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0} \int_I \mathrm{d}t\ l(\varphi_\epsilon, \dot\varphi_\epsilon, \ddot\varphi_\epsilon, \ldots) \\ &= \int_I \frac{\partial l}{\partial x^i}\Big|_{x^i_\beta = \frac{\mathrm{d}^\beta \varphi^i_\epsilon}{\mathrm{d}t^\beta}} \frac{\mathrm{d}}{\mathrm{d}\epsilon}\varphi^i_\epsilon + \frac{\partial l}{\partial x^i_t}\frac{\mathrm{d}}{\mathrm{d}\epsilon}\dot\varphi^i_\epsilon + \ldots\end{aligned}$$

We assume for simplicity that the function l depends only on derivatives up to order one, i.e. restricts to a function on $\mathrm{J}^1\mathbb{R}^N$. We notice that

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\dot\varphi^i_\epsilon = \frac{\partial^2 \varphi(t,\epsilon)}{\partial\epsilon\partial t} = \frac{\partial^2\varphi(t,\epsilon)}{\partial t\partial\epsilon}$$

By partial integration with respect to the variable t, we find

$$\begin{aligned}&= \int_I \Big(\frac{\partial l}{\partial x^i}\Big|_{x^i_\beta = \frac{\mathrm{d}^\beta\varphi^i}{\mathrm{d}t^\beta}} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial l}{\partial x^i_t}\Big|_{x^i_\beta = \frac{\mathrm{d}^\beta\varphi^i_\epsilon}{\mathrm{d}t^\beta}}\Big)\frac{\mathrm{d}}{\mathrm{d}\epsilon}\varphi^i_\epsilon + \int_I \mathrm{d}\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial l}{\partial x^i}\frac{\mathrm{d}\varphi^i_\epsilon}{\mathrm{d}\epsilon}\Big) \\ &= \int_I \Big(\underbrace{\frac{\partial l}{\partial x^i} - \sum_{j,\beta}\frac{\partial^2 l}{\partial x^j_\beta \partial x^i_t}x^j_{\beta+1}}_{\textit{important expression}}\Big) \circ \mathrm{j}^1(\varphi)\cdot\frac{\mathrm{d}}{\mathrm{d}\epsilon}\varphi^i_\epsilon + \underbrace{\int_I \mathrm{d}t\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial l}{\partial x^i}\frac{\mathrm{d}\varphi^i_\epsilon}{\mathrm{d}\epsilon}\Big)}_{\textit{boundary term}}.\end{aligned}$$

8. *Notice that the last term is an integral over a total derivative and thus depends only on the values of the function*

$$\frac{\partial l}{\partial x^i}\frac{\mathrm{d}\varphi^i_\epsilon}{\mathrm{d}\epsilon}$$

at the end points of the interval I. Our conditions on the trajectories in $T^{(\alpha)}_{(a_i)(b_i)}$ guarantee that at these points $\frac{\mathrm{d}\varphi^i_\epsilon}{\mathrm{d}\epsilon} = 0$. Thus the boundary term vanishes.

We have thus to deal with the first term. To this end, we introduce more terminology:

Definition 2.1.2
Let

$$l:\ \mathrm{J}^1(\mathbb{R}^N) \to \mathbb{R}$$

be a real-valued function on jet space of order 1. Denote by $\mathrm{E}(l)$ the $\mathbb{R}^N$-valued function on the jet space of order 2:

$$\mathrm{E}(l):\ \mathrm{J}^2(\mathbb{R}^N) \to \mathbb{R}^N$$

with components

$$\mathrm{E}(l)^i = \frac{\partial l}{\partial x^i} - \sum_{j,\beta}\frac{\partial^2 l}{\partial x^j_\beta \partial x^i_t}x^j_{\beta+1}$$

We call the operator that associates to l the function $\mathrm{E}(l)$ the Euler-Lagrange operator.

Remarks 2.1.3.

1. *We can rewrite the Euler-Lagrange operator as follows: the total derivative operator*

$$D := \sum_{i\beta} x^i_{\beta+1} \frac{\partial}{\partial x^i_\beta}$$

maps smooth functions on $\mathrm{J}^\alpha(\mathbb{R}^N)$ *to smooth functions on* $\mathrm{J}^{\alpha+1}(\mathbb{R}^N)$*. We then have*

$$\mathrm{E}(l)^i = \frac{\partial l}{\partial x^i} - D\frac{\partial l}{\partial x^i_t} \ .$$

2. *For any smooth trajectory, we obtain an* $\mathbb{R}^N$*-valued function on* $I \subset \mathbb{R}$ *with* i*-th component*

$$\varphi \mapsto \mathrm{E}(l)^i \circ \mathrm{j}^\alpha(\varphi) =: \mathrm{E}(l)\,[\varphi]$$

In terms of this function, the derivative of a variational family with respect to the variational parameter ϵ *becomes*

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathrm{L}(\varphi_\epsilon) = \int_I \big\langle \mathrm{E}(l)\,[\varphi_{\epsilon=0}]\,,\, \frac{\mathrm{d}}{\mathrm{d}\epsilon}\,\varphi_\epsilon \big\rangle.$$

We are now ready to apply the next lemma which uses standard facts from real analysis:

Lemma 2.1.4.
Suppose that the continuous real-valued function

$$f:\ I = [t_1, t_2] \to \mathbb{R}$$

has the property that for any smooth function

$$h:\ I \to \mathbb{R}$$

vanishing at the end points of the interval, $h(t_1) = h(t_2) = 0$*, the integral over the product vanishes,*

$$\int_{t_1}^{t_2} f \cdot h \ \mathrm{d}t = 0 \ .$$

Then f *vanishes identically, i.e.* $f(t) = 0$ *for all* $t \in [t_1, t_2]$*.*

Proof:
Suppose that f does not vanish identically. After possibly replacing f by $-f$, we find $t^* \in (t_0, t_1)$ such that $f(t^*) > 0$. Since f is continuous, we find a neighbourhood U of t^* such that $f(t) \geqslant c > 0$ for all $t \in U$.

Using standard arguments from real analysis, we find a smooth function h with support in U such that $h|_{\tilde{U}} = 1$ for some neighborhood $\tilde{U}$ of t^* contained in U. We thus find the inequalities

$$\int_{t_0}^{t_1} f \cdot h = \int_U f \cdot h \geqslant \int_{\tilde{U}} f \geqslant |\tilde{U}| c > 0$$

contradicting our assumption. □

From this we deduce the following

Proposition 2.1.5.
Let L be a functional given on trajectories given by the smooth function

$$l : \mathrm{J}^\alpha(\mathbb{R}^N) \to \mathbb{R}.$$

Then the trajectory $\varphi : I \to \mathbb{R}^N$ is a stationary point for L, if and only if the N ordinary differential equations

$$\mathrm{E}(l)\,[\varphi] = 0$$

hold. This set of ordinary differential equations is called Euler-Lagrange equations for the function l on the trajectory φ.

Examples 2.1.6.

1. *We introduce so called natural systems of classical mechanics. Endow $\mathbb{R}^N$ with the standard Euclidean structure. Choose a smooth function $V : \mathbb{R}^N \to \mathbb{R}$. As we will see, in practice it is quite important to allow V to have singularities. We will not discuss the type of singularities involved, but rather take the perspective that in this case, the system is defined on the manifold obtained from $\mathbb{R}^N$ by removing the points at which V becomes singular.*

   Then the system is defined by the following function on the first order jet space:

$$l(x, x_t) = \frac{1}{2}\sum_{i=1}^{N}(x_t^i)^2 - V(x^i)$$

   The first summand is frequently called the kinetic term, the second summand the potential term. We compute relevant expressions

$$\frac{\partial l}{\partial x^i} = -\mathrm{grad}_i V \qquad \frac{\partial l}{\partial x_t^i} = x_t^i$$

   and find for the Euler-Lagrange operator

$$\mathrm{E}(l)^i = -\mathrm{grad}_i V - Dx_t^i \ .$$

   This gives us a system of N equations

$$x_{tt}^i = -\mathrm{grad}_i V$$

   which gives the following Euler-Lagrange equations on a trajectory $\varphi : I \to \mathbb{R}^N$:

$$Dx_t^i(\varphi) = \ddot{\varphi}^{\,i} = -\mathrm{grad}_i V\big(\varphi(\tau)\big) \ .$$

   These are Newton's equations of motion in the presence of a force given by the gradient of the potential V. If there is a potential for the forces of a system, all information about the equations of motion is thus contained in the real-valued function

$$l : \mathrm{J}(\mathbb{R}^N) \to \mathbb{R}$$

   which is also called the the Lagrangian function of the system.

Natural systems of particular importance include the langrangian for a one-dimensional harmonic oscillator

$$\begin{array}{rcl} l: \ \mathrm{J}^1\mathbb{R} & \to & \mathbb{R} \\ (x, x_t) & \mapsto & \frac{m}{2}(x_t)^2 - \frac{D}{2}x^2 \end{array}$$

and for the Kepler system

$$\begin{array}{rcl} l: \ \mathrm{J}^1\mathbb{R}^3 & \to & \mathbb{R} \\ (x, x_t) & \mapsto & \frac{m}{2}|x_t|^2 - \frac{k}{|x|} \end{array}$$

2. *Let $\varphi: \ [t_0, t_1] \to \mathbb{R}$ be a real valued function. The length of the curve*

$$\gamma = \left\{(t, x): \ x = \varphi(t) \text{ mit } t_0 \leqslant t \leqslant t_1\right\} \subset \mathbb{R}^2$$

is given by

$$\mathrm{L}(\gamma) = \int_{t_0}^{t_1} \sqrt{1 + \dot{\varphi}^2}\mathrm{d}t \ .$$

We thus consider the function

$$l(x, x_t) = \sqrt{1 + x_t^2} \ .$$

To find the Euler-Lagrange equations, we compute

$$\frac{\partial l}{\partial x} = 0 \quad \text{and} \quad \frac{\partial l}{\partial x_t} = \frac{x_t}{\sqrt{1 + x_t^2}}$$

and find

$$\frac{\partial}{\partial x_t}\frac{\partial l}{\partial x_t}x_{tt} = \frac{x_{tt}}{(1 + x_t^2)^{\frac{3}{2}}} = 0$$

which reduces to $x_{tt} = 0$. Hence we get the differential equation $\ddot{\varphi} = 0$ on the trajectories which have as solutions $\varphi(t) = ct + d$. We find that the stationary points are straight lines.

Definition 2.1.7

1. *We call a smooth function $l: \mathrm{J}^1\mathbb{R}^N \to \mathbb{R}$ the Lagrangian function of a classical mechanical system.*

2. *Given a Lagrangian $l(x^i, x_t^i, t)$ and a trajectory $\varphi: \ I \to \mathbb{R}^N$, we call $\mathrm{L}(\varphi) = \int_I \mathrm{d}t \ l \circ \mathrm{j}\varphi$ the corresponding action for the trajectory φ.*

Observation 2.1.8.

1. *We want to discuss in an example how changes of coordinates affect a Lagrangian. We choose a two-dimensional natural system with vanishing potential. In Cartesian coordinates, the Lagrangian is:*

$$l(x^1, x^2, x_t^1, x_t^2) = \frac{1}{2}\left((x_t^1)^2 + (x_t^2)^2\right) \ .$$

We find

$$\frac{\partial l}{\partial x^i} = 0 \quad \textit{and} \quad \frac{\partial l}{\partial x^i_t} = x^i_t$$

which implies the equation for the trajectory

$$\frac{\mathrm{d}}{\mathrm{d}t}(x^i_t(\varphi(t)) = 0$$

The Euler Lagrange equations thus read $\ddot{\varphi}^i = 0$, we thus find constant velocities for the solutions which are straight lines.

We can also consider the system in polar coordinates (r, φ) with $r \in \mathbb{R}_{>0}$ and $\phi \in [0, 2\pi)$. The relation is

$$x = r\cos\phi \quad \textit{and} \quad y = r\sin\phi \ .$$

2. *We next have to derive the transformation of the coordinates $x_{i,t}$ on jet space. This will be done more systematically later when we introduce jet spaces in a coordinate free way. For now, we remark that they should describe velocities. We find for an arbitrary trajectory $\varphi : I \to \mathbb{R}^N$ for two coordinates $(x^i)_{i=1,\dots,N}$ and $(y^i)_{i=1,\dots,N}$*

$$x^i_t(j^1\varphi) = \frac{\mathrm{d}}{\mathrm{d}t}x^i(\varphi(t)) = \frac{\partial x^i}{\partial y^j}\frac{\mathrm{d}}{\mathrm{d}t}y^j(\varphi(t)) = \frac{\partial x^i}{\partial y^j}y^j_t(j^1\varphi)$$

and thus

$$x^i_t = \frac{\partial x^i}{\partial y^j}y^j_t \ .$$

In our example, we get

$$\begin{aligned} x_{1,t} &= (r\cos\phi)_t = r_t\cos\phi - r\sin\phi\cdot\phi_t \\ x_{2,t} &= (r\sin\phi)_t = r_t\sin\phi + r\cos\phi\cdot\phi_t \end{aligned} \tag{1}$$

and thus for the Lagrangian in polar coordinates

$$\begin{aligned} l(r,\phi,r_t,\phi_t) &= \tfrac{1}{2}\big(r_t\cos\phi - r\sin\phi\phi_t)^2 + \tfrac{1}{2}\big(r_t\sin\phi + r_t\cos\phi\phi_t\big)^2 \\ &= \tfrac{1}{2}(r_t^2 + r^2\phi_t^2). \end{aligned}$$

For the Euler Lagrange equations, we compute

$$\frac{\partial l}{\partial \phi} = 0 \quad \textit{and} \quad \frac{\partial l}{\partial \phi_t} = r^2\phi_t$$

which implies the equation for the trajectory

$$\frac{\mathrm{d}}{\mathrm{d}t}(r(\varphi)^2\dot{\phi}(\varphi)) = 0 \ .$$

This means that for any trajectory φ that extremizes the functional, the quantity

$$r^2\cdot\phi(\varphi)$$

does not depend on the eigentime t of the trajectory. that expresses the conservation of angular momentum with respect to the origin. For the radial coordinate f, we find

$$\frac{\partial l}{\partial r} = r\dot{\phi}^2 \quad \textit{and} \quad \frac{\partial l}{\partial r_t} = \dot{r} \quad \textit{and thus} \qquad \ddot{r} = r\dot{\phi}^2.$$

This motivates us to introduce the following terminology:

Definition 2.1.9

1. *Any local coordinate x^i on $\mathbb{R}^N$ is called a generalized coordinate, x^i_t a (generalized) velocity. The function $\dfrac{\partial l}{\partial x^i_t}$ on jet space is called the generalized momentum canonically conjugate to the coordinate x^i. $\frac{\partial l}{\partial x^i}$ is called the generalized force.*

2. *A coordinate is called cyclic if the Lagrangian does not depend on it, i.e.*

$$\frac{\partial l}{\partial x^i} = 0.$$

Proposition 2.1.10.
The momentum canonically conjugate to a cyclic coordinate is constant for any solution of the Euler-Lagrange equations.

Proof:
For any trajectory

$$\varphi : \quad I \to \mathbb{R}^N$$

that is a solution of the Euler-Lagrange equations, we have for a cyclic coordinate x^i

$$\frac{\mathrm{d}}{\mathrm{d}\, t}\frac{\partial l}{\partial x^i_t} \circ \mathrm{j}^1\varphi = \frac{\partial l}{\partial x^i} \circ \mathrm{j}^1\varphi = 0.$$

□

Example 2.1.11.
To investigate the motion of earth around the sun, we consider Kepler's problem that is defined by the following Lagrangian

$$l_k : \mathrm{J}^1(\mathbb{R}^3) \to \mathbb{R}$$
$$l_k(x, x_t) = \tfrac{1}{2}\|x_t\|^2 - \tfrac{k}{\|x\|} \ .$$

Here k is a constant that we keep for convenience. If we allow for a mass parameter in front of the kinetic term, $\frac{1}{2}m\|x_t\|^2$ with m the mass of the earth, then k is a universal constant describing the strength of Newtonian gravity.

This problem can be solved explicitly with conics in the plane as trajectories; in the case of an ellipsis, one recovers Kepler's first law that the sun is in the focus of an ellipsis. This law only holds for a potential of the form $-\frac{k}{\|x\|}$. Details can be found in any book on classical mechanics.

Kepler's second law is just the fact that similarly to the case of observation 2.1.8, the quantity $r(\varphi)^2\dot{\phi}(\varphi)$ is conserved on any trajectory. It rephrased by saying that a line joining a planet and the sun sweeps out equal areas during equal intervals of time. Kepler's second law actually holds for any potential of the form $V(x) = f(\|x\|)$, i.e. a potential that depends only on the norm of x.

We finally show Kepler's third law, using a scaling argument and without solving the problem explicitly. We consider for any $\lambda > 0$ the map

$$\begin{aligned} s_\lambda :\ & \mathrm{J}^1(\mathbb{R}^3) \to \mathrm{J}^1(\mathbb{R}^3) \\ & s_\lambda(x, x_t) := (\lambda^2 x, \lambda^{-1} x_t). \end{aligned}$$

It is chosen in such a way that for the the specific Lagrangian we are interested in we have

$$l \circ s_\lambda = \lambda^{-2} l$$

and therefore for the Euler-Lagrange operators

$$\mathrm{E}\left[l \circ s_\lambda\right] = \lambda^{-2} \mathrm{E}\left[l\right].$$

Suppose we have a trajectory $\varphi :\ \mathbb{R} \to \mathbb{R}^N$ that is a global solution of the Euler-Lagrange equations

$$\mathrm{E}\left[l\right] \circ \mathrm{j}^1\varphi = 0\ .$$

Consider for any $\lambda \in \mathbb{R}_{>0}$ the trajectory

$$\begin{aligned} \varphi_\lambda :\ & \mathbb{R} \to \mathbb{R}^3 \\ & \varphi_\lambda(t) := \lambda^2 \varphi(\lambda^{-3} t). \end{aligned}$$

We then find

$$\mathrm{j}^1\varphi_\lambda(t) = (\lambda^2 \varphi(\lambda^{-3}t), \lambda^2 \lambda^{-3} \dot{\varphi}(\lambda^{-3}t)) = s_\lambda\, \mathrm{j}^1\varphi(\lambda^{-3}t)).$$

and thus

$$\mathrm{E}\left[l\right] \circ \mathrm{j}^1\varphi_\lambda(t) = \mathrm{E}\left[l\right] \circ \mathrm{s}_\lambda\ \mathrm{j}^1\varphi(\lambda^3 t) = \lambda^{-1} E\left[l\right] \mathrm{j}^1\varphi(\lambda^3 t) = 0.$$

and conclude that the trajectory φ_λ is a solution of the equations of motion as well.

We fix a solution that describes an ellipsis and consider the corresponding scaled family. The scaling affects the semimajor axis as

$$a_\lambda = \lambda^2 a.$$

If T is the time for one period,

$$\varphi(t + T) = \varphi(t),$$

then

$$\varphi_\lambda(t + \lambda^3 T) = \lambda^2 \varphi(\lambda^{-3} t + T) = \lambda^2 \varphi(\lambda^{-3} t) = \varphi_\lambda(t),$$

and hence we find for the periods

$$T_\lambda = \lambda^3 T.$$

We have found from scaling considerations Kepler's third law:

$$\frac{a_\lambda^3}{T_\lambda^2} = \frac{\lambda^6 a^3}{\lambda^6 T^2} = \frac{a^3}{T^2}$$

is independent of λ. It should be stressed that it only holds for a potential of the form $-\frac{k}{\|x\|}$.

2.2 Systems with constraints

Frequently, one observes that N particles can only move on a submanifold of $\mathbb{R}^{3N}$. Examples include:

- A roller coaster is (hopefully) more or less obliged to move on a one-dimensional submanifold.
- The points of a rigid body are required to move in such a way that their distances remain constant.

This is achieved by forces that are not known explicitly, but whose effect is known.

Hence we consider an embedded submanifold $M \subset \mathbb{R}^N$ which we can at least locally describe as the zero locus of r functions

$$g_\alpha : \mathbb{R}^N \to \mathbb{R}, \quad \alpha = 1, ..., r \ ,$$

i.e. as

$$M = \big\{ x \in \mathbb{R}^N | g_\alpha(x) = 0 \ \text{ for } \ \alpha = 1, \dots, r \big\}.$$

For example, a roller coaster moving on a circle of radius R in the $x-z$-plane is described by two equations

$$M = \big\{ x \in \mathbb{R}^3 \,|\, x^2 + z^2 = R^2 \quad \text{and} \quad y = 0 \big\} \ .$$

A rigid body consisting of N distinct particles is described by $\frac{N(N-1)}{2}$ functions, one for each pair of distinct points,

$$\|x_i - x_j\|^2 - c_{ij} = g_{ij}(x_1, ..., x_N)$$

with constants $c_{ij} \in \mathbb{R}_{>0}$. As a regularity assumption we require the rank of the Jacobian to be maximal,

$$\operatorname{rang} \frac{\partial g_\alpha}{\partial x^i}\Big|_{x \in M} = r = \text{maximal}.$$

As a warm-up problem, we restrict a smooth function $f : \mathbb{R}^N \to \mathbb{R}$ to the submanifold M and describe the stationary points of the restriction $f|_M$.

Proposition 2.2.1.
Consider the auxiliary function

$$\begin{array}{rccl} \tilde{f} : & \mathbb{R}^N \times \mathbb{R}^r & \to & \mathbb{R} \\ & (x, \lambda) & \mapsto & f(x) + \sum_{\alpha=1}^r \lambda_\alpha g_\alpha(x) \ . \end{array}$$

The additional parameters $\lambda \in \mathbb{R}^r$ are called Lagrangian multipliers. Then the restriction $f|_M$ has a stationary point in $x_0 \in M$, if and only if the function $\tilde{f}$ has a stationary point in (x_0, λ_0) for some $\lambda_0 \in \mathbb{R}^r$.

Proof:
The function $\tilde{f}$ has a stationary point $(x_0, \lambda_0) \in \mathbb{R}^N \times \mathbb{R}^r$, if and only if the two equations hold

$$0 = \frac{\partial \tilde{f}}{\partial \lambda^\alpha}\Big|_{(x_0, \lambda_0)} = g_\alpha(x_0) \quad \text{and} \quad \frac{\partial f}{\partial x^i}\Big|_{x_0} = -\sum_{\alpha=1}^r \lambda_\alpha \frac{\partial}{\partial x^i} g_\alpha(x_0) \ .$$

The first equation is equivalent to $x_0 \in M$. The second equation requires the gradient of f in x_0 to be normal to M, ensuring that x_0 is a stationary point of the restriction $f|_M$. $\square$

Observation 2.2.2.

1. *We now consider the natural system on $\mathbb{R}^N$ given by a potential $V : \mathbb{R}^N \to \mathbb{R}$,*

$$l(x, x_t) = \frac{1}{2}\sum_{i=1}^{N}(x_t^i)^2 - V(x) \ ,$$

that is constrained by unknown forces to a submanifold $M \subset \mathbb{R}^N$ of dimension $N-r$. We assume that M is given by r smooth functions $g_\alpha = g_\alpha(x^i)$.

We are only interested in trajectories

$$\varphi : I \to \mathbb{R}^N$$

with $\operatorname{im}\ \varphi \subset M$. In a variational family, only trajectories satisfying the constraints $g_\alpha(\varphi_\epsilon(t)) = 0$ for all $\alpha = 1, ..., r$ and all ϵ, t are admitted. For such a family, we are looking for the stationary points of

$$\int_I \mathrm{d}t \ l \circ \mathrm{j}^1\varphi_\epsilon \ .$$

We introduce Lagrangian multipliers and minimize the functional

$$f(\epsilon, \lambda) = \int_I \mathrm{d}t \left(l \circ \mathrm{j}^1\varphi_\epsilon + \sum_{\alpha=1}^{r} \lambda_\alpha g_\alpha(\varphi_\epsilon) \right) \ .$$

The derivative with respect to ϵ yields the following additional term

$$\sum_{\alpha=1}^{r} \lambda_\alpha \ \frac{\partial g_\alpha}{\partial x^i} \ \frac{\mathrm{d}\varphi_\epsilon^i}{\mathrm{d}\epsilon}\Big|_{\epsilon=0}$$

so that the equation of motions for a trajectory φ become

$$\mathrm{E}(l) \circ \mathrm{j}^1(\varphi) = \ddot{\varphi} + \mathrm{grad}V(\varphi) = -\sum \lambda_\alpha \ \mathrm{grad}\, g_\alpha(\varphi) \ . \tag{2}$$

The right hand side describes additional forces constraining the motion to the submanifold M. The concrete form of the forces described by the gradients of the functions g_α is, in general, unknown.

2. *We describe the local geometry of the situation in more detail: we consider a local coordinate*

$$q : M \supset U \to \mathbb{R}^{N-r}$$

of the submanifold, where $U \subset M$ is open. We use the embedding $M \hookrightarrow \mathbb{R}^N$ to identify T_pM with a vector subspace of $T_p\mathbb{R}^N$ and express the basis vectors as

$$\frac{\partial}{\partial q^\alpha} = \frac{\partial x^i}{\partial q^\alpha}\frac{\partial}{\partial x^i}, \qquad \alpha = 1, ..., N - r.$$

In subsequent calculations, the notation is simplified by introducing the vector $x \in \mathbb{R}^N$ with coordinates x^i, $i = 1, \dots N$. We then write for the basis vector of T_pM

$$\frac{\partial}{\partial q^\alpha} = \frac{\partial x}{\partial q^\alpha} \ . \tag{3}$$

3. *Next, we have to relate the jet spaces* J^2M *and* $\mathrm{J}^2\mathbb{R}^N$*. Since they are designed as recipients of trajectories and their derivatives, we consider a trajectory with values in* $U \subset M$

$$\varphi : \ I \to U \subset M \ .$$

Using the embedding $U \to \mathbb{R}^N$*, we can see this also as a trajectory* $\tilde{\varphi}$ *in* $\mathbb{R}^N$*: The chain rule yields*

$$\frac{\mathrm{d}\tilde{\varphi}^i}{\mathrm{d}t} = \frac{\partial x^i}{\partial q^\alpha} \cdot \frac{\mathrm{d}\varphi^\alpha}{\mathrm{d}t}.$$

This is, of course, just the usual map of tangent vectors induced by a smooth map of manifolds, in the case the embedding $M \hookrightarrow \mathbb{R}^N$*. From this, we deduce the following expression for the coordinates* x_t *of* $\mathrm{J}^1\mathbb{R}^N$*, seen as a function on jet space* J^1M*:*

$$x_t^i = x_t^i(q^\alpha, q_t^\alpha, t) = \frac{\partial x^i}{\partial q^\alpha} q_t^\alpha \tag{4}$$

and expressed in coordinates q^α, q_t^α *on* J^1M*. We find as an obvious consequence*

$$\frac{\partial x_t^i}{\partial q_t^\alpha} = \frac{\partial x^i}{\partial q^\alpha} \ . \tag{5}$$

Similarly, to find the transformation rules for the coordinates x_{tt}^i *describing the second derivative, we compute the second derivative of a trajectory:*

$$\frac{\mathrm{d}^2\tilde{\varphi}}{\mathrm{d}t^2} = \frac{\partial^2 x^i}{\partial q^\beta \partial q^\alpha} \cdot \frac{\mathrm{d}\varphi^\beta}{\mathrm{d}t}\frac{\mathrm{d}\varphi^\alpha}{\mathrm{d}t} + \frac{\partial x^i}{\partial q^\alpha}\frac{\mathrm{d}^2\varphi^\alpha}{\mathrm{d}t^2}.$$

This yields the following expression of the coordinate function x_{tt}^i *on* $\mathrm{J}^2\mathbb{R}^N$ *as a function on* J^2M*, in terms of the coordinates* q_{tt}^α *on* J^2M*:*

$$x_{tt}^i = x_{tt}^i(q^\alpha, q_t^\alpha, q_{tt}^\alpha, t) = \frac{\partial x^i}{\partial q^\alpha} q_{tt}^\alpha + \frac{\partial^2 x^i}{\partial q^\alpha \partial q^\beta} q_t^\beta q_t^\alpha \ . \tag{6}$$

It should be appreciated that this transformation rule is not linear in q_{tt}*.*

4. *Since we do not know the right hand side of the equation of motion (2), we take the scalar product* $\langle \cdot, \cdot \rangle$ *in* $\mathbb{R}^N$ *with each of the tangent vector (3) in* T_pM*. This yields* r *equations*

$$\left\langle x_{tt}, \frac{\partial x}{\partial q^\alpha} \right\rangle = -\left\langle \operatorname{grad} V, \frac{\partial x}{\partial q^\alpha} \right\rangle \qquad (*) \quad \textit{with} \quad \alpha = 1, \dots r \ .$$

Our goal is to rewrite these equations as the Euler-Lagrange equations for an action function on the jet space of the submanifold J^1M*. The right hand side is easily rewritten using the chain rule:*

$$\left\langle \operatorname{grad} V, \frac{\partial x}{\partial q^\alpha} \right\rangle = \frac{\partial V}{\partial x^i}\frac{\partial x^i}{\partial q^\alpha} = \frac{\partial V}{\partial q^\alpha}$$

This suggests to take the restriction of V *to* M *as the potential for the Lagrangian system on* M*.*

5. *To rewrite the left hand side of $(*)$, we use the total differential operator on smooth functions of $x^i, x^i_t, x^i_{tt}, \ldots$ introduced in definition 2.1.2:*

$$\mathrm{D}_x = \sum_{i,\beta} x^i_{\beta+1} \frac{\partial}{\partial q^i_\beta}.$$

Note that the operator is linear and obeys the Leibniz rule. Recall that the Euler-Lagrange equations read in terms of this operator

$$\mathrm{D}_x \frac{\partial l}{\partial x^i_t} = \frac{\partial l}{\partial x^i} \ .$$

Similarly, we introduce the differential operator

$$\mathrm{D}_q = \sum_{\alpha,\beta} q^\alpha_{\beta+1} \frac{\partial}{\partial q^\alpha_\beta}.$$

on smooth functions of $q^\alpha, q^\alpha_t, q^\alpha_{tt}, \ldots$

6. *We need the following differential identity for functions on J^2M:*

$$\mathrm{D}\left(\left\langle x_t, \frac{\partial x}{\partial q_\alpha}\right\rangle\right) = \left\langle x_{tt}, \frac{\partial x}{\partial q^\alpha}\right\rangle + \left\langle x_t, \frac{\partial x_t}{\partial q^\alpha}\right\rangle \ .$$

To see this, we first use the Leibniz rule for D to find

$$\mathrm{D}\left\langle x_t, \frac{\partial x}{\partial q^\alpha}\right\rangle = \left\langle \mathrm{D}x_t, \frac{\partial x}{\partial q^\alpha}\right\rangle + \left\langle x_t, \mathrm{D}\frac{\partial x}{\partial q^\alpha}\right\rangle$$

and compute for the first summand

$$\mathrm{D}\, x_t = \frac{\partial x_t}{\partial q^\alpha} q^\alpha_t + \frac{\partial x_t}{\partial q^\alpha_t} q^\alpha_{tt} = \frac{\partial^2 x}{\partial q^\beta \partial q^\alpha} q^\alpha_t q^\beta_t + \frac{\partial x}{\partial q^\alpha} q^\alpha_{tt} = x_{tt}$$

where we first used (4) and then (5). The last identity is (6).

For the second summand, we compute

$$\mathrm{D}\left(\frac{\partial x^i}{\partial q^\alpha}\right) = \frac{\partial^2 x^i}{\partial q^\beta \partial q^\alpha}\, q^\beta_t,$$

which turns out to be the same as

$$\frac{\partial x^i_t}{\partial q_\alpha} = \frac{\partial}{\partial q^\alpha}\left(\frac{\partial x^i}{\partial q^\beta}\, q^\beta_t\right) = \frac{\partial^2 x^i}{\partial q^\alpha \partial q^\beta}\, q^\beta_t.$$

7. *We now introduce the following function on the jet space J^1M depending on the coordinates q^α and q^α_t:*

$$T = \frac{1}{2}\sum_{i=1}^N (x^i_t)^2 = \frac{1}{2}\sum_{i=1}^N x^i_t(q^\alpha, q^\alpha_t)^2$$

To compute Euler-Lagrange equations for the function T on M, we first note that the chain rule implies

$$\frac{\partial T}{\partial q^\alpha} = \left\langle x_t, \frac{\partial x_t}{\partial q^\alpha} \right\rangle . \tag{7}$$

Next, we calculate, again using the chain rule,

$$\frac{\partial T}{\partial q_t^\alpha} = \left\langle x_{tt}, \frac{\partial x_t}{\partial q_t^\alpha} \right\rangle = \left\langle x_{tt}, \frac{\partial x}{\partial q^\alpha} \right\rangle$$

where we have used the relation (6). The left hand side of the Euler-Lagrange equation for T then reads

$$\mathrm{D}\left(\frac{\partial}{\partial q_t^\alpha} T\right) = \mathrm{D}\left\langle x_t, \frac{\partial x}{\partial q^\alpha} \right\rangle = \left\langle x_{tt}, \frac{\partial x}{\partial q^\alpha} \right\rangle + \left\langle x_t, \frac{\partial x_t}{\partial q^\alpha} \right\rangle ,$$

where we used in the last equalities identities derived in step 6. Together with equation (7) this gives us the identity

$$\left\langle x_{tt}, \frac{\partial x}{\partial q^\alpha} \right\rangle = \mathrm{D}\left(\frac{\partial}{\partial q_t^\alpha} T\right) - \frac{\partial T}{\partial q^\alpha}$$

8. *We substitute this in the equation of motion $(*)$ and find*

$$\mathrm{D}\left(\frac{\partial}{\partial q_t^\alpha} T\right) = -\frac{\partial V}{\partial q^\alpha} + \frac{\partial T}{\partial q^\alpha} .$$

Taking into account that V is independent of x_t^i and thus of q_t^α, we realize that these are the Euler-Lagrange equations for the Lagrangian

$$\mathrm{L}(q^\alpha, q_\tau^\alpha) = T(x_t(q, q_t)) - V(x(q)),$$

on $\mathrm{J}^1 M$.

We have thus shown:

Proposition 2.2.3.
Suppose a natural system on $\mathbb{R}^N$ with potential $V(x)$ is constrained by unknown forces to a submanifold $M \subset \mathbb{R}^N$. Then the equations of motion on M are the Euler-Lagrange equations for a Lagrangian $l(q, q_t)$ on $\mathrm{J}^1 M$ obtained from the Lagrangian on $\mathrm{J}^1\mathbb{R}^N$ by restricting V to M and transforming the coordinates x_t as in (4).

This result shows that, for the description of derivatives, we should associate bundles to a Lagrangian system that can be pulled back consistently.

2.3 Lagrangian systems: jet bundles as the kinematical setup

We will now look at Lagrangian mechanical systems in a more geometric way. A canonical system consists of a smooth Riemannian manifold M and a potential function $V : M \to \mathbb{R}$ from which we can build a Lagrangian. Trajectories are then smooth maps $I \to M$ or, equivalently, smooth sections of the surjective submersion $E := I \times M \to I$ given by the projection on the first factor. Now configuration space might depend on time; hence we do not ask E to be a product.

The leads us to the following formalizations:

Definition 2.3.1

1. *The kinematical description of a Lagrangian mechanical system is given by a surjective submersion $\pi : E \to I$, where $I \subset \mathbb{R}$ is an interval of "eigentime". E is called the extended configuration space. The fibre $\pi^{-1}(t) =: E_t$ for $t \in I$ is the space of configurations of the physical system at time t. If E is of the form $E = I \times M$, then M is called the configuration space. The fibre $\pi^{-1}(t) =: E_t$ for $t \in I$ is the space of configurations of the physical system at time t.*

2. *Global trajectories are global sections of π, i.e. smooth maps $\varphi : I \to E$ such that $\pi \circ \varphi = \mathrm{id}_I$. Local sections give local trajectories. For an open subset $U \subset I$, a local section is a smooth map $\varphi : U \to E$ such that $\pi \circ \varphi = \mathrm{id}_U$. The system of local sections forms a sheaf on I.*

3. *A system defined by constraints on the system $\pi : E \to I$ is a submanifold $\iota : E' \hookrightarrow E$ such that $\pi' := \pi \circ \iota$ is a surjective submersion $E' \to I$.*

4. *We also define the composition of two systems $\pi_1 : E_1 \to I$ and $\pi_2 : E_2 \to I$ defined over the same interval of eigentime as the fibre product*

$$\begin{array}{ccc} E_1 \times_I E_2 & \longrightarrow & E_2 \\ \downarrow & & \downarrow{\scriptstyle \pi_2} \\ E_1 & \xrightarrow{\pi_1} & I \end{array}$$

   If there is a configuration space for both systems, $E_i = I \times M_i$, then $E_1 \times_I E_2 = (I \times M_1) \times_I (I \times M_2) \cong I \times (M_1 \times M_2)$ so that the configuration space of the composite space is the Cartesian product of the configuration spaces.

Remarks 2.3.2.

1. *Later on, we will consider in this framework variational problems. In the present setup, not just the image of the trajectories, but also their parametrization is kept as an important piece of data. In this respect, the situation is frequently different for varational problems whose origin is in geometry: if we are looking e.g. for a shortest curve connecting two points, independently of its parametrization.*

2. *Let us relate this definition to the situation of a particle moving in Galilei space $\mathbb{A}$ that has been discussed in the first chapter.*

   We first construct a surjective submersion from $\mathbb{A}$ to the real line. This can be done by fixing a space-time point $p_0 \in \mathbb{A}$ and taking the absolute time difference to the reference point p_0

$$\begin{array}{ccc} \mathbb{A} & \to & \mathbb{R} \\ p & \mapsto & t(p - p_0) \end{array}$$

3. *We do not require E to be a bundle over I, i.e. to be locally a product. Any smooth bundle over an interval I is globally trivial as a smooth manifold. The extended configuration space E, however, in general carries additional structure that depends on I and that does not necessarily have a product structure.*

4. *In case the configuration space is a smooth manifold M, it is natural to consider a force field that is independent of the velocities as a one-form $f \in \Omega^1(M)$. Then the condition of the force field being conservative is only sufficient for the existence of a potential, if M is simply connected.*

Remarks 2.3.3.

1. *We will generalize the situation by considering a general surjective submersion $\pi : E \to M$, where M is not necessarily an interval.*

   This has applications for the kinematical setup of classical field theory: here trajectories are not parametrized any longer by a one-dimensional manifold, i.e. an interval, but by some manifold M of higher dimension. The dimension of M is then called the dimension of the classical field theory. The (local) sections of $\pi : E \to M$ are then not called trajectories any longer, but (local) field configurations.

2. *The following crucial difference of interpretation between field theory and mechanics should be noted: in mechanics, there is one parameter which can be chosen to be identical to time. Local spacial coordinates are coordinate functions on the manifold E of the extended configuration space. In a field theory, on the other hand, space and time are locally defined parameters and are to be considered as local coordinate functions on M.*

3. *In certain situations, it might be necessary to endow either of the manifolds E or M with more structure. Examples in field theory include metrics endowing E with the structure of a Riemannian manifold or, to be able to include fermions, a spin structure on M.*

To be able to formulate a dynamical principle in terms of a partial differential equation, we need to work with derivatives of any order of local sections. This is achieved by the following construction.

Definition 2.3.4

1. *Let E, M be smooth manifolds, $\dim M = m$ and let $\pi : E \to M$ be a surjective submersion. We say that π defines a fibred manifold.*

2. *For any open subset $U \subset M$, we denote by $\Gamma_\pi(U)$ the set of local smooth sections of π, i.e. the set of smooth functions $s_U : U \to E$ such that $\pi \circ s_U = \mathrm{id}_U$. These sets form a sheaf on M.*

3. *Consider a point $p \in M$ and two open neighbourhoods U, U' of p. We identify two local sections $\varphi \in \Gamma_\pi(U)$ and $\varphi' \in \Gamma_\pi(U')$ if they agree on a common open refinement $\tilde{U} \subset U \cap U'$. The equivalence classes are called germs of local sections in the point p. The set of equivalence classes is denoted by $\Gamma_\pi(p)$.*

4. *For any multi-index $I = (I(1), ..., I(m))$, we introduce its length*

$$|I| := \sum_{i=1}^{m} I(i)$$

   and the derivative operators

$$\frac{\partial^{|I|}}{\partial x^I} := \prod_{i=1}^{m} \left(\frac{\partial}{\partial x^i} \right)^{I(i)}$$

acting on functions of m variables $x_1, \dots x_m$. The multi-index $I = \emptyset$ is admitted as a value, $|\emptyset| = 0$ and the corresponding derivative operator is the identity.

5. *Let $p \in M$ and U, U' be open neighborhoods of M. Consider two local sections $\varphi \in \Gamma_\pi(U)$ and $\varphi' \in \Gamma_\pi(U')$ whose value in p agrees. Possibly after a further restriction, we can assume that the two sections are defined on the same coordinate neighborhood of p that we call for simplicity U. We fix local coordinates x^i on M around p and y^α on E around $\varphi(p) = \varphi'(p)$.*

   We say that φ and φ' have the same r-jet in p, if in local coordinates all partial derivatives up to order r coincide,

$$\frac{\partial^{|I|}\varphi^\alpha}{\partial x^I}\Big|_p = \frac{\partial^{|I|}(\varphi')^\alpha}{\partial x^I}\Big|_p, \quad 0 \leqslant |I| \leqslant r \ .$$

   Since differentiation is a local operation, local sections with the same germ in $p \in M$ give rise to the same r-jet. We have thus defined an equivalence relation on germs of sections that is moreover independent of the choice of local coordinates.

6. *An r-jet with representative φ will be denoted by $j^r_p\varphi$; the natural number r is called the order of the jet. The point $p \in M$ is called the source and $\varphi(p) \in E$ the target of the jet.*

7. *We introduce the sets*

$$\mathrm{J}^r_p\pi := \left\{ \mathrm{j}^r_p\varphi \middle| \varphi \in \Gamma_\pi(p) \right\} \qquad \text{and} \qquad \mathrm{J}^r\pi := \cup_{p\in M}\mathrm{J}^r_p\pi \ .$$

   (The notation $\mathrm{J}^r E$ is also in use, but we wish to stress that the set is determined by a morphism π and not only by the manifold E.)

So far, we have just introduced a way to address not only local sections of a fibred manifold, but also their derivatives up to some finite order or, put differently, their Taylor polynomial up to some finite order.

This will help us to deal with velocities and accelerations in the case of a mechanical system and, more generally, with the dynamics of fields. We wish to endow this with more geometric structure.

Observation 2.3.5.

1. *We have the following obvious projections:*

$$\begin{array}{rccll} \pi_r : & \mathrm{J}^r\pi & \to & M & (\textit{source projection}) \\ & \mathrm{j}^r_p(\varphi) & \mapsto & p & \\ \pi_{r,0} : & \mathrm{J}^r\pi & \to & E & (\textit{target projection}) \\ & \mathrm{j}^r_p(\varphi) & \mapsto & \varphi(p) & \end{array}$$

   and for $r \geqslant k \geqslant 1$

$$\begin{array}{rccll} \pi_{r,k} : & \mathrm{J}^r\pi & \to & \mathrm{J}^k\pi & (\textit{jet projections}) \\ & \mathrm{j}^r_p\varphi & \mapsto & \mathrm{j}^k_p\varphi \ . & \end{array}$$

Note that the jet projections just amount to forgetting derivatives of order $k+1, k+2, \ldots, r$. One should check that the map $\pi_{r,r-1} : \mathrm{J}^r\pi \to : \mathrm{J}^{r-1}\pi$ actually defines the structure of an affine bundle. Such that the following diagram commutes:

$$\begin{array}{ccccccccc} \mathrm{J}^r\pi & \xrightarrow{\pi_{r,r-1}} & \mathrm{J}^{r-1}\pi & \longrightarrow & \ldots & \longrightarrow & \mathrm{J}^1\pi & \xrightarrow{\pi_{1,0}} & E \\ \downarrow{\scriptstyle \pi_r} & & \downarrow{\scriptstyle \pi_{r-1}} & & \downarrow & & \downarrow{\scriptstyle \pi_1} & & \downarrow{\scriptstyle \pi} \\ M & = & M & = & \ldots & = & M & = & M \end{array}$$

2. *In a next step, we introduce local coordinates to endow $\mathrm{J}^r\pi$ with the structure of a smooth manifold and the surjection $\pi_r : \mathrm{J}^r\pi \to M$ with the structure of a bundle map, the jet bundle of π with fibre $\mathrm{J}^r_p\pi$. We include the case $\mathrm{J}^0\pi := E$.*

   We start with a bundle chart (U, u) for the bundle E. This is a pair, consisting of a local coordinate chart $x : U \subseteq M \to \mathbb{R}^m$ of M, where $U \subset M$ is an open subset of M and a local coordinate chart $u : \pi^{-1}(U) \to \mathbb{R}^n$ of E.

   A common terminology is to call the local coordinates x of M the independent coordinates and the local coordinates u of E the dependent coordinates. This language is motivated by the situation where we describe local sections $s : M \supset U \to E$. Each section leads to an expression of the local coordinates u^α of E in terms of the local coordinates x^i of E.

3. *The induced bundle chart of $\mathrm{J}^r\pi$ is defined on*

$$U^r := \left\{ \mathrm{j}^r_p\varphi \middle| \varphi \in \Gamma_\pi(U) \right\}$$

   and consists of maps $u^r = (x^i, u^\alpha_I)$ with $|I| \leq r$ defined by the coordinates x^i of the base point $p \in M$

$$x^i\left(\mathrm{j}^r_p\varphi\right) = x^i(p)$$

   and the values of the partial derivatives of the jet in local coordinates (u^α, x^i) at the point p,

$$u^\alpha_I\left(\mathrm{j}^r_p\varphi\right) = \left.\frac{\partial^{|I|}u^\alpha(\varphi)}{\partial x^I}\right|_p .$$

   In the local coordinates (x^i, u^α_I) on $\mathrm{J}^r\pi$ (x^i) on $U \subset M$, the map π_r reads $(x^i, u^\alpha_I) \mapsto (x^i)$. Thus all projections π_r are surjective submersions so that the jet bundles $\mathrm{J}^r\pi$ are indeed finite-dimensionale bundles.

   One should also notice that if $\pi : E \to M$ is a (real) vector bundle, i.e. if any fibre $\pi^{-1}(m) =: E_m$ has a natural structure of a (real) vector space, then the jet bundle $\mathrm{J}^r\pi$ is a vector bundle as well.

4. *Let φ be a local section of π defined on an open subset $U \subset M$. Its r-th jet prolongation is the local section*

$$\mathrm{j}^r\varphi : U \to \mathrm{J}^r\pi$$

   given by

$$\left(\mathrm{j}^r\varphi\right)(p) = \mathrm{j}^r_p\varphi \ .$$

   In local coordinates, it is given by

$$\left(\varphi^\alpha, \frac{\partial^{|I|}}{\partial x^I}\varphi^\alpha\right).$$

Example 2.3.6.
We consider the situation relevant for classical mechanics with time-independent configuration space a smooth manifold M. This is the trivial bundle $p_1 : \mathbb{R} \times M \to \mathbb{R}$.

A section of p_1

$$\begin{aligned} \varphi : \ & \mathbb{R} \to \mathbb{R} \times M \\ & t \mapsto (t, \tilde{\varphi}(t)) \end{aligned}$$

is given by a smooth trajectory

$$\tilde{\varphi} : \mathbb{R} \to M \ .$$

The 1-jet of the section φ in the point $t_0 \in \mathbb{R}$ reads in local coordinates x^i on M

$$\left(t_0, x^i, x^i_t\right) \circ \mathrm{j}^1_t \varphi = \left(t_0, x^i \tilde{\varphi}(t), \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=t_0} x^i \circ \tilde{\varphi}(t) \right) \ .$$

Thus, all information that appears is the tangent vector

$$\left. \frac{\mathrm{d}\tilde{\varphi}}{\mathrm{d}t} \right|_{t=t_0}$$

at the point $\tilde{\varphi}(t_0) \in M$ There is a canonical isomorphism

$$\begin{array}{ccc} \mathrm{J}^1\pi & \to & \mathbb{R} \times TM \\ \mathrm{j}^1_{t_0}\varphi & \mapsto & \left(t_0, \left. \frac{\mathrm{d}\tilde{\varphi}}{\mathrm{d}t} \right|_{t=t_0} \right) \end{array} \ .$$

We add a few remarks:

- *If one is only interested in 1-jets and time independent configuration spaces, one can equally well work with tangent bundles. This is done in many, if not most books of classical Lagrangian mechanics. One will therefore frequently find the statement that the configuration space in Lagrangian mechanics is the tangent bundle of a smooth manifold.*
- *Tangent bundles have the disadvantage that they are not appropriate to describe geometrically equations of motions which involve second order derivatives. In fact, the second jet bundle $\mathrm{J}^2\pi$ can be related to a subbundle of the tangent bundle TTM of the tangent bundle TM which is, however, not a vector subbundle, cf. equation (6).*
- *The language also does not generalize to field theory, i.e. higher dimensional parameter spaces.*

The r-th jet bundle $\mathrm{J}^r\pi$ keeps information on the derivatives of local sections up to order r. Our whole setting involves smooth functions; in fact, structures will become much more transparent if we keep derivatives of all order.

Definition 2.3.7
The infinite jet bundle $\pi_\infty : \mathrm{J}^\infty\pi \to M$ is defined as the projective limit (in the category of topological spaces) of the jet bundles

$$M \underset{\pi}{\leftarrow} E \underset{\pi_{1,0}}{\longleftarrow} \mathrm{J}^1\pi \underset{\pi_{2,1}}{\longleftarrow} \mathrm{J}^2\pi \leftarrow \dots$$

It comes with a family of natural projections

$$\pi_{\infty,r} : \mathrm{J}^\infty\pi \to \mathrm{J}^r\pi \ .$$

Remarks 2.3.8.

1. *We again fix bundle coordinates for $\pi : E \to M$ consisting of local coordinates x on $U \subset M$ and u^α on $\pi^{-1}(U) \subset E$. Then local coordinates for J^∞ are given by x^i, u^α_I with I running over all multiindices of any rank, including $I = \emptyset$. In particular, the jet bundle is infinite-dimensional. In this sense, it is an infinite-dimensional smooth manifold (we do not impose for such manifolds the conditions of being Hausdorff or paracompact). It is therefore quite important to keep in mind its structure of a projective limit of finite-dimensional bundles. It is a vector bundle, if $\pi : E \to M$ is a vector bundle.*

2. *The manifold J^∞ is actually not in a natural way a Banach manifold, i.e. modelled over a Banach space, but rather a Fréchet manifold. The model vector space is the space $\mathbb{R}[[X_1, \ldots X_m]]$ of formal power series in m variables, times some finite dimensional vector space. The topology of the vector spaces is generated by the family of seminorms generated by taking the pullback of the norms on each polynomial space. The space can be shown to be a complete metrizable Hausdorff space.*

3. *Any local section $s : U \to E$ can be extended to a section $\mathrm{j}^\infty(s) = \mathrm{j}^\infty s$ of $\mathrm{J}^\infty\pi$. Again, we have in local coordinate just the partial derivative functions*

$$u^\alpha_I\left(\mathrm{j}^\infty_p s\right) = \left.\frac{\partial^{|I|} u^\alpha(s)}{\partial x^I}\right|_p .$$

   However, not every local section ψ of a finite jet bundle is a prolongation: there is simply no reason why, in general, the coordinate representation of ψ should satisfy

$$u^\alpha_I(\psi) = \frac{\partial^{|I|}(u^\alpha \circ \psi)}{\partial x^I} \ .$$

   A section satisfying this is called a holonomic section.

4. *In other words, the extension to $\mathrm{j}^\infty(s)$ amounts to keeping all partial derivatives of the section s, or put differently, its Taylor series as a formal series. One should keep in mind that two different smooth functions can have the same Taylor series, though. Given a local section s, knowing $\mathrm{j}^\infty(s)$ at a single point $p \in M$ is thus not enough to determine s.*

The operation of taking jets behaves in a nice way under local diffeomorphisms:

Remark 2.3.9.

1. *We consider the category* Fib *of fibred manifolds: its objects are surjective submersions $\pi : E \to M$. As morphisms, we take fibred maps whose projections are local diffeomorphisms on the base, i.e. a morphism from $\pi : E \to M$ to $\pi' : E' \to M'$ is a pair of smooth maps*

$f_E : E \to E'$ *and* $f_M : M \to M'$*, where* $f_M : M \to M'$ *is a local diffeomorphism, such that the diagram*

$$\begin{array}{ccc} E & \xrightarrow[f_E]{} & E' \\ \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi'} \\ M & \xrightarrow[f_M]{} & M' \end{array}$$

commutes.

2. *Then for each* r*, we have a covariant functor*

$$\begin{array}{rcl} \mathrm{J}^r : \ \mathrm{Fib} & \to \mathrm{Fib} & \\ (\pi : E \to M) & \mapsto & (\pi_r : \mathrm{J}^r\pi \to M) \end{array}$$

We next need functions on the infinite jet bundle.

Definition 2.3.10
A real-valued function $f : \mathrm{J}^\infty\pi \to \mathbb{R}$ *is called* <u>smooth</u> *or* <u>local</u>*, if it factorizes through a smooth function on a finite jet bundle* $\mathrm{J}^r\pi$*, i.e. there exists* $r \in \mathbb{N}$ *and a smooth function* $f_r : \mathrm{J}^r\pi \to \mathbb{R}$ *such that*

$$f = f_r \circ \pi_{\infty r} \ .$$

The ring of smooth functions is denoted by $C^\infty(\mathrm{J}^\infty\pi)$ *or by* $\mathrm{Loc}(E)$*.*

Remarks 2.3.11.

1. *In plain terms, local functions are those functions on jet space which depend only on a finite number of derivatives.*

2. *The algebra* $\mathrm{Loc}(E)$ *of smooth functions on* $\mathrm{J}^\infty\pi$ *is thus defined as the inductive limit of the injections*

$$\pi^*_{k+1,k} : C^\infty\left(\mathrm{J}^k\pi\right) \to C^\infty\left(\mathrm{J}^{k+1}\pi\right)$$

of algebras. The embedding is given by considering a function in $C^\infty\left(\mathrm{J}^k\pi\right)$ *that depends on the derivatives of local sections up to order* k *as a function that depends on the derivatives up to order* $k+1$*, but in a trivial way on the* $k+1$*-th derivatives.*

3. *The algebra* $C^\infty\left(\mathrm{J}^\infty\right)$ *has the structure of a filtered commutative algebra.*

4. *A smooth function* $f \in \mathrm{Loc}(E)$ *can be seen as a non-linear differential operator* Δ_f *with values in smooth functions on* M*, acting on smooth sections* $s \in \Gamma_\pi$ *by*

$$\Delta_f(s) = f(\mathrm{j}^\infty(s)) = \mathrm{j}^\infty(s)^*(f) \ .$$

For example, in the notation used in the discussion of Lagrangian systems of classical mechanics, the smooth function

$$l(x, x_t) = \frac{1}{2}|x_t|^2 - V(x)$$

describing the Lagrangian of a natural system acts on a section $s : I \to M$ *by yielding the smooth function*

$$s \mapsto \frac{1}{2}\left|\frac{\mathrm{d}s}{\mathrm{d}t}\right|^2 - V(s(t))$$

on the interval I*.*

5. *More generally, consider another surjective submersion $\pi' : E' \to M$ and its pullback E'_π to $\mathrm{J}^\infty\pi$:*

$$\begin{array}{ccc} E'_\pi & \longrightarrow & E' \\ \downarrow & & \downarrow{\scriptstyle \pi'} \\ \mathrm{J}^\infty\pi & \longrightarrow & M \end{array}$$

If $E' := M \times \mathbb{R}$, we have $E'_\pi \cong \mathrm{J}^\infty\pi \times \mathbb{R}$ and a smooth section of E'_π is a smooth function on $\mathrm{J}^\infty\pi$. In general, smooth sections of E'_π can be seen as non-linear differential operators on sections in Γ_π with values in sections of $\Gamma_{\pi'}$.

Given a section $s \in \Gamma_\pi(M)$, its prolongation $\mathrm{j}^\infty s$ is a section of $\pi_\infty : \mathrm{J}^\infty\pi \to M$ and thus provides us in particular with a map

$$\mathrm{j}^\infty s : \ M \to \mathrm{J}^\infty\pi \ .$$

Any local function $l \in \mathrm{Loc}(E)$ can be pulled back along this map to get a smooth function $(\mathrm{j}^\infty s)^* l$ on M. In plain terms, by plugging in the section s and its derivatives into l, we get a function on M. Suppose, we have a volume form dvol_M on M. Then

$$\int_M \mathrm{dvol}_M (\mathrm{j}^\infty s)^* l$$

is a real number associated to s. To have well-defined expressions, one might have to restrict to sections s with compact support. Later, we will avoid the postulate of having a volume form on M by pulling back m-forms on jet space.

This discussion motivates the following definition:

Definition 2.3.12
Let $\pi : \ E \to M$ be a fibred manifold. A local functional on the space of smooth sections $\Gamma_\pi(M)$ is a function

$$S : \Gamma_\pi(M) \to \mathbb{R}$$

which can expressed as the integral over the pullback of a local function $l \in \mathrm{Loc}(E)$.

We also explain terminology of classical Lagrangian field theory described by the surjective submersion $\pi : \ E \to M$. Sections $s \in \Gamma_\pi(U)$ for $U \subset M$ are also called local field configurations. A local function $l \in \mathrm{Loc}(E)$ is then called a Lagrangian for the theory and the functional S the action functional.

We also comment on the case of classical mechanics: The notion of a natural system can now be slightly generalized. Let (M, g) be a smooth Riemannian manifold and $V : M \to \mathbb{R}$ be a potential function. Consider the surjective submersion

$$\mathrm{pr}_1 : \ I \times M \ \to I$$

and the Lagrangian function given by

$$l(x, x_t) \ = \ \frac{1}{2} g_x(x_t, x_t) - V(x) \ .$$

If $V = 0$, the system is called a free system. Its solutions are geodesics.

We have not discussed equations of motions so far. A first attempt for a definition could be as follows: A (geometric) partial differential equation of order r is a closed fibred submanifold $\mathcal{E} \subset \mathrm{J}^r\pi$. As an example, consider the bundle $p_1 : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ whose sections are described by real-valued functions on $\mathbb{R}^n$. There are natural global bundle coordinates $(x^1, ..., x^n, u)$.

A (local) section $s : U \to E$ on an open subset $U \subset M$ is then called a (local) solution, if its prolongation is in the fibred submanifold, $\mathrm{j}^r_p\varphi \in \mathcal{E}$ for all $p \in U$. Physicists sometimes call the field configurations corresponding to such sections "on shell".

As an example, consider as a submanifold $\mathcal{E}_{Lap} \subset \mathrm{J}^2\pi$ the zero locus of the smooth function

$$\begin{array}{rcl} F: \ \mathrm{J}^2\pi & \to & \mathbb{R} \\ (x^i, u, u_i, u_{ij}) & \mapsto & \sum_{i=1}^n u_{ii} \end{array}$$

Local solutions are described by real-valued functions $s : U \to \mathbb{R}$ such that

$$\sum_{i=1}^{n} \frac{\partial^2 s}{(\partial x^i)^2} = 0 \ .$$

The submanifold $\mathcal{E}_{Lap}$ thus describes Laplace's equation which is is a second order partial differential equation for a real-valued function in n variables.

We have emphasized that we prefer to work with the infinite jet space. Then, our preliminary notion of a geometric differential equation is not enough. The reason is that the fact that a section is a local solution implies relations between higher derivatives as well, e.g. for solutions of the two-dimensional Laplace equation the relation

$$\frac{\partial^3 s}{(\partial x)^2 \partial y} + \frac{\partial^3 s}{(\partial x)^3} = 0 \ .$$

To deal with these equations, we need more geometry on the jet bundle $\mathrm{J}^\infty\pi$.

Definition 2.3.13
A vector field on the jet bundle $\mathrm{J}^\infty\pi$ is defined as a derivation on the ring $C^\infty(\mathrm{J}^\infty\pi)$ of smooth functions.

Remark 2.3.14.
In local coordinates (x^i, u^α_I) on $\mathrm{J}^\infty\pi$, a vector field can be described by a formal series of the form

$$X = \sum_i A_i \frac{\partial}{\partial x^i} + \sum_{\alpha, I} B^\alpha_I \frac{\partial}{\partial u^\alpha_I} \ .$$

The word "formal" means that we need infinitely many smooth functions $A_i = A_i(x^j, u^\alpha_I)$ and $B^\alpha_I = B^\alpha_I(x^j, u^\alpha_I)$ to describe the vector field. But once we apply it to a smooth function in $C^\infty(\mathrm{J}^\infty\pi)$, only finitely many of the derivations $B^\alpha_I \frac{\partial}{\partial u^\alpha_I}$ yield a non-zero result, since smooth functions only depend on derivatives up to a finite order.

We next show that one can lift any vector field $X \in \Gamma(U)$ on an open subset $U \subset M$ to a vector field $\mathrm{pr}^\infty X \in \Gamma(\mathrm{J}^\infty\pi)$ on jet space. Fix a point $\mathrm{j}^\infty(s)(p) \in \mathrm{J}^\infty\pi$ given by the prolongation of the local section $s \in \Gamma_\pi(U)$. To define a vector field on jet space, we have to determine a derivation on any germ of a smooth function $f \in C^\infty(\mathrm{J}^\infty\pi)$ defined in the neighborhood of this point.

To this end, note that the composition $f \circ \mathrm{j}^\infty s$ is a real-valued function on the open subset $U \subset M$. We can thus apply the derivation X_p given by the vector field X on M on it and obtain a derivation on smooth functions on the jet bundle $\mathrm{J}^\infty\pi$. Therefore, the following definition makes sense:

Definition 2.3.15
The prolongation of a (local) vector field $X \in \Gamma(U)$ is the (local) vector field $\mathrm{pr}^\infty X$ on $\mathrm{J}^\infty\pi$ acting on a smooth function $f \in C^\infty(\mathrm{J}^\infty\pi)$ in the point $\mathrm{j}^\infty s(p) \in \mathrm{J}^\infty\pi$ as

$$\mathrm{pr}^\infty X_{\mathrm{j}^\infty s(p)} f := X_p \left(f \circ \mathrm{j}^\infty s\right) \ .$$

Proposition 2.3.16.
Let (x, u) be local bundle coordinates for a bundle $\pi : E \to M$. Then

$$\mathrm{D}_j := \mathrm{pr}^\infty \left(\frac{\partial}{\partial x^j}\right) = \frac{\partial}{\partial x^j} + \sum_I u^\alpha_{I+(j)} \frac{\partial}{\partial u^\alpha_I}.$$

Proof:
We have to show that the derivations on the left hand side and on the right hand side act in the same way on the germ of any smooth function $f \in C^\infty(\mathrm{J}^\infty\pi)$. Using the definition and then the chain rule, we find

$$\left(\mathrm{pr}^\infty \frac{\partial}{\partial x^j}\right)_{\mathrm{j}^\infty s(p)} \cdot f(x^i) = \frac{\partial}{\partial x^j}\left(f \circ \mathrm{j}^\infty s\right)(x^i) = \frac{\partial f}{\partial x^j} + \sum_{\alpha,i} \frac{\partial f}{\partial u^\alpha_I} \frac{\partial^{|I|+1} s(x^i)}{\partial x_{|I|+(j)}}$$

which gives the expression on the right hand side. □

Definition 2.3.17
For a multi-index $I = (i_1, \ldots i_n)$, we introduce the following operator acting on local functions defined on a coordinate patch:

$$D_I := \ D_{i_1} \circ D_{i_2} \circ \cdots \circ D_{i_n}$$

A total differential operator is a mapping from $\mathrm{Loc}(E)$ to itself which can be written in local coordinates in the form $Z^I D_I$ where the sum goes over symmetric multi-indices I where $Z^I \in \mathrm{Loc}(E)$ is a local function.

Remark 2.3.18.
In the special case of a time-independent classical Lagrangian mechanical system, we encountered in proposition 2.1.2 the operator

$$\sum_{i,\beta} x^i_{\beta+1} \frac{\partial}{\partial x^i_\beta}$$

acting on local functions. Had we included also time dependent Lagrangians in our discussion, we would have naturally worked with the total differential operator

$$D = \mathrm{pr}^\infty \left(\frac{\mathrm{d}}{\mathrm{d}t}\right) = \sum_{i,\beta} x^i_{\beta+1} \frac{\partial}{\partial x^i_\beta}$$

Our next goal is to endow the jet bundle $\mathrm{J}^\infty\pi$ with a connection. To this end we need a sufficiently general notion of a connection.

Let $\pi : E \to M$ be a fibred manifold. Then $V = V(\pi) := \ker(\mathrm{d}\pi : TE \to TM)$ is vector subbundle of rank $\dim E - \dim M$ of TE over E, called the vertical bundle of π. The map

$$TE \overset{\mathrm{d}\pi}{\to} TM \ \to M$$

endows it with a natural submersion to M.

Definition 2.3.19

1. *An Ehresmann connection on the fibred manifold $\pi : E \to M$ is a smooth vector subbundle H of the tangent bundle TE over E such that*

   $$TE = H \oplus V$$

   where the direct sum of vector bundles over E is defined fibrewise.

2. *The fibres of H are called the horizontal subspaces of the connection.*

3. *A vector field on E is called horizontal, if it takes its values in the horizontal subspaces.*

Remarks 2.3.20.

1. *The horizontal bundle H is a vector bundle of rank $\dim M$. The restriction $\mathrm{d}\pi\Big|_H : H \to TM$ is an isomorphism.*

2. *An equivalent definition of an Ehresmann connection is in terms of a TE-valued 1-form ν on E. The 1-form ν can be applied to a tangent vector of E to yield another tangent vector. It can thus be seen as an endomorphism of the vector bundle TE. To define an Ehresmann connection, this one-form should fulfill the following conditions:*

   - $\nu^2 = \nu$ *(all endomorphisms are idempotents)*
   - $\operatorname{im}\nu = V$

   Then $H := \ker\nu$ is the horizontal subbundle.

3. *Given an Ehresmann connection on a fibred manifold $\pi : E \to M$, we can define a covariant derivative for a smooth (local) section $s : M \to E$ and a vector field $X \in \mathrm{vect}M$ by*

   $$\nabla_X s : M \underset{X}{\to} TM \underset{Ts}{\to} TE \underset{K}{\to} H.$$

   In a fibred chart (x^i, u^α) on E, we can described this locally by a one-form Γ on M with values in TE:

   $$\Gamma = \mathrm{d}x^i \otimes \left(\frac{\partial}{\partial x^i} + \Gamma^\alpha_i \frac{\partial}{\partial u^\alpha} \right)$$

   so that

   $$\nabla_{\partial_i} = \frac{\partial}{\partial x^i} + \Gamma^\alpha_i \frac{\partial}{\partial u^\alpha}$$

   and the Γ^α_i are generalizations of Christoffel symbols.

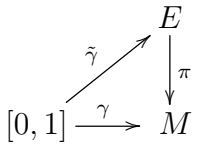
Given an Ehresmann connection, we can uniquely decompose any vector field $X \in \mathrm{vect}(E)$ on E in its horizontal part X_H and its vertical part X_V:

$$X = X_H + X_V \ .$$

We need a few properties of Ehresmann connections:

Remarks 2.3.21.
Let $\pi : E \to M$ be a smooth fibred manifold with an Ehresmann connection.

(a) Let $\gamma : [0,1] \to M$ be a smooth curve in M. A lift of γ is a smooth curve $\tilde{\gamma} : [0,1] \to E$ such that $\pi \circ \tilde{\gamma} = \gamma$. As a diagram:



(b) For an Ehresmann connection, we have an isomorphism of vector spaces between H_x and $T_{\pi x}M$ for any $x \in E$. Given a smooth curve γ in M, denote for any $x_t \in \pi^{-1}(\gamma(t))$ by

$$\tilde{V}(t) \in H_{x_t} \quad \textit{for all} \quad t \in [0,1]$$

the unique preimage of the derivative $\gamma'(t) \in T_{\gamma(t)}M$.
A lift $\tilde{\gamma} : [0,1] \to M$ is called horizontal, if we have for the derivative vector

$$\tilde{\gamma}'(t) = \tilde{V}(t) \quad \textit{for all} \quad t \in [0,1].$$

By standard theorems about ordinary differential equations, horizontal lifts locally exist for a given initial condition $\tilde{\gamma}(0) \in \pi^{-1}(\gamma(0))$. If global lifts exist for all curves, the connection is called complete.

(c) Given an Ehresmann connection, we define a 2-form R with on E with values in TE which is called the curvature of the Ehresmann connection. The 2-form is defined by its value on two vector fields $X, Y \in \mathrm{vect}(E)$ by taking the Lie bracket of their horizontal projections and measuring their derivation from being horizontal:

$$R(X,Y) = [X_H, Y_H]_V = [\Gamma(X), \Gamma(Y)] - \Gamma([\Gamma(X), \Gamma(Y)]) \ .$$

This is a vertical tangent valued 2-form which reads in coordinates

$$R = R^\alpha_{ij} \mathrm{d}x^i \wedge \mathrm{d}x^j \frac{\partial}{\partial u^\alpha} = \frac{1}{2}\left(\frac{\partial \Gamma^\alpha_j}{\partial x^i} + \Gamma^\beta_i \frac{\partial \Gamma^\alpha_j}{\partial u^\beta} - \frac{\partial \Gamma^\alpha_i}{\partial x^j} - \Gamma^\beta_j \frac{\partial \Gamma^\alpha_i}{\partial u^\beta} \right) \mathrm{d}x^i \wedge \mathrm{d}x^j \frac{\partial}{\partial u^\alpha}$$

(d) An Ehresmann connection is called flat, if the 2-form R vanishes. Flat connections are exactly Frobenius integrable connections.

We one can also show that the notion of Ehresmann connection comprises other notions of connection if we combine it with suitable compatibility conditions, e.g. linearity conditions in the case of vector bundles $\pi : E \to M$. In this case, one can in particular recover the ordinary covariant derivative of vector fields.

The crucial geometric structure in our setup is a natural Ehresmann connection on the infinite jet bundle $\mathrm{J}^\infty\pi \to M$.

Proposition 2.3.22.
Let $\pi : E \to M$ be a fibred manifold and $\mathrm{J}^\infty\pi$ be the corresponding jet bundle. For any $y \in \mathrm{J}^\infty\pi$ consider the subspace of $T_y\mathrm{J}^\infty\pi$ spanned by all prolongations $\mathrm{pr}^\infty X$ of tangent vectors $X \in T_{\pi_\infty y}M$. The following holds:

1. *The subspaces endow $\mathrm{J}^\infty\pi$ with an Ehresmann connection, the so-called Cartan connection.*
2. *One has*
$$\mathrm{pr}^\infty [X_1, X_2] = [\mathrm{pr}^\infty X_1, \mathrm{pr}^\infty X_2] \ ;$$
hence the Cartan connection is flat.

We have already seen that a Lagrangian for an m-dimensional field theory, $\dim M = m$, might be seen more naturally as an m-form on jet space. We therefore introduce differential forms on the infinite jet space $\mathrm{J}^\infty\pi$.

Definition 2.3.23
1. *Local differential forms on the jet bundle $\mathrm{J}^\infty\pi$ are defined as the inductive limit of the system*
$$\pi^*_{k+1,k} : \ \Omega^\bullet\left(\mathrm{J}^k\pi\right) \to \Omega^\bullet\left(J^{k+1}\pi\right)$$
i.e.
$$\Omega^\bullet\left(\mathrm{J}^\infty\pi\right) = \limind \Omega^\bullet\left(\mathrm{J}^k\pi\right).$$
A local p-form on $\mathrm{J}^\infty\pi$ is thus a finite *linear combination of expressions of the form*
$$A\, \mathrm{d}x^{i_1} \wedge ... \wedge \mathrm{d}x^{i_p} \wedge \mathrm{d}u^{\alpha_1}_{j_1 ... j_{p_q}} \wedge ... \wedge \mathrm{d}u^{\alpha_q}_{l_1 ... l_{p_q}}$$
with $p + q = k$ and with A a smooth function on $\mathrm{J}^\infty\pi$.
2. *A differential form $\omega \in \Omega^\bullet\left(\mathrm{J}^\infty\pi\right)$ is called a contact form, if*
$$\mathrm{j}^\infty(s)^*\omega = 0 \ \ \in \ \Omega^\bullet(M)$$
for all local sections $s \in \Gamma(E)$.

Remarks 2.3.24.
1. *The subspace of contact forms is a differential ideal of $\Omega^\bullet\left(\mathrm{J}^\infty\pi\right)$, the contact ideal $\mathfrak{I}$.*
2. *One can check that a vector field on the jet bundle $X \in \mathrm{Vect}\left(\mathrm{J}^\infty\pi\right)$ is horizontal, if and only if its contraction with all contact forms vanishes,*
$$\iota_X\omega = 0 \quad \textit{for all} \ \ \omega \in \mathfrak{I}$$
3. *Locally, the contact ideal $\mathfrak{I}$ is generated as a differential ideal by the one-forms*
$$\Theta^\alpha_I := \mathrm{d}u^\alpha_I - u^\alpha_{I+(j)}\mathrm{d}x^j \ ,$$
where I runs over all multi-indices. These forms are sometimes called Cartan forms. Here, it is essential that we do not impose any upper bound on the degree of the multi-index.

We convince ourselves that these differential forms are indeed contact forms: given a local smooth section $s : U \to E$, we consider its prolongation

$$f := \mathrm{j}^\infty s : U \to \mathrm{J}^\infty \pi$$

and pull back using the usual formula for a one-form $\omega_\alpha \mathrm{d}y^\alpha \in \Omega^\bullet(\mathrm{J}^\infty\pi)$ on $\mathrm{J}^\infty\pi$

$$f^*\omega = \omega_\alpha \frac{\partial f^\alpha}{\partial x^j} \mathrm{d}x^j \in \Omega^\bullet(U) \ .$$

This yields

$$(\mathrm{j}^\infty s)^* \Theta^\alpha_I = \frac{\partial (\mathrm{j}^\infty s)^\alpha_I}{\partial x^j} \mathrm{d}x^j - u^\alpha_{I+(j)}(\mathrm{j}^\infty s)\mathrm{d}x^j = 0 \ ,$$

where in the last step we used the definition of jet coordinates, $u^\alpha_{I+(j)}(\mathrm{j}^\infty s) = \frac{\partial^{|I|+1}}{\partial x^I \partial x^j} s$.

We now introduce a natural splitting on the complex of local differential forms on a jet space.

Definition 2.3.25

1. *In local bundle coordinates, any p-form $\omega \in \Omega^p(\mathrm{J}^\infty\pi)$ can be uniquely written as a sum of terms of the form*
$$f \ \mathrm{d}x^{i_1} \wedge ... \wedge \mathrm{d}x^{i_r} \wedge \Theta^{\alpha_1}_{I_1} \wedge \ldots \wedge \Theta^{\alpha_s}_{I_s}$$
with $r + s = p$ and with a smooth function f. We say that such a local differential form is of type (r, s).

2. *The subspace of local differential forms of type (r, s) is denoted by $\Omega^{r,s}(\mathrm{J}^\infty\pi)$. We decompose the space of local p-forms into a direct sum*
$$\Omega^p(\mathrm{J}^\infty) = \bigoplus_{r+s=p} \Omega^{r,s}(\mathrm{J}^\infty\pi) \ .$$

3. *Correspondingly, the exterior derivative*
$$\mathrm{d} : \Omega^p(\mathrm{J}^\infty\pi) \to \Omega^{p+1}(\mathrm{J}^\infty\pi)$$
decomposes into horizontal and vertical derivatives
$$\mathrm{d} = \mathrm{d}_H + \mathrm{d}_V$$
with
$$\mathrm{d}_H : \Omega^{r,s}(\mathrm{J}^\infty\pi) \to \Omega^{r+1,s}(\mathrm{J}^\infty\pi)$$
$$\mathrm{d}_V : \Omega^{r,s}(\mathrm{J}^\infty\pi) \to \Omega^{r,s+1}(\mathrm{J}^\infty\pi)$$
The relation $\mathrm{d}^2 = 0$, together with the bigrading, implies the three identities
$$\mathrm{d}_H^2 = 0, \ \ \mathrm{d}_V^2 = 0, \ \ \mathrm{d}_H\mathrm{d}_V = -\mathrm{d}_V\mathrm{d}_H.$$

Remarks 2.3.26.

1. *One can convince oneself that a p-form $\omega \in \Omega^p(\mathrm{J}^\infty\pi)$ is of type (r,s), if at all points $\sigma = \mathrm{j}^\infty s \in \mathrm{J}^\infty\pi$ one has*

$$\omega(X_1, ..., X_p) = 0,$$

as soon as either more than s tangent vectors are vertical or more than r tangent vectors are horizontal.

2. *The vertical and the horizontal differential can in particular be applied to smooth functions $f \in \mathrm{Loc}(E)$. For this situation, the following formula holds in local coordinates:*

$$\mathrm{d}_V f = \frac{\partial f}{\partial u^\alpha_I}\Theta^\alpha_I$$
$$\mathrm{d}_H f = (\mathrm{D}_i f)\,\mathrm{d}x^i = \left[\frac{\partial f}{\partial x^i} + u^\alpha_{I+(i)}\frac{\partial f}{\partial u^\alpha_I}\right]\mathrm{d}x^i$$

Moreover, one checks

$$\mathrm{d}_H(\mathrm{d}x^i) = 0 \quad \textit{and} \quad \mathrm{d}_V(\mathrm{d}x^i) = 0$$

as well as

$$\mathrm{d}_H\Theta^\alpha_I = \mathrm{d}x^j \wedge \Theta^\alpha_{I+(j)} \quad \textit{and} \quad \mathrm{d}_V\Theta^\alpha_I = 0.$$

Remarks 2.3.27.

1. *We have thus associated to a fibred manifold $\pi : E \to M$ the following double complex of differential forms on $\mathrm{J}^\infty\pi$:*

$$\begin{array}{ccccccccccccc}
 & & & & \uparrow & & \uparrow & & & & & & \uparrow \\
 & & 0 & \to & \Omega^{0,2} & \underset{\mathrm{d}_H}{\to} & \Omega^{1,2} & \to & ... & & & \underset{\mathrm{d}_H}{\to} & \Omega^{n,2} \\
 & & & & \uparrow \mathrm{d}_V & & \uparrow \mathrm{d}_V & & & & & & \uparrow \mathrm{d}_V \\
 & & 0 & \to & \Omega^{0,1} & \underset{\mathrm{d}_H}{\to} & \Omega^{1,1} & \to & ... & & & \underset{\mathrm{d}_H}{\to} & \Omega^{n,1} \\
 & & & & \uparrow \mathrm{d}_V & & \uparrow \mathrm{d}_V & & & & & & \uparrow \mathrm{d}_V \\
0 & \to & \mathbb{R} & \to & \Omega^{0,0} & \underset{\mathrm{d}_H}{\to} & \Omega^{1,0} & \underset{\mathrm{d}_H}{\to} & \Omega^{2,0} & \to & ... \to \Omega^{n-1,0} & \underset{\mathrm{d}_H}{\to} & \Omega^{n,0}
\end{array}$$

One should be careful: with this convention, the squares do not commute, but rather anticommute.

2. *We have already realized the forms in $\Omega^{n,0}$ as the candidates for Lagrangians. We will soon see how forms of type $\Omega^{n,1}$ are related to Euler Lagrange equations.*

3. *The algebraic Poincaré lemma states that on good open subsets of $\mathrm{J}^\infty\pi$, the lower line is exact at $\Omega^{p,0}$ for $p = 0, 1, \ldots n-1$. This statement can be generalized to $\Omega^{p,s}$ for $p = 0, 1, \ldots n-1$. For $\Omega^{n,s}$, this statement does not even hold locally.*

2.4 Lagrangian dynamics

We now turn to variational problems and described our setup:

Definition 2.4.1

1. *A Lagrangian system of dimension m consists of a smooth fibred manifold $\pi : E \to M$ with M a smooth m-dimensional manifold, together with a smooth differential form $l \in \Omega^{n,0}\left(\mathrm{J}^\infty\pi\right)$, where $n := \dim M$.*

2. *The manifold E is called the (extended) configuration space of the system.*

   The differential form l is called the Lagrangian density of the system.

3. *Suppose that $\mathrm{vol}_M \in \Omega^n(M)$ is a volume form on M and that $L \in \mathrm{Loc}(E)$ is a local function. Then*
$$l := L \cdot \pi_\infty^*(\mathrm{vol}_M) \in \Omega^{n,0}\left(\mathrm{J}^\infty\pi\right)$$
   is a Lagrangian density and the local function L is called the Lagrange function of the system.

4. *A Lagrangian system is called mechanical, if M is an interval $I \subset \mathbb{R}$.*

In local coordinates, a Lagrangian density looks like

$$l = l\left(x^i, u_I^\alpha\right) \mathrm{d}x^1 \wedge ... \wedge \mathrm{d}x^n \in \Omega^{n,0}\left(\mathrm{J}^\infty\pi\right) \ .$$

Given any local section $s : U \to E$, we can pull back l to an n-form $(\mathrm{j}^\infty s)^* l$ on $U \subset M$ that depends on the section s and its derivatives. This n-form can be integrated over U.

Definition 2.4.2
Let $\pi : E \to M$ be a fibred manifold and l a Lagrangian density. A (local) extremal of the variational problem defined by the Lagrangian density $l \in \Omega^{n,0}(\mathrm{J}^\infty\pi)$ is a (local) section $s : U \to E$, with $U \subset M$ open, satisfying

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0} \int_C (\mathrm{j}^\infty s_\epsilon)^* l = 0$$

whenever $C \subset U$ is a compact submanifold and whenever s_ϵ is a one-parameter family of local sections of π satisfying $s_0 = s$ and $s_\epsilon\Big|_{\partial C} = s\Big|_{\partial C}$.

Local extremals for a Lagrangian density are also called "on shell configurations" in the case of classical field theories or motions in the case of a mechanical system.

In local coordinates, we consider

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0} l\left(\mathrm{j}^\infty s_\epsilon\right) = \frac{\partial l}{\partial x^i}\Big|_{\mathrm{j}^\infty s} + \sum_{\alpha,I} \frac{\partial l}{\partial u_I^\alpha}\Big|_{\mathrm{j}^\infty s} \cdot \frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0} u_I^\alpha(s_\epsilon)$$

This expression suggests to write the variation as a pairing between expressions involving l and derivatives of the family s_ϵ with respect to ϵ. For the first aspect, recall the expression

$$\mathrm{d}_V l = \frac{\partial l}{\partial u_I^\alpha} \Theta_I^\alpha$$

which is a vertical n-form on $\mathrm{J}^\infty\pi$.

It is natural to contract this vertical n-form on jet space with a vertical vector field on jet space $\mathrm{J}^\infty E$. We first obtain a vector on E. Therefore, we assume that our variational family of

sections $s_\epsilon : I \times U \to E$ is generated by a vertical vector field X on E in the following sense: if ψ_ϵ is the local flow on E generated by the vector field X, then

$$s_\epsilon : M \overset{s_0}{\to} E \overset{\psi_\epsilon}{\to} E$$

is the variational family. The boundary condition then translates to $X\Big|_{\pi^{-1}\partial C} = 0$.

In local coordinates on E, this "infinitesimal version" of the variation reads

$$X := \left[\frac{\mathrm{d}s^\alpha_\epsilon}{\mathrm{d}\epsilon}\Big|_{\epsilon=0}\right]\frac{\partial}{\partial u^\alpha} \ .$$

To be able to contract with the vector field $\mathrm{d}_V l$ on $\mathrm{J}^\infty E$, we have to go from vector fields on E to vector fields on $\mathrm{J}^\infty\pi$.

Definition 2.4.3
Let X be a vector field on E. Then there is a unique vector field on $J^\infty(E)$, also called the prolongation of X and denoted by $\mathrm{pr}_E(X)$, such that

1. *X and $\mathrm{pr}_E(X)$ agree on functions on E.*
2. *The Lie derivative of $\mathrm{pr}_E(X)$ preserves the contact ideal: $L_{\mathrm{pr}_E(X)}C(\mathrm{J}^\infty\pi) \subset C(\mathrm{J}^\infty\pi)$.*

Remark 2.4.4.

1. *We present expressions in local adapted coordinates. Consider a vector field on E,*

$$X = a^i\frac{\partial}{\partial x^i} + b^\alpha\frac{\partial}{\partial u^\alpha} \ .$$

With hindsight, we generalize the situation and allow for a^i and b^α not only functions on E, but even functions in $\mathrm{Loc}(E)$. Such vector fields are also called generalized vector fields.

2. *We then write the prolongation as*

$$\mathrm{pr}_E(X) = Z^i\frac{\partial}{\partial x^i} + Z^\alpha_I\frac{\partial}{\partial u^\alpha_I} \ .$$

The first condition in the definition immediately yields

$$Z^i = a^i \quad \text{and} \quad Z^\alpha = b^\alpha$$

One can determine the coefficients to be

$$Z^\alpha_I = D_I(b^\alpha - u^\alpha_j) + u^\alpha_{jI}a^j$$

Definition 2.4.5
Given a Lagrangian density

$$l = l\left(x^i, u^\alpha_I\right)\mathrm{d}x^1 \wedge ... \wedge \mathrm{d}x^n \in \Omega^{n,0}\left(\mathrm{J}^\infty\pi\right) \ ,$$

the Euler-Lagrange form $E(l) \in \Omega^{n,1}\left(\mathrm{J}^\infty\pi\right)$ *is defined as*

$$E(l) = E_\alpha(l)\Theta^\alpha \wedge \mathrm{d}x^1 \wedge ... \wedge \mathrm{d}x^n$$

with

$$E_\alpha(l) = \frac{\partial l}{\partial u^\alpha} - \mathrm{D}_i \frac{\partial l}{\partial u^\alpha_i} + \mathrm{D}_{ij}\frac{\partial l}{\partial u^\alpha_{ij}} + - \ldots = (-D)_K \frac{\partial l}{\partial u^\alpha_K} \ .$$

Let M be a smooth manifold of dimension n and $\pi : E \to M$ a fibred manifold. Because of the Euler-Lagrange form, we are interested in forms in $\Omega^{n,1}\left(\mathrm{J}^\infty\pi\right)$. Forms on $\mathrm{J}^\infty\pi$ of type (n, s) with $s \geq 1$ are d_H-closed, but turn out to be not even locally d_H-exact.

To improve the situation, we introduce for $s \geq 1$ coaugmentations, called the inner Euler operators

$$I : \ \Omega^{n,s}\left(\mathrm{J}^\infty\pi\right) \to \Omega^{n,s}\left(\mathrm{J}^\infty\pi\right)$$

by taking the contractions

$$I(\omega) := \frac{1}{s}\Theta^\alpha \wedge \sum_I (-1)^{|I|} \iota_{\frac{\partial}{\partial u^\alpha_I}} \omega.$$

Lemma 2.4.6.
We have the following identities for the inner Euler operators:

1. *For* $\eta \in \Omega^{n-1,s}\mathrm{J}^\infty\pi J$, *we have*
$$I\left(\mathrm{d}_H\eta\right) = 0 \ .$$

2. *For any* $\omega \in \Omega^{n,s}\left(\mathrm{J}^\infty\pi\right)$ *there exists* $\eta \in \Omega^{n-1}\left(\mathrm{J}^\infty\pi\right)$ *such that*
$$\omega = I(\omega) + \mathrm{d}_H\eta.$$

3. I *is an idempotent,* $I^2 = I$.

4. *For any Lagrangian density* $\lambda \in \Omega^{n,0}\left(\mathrm{J}^\infty\pi\right)$, *we have*
$$E(\lambda) = I\left(\mathrm{d}_V\lambda\right)$$

This suggests to extend the bicomplex. We introduce the following subspace of $\Omega^{n,s}(\mathrm{J}^\infty\pi)$:

$$\mathcal{F}^s\left(\mathrm{J}^\infty\pi\right) := \mathrm{im}\left\{I : \Omega^{n,s} \to \Omega^{n,s}\right\} = \left\{\omega \in \Omega^{n,s}\left(\mathrm{J}^\infty\pi\right) : I\omega = \omega\right\}$$

The elements of $\mathcal{F}^s\left(\mathrm{J}^\infty\pi\right)$ are called source forms.

Moreover, we introduce maps

$$\begin{aligned} \delta_V : \ \mathcal{F}\left(\mathrm{J}^\infty\pi\right) &\to \mathcal{F}^{s+1}\left(\mathrm{J}^\infty\pi\right) \\ \delta_V(\omega) &= I(\mathrm{d}_V\omega) \end{aligned}$$

One can check that the following identity $\delta_V^2 = 0$ holds.

Definition 2.4.7

1. *The augmented variational bicomplex for a fibred manifold $\pi : E \to M$ is the double complex*

$$\begin{array}{ccccccccccccccc}
 & & 0 & \to & \Omega^{0,3} & & & & & & & & & & \\
 & & & & \uparrow & & \uparrow & & & & \uparrow & & \uparrow & & \\
 & & 0 & \to & \Omega^{0,2} & \to & \Omega^{1,2} & \to & \dots & \to & \Omega^{n,2} & \underset{I}{\to} & \mathcal{F}^2 & \to & 0 \\
 & & & & \uparrow & & \uparrow & & & & \uparrow & & \uparrow \delta_V & & \\
 & & 0 & \to & \Omega^{0,1} & \to & \Omega^{1,1} & \to & \dots \quad \Omega^{n-1,1} & \to & \Omega^{n,1} & \underset{I}{\to} & \mathcal{F}^1 & \to & 0 \\
 & & & & \uparrow \mathrm{d}_V & & \uparrow & & \uparrow & & \uparrow & \nearrow_E & & & \\
0 & \to & \mathbb{R} & \to & \Omega^{0,0} & \underset{\mathrm{d}_H}{\to} & \Omega^{1,0} & \to & \dots \to \Omega^{n-1,0} & \to & \Omega^{n,0} & & & &
\end{array}$$

2. *The Euler-Lagrange complex is the edge complex*

$$0 \to \mathbb{R} \to \Omega^{0,0} \xrightarrow{\mathrm{d}_H} \Omega^{1,0} \to \dots \xrightarrow{\mathrm{d}_H} \Omega^{n,0} \xrightarrow{E} \mathcal{F}^1 \xrightarrow{\delta_V} \mathcal{F}^2 \to \dots$$

where δ_V is called the Helmholtz operator.

Remarks 2.4.8.

1. *We recognize $\Omega^{n,0}$ as the recipient for the Lagrangian densities and $\mathcal{F}^1$ as the recipient for the Euler Lagrange equations, i.e. the equations of motion.*

2. *Lagrangians in the kernel of E have trivial equations of motion and are called trivial Lagrangians. If the Euler-Lagrange complex is exact at $\Omega^{n,0}$, this means that trivial Lagrangians are of the form d_H of an $n-1$-form.*

3. *If the Euler-Lagrange complex is exact at $\mathcal{F}^1$, then the kernel of δ_V describes all Euler-Lagrange operators. We can find a corresponding Lagrange density thus for all Langrange operators in the kernel of the Helmholtz operator.*

There are exactness results on the Euler-Lagrange complex in the literature. We quote some of them without proof.

Proposition 2.4.9.

1. *Consider $\pi : E = \mathbb{R}^{n+m} \to \mathbb{R}^n$ to be the trivial bundle. Then the rows and columns of the augmented variational bicomplex are exact. The Euler- Lagrange complex (2.14) is also exact. This can be shown by constructing appropriate homotopy operators.*

2. *For a* vector *bundle $\pi : E \to M$, the Euler-Lagrange sequence is not exact any more. Still, Vinogradov's one-line theorem tells us that in degree less or equal to $n-1$, the cohomology of the Euler-Lagrange complex is equal to the de Rham cohomology of M.*

Observation 2.4.10.

1. *Let $l \in \Omega^{n,0}(\mathrm{J}^\infty\pi)$ be any Lagrangian. We claim that then there always exists $\eta \in \Omega^{n-1,1}(\mathrm{J}^\infty\pi)$ such that*

$$\mathrm{d}_V l = E(l) + \mathrm{d}_H(\eta) \ .$$

This can be seen as follows: by definition of the interior Euler operator, we have $I(\mathrm{d}_V l) = E(l)$. Since I is an idempotent, it follows that $I(E(l)) = E(l)$ and hence $I(\mathrm{d}_V\lambda - E(l)) = 0$. From the exactness of the rows of augmented bicomplex, applied to the row with $s = 1$, we deduce that there is a form η with the desired properties.

2. *Using a Cartan-like calculus for differential forms on jet space, one can use this to show that, if X is a vertical vector field on E, then there exists a form$\sigma \in \Omega^{n-1,0}$ such that*

$$L_{\mathrm{pr}_E(X)} l = \iota_{\mathrm{pr}X}(E(l)) + \mathrm{d}_H \sigma \ .$$

This is a global first variational formula for any variational problem on E. We find altogether

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0} l\,(\mathrm{j}^\infty s_\epsilon) = L_{\mathrm{pr}_E(X)} l = \iota_{\mathrm{pr}X}(E(l)) + \mathrm{d}_H \sigma \ .$$

2.5 Symmetries and Noether identities

The subject of this subsection are Noether's two theorems. The first theorem establishes a relation between symmetries of the Lagrangian and conserved quantities. The second theorem shows that gauge symmetries yields dependencies between the equations of motion. We will give a precise mathematical definition of all these notions in this subsection.

We start to describe symmetries of the Cartan H distribution on jet space $\mathrm{J}^\infty\pi$. In a first step, we do not take any Lagrangian into account. We introduce the notion vect_H for vector fields on jet space $\mathrm{J}^\infty\pi$ taking their values in the horizontal subbundle of the tangent space to $\mathrm{J}^\infty\pi$. The integrability of the Cartan connection implies that this is a Lie subalgebra.

Definition 2.5.1

1. *Let $\pi : E \to M$ be a fibred manifold. A vector field X on jet space $\mathrm{J}^\infty\pi$ is called a symmetry of the fibred manifold, if $[X, Z] \in H$ for all $Z \in \mathrm{vect}_H$.*

2. *Due to the integrability of the Cartan distribution H, all vector fields in vect_H are symmetries. The Lie algebra of vector fields vect_H is, by definition of $\mathrm{vect}_{\mathrm{sym}}(\pi)$, an ideal in the Lie algebra $\mathrm{vect}_{\mathrm{sym}}(\pi)$ of all symmetries. We introduce the Lie algebra of non-trivial symmetries as the quotient*

$$\mathrm{sym}(\pi) := \mathrm{vect}_{\mathrm{sym}}(\pi)/\ \mathrm{vect}_H \ .$$

Using the split of vector fields into horizontal and vertical vector fields, we can restrict to vertical *vector fields on $\mathrm{J}^\infty\pi$ for the description of symmetries.*

We will need a different description of symmetries.

Observation 2.5.2.

1. *A vector field $X \in \mathrm{vect}_{\mathrm{sym}}$ acts as a derivation on the filtered algebra of local functions $\mathrm{Loc}(E)$ and yields a local function. The algebra $C^\infty(E)$ of smooth functions on E is a subalgebra of the algebra $\mathrm{Loc}(E)$. We can thus restrict the action of any vector field $X \in \mathrm{vect}_{\mathrm{sym}}$ to the subalgebra $C^\infty(E)$. We get a derivation φ_X on $C^\infty(E)$ which takes its values in $\mathrm{Loc}(E)$.*

2. *Since we assumed that vector fields in $\mathrm{vect}_{\mathrm{sym}}$ are vertical, we can write φ_X in local coordinates as*

$$Q^\alpha \frac{\partial}{\partial u^\alpha} \qquad \text{with } Q^\alpha \in \mathrm{Loc}(E) \ .$$

We can think of this vector field as an "infinitesimal variation depending on the fields and their derivatives".

3. *We can thus informally see φ_X as a vertical vector field on E with coefficients in local functions. More precisely, the differential operator φ_X is a section $\mathrm{J}^\infty\pi \to \pi^*(\pi)$ of the pullback bundle*

$$\begin{array}{ccc} \pi^*(\pi) & \longrightarrow & E \\ \downarrow & & \downarrow{\scriptstyle \pi} \\ \mathrm{J}^\infty\pi & \xrightarrow[\pi_\infty]{} & M \end{array}$$

We call $\kappa(\pi)$ the space of smooth sections of the bundle $\pi^*(\pi) \to \mathrm{J}^\infty\pi$ on jet space.

4. *One can show that the map*

$$X \mapsto \varphi_X$$

factors through vector fields with values in H and defines a bijection between the Lie algebra of symmetries of the fibred manifold $\mathrm{sym}(\pi)$ and the space of sections $\kappa(\pi)$.

Definition 2.5.3
An element φ_X is called a generating section of a symmetry X or also an evolutionary vector field.

Remarks 2.5.4.
1. *We denote the symmetry of the fibred manifold corresponding to a section $\varphi \in \kappa(\pi)$ by E_φ. Some authors reserve the term evolutionary vector field for this symmetry.*

2. *If $\varphi \in \kappa(\pi)$ is described in local coordinates as*

$$\varphi = \sum_{i=\alpha}^{m} \varphi^\alpha \frac{\partial}{\partial u^\alpha}$$

with local functions $(\varphi^i, \ldots, \varphi^m)$, then E_φ is its prolongation pr_E

$$E_\varphi = \sum_{I,\alpha} D_I(\varphi^\alpha) \frac{\partial}{\partial u^\alpha_I} \ .$$

The first goal of this subsection is to explore symmetries of Lagrangian systems $(\pi : E \to M, l)$.

Observation 2.5.5.
1. *Suppose a Lie group G acts by automorphisms of fibred manifold $E \longrightarrow M$, i.e. for each $g \in G$ there is a diffeomorphism g_E of E covering a diffeomorphism g_M of M:*

$$\begin{array}{ccc} E & \xrightarrow{g_E} & E \\ {\scriptstyle \pi}\downarrow & & \downarrow{\scriptstyle \pi} \\ M & \xrightarrow[g_M]{} & M \end{array}$$

which depends smoothly on $g \in G$. This contains as a subset diffeomorphisms g_E covering the identity on M, i.e. for which $g_M = \mathrm{id}_M$.

It is helpful not to impose this restriction: in the case of mechanical systems, we might want to consider translations by time: the field theories, translations by space and time which are all non-trivial diffeomorphisms on M.

2. *If $s : M \to E$ is a section of π, then $g.s$ with*

$$(g.s)(m) := (g_E).(s(g_M^{-1}m))$$

is another section of π. Indeed, we compute

$$\pi(g.s)(s) = \pi\left((g_E).(s(g_M^{-1}m))\right) = g_M \circ \pi \circ s \circ g_M^{-1}(m) = g_M \circ g_M^{-1}(m) = m \ .$$

We assume that the group action of G leaves the action S of a Lagrangian $l : \mathrm{J}^\infty\pi \longrightarrow \mathbb{R}$ invariant in the sense that $S(g.s) = S(s)$ for all sections s and all group elements $g \in G$.

3. *As an example, consider a natural mechanical system on $E = I \times \mathbb{R}^N$, with Lagrangian $l(x, x_t) = \frac{1}{2}|x_t|^2 - V(x)$. Consider the (time-independent) action of the Lie group $\mathbb{R}$ given by translations by a scalar multiple of a fixed vector $T \in \mathbb{R}^N \setminus \{0\}$, i.e.*

$$\begin{array}{rcl} T_T : \mathbb{R} \times \mathbb{R}^N & \to & \mathbb{R}^N \\ (t, x) & \mapsto & x + tT \ , \end{array}$$

Suppose that potential V is invariant under this family of translations, i.e. $T^i \frac{\partial V}{\partial x^i} = 0$.

4. *The group action of G on $\pi : E \to M$ induces, for each element η of the Lie algebra of G, vectors*

$$\tilde{\eta}_e := \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \exp(t\eta)e \in T_eE$$

which compose into a vector field on E. It can be thought of as the infinitesimal version of the group action.

In our example, it suffices to consider a single generator of $\mathbb{R}$ and thus a single vector field. Explicitly, we have the constant vector field $(0, T)$ representing a "constant infinitesimal translation".

5. *We have to translate the invariance of the action S for a Lagrangian density into a property of the Lagrangian density l which is a $(n, 0)$-form on $\mathrm{J}^\infty\pi$. The prolongation $\mathrm{pr}_E(\tilde{\eta})$ of the vector field $\tilde{\eta}$ on E is a vector field on $\mathrm{J}^\infty\pi$. The jet prolongation $\mathrm{j}^\infty s$ of a section $s : M \to E$ is a section of $\mathrm{J}^\infty s$; the infinitesimal transformation of $\mathrm{j}^\infty s$ under the action of $\eta \in \mathrm{Lie}(G)$ is then given by the vector field $\mathrm{pr}_E(\tilde{\eta})$.*

Invariance of the Lagrangian then means that the Lie derivative of the Lagrangian density l in the direction of the vector field $\mathrm{pr}_E\tilde{\eta}$ is a total derivative, i.e. that the n-form l on jet space

$$L_{\mathrm{pr}_E\tilde{\eta}}l$$

is d_H exact.

If we have a system with a volume form whose dynamics is described by a Lagrangian function L on jet space, this translates into the condition that the derivative $\mathrm{pr}_E\tilde{\eta}(L)$ is a d_H-divergence.

We only need infinitesimal aspects of the group action and hence formulate symmetries in terms of vector fields. We introduce the following notions:

Definition 2.5.6

Let $(\pi : E \to M, l)$ be a Lagrangian system. An evolutionary vector field Q on E is called a variational symmetry of a Lagrangian l, if it has the property that the Lie derivative

$$L_{\mathrm{pr}_E \tilde{\eta}} l$$

of the Lagrangian density is d_H exact.

Remark 2.5.7.
Suppose that there are volume forms given and we work with a Lagrangian function L. We can then formulate a variational family for the Lagrangian function as follows: this is a pair (Q, j^K), consisting of an evolutionary vector field Q and m local functions j^i, with $i = 1, \ldots m$,

$$\mathrm{pr}_E(Q_E)(l) = D_K(Q_E^\alpha)\frac{\partial l}{\partial u_K^\alpha} = D_i j^i \ ,$$

where D_i is the total derivative in the direction of x^i.

We also need the definition of a conserved quantity.

Definition 2.5.8
Let $(\pi : E \to M, l)$ be a Lagrangian system. A $m-1$-form $\alpha \in \Omega^{m-1,0}(\mathrm{J}^\infty\pi)$ is called a conserved quantity for the Lagrangian system, if for every solution $s : M \to E$ of the equations of motion given by l the $m-1$-form $(\mathrm{j}^\infty s)^*\alpha \in \Omega^{n-1}(M)$ is closed.

Remarks 2.5.9.

1. *In the case of a mechanical system over an interval $I = [t_0, t_1]$, we have $m = 1$ and for every solution of the equations of motions a function $\alpha_s := \alpha \circ \mathrm{j}^\infty(t) = \alpha(t, s(t), s_t(t), \ldots)$ on the interval I such that $\frac{\mathrm{d}}{\mathrm{d}t}\alpha_s = 0$. This implies*

$$\alpha_s(t_1) - \alpha_s(t_0) \ = \int_{t_0}^{t_1} \mathrm{d}t \frac{\mathrm{d}}{\mathrm{d}t}\alpha_s(t) = 0 \ ,$$

   which justifies the term "conserved quantity". Notice that for a given section s, the value of this conserved quantity can depend on the section and its derivatives.

2. *In the case of a field theoretical system, we obtain a $m-1$ form which, in case a Hodge star exists, can be identified with a 1-form*

$$\alpha_s = \sum_{i=1}^{m} (\alpha_s)_i \mathrm{d}x^i \ .$$

   One speaks of a conserved current. Here α_s are pullbacks of local functions along the solution s of the equations of motion. The fact that they are conserved means that $\sum_{i=1}^m \frac{\partial}{\partial x^i}(\alpha_s)_i = 0$.

3. *One speaks of a conserved "current". Let us explain this and consider the situation of a Galilei space of any dimension n. This is really a fibred manifold of affine spaces, $\mathbb{A}^n \to \mathbb{A}^1$. Like for any fibred manifold, we can split differential forms into a horizontal component and a vertical component.*

For any solution s of the equations of motion, we have a conserved $n-1$-form j_s on $\mathbb{A}^n$ which we write as

$$\mathrm{j}_s = \rho_s + \mathrm{d}t \wedge j_s$$

with $\rho_s \in \Omega^{0,n-1}(\mathbb{A}^n)$ and $j_s \in \Omega^{0,n-2}(\mathbb{A}^n)$. We then get the equation

$$0 = \mathrm{dj} = \mathrm{d}t \wedge \left(\frac{\partial \rho_s}{\partial t} + \mathrm{d}_V j_s \right) ,$$

where d_V is the vertical (i.e. here: spacial) differential. Fix a certain $n-1$-dimensional volume at fixed time. We then have

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_V \rho_s = \int_V \frac{\partial}{\partial t} \rho_s = - \int_V \mathrm{d}_V j_s = - \int_{\partial V} j_s ,$$

where in the last step we used Stokes' theorem. This has the following interpretation: for ever solution s of the equations of motion, we find a quantity ("charge") ρ_s that can be assigned to any spacial volume V and a "current" j_s whose flux across the boundary ∂V describes the loss or gain of this quantity. Together, they form the m components of the conserved $m-1$-form.

We need one mathematical preparation:

Definition 2.5.10
Given a total differential operator Z, we define its adjoint Z^+ as the total differential that obeys

$$\int_M (j^\infty s)^*(FZ(G))\mathrm{dvol}_M = \int_M (j^\infty s)^*(Z^+(F)G)\mathrm{dvol}_M$$

for all sections $s : M \to E$ and all local functions $F, G \in Loc_E$. It follows that

$$FZ(G)\mathrm{dvol}_M = Z^+(F)G\mathrm{dvol}_M + \mathrm{d}_H \zeta$$

for some $\zeta = \zeta(Z, F, G) \in \Omega^{n-1,0}(\mathrm{J}^\infty\pi)$ that depends on F and G and the precise form of the total differential operator Z.

If the total differential operator reads $Z = Z^J D_J$ in local coordinates, integration by parts yields the explicit formula

$$Z^+(F) = (-D)_J(Z^J F) .$$

Remark 2.5.11.
The divergence term ζ can look rather complicated and has to be worked out by doing the integrations explicitly. It has, however, a rather simple form for a first order total differential operator $Z = a^i D_i$ with local functions a^i. We have

$$\int_M F a^i D_i(G) \circ \mathrm{j}^\infty s = \int_M (-D_i)[a^i F] \cdot G \circ \mathrm{j}^\infty s + \int_M D_i[a^i F \cdot G] \circ \mathrm{j}^\infty s$$

We are now ready to formulate the relation between symmetries of a Lagrangian system and conserved quantities. We will choose to work with a Lagrange function and describe the conserved quantities in terms of a one-form.

Theorem 2.5.12 (Noether's first theorem)**.**

1. *Let $(\pi : E \to M, l)$ be a Lagrangian system with a Lagrange function l that, for simplicity does not explicitly depend on M, i.e. $\frac{\partial}{\partial x^i} L = 0$. Let (Q, j) be a variational symmetry in the sense of remark 2.5.7. Then there is a conserved one-form on jet space which can be worked out by doing repeated integrations by parts.*

2. *If the Lagrangian depends only on first order derivatives, the conserved one-form on jet space reads explicitly*
$$(Q^\alpha \frac{\partial l}{\partial u^\alpha_i} - j_i) \mathrm{d}x^i \ .$$

Proof:

1. Let s be a solution of the equations of motion. Since Q is a variational symmetry, we have
$$0 = \int_M \left(D_i j^i\right) \circ \mathrm{j}^\infty s = \int_M \left(\frac{\partial l}{\partial u^\alpha_I} (D_I Q^\alpha) \right) \circ \mathrm{j}^\infty s$$
Repeated integration by parts on M yields by the previous comment on adjoint operators and local functions ζ^i on jet space such that
$$0 = \int_M \left((-D)_I \frac{\partial l}{\partial u^\alpha_I} \right) Q^\alpha \circ \mathrm{j}^\infty s + \int_M D_i \zeta^i \circ \mathrm{j}^\infty s \ .$$
Since the Lagrangian does explicitly depend on M, the Euler-Lagrange equations for the section s take the form
$$\left((-D)_I \frac{\partial l}{\partial u^\alpha_I} \right) \circ \mathrm{j}^\infty s = 0 \ .$$
We thus learn that
$$0 = \int_M D_i (\zeta^i - j^i) = 0 \ .$$

2. If the Lagrangian depends only on first order derivatives, the divergence term is explicitly
$$\int_M D_i \left(\frac{\partial l}{\partial u^\alpha_i} Q^\alpha \right) \circ \mathrm{j}^\infty s$$
so that the conserved quantity is given by the one-form on jet space
$$\left[\frac{\partial l}{\partial u^\alpha_i} Q^\alpha - j^i \right] \mathrm{d}x^i \ .$$

□

Example 2.5.13.
Consider the natural system on $E = I \times \mathbb{R}^N$ with Lagrangian $l(x, x_t) = \frac{1}{2}|x_t|^2 - V(x)$ invariant

under translations by scalar multiples of a given vector $T \in \mathbb{R}^N$. There is a constant evolutionary vector field $(0, T)$ on E. It gives rise to the conserved quantity described by the function

$$T^i \frac{\partial l}{\partial x_t^i} = T^i (x_t)_i$$

on jet space. The conserved quantity called momentum in the direction T.

Consider a solution $s : I \to \mathbb{R}^N$ of the equations of motion,which are

$$\ddot{s} = -\mathrm{grad} V(s(t)) \ .$$

The conserved function on I is

$$\alpha_s = T^i (x_t)_i (\mathrm{j}^\infty s) = \ T^i \dot{s}_i \ .$$

We check that it is indeed conserved

$$\frac{\mathrm{d}}{\mathrm{d}t} \alpha_s(t) = \frac{\mathrm{d}}{\mathrm{d}t} \left(T^i \dot{s}_i(t) \right) \ = T^i \ddot{s}_i(t) = -T^i \mathrm{grad}_i V(s(t)) = 0 \ ,$$

where we used the equations of motion and then the fact that the potential V does not depend on the direction given by T.

Our next example illustrating the use of symmetries of a mechanical system is the rigid body:

Example 2.5.14.

- *A rigid body is a mechanical system of $N \geq 3$ mass points moving in $\mathbb{R}^3$ with the standard Euclidean structure, subject to the time-independent constraints*

$$\|x_i - x_j\| = r_{ij} = \ const. \quad for\ all\ 1 \leq i, j \leq N \ .$$

To avoid degenerate cases, we require that not all points are on a line.

We describe the configuration space.
To this end, we choose three mass points x_1, x_2, x_3 that are not on a line. Their position clearly determines the position of the body. We keep the position x_1 of the first mass point in our description.

Next consider the two difference vectors $v_1 := x_2 - x_1, \quad v_2 := x_3 - x_1$. These are two linearly independent vectors in $\mathbb{R}^3$. The scalar products of these two vectors, however, i.e. their lengths and the angle they include, are constant for a rigid body. For reference, we fix a pair (w_1, w_2) of vectors in $\mathbb{R}^3$ having the same scalar products as lengths and including the same angle:

$$\langle v_i, v_j \rangle = \langle w_i, w_j \rangle \quad for \ \ i, j = 1, 2 \ .$$

For each possible configuration of the rigid body with given v_1, v_2, there is a unique rotation $g(t) \in \mathrm{SO}(3)$ such that

$$v_1 = g w_1 \quad and \quad v_2 = w_2 \ .$$

With these choices, we therefore identify the configuration space of the rigid body with the semi-direct product manifold $(\mathbb{R}^3, +) \rtimes \mathrm{SO}(3)$.

- *We consider the system in the absence of external forces, i.e. we assume that it is described by the natural Lagrangian with vanishing potential:*

$$l = \sum_{i=1}^{N} \frac{1}{2}\, m_\alpha \langle x_{i,t}, x_{i,t} \rangle \quad \textit{with} \;\; m_i > 0$$

Any translation T of $\mathbb{R}^3$ gives a symmetry of the system. The varitional symmetry is given by the vertical constant vector field on $E = I \times \mathbb{R}^{3N}$ which is equal to T on each of the N components,

$$Q_E^{i,\alpha} = T^\alpha \ .$$

with $\alpha = 1, 2, 3$ and for all $i = 1, ..., N$. For the derivative of the action, we find

$$\frac{\partial l}{\partial x_t^{i\alpha}} = m_i x_t^{i\alpha}$$

Noether's theorem gives three conserved quantities for the three translations in $\mathbb{R}^3$: Hence the conserved quantity is

$$I = \sum_{i,\alpha} \frac{\partial l}{\partial x_t^{i\alpha}} Q_E^{i,\alpha} = \sum_{i,\alpha} m_\alpha x_t^{i,\alpha} T^i \ .$$

Since this applies to all translations $T \in \mathbb{R}^3$, we have a vector valued conserved quantity

$$I^i = \sum_{i=1}^{N} m_i x_t^{i,\alpha} \qquad \alpha = 1, ..., 3$$

corresponding to the three conserved components of total momentum.

Denote by $M := \sum_{i=1}^{N} m_i > 0$ the total mass of the body and introduce the center of mass

$$x_{cm} := \frac{1}{M} \sum_{i=1}^{N} m_i x^i$$

of the rigid body, then the conservation law implies

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} x_{cm} = 0.$$

In other words, the center of mass is in uniform motion.

- *This allows us to continue our discussion in a reference frame where the center of mass is at rest, $x_{cm} = \text{const}$.*

  The invariance under rotations gives, again by Noether's theorem three further conserved quantities; one more conserved quantity is energy, as follows from the time independence of the action.

  In the situation at hand, we can limit our discussion to first derivatives and thus to the phase space which is the tangent bundle $T\,\mathrm{SO}(3)$.

  The four conserved quantities determine for any given solution s of the equations of motion a submanifold C_s of the six-dimensional phase space $T\,\mathrm{SO}(3)$. Only points on C_s qualify as configurations for a later time. We list properties of this two-dimensional submanifold C:

– *Bounds on the tangent vectors from energy conservation imply that the submanifold C is a compact submanifold of the non-compact manifold $T\mathrm{SO}(3)$.*

– *Tangent bundles of Lie groups are parallelizable and thus in particular orientable. Any submanifold of an orientable manifold that is defined by global regular equations is orientable as well. Thus C is orientable.*

– *The derivative of the motion gives a vector field on C that for non-vanishing energy vanishes nowhere.*

A standard theorem of topology then implies that C is topologically a torus. One can choose coordinates $\varphi_1, \varphi_2 \in \mathbb{R}/2\pi\mathbb{Z}$ on C such that the equations of motion have the form

$$\dot{\varphi}_1 = \omega_1 \qquad \dot{\varphi}_2 = \omega_2 \ .$$

Obviously, the two angular frequencies ω_i depend on the initial conditions.

We now turn to Noether's second variational theorem which establishes a correspondence between gauge symmetries and differential algebraic relations among Euler Lagrange equations.

Noether's second theorem deals with gauge symmetries rather than just symmetries.

Definition 2.5.15
A gauge symmetry of a Lagrangian system $(\pi : E \to M, l)$ consists of a family of local functions, for $\alpha = 1, \dots n$ and all multi-indices I,

$$R^{\alpha I} : \mathrm{J}^\infty\pi \longrightarrow \mathbb{R}$$

for $\alpha = 1, \dots n$ and all multi-indices I, such that for any *local function*

$$\epsilon : \mathrm{J}^\infty\pi \longrightarrow \mathbb{R}$$

the evolutionary vector field $R^{\alpha I}(D_I\epsilon)\frac{\partial}{\partial u^\alpha}$ on E is a variational symmetry of l.

Remarks 2.5.16.

1. *Loosely speaking, a gauge symmetry is a linear mapping from local functions on $\mathrm{J}^\infty\pi$ into the evolutionary vector fields on E preserving the Lagrangian. It is crucial for a gauge symmetry that there is a symmetry for* every *local function.*

2. *Notice that the coefficients of the vector field depend linearly on ϵ and on all its total derivatives.*

3. *By the results just obtained, it follows that being a gauge symmetry is equivalent to requiring $(R^{\alpha I}D_I\epsilon)E_\alpha(l)$ to be a divergence for each local function ϵ on $J^\infty\pi$.*

4. *This in turn is equivalent to saying that $\epsilon(R^{\alpha I}D_I)^+(E_\alpha(l))$ is a divergence for each local function ϵ, where $(R^{\alpha I}D_I)^+$ is the adjoint of the differential operator $R^{\alpha I}D_I$.*

   Since the adjoint of a total differential operator is again a total differential operator, there exist local functions $R^{+\alpha I} : J^\infty E \longrightarrow \mathbb{R}$ such that $(R^{\alpha I}D_I)^+ = R^{+\alpha I}D_I$. These functions can be found by working out the iterated total derivatives $(-D)_I(R^{\alpha I}F)$. In many cases it is easier to use an "integration by parts" procedure to obtain the coefficient functions on $\mathrm{J}^\infty\pi$, $\{R^{+\alpha I}\}$.

5. *It follows that* $\epsilon \mapsto R^{\alpha I}(D_I\epsilon)\frac{\partial}{\partial u^\alpha}$ *defines a gauge symmetry iff* $\epsilon R^{+\alpha I}D_I(E_\alpha(l))$ *is a divergence for each* ϵ*. This condition can be shown to be equivalent to the fact that* $R^{+\alpha I}D_I(E_\alpha(l))$ *is identically zero on the jet bundle.*

We have thus found a one-one correspondence between gauge symmetries of a Lagrangian and differential dependencies between the Euler Lagrange equations, so-called Noether identities:

Theorem 2.5.17 (Noether's second theorem)**.**
For a given Lagrangian system $(\pi : E \to M, l)$ *and for local real-valued functions* $\{R^{\alpha I}\}$ *defined on* $\mathrm{J}^\infty\pi$*, the following statements are equivalent:*

1. *The functions* $\{R^{\alpha I}\}$ *define a gauge symmetry of* l*, i.e.,* $R^{\alpha I}(D_I\epsilon)\frac{\partial}{\partial u^\alpha}$ *is a variational symmetry of* l *for* any *local function* $\epsilon : \mathrm{J}^\infty\pi \longrightarrow \mathbb{R}$*.*

2. $R^{\alpha I}(D_I\epsilon)E_\alpha(l)$ *is a divergence for* any *local function* ϵ*,*

3. *The functions* $\{R^{\alpha I}\}$ *define Noether identities of* l*, i.e.* $R^{+\alpha I}D_I(E_\alpha(l))$ *is identically zero on the jet bundle.*

Gauge symmetries induce differential identities between the equations of motions, i.e. dependencies between the equations of motions and their derivatives. They can be thought of a differential versions of syzygies and are in fact the starting point for homological methods in field theory like the BV or BRST formalism.

2.6 Natural geometry

Disclaimer: this section is not ready to be presented in lectures and will be omitted

A comment on frames:
First order jets at zero of maps $\mathbb{R} \to N$ are tangent vectors. First order jets at zero of *non-degenerate* maps from $\mathbb{R}^n \to N$ with $n = \dim N$ are frames in N. Non-degenerate means that the Jacobian at zero is non-zero. The set of all frames is the frame bundle.

A natural differential operator is a recipe that constructs from a geometric object another one, in a natural fashion, and which is locally a function of coordinates and their derivatives. They are thus intimately related to jet bundles and have to be compatible with smooth maps between manifolds.

Examples 2.6.1.

1. *Let* M *be a* n*-dimensional smooth manifold. The classical* Lie bracket $X, Y \mapsto [X, Y]$ *is a natural operation that constructs from two vector fields on* M *a third one. Given a local coordinate system* $(x^1, \ldots, x^n)$ *on* M*, the vector fields* X *and* Y *are locally expressions* $X = \sum_{1\le i\le n} X^i\partial/\partial x^i$*,* $Y = \sum_{1\le i\le n} Y^i\partial/\partial x^i$*, where* X^i, Y^i *are smooth functions on* M*. If we define* $X^i_j := \partial X^i/\partial x^j$ *and* $Y^i_j := \partial Y^i/\partial x^j$*,* $1 \le i, j \le n$*, then the Lie bracket is locally given by the formula* $[X, Y] = \sum_{1\le i,j\le n}\left(X^jY^i_j - Y^jX^i_j\right)\partial/\partial x^i$*.*

2. *The* <u>covariant derivative</u> $(\Gamma, X, Y) \mapsto \nabla_X Y$ *is a natural operator that constructs from a linear connection* Γ *and vector fields* X *and* Y*, a vector field* $\nabla_X Y$*. In local coordinates,*

$$\nabla_X Y = \left(\Gamma^i_{jk}X^jY^k + X^jY^i_j\right)\frac{\partial}{\partial x^i}, \tag{8}$$

where Γ^i_{jk} *are Christoffel symbols.*

3. *Natural operations can be composed into more complicated ones. Examples of 'composed' operations are the* torsion $T(X,Y) := \nabla_X Y - \nabla_Y X - [X,Y]$ *and the* curvature $R(X,Y)Z := \nabla_{[X,Y]}Z - [\nabla_X, \nabla_Y]Z$ *of the linear connection* Γ.

4. *Let* X *be a vector field and* ω *a 1-form on* M. *Denote by* $\omega(X) \in C^\infty(M)$ *the evaluation of the form* ω *on* X. *Then* $(X,\omega) \mapsto \exp(\omega(X))$ *defines a natural differential operator with values in smooth functions. Clearly, the exponential can be replaced by an arbitrary smooth function* $\phi : \mathbb{R} \to \mathbb{R}$, *giving rise to a natural operator* $\mathfrak{O}_\phi(X,\omega) := \phi(\omega(X))$.

5. *'Randomly' generated local formulas need not lead to natural operators. As we will see later, neither* $O_1(X,Y) = X_3^1 Y^4 \partial/\partial x^2$ *nor* $O_2(X,Y) = X^j Y_j^i \partial/\partial x^i$ *behave properly under coordinate changes, so they do not give rise to vector-field valued natural operators.*

In the examples, the natural differential operators are recipes given as a smooth function in coordinates and derivatives that are covariant under changes of local coordinates.

Definition 2.6.2

1. *Denote by* $\mathtt{Man}_n$ *the category of* n*-dimensional manifolds and open embeddings. Let* $\mathtt{Fib}_n$ *be the category of smooth fiber bundles over* n*-dimensional manifolds with morphisms differentiable maps covering morphisms of their bases in* $\mathtt{Man}_n$.

2. *A* <u>natural bundle</u> *is a functor* $\mathfrak{B} : \mathtt{Man}_n \to \mathtt{Fib}_n$ *such that for each* $M \in \mathtt{Man}_n$, $\mathfrak{B}(M)$ *is a bundle over* M. *Moreover,* $\mathfrak{B}(M')$ *is the restriction of* $\mathfrak{B}(M)$ *for each open submanifold* $M' \subset M$, *the map* $\mathfrak{B}(M') \to \mathfrak{B}(M)$ *induced by* $M' \hookrightarrow M$ *being the inclusion* $\mathfrak{B}(M') \hookrightarrow \mathfrak{B}(M)$.

For each $s \geq 1$ we denote by $\mathrm{GL}_n^{(s)}$ the group of s-jets of local diffeomorphisms $\mathbb{R}^n \to \mathbb{R}^n$ at 0, so that $\mathrm{GL}_n^{(1)}$ is the ordinary general linear group GL_n of linear invertible maps $A : \mathbb{R}^n \to \mathbb{R}^n$. Let $\mathrm{Fr}^s(M)$ be the bundle of s-jets of frames on M whose fiber over $z \in M$ consist of s-jets of local diffeomorphisms of neighborhoods of $0 \in \mathbb{R}^n$ with neighborhoods of $z \in M$. It is clear that $\mathrm{Fr}^s(M)$ is a principal $\mathrm{GL}_n^{(s)}$-bundle and $\mathrm{Fr}^1(M)$ the ordinary GL_n-bundle of frames $\mathrm{Fr}(M)$.

Theorem 2.6.3 (Krupka, Palais, Terng)**.**
For each natural bundle $\mathfrak{B}$, *there exists* $l \geq 1$ *and a manifold* B *with a smooth* $\mathrm{GL}_n^{(l)}$*-action such that there is a functorial isomorphism*

$$\mathfrak{B}(M) \cong \mathrm{Fr}^l(M) \times_{\mathrm{GL}_n^{(l)}} B := (\mathrm{Fr}^l(M) \times B)/\mathrm{GL}_n^{(l)}. \tag{9}$$

Conversely, each smooth $\mathrm{GL}_n^{(l)}$-manifold B induces, a natural bundle $\mathfrak{B}$. We will call B the *fiber* of the natural bundle $\mathfrak{B}$. If the action of $\mathrm{GL}_n^{(l)}$ on B does not reduce to an action of the quotient $\mathrm{GL}_n^{(l-1)}$ we say that $\mathfrak{B}$ has *order* l.

Examples 2.6.4.

1. *Vector fields are sections of the* tangent bundle $T(M)$. *The fiber of this bundle is* $\mathbb{R}^n$, *with the standard action of* GL_n. *The description* $T(M) \cong \mathrm{Fr}(M) \times_{\mathrm{GL}_n} \mathbb{R}^n$ *is classical.*

2. *De Rham* m*-forms are sections of the bundle* $\Omega^m(M)$ *whose fiber is the space of anti-symmetric* m*-linear maps* $\mathrm{Lin}(^m(\mathbb{R}^n), \mathbb{R})$, *with the obvious induced* GL_n*-action. The presentation*

$$\Omega^m(M) \cong \mathrm{Fr}(M) \times_{\mathrm{GL}_n} \mathrm{Lin}(\Lambda^m(\mathbb{R}^n), \mathbb{R})$$

is also classical. A particular case is $\Omega^0(M) \cong \mathrm{Fr}(M) \times_{\mathrm{GL}_n} \mathbb{R} \cong M \times \mathbb{R}$*, the bundle whose sections are smooth functions. We will denote this natural bundle by* $\mathbb{R}$*, believing there will be no confusion with the symbol for the reals.*

3. *Linear connections are sections of the* bundle of connections $Con(M)$ *which we recall below. Let us first describe the group* $\mathrm{GL}_n^{(2)}$*. Its elements are expressions of the form* $A = A_1 + A_2$*, where* $A_1 : \mathbb{R}^n \to \mathbb{R}^n$ *is a linear invertible map and* A_2 *is a linear map from the symmetric product* $\mathbb{R}^n \odot \mathbb{R}^n$ *to* $\mathbb{R}^n$*. The multiplication in* $\mathrm{GL}_n^{(2)}$ *is given by*

$$(A_1 + A_2)(B_1 + B_2) := A_1(B_1) + A_1(B_2) + A_2(B_1, B_1).$$

The unit of $\mathrm{GL}_n^{(2)}$ *is* $id_{\mathbb{R}^n} + 0$ *and the inverse is given by the formula*

$$(A_1 + A_2)^{-1} = A_1^{-1} - A_1^{-1}(A_2(A_1^{-1}, A_1^{-1})).$$

Let $\mathcal{C}$ *be the space of linear maps* $\mathrm{Lin}(\mathbb{R}^n \otimes \mathbb{R}^n, \mathbb{R}^n)$*, with the left action of* $\mathrm{GL}_n^{(2)}$ *given as*

$$(Af)(u \otimes v) := A_1 f(A_1^{-1}(u), A_1^{-1}(v)) - A_2(A_1^{-1}(u), A_1^{-1}(v)), \tag{10}$$

for $f \in \mathrm{Lin}(\mathbb{R}^n \otimes \mathbb{R}^n, \mathbb{R}^n)$*,* $A = A_1 + A_2 \in \mathrm{GL}_n^{(2)}$ *and* $u, v \in \mathbb{R}^n$*. The bundle of connections is then the order* 2 *natural bundle represented as*

$$Con(M) := \mathrm{Fr}^2(M) \times_{\mathrm{GL}_n^{(2)}} \mathcal{C} \ .$$

Observe that, while the action of $\mathrm{GL}_n^{(2)}$ *on the vector space* C *is not linear, the restricted action of* $\mathrm{G}L_n \subset \mathrm{GL}_n^{(2)}$ *on* C *is the standard action of the general linear group on the space of bilinear maps.*

For $k \geq 0$ we denote by $\mathfrak{B}^{(k)}$ the bundle of k-jets of local sections of the natural bundle $\mathfrak{B}$ so that $\mathfrak{B}^{(0)} = \mathfrak{B}$. If $\mathfrak{g}$ is represented in this way, then

$$\mathfrak{B}^{(k)}(M) \cong \mathrm{Fr}^{(k+l)}(M) \times_{\mathrm{GL}_n^{(k+l)}} \mathfrak{B}^{(k)} \ ,$$

where $B^{(k)}$ is the space of k-jets of local diffeomorphisms $\mathbb{R}^n \to B$ defined in a neighborhood of $0 \in \mathbb{R}^n$.

Definition 2.6.5
Let $\mathfrak{F}$ *and* $\mathfrak{G}$ *be natural bundles. A (finite order)* natural differential operator $\mathfrak{O} : \mathfrak{F} \to \mathfrak{G}$ *is a natural transformation (denoted by the same symbol)* $\mathfrak{O} : \mathfrak{F}^{(k)} \to \mathfrak{G}$*, for some* $k \geq 1$*. We denote the space of all natural differential operators* $\mathfrak{F} \to \mathfrak{G}$ *by* $\mathfrak{Nat}(\mathfrak{F}, \mathfrak{G})$.

If $\mathfrak{F}$ and $\mathfrak{G}$ are natural bundles of order $\leq l$, with fibers F and G, respectively, then each natural operator in Definition is induced by an $\mathrm{GL}_n^{(k+l)}$-equivariant map $O : F^{(k)} \to G$, for some $k \geq 0$. Conversely, such an equivariant map induces an operator $\mathfrak{D} : \mathfrak{F} \to \mathfrak{G}$. This means that the study of natural operators is reduced to the study of equivariant maps. The procedure described above is therefore called the *IT reduction* (from invariant-theoretic).

Examples 2.6.6.

1. *Given natural bundles* $\mathfrak{B}'$ *and* $\mathfrak{B}''$ *with fibers* B' *resp.* B''*, there is an obviously defined natural bundle* $\mathfrak{B}' \times \mathfrak{B}''$ *with fiber* $B' \times B''$*. With this notation, the Lie bracket is a natural operator* $[-,-] : T \times T \to T$ *and the covariant derivative an operator* $\nabla : \mathrm{Con} \times T \times T \to T$*, where* T *is the tangent space functor and* Con *the bundle of connections. The corresponding equivariant maps of fibers can be easily read off from local formulas given in Examples.*

2. *The operator* $\mathfrak{D}_\phi : T \times \Omega^1 \to C^\infty$ *from the Example above is induced by the* $\mathrm{G}L_n$*-equivariant map* $O_\phi : \mathbb{R}^n \times (\mathbb{R}^n)^* \to \mathbb{R}$ *given by* $o_\phi(v, \alpha) := \phi(\alpha(v))$.

3 Classical field theories

3.1 Maxwell's equations

We start explaining Maxwell's equation on a Galilei space $\mathbb{A}$. We will see only later that this is not an appropriate conceptual setting, because the laws of electrodynamics are incompatible with the Galilean principle of relativity.

Galilei space is actually a very simple fibred manifold of affine spaces

$$t : \mathbb{A} \to \mathbb{A}^1$$

where t is essentially given by the global time difference function and $\mathbb{A}^1$ a one-dimensional affine space. The fibres $\mathbb{A}_\tau := t^{-1}(\tau)$ for $\tau \in \mathbb{A}^1$ are the hypersurfaces of simultaneous events or "time slices". By the axioms for Galilei space, each hypersurface $\mathbb{A}_t$ has the structure of a three-dimensional Euclidean space. We endow it with the structure of an oriented Euclidean space.

As an affine space over $\mathbb{R}^4$, the standard basis over $\mathbb{R}^4$ gives a canonical basis $\mathrm{d}t, \mathrm{d}x, \mathrm{d}y, \mathrm{d}z$ for the cotangent space T_p^*M of every point $p \in \mathbb{A}$. Consider the subspace $\Omega^{0,p}(\mathbb{A})$ of p-forms that vanish on ∂_t. Put differently, these so-called vertical forms are linear combinations of wedges of $\mathrm{d}x, \mathrm{d}y$ and $\mathrm{d}z$. Any p-form ω on $\mathbb{A}$ can be written uniquely as a sum

$$\omega = \mathrm{d}t \wedge \omega_1 + \omega_2$$

where $\omega_1 \in \Omega^{0,p-1}$ and $\omega_2 \in \Omega^{0,p}$. The complex of differential forms thus splits,

$$\Omega^p(\mathbb{A}) \ = \oplus_{r+s=p} \Omega^{r,s}(\mathbb{A}) \ .$$

We also note for later use that the vertical differential d_V with

$$\mathrm{d}_V \omega = \mathrm{d}x \partial_x \omega + \mathrm{d}y \partial_y \omega + \mathrm{d}z \partial_z \omega$$

is just the spacial differential.

The orientation of $\mathbb{A}_t$ and its Euclidean structure give us a three-dimensional Hodge operator

$$*_V : \ \Omega^{\cdot,p}(\mathbb{A}) \xrightarrow{\sim} \Omega^{\cdot,3-p}(\mathbb{A}) \ .$$

We base our discussion of electrodynamics on two conservation laws which condense a lot of empirical observations to which we will have to add a principle describing properties of matter (or the vacuum, seen here as the absence of matter).

Observation 3.1.1.

1. *The first new quantity is electric charge. It can be observed e.g. in processes like discharges. Charge is measured at fixed time, so for any measurable subset $U \subset \mathbb{A}_t$, we should be able to determine the electric charge by an integral over U. It is therefore natural to consider the charge density ρ at time t as a three-form on $\mathbb{A}_t$:*

$$\rho_t(x) \, \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3 \quad \in \Omega^3(\mathbb{A}_t)$$

The charge density is not constant in time and we will study its time dependence. We thus define charge density as a differential form $\rho \in \Omega^{0,3}(\mathbb{A})$. We are, deliberately, vague about smoothness properties of ρ since many important idealizations of charge distributions – point charges, charged wires or charged plates – involve singularities.

2. *If we take a volume $V \subset \mathbb{A}_t$ at fixed time with smooth boundary ∂V, then electric charge can pass through its boundary ∂V. The amount of charge passing per time should be described by the integral over a two form $j_t \in \Omega^2(\mathbb{A}_t)$. Again, this two-form will depend on time so that we introduce the current density as a 3-form*

$$j(x,t) = \mathrm{d}t \wedge j_t(x) \in \Omega^{1,2}(\mathbb{A}) \ .$$

We then have the natural conservation law for any closed volume $V \in \mathbb{A}_t$

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_V \rho = -\int_{\partial V} j_t$$

which by Stokes' theorem takes the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_V \rho = -\int_{\partial V} j_t = -\int_V \mathrm{d}_V j_t \ .$$

Since this holds for all volumes $V \subset \mathbb{A}_t$, we have the infinitesimal form of the conservation law

$$\frac{\partial}{\partial t}\rho_t + \mathrm{d}_V j_t = 0$$

which is an equality of three-forms on $\mathbb{A}_t$ for all t.

3. *We can write the conservation law more compactly in terms of the three-form, the charge-current density* j*:*

$$\mathrm{j} := \rho - j = \rho_t \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z - \mathrm{d}t \wedge j_t \in \Omega^3(\mathbb{A})$$

defined on the four-dimensional space $\mathbb{A}$. We find

$$\mathrm{dj} = \frac{\partial}{\partial t}\rho_t \mathrm{d}t \wedge \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z + \mathrm{d}t \wedge \mathrm{d}_V j_t = 0 \ .$$

4. *To make contact with standard literature, we make a side remark: Using the Hodge star on oriented Euclidean space $\mathbb{A}_t$, we can relate the two-form j_t on $\mathbb{A}_t$ to the 1-form $*_V j$ which in turn can be identified with a vector $\vec{j}$. Similarly, the three-form ρ can be related to the function $\rho_t(x)$. With appropriate conventions, we recover the conservation law in the classical form*

$$\frac{\partial \rho_t(x)}{\partial t} + \mathrm{div}\vec{j} = 0 \ .$$

5. *Since the charge-current density 3-form is closed,* $\mathrm{dj} = 0$*, it is exact by the Poincaré lemma. We can find a two-form* $\mathrm{H} \in \Omega^2(\mathbb{A})$*, the excitation 2-form such that*

$$\mathrm{dH} \ = \mathrm{j} \ .$$

The fact that such a 2-form exists even globally *is the content of the inhomogeneous Maxwell equation.*

6. *Using the bigrading of 2-forms on $\mathbb{A}$, we can write*

$$\mathrm{H} = \mathrm{d}t \wedge H + D$$

with $H \in \Omega^{0,1}(\mathbb{A})$ and $D \in \Omega^{0,2}(\mathbb{A})$. One calls H the magnetic "excitation" and D the "electric excitation". Using the three-dimensional metric, the field H can be identified with a vector field. For the field D, we first need to apply a three-dimensional Hodge star to get a one-form which then, in turn, can be identified with a vector field. The Hodge star and thus the vector field depend on the orientation chosen on three-dimensional space. For this reason, D is sometimes called a pseudo vector field.

Then the inhomogeneous Maxwell-equations

$$\mathrm{j} = \rho - \mathrm{d}t \wedge j = \mathrm{dH} = -\mathrm{d}t \wedge \mathrm{d}_V H + \mathrm{d}t \wedge \partial_t D + \mathrm{d}_V\ D$$

are equivalent to

$$\mathrm{d}_V D\ =\ \rho \quad \text{and} \quad \partial_t D\ =\ \mathrm{d}_V H - j\ ,$$

where the vertical derivative d_V is just the spacial exterior derivative.

The first equation is the Coulomb-Gauss law, the second equation the Oersted-Ampère equation. The term containing the time derivative $-\partial_t\ D$ of the excitation is sometimes called Maxwell's term.

We can see the Gauss-equation as a constraint equation; the Ampère equation are then three time evolution equations for the excitation D.

Example 3.1.2.
Gauß' law reads in integral form for a volume V in space

$$Q = \int_V \rho\ \mathrm{d}^3x = \int_{\partial V} D\ .$$

The electric flux through the surface ∂V is thus proportional to the electric charge included by the surface. It should be appreciated that this holds even for time dependent electric fields.

We use Gauß' law to determine the electric field of a static point charge in the origin:

$$\rho(\vec{x}, t) = q\delta_0(\vec{x})\ .$$

To make contact with venerable formulae, we identify the 2-form D on three-dimensional space with a vector field $\vec{D}$. Since radial symmetry should be preserved, we make the ansatz

$$\vec{D}(\vec{x}, t) = f(r)\vec{e_r}$$

with

$$\vec{e_r} := \frac{1}{\sqrt{x^2+y^2+z^3}}(x, y, z) \in T_{(x,y,z)}\left(\mathbb{R}^3 \setminus \{0\}\right)$$

the radial unit vector field on $\mathbb{R}^3 \setminus \{0\}$.

Integrating over the two-sphere with center 0 and radius r, we find

$$q = \int_{S_r^2} \vec{D}\,\mathrm{d}\vec{f} = \ 4\pi\, r^2\, f(r)$$

and thus Coulomb's law:

$$\vec{D}(\vec{x},t) = \frac{q}{4\pi}\frac{1}{r^2}\vec{e_r} \ .$$

In the case of a static field, Coulombs law and the principle of superposition of charges conversely implies the first inhomogeneous Maxwell equation $\mathrm{d}_V D \ = \ \rho$.

Our second second observation is the fact that in an electric field, certain particles experience a force, the so-called Lorentz force. The strength of this force depends on their electric charge – which is a possible property of particles, like mass – and their velocity.

Observation 3.1.3.

1. *Keeping in mind that in natural systems forces on a point particle are gradients of a potential* V *and that it is more natural to replace the gradient by the exterior derivative* $\mathrm{d}V$*, we see a force not as a vector field on* $\mathbb{A}$ *but rather as a one-form* F*. A section* $s : I \to \mathbb{A}$ *is described in global spacial coordinates by a trajectory* $\varphi : \ I \to \mathbb{R}^3$ *such that*

$$s(t) \ = \ (t, \varphi(t)) \ .$$

Newton's equation of motion for the trajectory is then

$$\ddot{\varphi} = \iota^{-1}(\mathrm{F}) \ ,$$

where $\iota : \mathrm{vect}(\mathbb{A}_t) \to \Omega^{0,1}(\mathbb{A}_t)$ *uses the Euclidean metric on* $\mathbb{A}_t$ *to identify vertical tangent vectors to* $\mathbb{A}_t$ *and one-forms in* $\Omega^{\cdot,1}(\mathbb{A}_t)$*.*

2. *For the sake of conciseness, we will not work with discrete particles, but rather with a charge distribution on space described by the three-form* $\mathrm{j} \in \Omega^3(\mathbb{A})$*. Correspondingly, the force will be replaced by a force density on space-time.*

   The force density is a 4-form on $\mathbb{A}$ *with values in vertical, i.e. spacial differential one-forms:*

$$\mathrm{f} \in \Omega^{0,3}(\mathbb{A}, \Omega^{0,1}) \ .$$

   In local coordinates $\xi = (\xi^\alpha)_{\alpha=0,\dots 3} = (t, x, z, y)$*, we have*

$$\mathrm{f} = \sum_{\alpha=1}^{3} \mathrm{d}\xi^\alpha f_\alpha(t,x,y,z)\mathrm{d}t \wedge \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z \, .$$

3. *The force density* f *is familiar from mechanics and can be measured. We relate it to two quantities: the charge density* j *we have encountered previously in the discussion of charge conservation and a new property of space time, described by a two-form*

$$\mathrm{F} \in \Omega^2(\mathbb{A}) \ ,$$

   called the electromagnetic field strength.

   Again this field strength has a decomposition induced by the bigrading of differential forms on $\mathbb{A}$*:*

$$\mathrm{F} = E \wedge \mathrm{d}t + B$$

   with $E \in \Omega^{0,1}(\mathbb{A})$*, the electric field, and* $B \in \Omega^{0,2}(\mathbb{A})$ *the magnetic field. The electric and the magnetic field are to be thought of as a space time property that is the source of the Lorentz force. The electric field is a two-form and thus a "vector" while the magnetic field is a 2-form and thus a "pseudo-vector".*

4. *Suppose now that $\omega \in \Omega^p$ is a smooth p-form on a smooth manifold M. Let $x : U \to \mathbb{R}^n$ be a local coordinate system. Then*

$$I(\omega) := \sum_{i=1}^{n} \mathrm{d}x^i \iota_{\frac{\partial}{\partial x^i}} \omega$$

is a smooth $(p-1)$-form with values in 1-forms. Using linearity of the contraction and the behaviour under changes of local coordinates, one sees that $I(\omega)$ does not depend on the choice of local coordinates.

5. *The Lorentz equation relates the force density* f*, the charge density* j *and the electromagnetic field strength* F

$$\mathrm{f} = I(\mathrm{F}) \wedge \mathrm{J} \quad .$$

This relation between 1-form valued 4-forms is our second axiom for electrodynamics.

6. *Let us write this in a maybe more familiar form: the charge density is*

$$\mathrm{j} = \ q(t,x,y,z)\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z - \epsilon_{pqr} j_p \mathrm{d}t \wedge \mathrm{d}x^q \wedge \mathrm{d}x^r$$

and the electric field strength is

$$\mathrm{F} = E_i \mathrm{d}t \wedge \mathrm{d}x^i + \epsilon_{pqr} B_p \mathrm{d}x^q \wedge \mathrm{d}x^r \quad .$$

Let us compute the component f^i *of the force density for $i = 1, 2, 3$. We get from* $I(\mathrm{F})$ *the one-form*

$$I(F)_i = E_i \mathrm{d}t + \epsilon_{p,i,r} B_p \mathrm{d}x^r$$

which we have to wedge with j *to get a four-form on $\mathbb{A}$ for each $i = 1, 2, 3$:*

$$I(F)_i \wedge \mathrm{j} = (qE_i + \epsilon_{i,p,q} j_p B_q)\, \mathrm{d}t \wedge \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z$$

which gives the expression for the Lorentz force in three dimensions:

$$\vec{F} = q\vec{E} \ + \vec{j} \wedge \vec{B} \ .$$

In four dimensional tensor language, one has $f_i = \ F_{ik} J^k$.

We have now to add equations describing the dynamics of the electromagnetic field F that is the source of the Lorentz force.

Observation 3.1.4.

1. *Based on empirical evidence, we impose on the electromagnetic field* $\mathrm{d}F$ *the condition to be closed,*

$$\mathrm{d}F \ = 0 \ .$$

These are the homogeneous Maxwell equations.

2. *In the three-dimensional language* $\mathrm{F} = \mathrm{d}t \wedge \mathrm{E} + \mathrm{B}$*, the equation*

$$0 = \mathrm{d}F \ = \mathrm{d}t \wedge \mathrm{d}_V \ E + \mathrm{d}t \wedge \partial_t B + \mathrm{d}_V B$$

is equivalent to the following two equations:

$$\mathrm{d}_V \mathrm{B} = 0 \quad \textit{and} \quad \partial_t B = -\mathrm{d}_V \mathrm{E} .$$

The first condition expresses the fact that the magnetic field does not have any source, i.e. the absence of magnetic charges, compare our previous calculation for the electric field. The second equation is Faraday's law of induction. It can be seen as three equations for the time evolution of the magnetic field.

Faraday's law of induction is part of the basis of all electrical engineering:

Example 3.1.5.
Faraday's law of induction yields a relation between the induced voltage along a loop ∂F bounding a surface F

$$U_{ind} := \int_{\partial F} \mathrm{d}x \, \vec{E}$$

and the magnetic flux

$$\Phi^{mag} := \int_F \vec{B}$$

through the surface which reads

$$U_{ind} \equiv \int_{\partial F} \mathrm{d}x \, \vec{E} = -\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_F \vec{B}\right) = -\frac{\mathrm{d}}{\mathrm{d}t}\Phi_f^{mag} = -\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_F \vec{B}\right)$$

This law is the basis of the electric motor and the electrical generator. The minus sign is quite famous: it is called Lenz' rule.

We have now arrived at the following system: field strength F or, equivalently (E, B), containing in total 6 degrees of freedom and excitations H or, equivalently (D, H), containing in total 6 degrees of freedom as well. For these 12 degrees of freedom, we have just 6 time evolution equations.

We therefore need another relation which expresses the excitation H in terms of the field strength F. This is a relation in which properties of space time enter and which looks different in material or in empty space. At this point, additional properties of space-time may enter. In particular, four-dimensional space-time may have a four-dimensional metric and an orientation and thus a four-dimensional Hodge star $*$.

An obvious relation one would like to impose between

$$\mathrm{H} = -\mathrm{d}t \wedge H + D \quad \text{and} \quad \mathrm{F} = \mathrm{d}t \wedge E + B$$

would be to equate the electric and magnetic degrees of freedom,

$$E = D \quad \text{and} \quad B = H .$$

To this end, we would like to use a Hodge star $*$ on Galilei space $\mathbb{A}$. We endow $\mathbb{A}$ with a full metric that restricts to the metric on the fibres $\mathbb{A}_t$ and for which the global vector field ∂_t giving the time direction is perpendicular to all spacial hypersurfaces $\mathbb{A}_t$. We put

$$g(\frac{\partial}{\partial t}, \frac{\partial}{\partial t}) = a$$

where we assume the metric to be normalized such that $a = \pm 1$. It should be noted that we admit pseudo-Riemannian metrics on $\mathbb{A}$ as well. This gives us two possible four-dimensional Hodge star operators on Galilei space $\mathbb{A}$:

Using the results in the appendix, we find

$$*(B_x \mathrm{d}y \wedge \mathrm{d}z) = B_x \mathrm{d}t \wedge \mathrm{d}z \qquad \text{and} \qquad *\left(E_x \mathrm{d}t \wedge \mathrm{d}x\right) = aE_x \mathrm{d}y \wedge \mathrm{d}z$$

and equations for all cyclic permutations of x, y and z. This means

$$*F = aE + \mathrm{d}t \wedge B$$

so that we can use the Hodge star to impose the condition if and only if we choose $a = -1$. Then we have

$$*F = -H \ .$$

This is a first indication that electrodynamics is nicely compatible with an indefinite metric on four-dimensional space-time.

Observation 3.1.6.

1. *The following space-time relations are important:*

   - *In empty space, one has the relation*

     $$\mathrm{H} \ = \ -\lambda * F$$

     which makes sense in four dimensions. The constant λ is a constant of nature. Sometimes $*F$ is called the dual field strength.
   - *In axion-electrodynamics, one has two constants of nature*

     $$-\mathrm{H} = \lambda_1 * \mathrm{F} + \lambda_2 \mathrm{F} \ .$$

   - *In realistic media, the constants typically depend on frequencies or, equivalent, wave lengths, and yield quite complicated relations between* F *and* H.

2. *We summarize the Maxwell equations in vacuo:*

   $$\mathrm{dF} \ = 0 \quad \textit{and} \quad \mathrm{d}(*\mathrm{F}) \ = \mathrm{J} \ ,$$

   which are the homogeneous and the inhomogeneous Maxwell equations. As a consequence, we have the continuity equation

   $$\mathrm{dJ} = \mathrm{dd}(*\mathrm{F}) \ = 0 \ .$$

   It is complemented by the equation for the Lorentz force

   $$\mathrm{f} = I(\mathrm{F}) \wedge \mathrm{J} \ .$$

   In this form, the equations are also valid in special and even in general relativity. The Maxwell equations in vacuo can be considered as truly fundamental laws of nature.

3. *We also present these equations in the classic notation of vector calculus:*

$$\begin{array}{rclcrcl} \mathrm{div} B & = & 0 & \quad \textit{and} \quad & \partial_t B + \mathrm{rot}\vec{E} & = & 0 \\ \mathrm{div} E & = & \rho & \quad \textit{and} \quad & \mathrm{rot} B - \partial_t E & = & \vec{j} \end{array}$$

We next describe an important consequence of the Maxwell equations in vacuo:

Observation 3.1.7.

1. *Consider the Maxwell equations in vacuo without external sources, i.e.* $\mathrm{j} = 0$. *We find*

$$\mathrm{dF} = 0 \quad \textit{and} \quad \mathrm{d} * F = 0 \ .$$

The last equation implies $\delta F = 0$. *We thus have for the Laplace operator on differential forms*

$$\Delta F = \delta \mathrm{d} F + \mathrm{d}\delta F = 0$$

so that in the vacuum without external source, $\mathrm{J} = 0$, *harmonic forms are a solution to the Maxwell equations.*

2. *To restore familiarity, we repeat this analysis in the language of vector analysis. Then the Maxwell equations read:*

$$\begin{array}{ll} \mathrm{div}\, \vec{E} = 0 & \mathrm{rot}\, \vec{E} = -\frac{\partial \vec{B}}{\partial t} \\ \mathrm{div}\, \vec{B} = 0 & \mathrm{rot}\, \vec{B} = \frac{\partial \vec{E}}{\partial t} \end{array}$$

Taking the rotation of the second equation yields

$$\begin{aligned} \triangle \vec{E} = \triangle \vec{E} - \mathrm{grad}(\mathrm{div} \vec{E}) = -\mathrm{rot}\, \mathrm{rot} \vec{E} = \\ = \mathrm{rot}\, \frac{\partial}{\partial t} \vec{B} = \frac{\partial}{\partial t} \mathrm{rot}\, \vec{B} = \frac{\partial^2 \vec{E}}{\partial t^2}, \end{aligned}$$

A dimensional analysis shows that we should restore a factor c^2 *with the dimension of the square of a velocity. Hence we get*

$$\Box \vec{E} := \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \triangle \right) \vec{E} = 0$$

with $\Box$ *the so-called d'Alembert operator. Similarly, one finds*

$$\Box \vec{B} := \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \triangle \right) \vec{B} = 0.$$

3. *To find solutions, we make the ansatz*

$$\begin{aligned} \vec{E}(t, \vec{x}) &:= \vec{E}_0 e^{-i(\omega t - \vec{k}\vec{x})} \\ \vec{B}(t, \vec{x}) &:= \vec{B}_0 e^{-i(\omega t - \vec{k}\vec{x})} \end{aligned}$$

of plane waves with direction of propagation $\frac{\vec{k}}{|\vec{k}|}$. *Waves are thus examples of harmonic functions; they describe the propagation of light or other electromagnetic radiation (radio*

waves, gamma rays, ...) in empty space. Hence c can be identified with the velocity of light.

Applying the d'Alembert operator to these plane waves, we find:

$$\Box\vec{E} = \left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \triangle\right)\vec{E} = \vec{E}_0 e^{-i(\omega t - \vec{k}\vec{x})}\left(\frac{\omega^2}{c^2} - \vec{k}^2\right) = 0$$

This implies the so-called dispersion relation

$$\omega^2 = c^2 k^2$$

for plane wave solutions which is an algebraic relation between wave length $\lambda = \frac{2\pi}{k}$ and frequency $\nu = \frac{\omega}{2\pi}$. This shows that plane wave solutions travel at the speed given by the constant c. Since this includes in particular light waves, the constant c is called speed of light.

Maxwell's equations imply moreover that

$$\begin{array}{ll} \vec{k}\vec{E}_0 = 0 & \vec{k}\vec{B}_0 = 0 \\ \vec{k}\wedge\vec{E}_0 = \kappa\omega\vec{B}_0 & \vec{k}\wedge\vec{B}_0 = -\frac{\omega}{\kappa c^2}\vec{E}_0. \end{array}$$

In other words, $(\vec{k}, \vec{E}, \vec{B})$ form an oriented basis of spacial $\mathbb{R}^3$: electromagnetic waves are transversal. Depending on the initial conditions, they can be polarized, either linearly or elliptically, in particular in a circular way.

A characteristic feature of electrodynamics is the existence of a quantity c having the dimension of a velocity and being the speed of light. Galileian invariance for electrodynamics would imply that the Maxwell equations have the same form in all inertial systems. In fact, very precise experiments, like e.g. the famous Michelson-Morley interferometrical experiment, in which the velocity of the earth is added and subtracted from the velocity of light confirm that the speed of light is the same in different inertial systems. But this is incompatible with the transformation of velocities in Galilei space which is simple vector addition. It turns out that we have to give up Galilei space; this is the starting point of Einstein's theory of relativity.

3.2 Special relativity

The existence of a distinguished velocity that is the same in all reference frames in relative uniform motion has been experimentally verified to very high precision. It is, however, incompatible with the transformation of velocities under the Galilei group. It thus forces us to revise our ideas about space and time: it gives rise to the theory of special relativity.

We discuss our postulates:

First postulate: Space time is homogeneous.

In mathematical terms, this means that our model for space time is still the four-dimensional affine space $\mathbb{A}$ over $\mathbb{R}^4$, but we drop the requirement of an absolute time difference function. As a consequence, the relevant symmetry group is not any longer the Galilei group which is a subgroup of the symmetry group $\mathbb{R}^4 \rtimes \mathrm{GL}(4,\mathbb{R})$ of affine space. We require the symmetry group to be of the form $\mathbb{R}^4 \rtimes G$ for some Lie subgroup of $\mathbb{R}^4 \rtimes \mathrm{GL}(4,\mathbb{R})$.

Second postulate: relativity. The laws of physics are of same form for all observers in relative uniform motion.

To this end, we have to specify such a set of observers. We require that their worldlines are a distinguished set of affine lines that is invariant under the action of $\mathbb{R}^4 \rtimes G$.

Third postulate: velocity of light as a limit velocity.

Since the velocity of light is fixed by the Maxwell equations, it should be same for all observers and independent of the motion of the source of the light. We require even more: the velocity of light should be a limit velocity that cannot be reached by any of the observers in the set.

The structure of electrodynamics suggests to take a metric of signature $(-1,1,1,1)$ on space time. The group G should then be the non-compact Lie group $\mathrm{O}(3,1)$ preserving a non-degenerate symmetric bilinear form of that signature. We thus arrive at the following definition:

Definition 3.2.1

1. *An affine space $\mathbb{M}$ over $\mathbb{R}^4$ together with a metric of signature $(-1,+1,+1,+1)$ on its difference space is called a Minkowski space.*

2. *Standard Minkowski space is $\mathbb{R}^4$ with the diagonal metric $\eta = (-1,1,1,1)$. It plays the role of standard Galilei space. Using the velocity of light c, we endow it with coordinates (ct, x^1, x^2, x^3) whose dimension is length.*

3. *A Lorentz system of $\mathbb{M}$ is an affine map*

$$\phi : \mathbb{M} \to (\mathbb{R}^4, \eta)$$

which induces an isometry on the difference space.

4. *The light cone in $T_p\mathbb{M}$ is the subset*

$$\mathrm{LC}_p := \{x \in T_p\mathbb{M} \big| \ \eta_p(x,x) = 0\}.$$

5. *A Poincaré transformation is a diffeomorphism of $\mathbb{M}$ whose differentials respect the metrics. A Lorentz transformation is the linear map induced on the difference space.*

   the light cones in all tangent spaces, the so-called "causal structure".

The importance of the causal structure for any relativistic theory cannot be overrated. In fact, the notion of a causal structure can be introduced for an Lorentz manifold, i.e. any smooth n-dimensional manifold M with a metric of signature $(n-1,1)$.

In the case of Minkowski space, the symmetries of the causal structure can be determined explicitly:

Lemma 3.2.2.
Let $\Lambda : \mathbb{M} \to \mathbb{M}$ be a diffeomorphism that preserves the light cones. Then Λ is an affine mapping. For the induced map on tangent space, one has

$$\eta_p(\Lambda_p x, \Lambda_p y) = a(\Lambda)\eta_p(x,y) \quad \text{for all } x, y \in T_p\mathbb{M}$$

with some positive constant $a(\Lambda)$. The symmetries are thus composed of a four-dimensional subgroup of translations, a one-dimensional subgroup of dilatations and the six-dimensional Lorentz group $\mathrm{SO}(3,1)$.

We list properties of the Lorentz group $\mathrm{L} = \mathrm{O}(3,1)$:

Remarks 3.2.3.

1. *One can show that the determinant of any element of* L *is* ± 1*. Elements of different determinant are in different connected components of* L*.*

   One can moreover show that for the matrix element, one has $\Lambda_0^0 = \eta|\Lambda_0^0|$ *with* $\eta \in \{\pm 1\}$*. Group elements with different sign* η *cannot be in the same connected component of* L*.*

   The subgroup
   $$\mathrm{L}^\uparrow := \{\Lambda \in \mathrm{L} \,|\, \Lambda_0^0 > 0\}$$
   is called the orthochronous Lorentz group. The subgroup
   $$\mathrm{L}_+^\uparrow := \{\Lambda \in \mathrm{L}^\uparrow \,|\, \det \Lambda = 1\}$$
   is called the proper orthochronous Lorentz group. It is the connected component of the identity of the Lorentz group.

   The Lorentz group has four connected components that are determined by the sign of the determinant and η*.*

2. *Consider the bijection of vectors in* $\mathbb{R}^4$ *to two-dimensional hermitian matrices*
   $$\Phi : (x^0, x^1, x^2, x^3) \;\mapsto\; \begin{pmatrix} x^0 + x^3 & x^1 - \mathrm{i}x^2 \\ x^1 + \mathrm{i}x^2 & x^0 - x^3 \end{pmatrix}$$
   and
   $$\det \Phi(x)\Phi(y) = -(x, y)$$
   Lie group $\mathrm{SL}(2,\mathbb{C})$ *acts on hermitian matrices as* $H \mapsto SHS^\dagger$*. This action induces an isomorphism of*
   $$\mathrm{PSL}(2,\mathbb{C}) := \mathrm{SL}(2,\mathbb{C})/\{\pm 1\}$$
   and the proper orthochronous Lorentz group.

3. *Using the parity transformation*
   $$P = \mathrm{diag}(+1, -1, -1, -1)$$
   and time reversal
   $$T = \mathrm{diag}(-1, +1, +1, +1)$$
   we can write every element of L *uniquely in the form*
   $$\Lambda = P^n T^m \Lambda_0 \quad \textit{with} \;\; n, m \in \{0, 1\},\; \Lambda_0 \in \mathrm{L}_+^\uparrow.$$

4. *The Lorentz group* $\mathrm{L}_{+1}^\uparrow$ *is not compact. A maximal compact subgroup is the group of rotations of the Euclidean subspace spanned by* $\mathrm{span}(e_1, e_2, e_3)$ *which is isomorphic to the group* $\mathrm{SO}(3)$ *of rotations. Since the Lorentz group is semi-simple, any other maximal compact subgroup is conjugate to this subgroup.*

The one parameter subgroup

$$\Lambda_1(\theta) = \begin{pmatrix} \cosh\theta & -\sinh\theta & 0 & 0 \\ -\sinh\theta & \cosh\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \in \Lambda_+^\uparrow.$$

obeys $\Lambda_1(\theta_1) \circ \Lambda(\theta_2) = \Lambda(\theta_1 + \theta_1)$. *An element of this non-compact subgroup is called a* <u>boost</u> e_1*-direction with* <u>rapidity</u> θ. *(Boost can be defined with respect to any direction in the spacial subspace spanned by* e_1, e_2 *and* e_3*.)*

For $y = \Lambda(\theta)x$, *we have* $y^2 = x^2$, $y^3 = x^3$ *and*

$$y^0 = x^0 \cosh\theta - x^1 \sinh\theta$$
$$y^1 = x^1 \cosh\theta - x^0 \sinh\theta$$

In particular, $y^1 = 0$ *is equivalent to* $x^1 = ct\tanh\theta = vt$ *so that the origin of the* y *coordinate system is in uniform relative motion, with* <u>velocity</u> $v = c\tanh\theta$.

We deduce:

- *While rapidities are additive, velocities in special relativity are not:*

$$v = \frac{v_1 + v_2}{1 + \frac{v_1 v_2}{c^2}}.$$

- *In the specific case* $v_2 = c$, *we find*

$$v = \frac{v_1 + c}{1 + \frac{v_1}{c}} = \ c\ .$$

  This explains in particular the result of the Michelson-Morley experiment: adding or subtracting the velocity v_1 *of the earth to the velocity of light again gives the velocity of light.*

- *Using* $|\tanh\theta| < 1$ *for any real parameter* θ, *we see that the velocity* c *of light is a limit that cannot be reached:* $|v| < |c|$.

5. *Every Lorentz transformation* Λ *can be written uniquely as a product of a boost* $B(\Lambda)$ *and a rotation* (Λ), *but the boosts are not a subgroup.*

6. *The Poincaré group turns out to be the semi-direct product of the Lorentz group and the translation group* $\mathbb{R}^4$. *It is a ten-dimensional non-compact Lie group.*

We reconsider the geometry of Minkowski space:

Definition 3.2.4

1. *A non-zero vector* x *in the difference space for Minkowski space* $\mathbb{M}$ *is called*

<u>time like</u>, *if* $x^2 = (x,x) < 0$
<u>light like</u> *or* <u>null</u>, *if* $x^2 = (x,x) = 0$
<u>space like</u>, *if* $x^2 = (x,x) > 0$

2. *A time like or light like vector $x = (x^0, x^1, x^2, x^3)$ in the difference space for Minkowski space is called future directed, if $x^0 > 0$ and past directed if $x^0 < 0$. The light cone at any point decomposes into the forward light cone consisting of light like future directed vectors, the backward light cone consisting of past directed light like vectors and the zero vector.*

   The notion of a light cone makes sense on an arbitrary Lorentz manifold, i.e. smooth manifold M of any dimension n and a metric of signature $(1, n-1)$. The notions of future or a past light cone, however, do not make sense on a general Lorentz manifold. If they make sense, the manifold is called time-orientable. A smooth choice of a future light cone in any point is also called a time orientation.

3. *The (absolute) future $I^+(x)$ and the (absolute) past $I^-(x)$ of a point $x \in \mathbb{M}$ are the sets*

$$I^{\pm}(x) := \{y \in \mathbb{M} | (y-x)^2 \leq 0, y - x \text{ future / past directed } \}.$$

   The future of a subset $S \subset \mathbb{M}$ is defined by

$$I^{\pm}(S) := \cup_{x \in S} I^{\pm}(x)$$

We can now formulate the **postulate of causality**:

- An event $x \in \mathbb{M}$ can influence an event $y \in \mathbb{M}$ (slang: a signal can be sent from x to y) if and only if y lies in the future of x.

This postulate really restricts the equations of motion for the physical fields living on the Minkowski space. It will be made mathematically precise in our discussion of general relativity. If $x \notin I^+(y)$ and $y \notin I^+(x)$, the events are called causally disconnected.

The following lemma states the appropriate generalization of the Cauchy-Schwarz identity. We leave its proof as an exercise:

Lemma 3.2.5.

1. *For any two* space like *vectors x, y, the Cauchy-Schwarz identity holds in its usual form,*

$$|(x,y)| \leq \sqrt{(x,x)}\sqrt{(y,y)}$$

   with equality if and only if the vectors x, y are linearly dependent.

2. *For any two* time like *vectors x, y in the difference space to Minkowski space, we have*

$$|(x,y)| \geq \sqrt{-(x,x)}\sqrt{-(y,y)}$$

   and equality if and only if the vectors x, y are linearly dependent.

3. *For any two lightlike or timelike vectors x, y, we have $(x, y) \leq 0$, if both are future directed or both are past directed. We have $(x, y) \geq 0$, if and only if one is future and one is past directed.*

Proof:
We restrict ourselves to the second statement. We work in the rest system of x, i.e. in local coordinates where x takes the form $x = (x^0, 0, 0, 0)$. Using $(y, y) = \vec{y}^2 - (y^0)^2$, we find

$$|(x, y)| = |x^0 y^0| = \sqrt{-x^2}\sqrt{-y^2 + \vec{y}^2} \geqslant \sqrt{-x^2}\sqrt{-y^2} \ .$$

□

We are now ready to specify our set of observers:

Definition 3.2.6

1. *A velocity unit vector $\hat{w}$ is a future directed time like vector normalized to $(\hat{w}, \hat{w}) = -1$.*

2. *Our set of observers – or, more precisely, their world lines, consists of the affine lines of the form $a + \mathbb{R}\hat{w}$, where $a \in \mathbb{M}^4$ is any point and $\hat{w}$ is any velocity unit vector. We abbreviate the observer or the corresponding affine lie respectively, with $\mathcal{O}_{a,\hat{w}}$ or sometimes even with $\mathcal{O}_{\hat{w}}$, when the base point a does not matter.*

3. *The time interval elapsed between two events $x, y \in \mathbb{R}^4$ as observed by the observer $\mathcal{O}_{\hat{w}}$ is given by the observer-dependent time function*

$$\Delta t_{\hat{w}}(x, y) := (\hat{w}, x - y) \ .$$

   Note that the difference vector from x to y is $y - x$, which is however, time like.

4. *Two events are simultaneous for the observer $\mathcal{O}_{\hat{w}}$, if $\Delta t_{\hat{w}}(x, y) = 0$. We say that x happens before y, if $\Delta t_{\hat{w}}(x, y) > 0$ and x happens after y if $\Delta t_{\hat{w}}(x, y) < 0$.*

Remarks 3.2.7.

1. *The set of observers is invariant under the symmetry group $\mathbb{R}^4 \rtimes \mathrm{O}(3, 1)$.*

2. *The time interval between two events $x, y \in \mathbb{M}$ depends on the observer. This fact might be called "relativity of time"; it is in sharp contrast to the absolute time difference function in Galilei space.*

   Physically, this can be traced back to an operational *definition of time as a measurement carried out by some observer. In a certain sense, the clocks of special relativity are "small" and are thought of as being connected to observers. This is in sharp contrast to the universal clock for the whole universe in Galilean theory.*

3. *The set of events simultaneous to $x \in \mathbb{M}$ for the observer $\mathcal{O}_{\hat{w}}$ is the affine hyperplane perpendicular to $\hat{w}$ containing x.*

   For any pair of events $x, y \in \mathbb{M}$ simultaneous for the observer, $\mathcal{O}_{\hat{w}}$, the relative position $y - x$ is space like so that the events are causally disconnected and cannot influence each other.

   This gives a foliation of Minkowski space into space-like affine hyperplanes of simultaneous events that depends on the velocity unit vector $\hat{w}$ and thus on the observer. This might be called "relativity of space".

4. *More generally, a massive of massless point particle is described by its trajectory*

$$\begin{aligned} x:\ I=[t_1,t_2] &\to \mathbb{M} \\ t &\mapsto x(t) \end{aligned}$$

where we require for physical motions the velocity vector to be future directed. It is required to be time like for the class of "massive particles" and light like for "massless particles",

$$\frac{\mathrm{d}x^0}{\mathrm{d}t} > 0 \qquad \left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 \leq 0 \ .$$

The Lorentz invariant quantity

$$T := \frac{1}{c}\int_{t_1}^{t_2} \mathrm{d}\,t \sqrt{-\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2}$$

is called eigen time of the particle elapsed between the two points $x(t_1)$ and $x(t_2)$ on its trajectory.

We discuss some more properties of the observer dependent time functions.

Remarks 3.2.8.

1. *For any two observers $\mathcal{O}_{\hat{w}}$ and $\mathcal{O}_{\hat{v}}$ with different velocity unit vectors $\hat{w}$ and $\hat{v}$, there exist events $x, y \in \mathbb{M}$ such that x happens before y for $\mathcal{O}_{\hat{w}}$ and y happens before x for $\mathcal{O}_{\hat{v}}$. This fact might be called the "relativity of simultaneity."*

   However, such events are always causally disconnected and cannot influence each other so that causal paradoxes are excluded.

2. *The future $I^+(x)$ of an event $x \in \mathbb{M}$ equals*

$$I^+(x) = \{y \in \mathbb{M} | (\hat{v}, y-x) < 0 \text{ for all observers } \mathcal{O}_{\hat{v}}\}$$

   i.e. the set of events that happens after x for all *initial observers. This justifies the qualifier* absolute *future for $I^+(x)$.*

Observation 3.2.9.
We discuss some famous relativistic phenomena:

1. *Time dilatation:*
   Consider two observers $\mathcal{O}_{\hat{w}}$ and $\mathcal{O}_{\hat{v}}$ which are in relative uniform motion. We take two events $x, y \in \mathbb{M}$ which occur for the observer $\hat{w}$ at the same place in space. This just means $x - y = t\hat{w}$ with some $t \in \mathbb{R}$. For the time elapsed between these two events, the observer $\hat{w}$ measures

$$\Delta t_{\hat{w}}(x,y) = |x-y| = \sqrt{-(x-y)^2} \ .$$

   You can imagine that observer $\hat{w}$ looks twice on his wrist watch, at the event x and at the event y.

   Let us compare this to the time difference measured by observer $\mathcal{O}_{\hat{v}}$ for the events $x \in \mathbb{M}$ and $y \in \mathbb{M}$ which is

$$\Delta t_{\hat{v}}(x,y) = |(\hat{v}, x-y)|$$

We introduce the vector

$$a := \ (x-y) \ + (x-y,\hat{v}) \ \hat{v}$$

which is, because of $(\hat{v},\hat{v}) = -1$*, orthogonal to* $\hat{v}$*,*

$$(a,\hat{v}) = (x-y,\hat{v}) \ + (x-y,\hat{v})(\hat{v},\hat{v}) = 0$$

and which is, because of the Cauchy-Schwarz inequality for future-directed time like vectors,

$$|(\hat{v},\hat{w})| \geq \sqrt{-(\hat{v},\hat{v})}\sqrt{-(\hat{w},\hat{w})} = 1 \ ,$$

a space-like vector:

$$\begin{aligned} a^2 &= (x-y)^2 + 2(x-y,\hat{v})^2 + (x-y,\hat{v})^2(\hat{v},\hat{v}) \\ &= (x-y)^2 + (x-y,\hat{v})^2 = -t^2 + t^2(\hat{v},\hat{w})^2 > 0 \ . \end{aligned}$$

We have thus

$$x - y = \ -\Delta t_{\hat{v}}(x,y) \ \hat{v} + a$$

which implies

$$\Delta t_{\hat{w}}(x,y)^2 = \ |(\hat{v},x-y)|^2 + a^2 = \ \Delta t_{\hat{v}}(x,y)^2 + a^2$$

Hence

$$|\Delta t_{\hat{v}}(x,y)| > \ |\Delta t_{\hat{w}}(x,y)| \ .$$

Thus the moving observer $\mathcal{O}_{\hat{v}}$ *measures a strictly longer time interval than the observer* $\mathcal{O}_{\hat{w}}$ *at rest. This effect has been measured with very high precision in observations of the life time of instable particles in motion.*

2. *Similarly, lengths in motion see to be contracted: one has the phenomenon of (length contraction).*

3. *We finally discuss the twin paradoxon as the geometry of triangles given by future-directed time like vectors in* $\mathbb{M}$*. Consider three future directed time like vectors in* $\mathbb{M}$ *that obey*

$$z \ = \ x + y \ .$$

Usually, this is given the following interpretation: at an event $m \in \mathbb{M}$*, space ship one travelling in uniform motion ejects another space ship, space ship two, which travels with a different velocity. At the point* $m + x$ *this ships sharply (i.e. instantaneously) changes direction and travels, again by uniform motion from* $m + x$ *to* $m + x + y$*. Here, it meets spaceship one that travelled directly from* m *to* $m + z$*. Usually, one staffs the space ships with twins. If one wishes, one can work in the inertial system where space ship one and twin one are at rest. Twin two then travelled to a star at* $m + x$ *and returned to the space ship one at rest.*

For twin one at rest, we write $z = t\hat{z}$ *with a unit velocity vector* $\hat{z}$ *and* $t = -(z,\hat{z})$ *and the time elapsed is*

$$\Delta t_{\hat{z}} = (\hat{z},-z) \ = t = \sqrt{-(z,z)} \ .$$

On the other hand, using the Cauchy-Schwarz identity for two future-directed time like vectors x, y*, one has*

$$-(x,y) \geq \sqrt{-(x,x)}\sqrt{-(y,y)}$$

and thus

$$-(z,z) = -(x,x) - 2(x,y) - (y,y) \geq -(x,x) + 2\sqrt{-(x,x)}\sqrt{-(y,y)} - (y,y)$$

It implies

$$-z^2 \geqslant \left(\sqrt{-x^2} + \sqrt{-y^2}\right)^2$$

and thus

$$\sqrt{-z^2} \geqslant \sqrt{-x^2} + \sqrt{-y^2}$$

so that twin one which is at rest is older. The twin paradox is thus a consequence of the Cauchy-Schwarz inequality for time like vectors.

4. *In fact, one can base the discussion of relativity on the principle of maximal proper time (see the book by Bertel Laurent for an elementary discussion):*
if two space ships part and meet and one of them is freely floating throughout its journey, then this ship will measure a longer travel time than the other one.

3.4 Electrodynamics as a gauge theory

We start our discussion with some comments on electrodynamics on Minkowski space $\mathbb{M}$.

Remarks 3.4.1.

1. *We consider Maxwell's equations on a star-shaped region $U \subset \mathbb{M}$ on which we can apply Poincaré's lemma. From the homogeneous Maxwell equations $\mathrm{d}F = 0$ and the fact that Minkowski space is contractible, we conclude that there exists a one-form $A \in \Omega^1(\mathbb{M})$ such that $\mathrm{d}A = F$. The one-form A is called a gauge potential for the electromagnetic field strength F.*

   The one-form A is not unique: taking any function $\lambda \in \Omega^0(\mathbb{M})$, we find that

   $$A' := A + \mathrm{d}\lambda \in \Omega^1(\mathbb{M})$$

   also obeys the equation $\mathrm{d}A' = F$. This change of gauge is also called a gauge transformation. This arbitrariness in choosing A is called the gauge freedom and choosing one A is called a gauge choice.

2. *One can impose additional gauge conditions on A to restrict the choice. For example, using the metric on $\mathbb{M}$, one can impose the condition $\delta A = 0$. A choice of A that obeys this equation is called a Lorentz gauge. This gauge is called a covariant gauge, since the gauge* condition *is covariant.*

   It should be appreciated that gauge conditions used in physics need not lead to a unique choice for the one-form A. For example, the Lorentz condition is preserved, if a gauge transformation given by a harmonic function λ is applied: $\delta\lambda = 0$ implies $\mathrm{d}(A+\mathrm{d}\lambda) = \mathrm{d}A$ and $\delta(A + \mathrm{d}\lambda) = 0$.

   Another type of conditions are Coulomb gauges: suppose a Lorentz manifold has global vector field V that nowhere vanishes, e.g. $\frac{\partial}{\partial t}$ on Minkowski space $\mathbb{M}$. For a Coulomb gauge, one requires $\iota_V A = 0$.

3. *We also write these equations in the three-dimensional language of vector calculus*

$$\vec{B} = \mathrm{rot}\,\vec{A}$$

where $\vec{A}$ is a 3-vector field Hodge-dual to the spacial part of the one-form $A \in \Omega^1(U)$. The field $\vec{A}$ is called a vector potential. For the electric field, the Faraday's equation implies

$$\mathrm{rot}(\vec{E}) \ = -\partial_t \vec{B} \ = -\mathrm{rot}\partial_t \vec{A}$$

so that $E + \partial_t A$ can be written as the gradient of a time dependent function Φ, the scalar potential:

$$\vec{E} = -\mathrm{grad}\Phi - \frac{\partial \vec{A}}{\partial t} \quad .$$

It is given by the time component of the one-form $A \in \Omega^1(U)$.

In this language, the gauge freedom is expressed by the fact that vector potential and scalar potential are not unique, but can be change for any smooth function $\chi(\vec{x}, t)$ to

$$\vec{A}' = \vec{A} + \mathrm{grad}\chi \quad \textit{and} \quad \Phi' = \Phi - \frac{\mathrm{d}\chi}{\mathrm{d}t}$$

which are again potentials for the same field strength.

4. *Typical gauge conditions then read:*

   - *for the Coulomb gauge on Minkowski space for the vector field ∂_t:* $\mathrm{div}\vec{A} = 0$
     for which the remaining gauge transformations are those with $\triangle\chi = 0$.
   - *for the Lorentz gauge:* $\mathrm{div}\vec{A} + \frac{1}{c^2}\frac{\partial \chi}{\partial t}$
     for which the remaining gauge transformations are those with $\Box\chi = \frac{1}{c^2}\frac{\partial x}{\partial t^2} - \triangle_\chi = 0$.

5. *We pause a moment to consider the special case of so called static situations in which fields, currents and charge densities are time independent, the Maxwell equations decouple:*

$$\begin{array}{ll} \mathrm{div}\,\vec{E} = \rho & \mathrm{div}\,\vec{B} = 0 \\ \mathrm{rot}\,\vec{E} = 0 & \mathrm{rot}\,\vec{B} = \vec{j} \quad \textit{with } \mathrm{div}\vec{j} = 0 \end{array}$$

On a star-shaped region, we find $E = -\mathrm{grad}\,\Phi$ with $\triangle\Phi = -\frac{\varrho}{\epsilon}$. Mathematically, electrostatics is thus the theory of the (inhomogeneous) Poisson equation. Similarly, we find for the magnetic field in terms of the vector potential $\vec{B} = \mathrm{rot}\,\vec{A}$ the equation

$$\triangle A - \mathrm{grad}\ \mathrm{div}A = -\kappa\mu_0\vec{j},$$

which reduces in the Coulomb gauge to $\triangle A = -\kappa\mu_0\vec{j}$ and thus to a vector valued Poisson equation.

Classical electrodynamics is a theory of two-forms. For this theory, the gauge potential is just a calculational tool. Since this is different for quantum theories, we study the underlying mathematical structure more closely.

Observation 3.4.2.
We study the homogenous Maxwell equation $\mathrm{d}F$ *on an arbitrary smooth four-dimensional manifold* M*. We do not required that the manifold is endowed with a metric.*

1. *We cover* M *with coordinate neighborhoods* $(U_\alpha)_{\alpha\in I}$ *such that the image of each* U_α *in* $\mathbb{R}^4$ *is star-shaped.*

   For generalizations of this discussion on a manifold of arbitrary dimension, one would fix a so-called good open cover (U_α)*; for such a cover, arbitrary finite intersections*

   $$U_{\alpha_1\ldots\alpha_n} := U_{\alpha_1}\cap\ldots\cap U_{\alpha_n}$$

   of coordinate neighbourhoods are either empty or contractible.

2. *By the Poincaré lemma for the contractible subset* U_α*, the restriction* $F|_{U_\alpha}\in\Omega^2(U)$ *of the closed 2-form* $F\in\Omega^2(M)$ *is exact. We can find* $A_\alpha\in\Omega^1(U_\alpha)$ *such that* $F\big|_{U_\alpha} = \mathrm{d}\,A_\alpha$*. We call these one-forms local gauge potentials.*

   On twofold intersections $U_{\alpha\beta} = U_\alpha\cap U_\beta$ *, we find for the restrictions of the one-forms a one-form*

   $$\mathrm{d}\left(A_\alpha\big|_{U_\alpha\cap U_\beta} - A_\beta\big|_{U_\alpha\cap U_\beta}\right) = F\big|_{U_\alpha\cap U_\beta} - F\big|_{U_\beta\cap U_\alpha} = 0.$$

   By the Poincaré lemma, this time applied to the intersection $U_\alpha\cap U_\beta$*, we find a real-valued function* $\lambda_{\alpha\beta}\in C^\infty(U_{\alpha\beta},\mathbb{R})$ *with*

   $$(\delta A)_{\alpha\beta} := A_\alpha - A_\beta = \mathrm{d}\,\lambda_{\alpha\beta} = \frac{1}{\mathrm{i}}\,\mathrm{e}^{-\mathrm{i}\,\lambda_{\alpha\beta}}\,\mathrm{d}\,\mathrm{e}^{\mathrm{i}\,\lambda_{\alpha\beta}} = \frac{1}{\mathrm{i}}\,\mathrm{d}\,\log\,\mathrm{e}^{\mathrm{i}\,\lambda_{\alpha\beta}}$$

   We prefer to work with smooth $\mathrm{U}(1)$*-valued functions defined on two-fold overlaps,* $g_{\alpha\beta} := e^{i\,\lambda_{\alpha\beta}}\in C^\infty\left(U_{\alpha\beta},\mathrm{U}(1)\right)$*.*

   On triple overlaps $U_{\alpha\beta\gamma} = U_\alpha\cap U_\beta\cap U_\gamma$*, we find*

   $$\begin{aligned} 0 &= (A_\alpha - A_\beta) + (A_\beta - A_\gamma) + (A_\gamma - A_\alpha)\\ &= \tfrac{1}{\mathrm{i}}\mathrm{d}\log\left(g_{\alpha\beta}g_{\beta\gamma}g_{\gamma\alpha}\right)\end{aligned}$$

   so that on triple overlaps the product $g_{\alpha\beta}g_{\beta\gamma}g_{\gamma\alpha}$ *of the* $\mathrm{U}(1)$*-valued functions is constant. We require this constant to be equal to one.*

3. *We take the point of view that the field strength* F *is a derived quantity and thus consider elements in*

   $$V^2 := \oplus_\alpha\,\Omega^1(U_\alpha)\oplus\oplus_{\alpha,\beta}C^\infty\left(U_{\alpha\beta},\mathrm{U}(1)\right)$$

   which are of the form $(A_\alpha,\, g_{\alpha\beta})$*, i.e. a 1-form for each open subset and a* $\mathrm{U}(1)$*-valued function on each twofold intersection. They are required to obey*

   $$\begin{aligned}(0,1) &= D_2\left(A_\alpha,\, g_{\alpha\beta}\right) := \left((\delta\,A)_{\alpha\beta} - \tfrac{1}{\mathrm{i}}\mathrm{d}\log g_{\alpha\beta}, (\delta\,g)_{\alpha\beta\gamma}\right)\\ &= \left(A_\alpha - A_\beta - \tfrac{1}{\mathrm{i}}\mathrm{d}\log g_{\alpha\beta}, g_{\alpha\beta}g_{\beta\gamma}g_{\gamma\alpha}\right)\end{aligned}$$

   where the first component is to be considered on twofold overlaps and the second component on threefold overlaps. Thus D_2 *takes its values in*

   $$V^3 := \oplus_{\alpha,\beta}\,\Omega^1(U_{\alpha\beta})\oplus\oplus_{\alpha,\beta,\gamma}C^\infty\left(U_{\alpha\beta\gamma},\mathrm{U}(1)\right)$$

In fact, one can combine the Čech complex for the cover $(U_\alpha)_{\alpha\in I}$ with the de Rham complex into a double complex and interpret this equation as the vanishing of the differential in the total complex of the double complex. As usual for Čech complexes, one gets rid, in a second step, of the choice of cover by taking a limit over all covers.

4. *It remains to take into account gauge transformations which correspond to exact elements in the total complex of this double complex. Gauge transformations are defined locally, so they are parametrized by a* $\mathrm{U}(1)$*-valued function λ_α on each open subset U_α. A gauge transformation acts by*

$$A_\alpha \mapsto A_\alpha + \frac{1}{\mathrm{i}}\mathrm{d}\log\lambda_\alpha$$

where λ_α is a smooth $\mathrm{U}(1)$*-valued function on U_α. Hence we set*

$$V^1 := \oplus_\alpha C^\infty\left(U_\alpha, \mathrm{U}(1)\right) \ .$$

The gauge transformed potentials

$$A'_\alpha := A_\alpha + \frac{1}{\mathrm{i}}\mathrm{d}\log\lambda_\alpha \quad \textit{and} \quad A'_\beta := A_\beta + \frac{1}{\mathrm{i}}\mathrm{d}\log\lambda_\beta$$

should then be related by smooth $\mathrm{U}(1)$*-valued functions on twofold overlaps as*

$$A'_\alpha - A'_\beta = \frac{1}{\mathrm{i}}\mathrm{d}\log g'_{\alpha\beta} \ .$$

Ths implies $g'_{\alpha\beta} = g_{\alpha\beta}\lambda_\alpha\lambda_\beta^{-1}$ up to a multiplicative constant on $U_{\alpha\beta}$ which we put to one. A gauge transformation thus acts by

$$A_\alpha \mapsto A_\alpha + \frac{1}{\mathrm{i}}\mathrm{d}\log\lambda_\alpha \quad \textit{and} \quad g_{\alpha\beta} \mapsto g_{\alpha\beta}\lambda_\alpha(\lambda_\beta)^{-1} \ .$$

We therefore consider the operator

$$\begin{array}{rcl} D_1 : V^1 & \to & V^2 \\ (\lambda_\alpha) & \mapsto & \left(\mathrm{d}\lambda_\alpha,\ \lambda_\alpha\lambda_\beta^{-1}\right). \end{array}$$

One easily verifies $D_2 \circ D_1 = 0$.

We are thus interested to gain a deeper understanding of the quotient

$$\hat{H}^2(M) := \ \ker D_2/\mathrm{Im}D_1 \ .$$

The abelian group $\hat{H}^2(M)$ is a so-called differential cohomology group assigned to the manifold M. These groups are in fact infinite-dimensional topological groups. We will first give a geometric interpretation in terms of line bundles with connection:

Observation 3.4.3.

1. *For a geometric interpretation, we consider on the disjoint union*

$$\tilde{\mathrm{L}} = \bigsqcup_{\alpha\in I}\left(U_\alpha \times \mathbb{C}\ \right) := \bigcup_{\alpha\in I}\left(U_\alpha \times \{\alpha\} \times \mathbb{C}\right)$$

the equivalence relation

$$(x, \alpha, g_{\alpha\beta}\, z) \;\sim\; (x, \beta, z)\ .$$

Then the quotient

$$\mathrm{L} := \tilde{\mathrm{L}}\big/_{\sim}$$

is a smooth manifold and comes with a natural projection to M which provides a complex line bundle $\pi :\ \mathrm{L} \to M$.

2. *This bundle comes with a linear connection which locally is*

$$\bigtriangledown_\alpha = \ \mathrm{d} \ + \ \frac{1}{\mathrm{i}}\, A_\alpha\ .$$

Locally, we can consider the two form

$$F_\alpha := \ \mathrm{d}\, A_\alpha$$

which is trivially closed. On twofold overlaps, we have

$$F_\alpha - F_\beta = \mathrm{d}\, A_\alpha - \mathrm{d}\, A_\beta \ = \ \mathrm{d}\,(\delta\, A)_{\alpha\beta} \ = \ \frac{1}{\mathrm{i}} \mathrm{d}\, \mathrm{dlog}\, g_{\alpha\beta} \ = \ 0$$

so the locally defined two-forms F_α patch together into a globally defined two-form which we interpret as the electromagnetic field strength.

3. *One can now check that isomorphism classes of line bundles with connection are in bijection to the classes of $\hat{H}^2(M)$ defined before.*

We now put the differential cohomology classes in $\hat{H}^2(M)$ into context:

Observation 3.4.4.

1. *We have already constructed above a map*

$$\begin{array}{rcl} F : \hat{H}^2(M) & \to & \Omega^2_{\mathrm{closed}}(M) \\ (A_\alpha, g_{\alpha\beta}) & \mapsto & F \end{array}$$

with values in closed two-forms which we call "field strength" or "curvature". Its kernel are so-called flat *connections. Their isomorphism classes are classified by the cohomology group $H^1(M, \mathbb{R}/\mathbb{Z})$ so that we have an exact sequence:*

$$0 \to H^1(M, \mathbb{R}/\mathbb{Z}) \to \hat{H}^2(M) \xrightarrow{F} \Omega^2_{\mathrm{integral}}(M) \to 0\ ,$$

where the field strength turns out to take values in closed 2-forms with integral periods, i.e. the integral $\int_\Sigma\ F$ over each closed oriented two-dimensional submanifold $\Sigma \subset M$ is an integer. This is the quantization of electric flux. It is not a phenomenon of classical electrodynamics seen as a theory of 2-forms.

Roughly, $\hat{H}^2(M)$ is bigger than integral 2-forms by the flat connections and thus provides additional degrees of freedom as compared to a theory of two-forms. They turn out to be measurable measurable in quantum mechanics, e.g. in Aharonov-Bohm effect.

2. *There is a second interesting short exact sequence: the family of* $\mathrm{U}(1)$*-valued functions* $(g_{\alpha\beta})_{\alpha,\beta\in I}$ *is closed,*

$$1 = (\delta g)_{\alpha\beta\gamma} = g_{\alpha\beta} g_{\beta\gamma} g_{\gamma\alpha} \quad \textit{for all triple overlaps,}$$

and thus determines a class in Čech cohomology $H^1(M, \mathrm{U}(1)) \cong H^2(M, \mathbb{Z})$*. This class is called the Chern class of the line bundle and measures its topological non-triviality.*

Since field strength is closed, it determines a cohomology class $[F] \in H^2_{dR}(M)$*. One has a natural map* $H^2(M, \mathbb{Z}) \to H^2_{dR}(M)$ *and it turns out that the image of the Chern class in* $H^2_{dR}(M)$ *is* $[F]$*. This is an observation due to Dirac that in quantum mechanical systems the electric field strength enjoys the integrality condition*

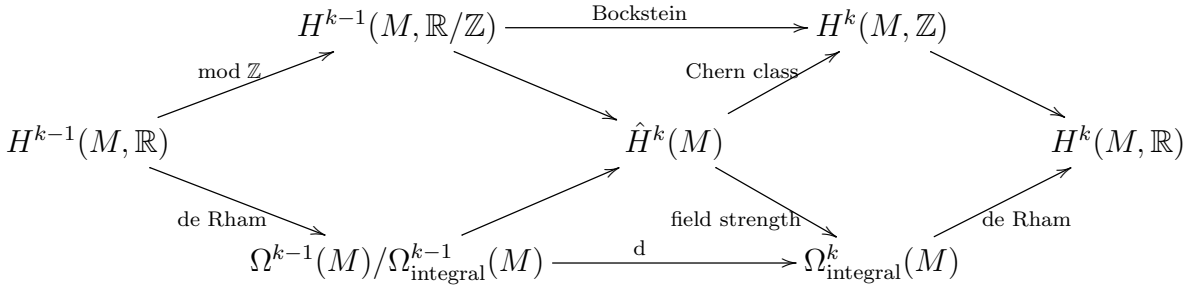
$$[F] \in \mathrm{H}^2(M, \mathbb{Z}) \subseteq \mathrm{H}^2_{\mathrm{dR}}(M, \mathbb{R}) \ .$$

It cannot be derived in classical electrodynamics.

3. *One can finally classify topologically trivial elements of* $\hat{H}^2(M)$*, i.e. elements of trivial Chern class. Topologically trivial bundles are those line bundles with connection, for which the underlying bundle is topologically a product,* $L \cong M \times \mathbb{C}$*. They admit representatives of the connection that are global* 1*-forms. These representatives are, however, not unique, but only determined up to a one-form whose integral over any closed oriented one-dimensional submanifold is an integer. One finds thus the exact sequence*

$$0 \to \Omega^1(M)/\Omega^1_{integral}(M) \to \hat{H}^2(M) \to H^2(M, \mathbb{Z}) \to 0 \ .$$

These two exact sequences fit together into the following diagram with $k = 2$:



containing the The commuting square containing the field strength map and the Chern class shows that differential cohomology $\hat{H}^2(M)$ combines both the cohomology class $H^k(M, \mathbb{Z})$ with *differential* information contained in the integral k-forms, which justifies the name differential cohomology.

Remarks 3.4.5.

1. *Remarkably enough, the cases for* $k \neq 2$ *are quite interesting for both mathematics and physics:*

   - *One can geometrically realize* $\hat{H}^0(M)$ *as smooth maps on* M *with values in* $\mathbb{Z}$*.*
   - *One can geometrically realize* $\hat{H}^1(M)$ *as smooth maps on* M *with values in* $\mathbb{R}/\mathbb{Z}$*. This is a periodic scalar field on* M*. On two-dimensional manifolds, it is a basic ingredient for many constructions in conformal field theory and string theory.*

- *As we have seen, one can geometrically realize $\hat{H}^2(M)$ as line bundles with connection. This is relevant for electrodynamics.*
- *One can geometrically realize $\hat{H}^3(M)$ as so-called abelian bundle gerbes with connection. They provide a natural framework to describe the B-field in string theory.*

2. *Differential cohomology can be pulled back. Consider thus a class $\eta \in \hat{H}^k(M)$ and a smooth map*

$$\varphi : \Sigma \to M$$

with Σ a $(k-1)$-dimensional smooth oriented closed manifold. Then the pullback $\varphi^*\eta$ is a class in $\hat{H}^k(\Sigma)$. Since Σ is k-dimensional, the corresponding Chern class $c_1(\varphi^*\eta)) \in H^k(\Sigma)$ vanishes. Hence the differential object is topologically trivial and by the exactness of the above diagram, it corresponds to a $(k-1)$-form $\tilde{\eta} \in \Omega^{k-1}(\Sigma)$ that is well-defined on all of Σ and determined up to a form with integral periods. To get rid of the latter ambiguity, we define the holonomy by taking the exponential

$$\mathrm{hol}_\varphi(\eta) := \exp(2\pi\mathrm{i}\int_\Sigma \tilde{\eta}) \in \mathrm{U}(1) \ .$$

3. *Let us discuss holonomy in special cases:*
   - *In the case $k = 1$, one has to consider maps of zero-dimensional manifolds to M, i.e. points in M and the holonomy is simply evaluation of the $\mathrm{U}(1)$-valued function in $\hat{H}^1(M)$.*
   - *In the case $k = 2$, one considers curves $\varphi : \ S^1 \to M$ which can be imagined as closed world lines of particles. For a charged particle, one obtains the exponential of a term that can be added to the action that physicists like to write in terms of local quantities as*

$$\exp(2\pi\mathrm{i}\oint_\varphi A_\mu \mathrm{d}s^\mu) \ .$$

   - *In the case $k = 3$, one obtains the notion of a surface holonomy that enters as the exponentiated Wess-Zumino term in the action of two-dimensional sigma models and in particular into world sheet actions of the string. Higher-dimensional objects in string theory, so-called branes, require the case of higher k.*

4. *Holonomy provides a map*

$$\chi : \ Z_{k-1} \to \mathbb{R}/\mathbb{Z}$$

from closed submanifolds (actually, cycles in homology) with values in $\mathbb{R}/\mathbb{Z}$. This map contains information that is equivalent to all the above data.

Remark 3.4.6.
The gauge potential also enters if one wants to obtain the Lorentz force for a charged particle from a lagrangian. Indeed, consider the simplest case of a particle moving in background fields

$$E^i = -\frac{\partial\Phi}{\partial x^i} - \partial_t A^i \quad \text{and} \quad B^i = \mathrm{rot}^i A$$

where we are using three-dimensional notation. Then the equations of motion for the Lagrangian

$$l(x, x_t, t) = \frac{m}{2}x_t^2 - q(x_t^i A^i(x,t) + \Phi(x,y))$$

are, because of

$$\frac{\partial l}{\partial x_t^i} = m x_t^i - q A^i$$

and

$$\frac{\partial l}{\partial x^i} = -q x_t^j \frac{\partial A^j}{\partial x^i} - q \frac{\partial \Phi}{\partial x^i}$$

given by

$$m\ddot{\varphi}^i - q\frac{\partial A^i}{\partial x^j}\dot{\varphi}^j = -q\dot{\varphi}^j \frac{\partial A^j}{\partial x^i} - q\frac{\partial \Phi}{\partial x^i} - q\partial_t A^i$$

which is just

$$m\ddot{\varphi}^i = q(\dot{\varphi} \wedge B)^i + qE^i$$

and thus the Lorentz force.

Let us finally comment on what one might wish to call generalized Maxwell theory.

Remarks 3.4.7.

1. *Let M be a smooth n-dimensional manifold with a (pseudo-)Euclidean metric. We set up a theory of l-forms for $0 \leq l \leq n$. As equations of motion for the l-form $\Omega^l(M)$, we take*

$$\mathrm{d}F = 0 \quad \textit{and} \quad \mathrm{d} * F = 0 \ .$$

(We consider here for simplicity the case without external sources; external source require relative cohomology.) It obviously has waves as classical solutions: any solution obeys $\mathrm{d}F = 0$ and $\delta F = 0$ and thus

$$(\mathrm{d} + \delta)^2 F = (\mathrm{d} + *^{-1}\mathrm{d}*)^2 F = 0$$

and is thus harmonic. Given any solution, one can define its electric flux

$$[F] \in H^l_{dR}(M)$$

and its magnetic flux

$$[*F] \in H^{n-l}_{dR}(M) \ .$$

At this point, there is a perfect symmetry ("duality") between electric and magnetic degrees of freedom and a theory of l-forms and of $(n-l)$-forms cannot be distinguished. Also an energy momentum tensor which is a section in symmetric cotangent fields

$$T_F \in \Gamma(\mathrm{Sym}^2 T^* M)$$

can be introduced by the quadratic form on the vector field $v \in \mathrm{vect}(M)$

$$T_F(v) := \ (\iota_v F, \iota_v F) \ - \frac{1}{2} v^2 F^2 \ .$$

It can be shown to be symmetric under magnetic-electric duality.

2. *If one wants to introduce an action, one has to break electric magnetic duality and use one of the equations, say* $\mathrm{d}F = 0$*, to introduce locally defined gauge potentials and transition functions* $(A_\alpha, g_{\alpha\beta})$*. These are to be considered as the fundamental degree of freedom; the field strength is a function of them,* $F = F[A, g] = \mathrm{d}A$*.*

   The following action is naturally defined:

$$S[A] = g \int_M F[A, g] \wedge *F[A, g]$$

   with g *a constant of the theory and can be shown to lead to the equations of motion* $\mathrm{d} * F[A, g] = 0$*. We have now a gauge symmetry in the sense of Noether's second theorem. For example, for* $n = 4$ *and* $k = 2$*, we have for every function* λ *the symmetry* $S[A + \mathrm{d}\lambda] = S[A]$*. For a general local function, we neglect its dependence on derivatives. Differential relations between the equations of motion are then introduced by the relation* $\mathrm{d}F[A] = 0$*.*

3.5 General relativity

We summarize in this subsection some important aspects of general relativity in the form of several postulates.

Observation 3.5.1.

1. *The mathematical model for space time, i.e. the collection of all events, is a four-dimensional smooth manifold* M *with a Lorentz metric* g*.*

2. *The topological space underlying a manifold is always required to be Hausdorff; it can be shown that the existence of a Lorentz metric implies that the space is paracompact.*

3. *The manifold structure is experimentally well established up to length scales of at least* $10^{-17}m$*.*

4. *For a general Lorentz manifold, non-zero tangent vectors* $X \in T_pM$ *fall in the classes of time-like, space-like or null vectors.*

5. *We say that a space-time* (M', g') *is an extension of* (M, g)*, if there is an isometric embedding* $\iota : (M, g) \to (M', g')$*. We require space-time to be inextendible, i.e. we require that no non-trivial extension exists. Stronger notions of local inextendibility have been discussed in the literature.*

One then includes matter fields which are typically sections in bundles over M. By "matter fields" we understand any potentially present other fields. Actually, this includes e.g. the electromagnetic field and other fields describing non-gravitational interactions which one would not call a matter field in different context.

We discuss the postulates.

First postulate: local causality:
The equations for the matter fields must be such that for any convex neighborhood $U \subset M$ a signal can be sent from $p \in U$ to $q \in U$, if and only if there is a curve in U with time-like or null tangent vectors that joins p and q.

We present a more mathematical formulation of this statement in terms of the so-called Cauchy-problem: let (U, x) be a coordinate neighborhood and $p \in U$ such that every non space-like curve through p intersects the points with $x^0 = 0$ in U. Let F be the subset of points in $x^0 = 0$ which can be reached by non space-like curves in U from p. We then require that the values of matter fields in p are uniquely determined by their values and the values of their derivatives up to finite order in F and that they are not determined by the values of any proper subset of F.

We add the comment that the causal structure, i.e. the collection of the light cones determines the metric up to a conformal factor.

Lemma 3.5.2.
Let M be a smooth manifold and $g^{(1)}$ and $g^{(2)}$ two Lorenz metrics on M with the same set of light cones. Then there is a smooth function $\lambda \in C^\infty(M, \mathbb{R}_+)$ with values in the positive real numbers such that $g^{(1)}_p = \lambda(p) \cdot g^{(2)}_p$ for all $p \in M$.

Proof:

- Let $X \in T_pM$ a time-like and $Y \in T_pM$ a space-like vector. The quadratic equation in λ

$$0 = g(X + \lambda Y, X + \lambda Y) = g(X, X) + 2\lambda g(X, Y) + \lambda^2 g(Y, Y)$$

  has two roots

$$\lambda_1, \lambda_2 = \frac{g(X, Y)}{g(Y, Y)} \pm \sqrt{\frac{g(X, Y)^2}{g(Y, Y)^2} - \frac{g(X, X)}{g(Y, Y)}}$$

  which are real since $g(Y, Y) > 0$ and $g(X, X) < 0$. We have

$$\lambda_1 \cdot \lambda_2 = \frac{g(X, X)}{g(Y, Y)}$$

  so that the ratio of the magnitudes of a space-like and a time-like vector can be derived from the light cone.

- We fix X and Y as above. Now suppose that $W, Z \in T_pM$ and $W + Z$ are not light-like. Then

$$g(W, Z) = \frac{1}{2}\left(g(W, W) + g(Z, Z) - g(W + Z, W + Z)\right) \ .$$

  Each of the terms on the right hand side can be compared to either X or Y and is thus fixed. If either of the vectors W, Z and $W + Z$ is a null vector, then work with other linear combinations like $W + 2Z$.

□

Second postulate: local conservation of energy momentum:
The equations governing the matter fields should be such that there is a symmetric tensor $T \in \Gamma(\mathrm{Sym}^2 T^*M)$ which depends on the fields and their covariant derivatives (with respect to the Levi-Cività connection) and which has the two properties:

- T vanishes on an open subset U, if and only if all matter field vanish on U.
  This principle expresses the fact that all matter fields have (positive) energy.
- T obeys the conservation equation $T^{ab}_{;b} = 0$ where the subscript indicates covariant differentiation with respect to the Levi-Cività connection.

If the equations of motion for the matter fields are given by the variation of an action I for a lagrangian density, there is a distinguished candidate for the energy momentum tensor.

Consider a classical Lagrangian field theory on the Lorentz manifold M with Lagrangian density l. This density typically depends on the metric g on M, e.g. via the volume form dvol_g associated to the metric. Consider a variational family of metrics $g(u)$ on M. We assume that the action can be written down for all metrics involved. We may then write

$$\frac{\partial I}{\partial u} = \frac{1}{2}\int T^{ab}(\psi, \partial\psi, \ldots)\frac{\partial g_{ab}}{\partial u}\big|_{u=0}$$

with some symmetric tensor field T whose components are local functions.

Proposition 3.5.3.
The tensor T is a symmetric covariantly conserved tensor: if evaluated on any field configuration ψ obeying the equations of motion implied by a Lagrangian function L, one has $T^{ab}(\psi, \partial\psi, \ldots)_{;b} = 0$.

Proof:
We have for any diffeomorphism $\Phi : M \to M$ that restricts to the identity outside a compact subset $D \subset M$ for the Lagrangian density $l = L\mathrm{dvol}_g$

$$I \ = \int_D l = \int_{\Phi(D)} l = \int_D \ \Phi^*(l)$$

and thus

$$\int_D l - \Phi^*(l) = 0 \ .$$

If the diffeomorphism Φ is generated by a vector field X, we have

$$\int_D \ L_X l = 0 \ .$$

We find

$$\int_D \ L_X(l\mathrm{dvol}_g) = \int_D \ \left(\frac{\partial L}{\partial \psi} - D_e(\frac{\partial L}{\partial \psi_e})\right) L_X\psi \mathrm{dvol}_g + \frac{1}{2}\int_D T^{ab} L_X g_{ab}\mathrm{dvol}_g \ .$$

The first term vanishes due to the equations of motion. For the second term, we need the Lie derivative of the metric

$$L_X g_{ab} = 2X_{(a;b)}$$

which equals the symmetrization of the covariant derivative.

Thus

$$0 = \int_D \ T^{ab} L_X g_{ab}\mathrm{dvol}_g = 2\int_D \left((T^{ab} \ X_a)_{;b} - \ T^{ab}_{;b} X_a\right)\mathrm{dvol}_g \ .$$

The first term can be transformed into an integral over ∂D which vanishes since the vector field X vanishes on ∂X. Since this identity holds for all vector fields X, we have covariant conservation of the energy momentum tensor, $T^{ab}_{;b} = 0$.

□

We have finally to formulate a dynamical principle for the metric. We take the metric as the expression of gravity and any potentially present other fields as "matter fields". The equations should relate the metric to the distribution of matter fields.

We generalize the Newtonian principle that active mass (i.e. the mass producing a gravitational field) equals passive mass (i.e. the mass that experiences gravity) by saying that all matter should influence the metric only via the energy-momentum tensor which is its response to the metric which is thus a generalization of passive mass.

We have to built a tensor from the metric g that has the same properties as the energy momentum tensor, i.e. it should be symmetric and covariantly conserved, since

$$T^{ab}_{;b} = 0 \ .$$

In a conservative spirit, we keep the feature of Newtonian theories that only the metric g and its first and second derivatives should enter the equations. (In effective theories, this feature is frequently generalized.)

This is in principle a problem of invariant theory. The most general tensor of this type formed from the metric can be expressed in terms of the Ricci curvature $R_{ab}[g]$ of the metric and its trace, the scalar curvature, $R = R^a_a$. We arrive at the following covariantly conserved symmetric tensor:

$$(R_{ab}[g] - \frac{1}{2}R[g]g_{ab}) + \Lambda g_{ab} \ ,$$

where $\Lambda \in \mathbb{R}$ is the so-called cosmological constant that has to be determined experimentally. (It was suspected to be zero for a long time.)

Third postulate: field equation:
Einstein's equations read:

$$(R_{ab}[g] - \frac{1}{2}R[g]g_{ab}) + \Lambda g_{ab} = \frac{8\pi G}{c^4}T_{ab} \ ,$$

where G and Λ are constants of nature.

Remarks 3.5.4.

1. *This equation can be compared for so-called static space times, i.e. space times with a time-like Killing field that is orthogonal to a family of space-like surfaces in a certain limit to Newtonian gravity. The constant G then becomes Newton's constant.*

2. *The equations can be derived from the so-called Einstein-Hilbert action*

$$S[g, \Phi] = \int \frac{1}{16\pi}(R[g] - 2\Lambda)\mathrm{vol}_g + l(\Phi, g) \ ,$$

where Φ stands symbolically for other fields in the theory.

4 Hamiltonian mechanics

4.1 (Pre-) symplectic manifolds

We now present different formulation of mechanical systems that turns out to be a good starting point for quantum mechanics as well. So far, our focus was on (partial) differential equations of arbitrary finite order. We now turn to geometric structures that are adapted to described first order *ordinary* differential equations. This will enable us to make *quantitative* statements about mechanical systems.

We start with some notions from linear algebra. We restrict to vector spaces of finite dimension.

Definition 4.1.1
Let k be a field of characteristic different from 2. A symplectic vector space is a pair (V, ω) consisting of a k-vector space V and a non-degenerate two-form $\omega \in \Lambda^2 V$, i.e. an antisymmetric bilinear map $\omega : V \times V \to k$ such that $\omega(v, w) = 0$ for all $w \in V$ implies $v = 0$.

Remarks 4.1.2.

1. *Symplectic vector spaces have even dimension.*
2. *To present a standard example, consider $V = \mathbb{R}^{2n}$ with the standard basis $(e_i)_{i=1,\dots 2n}$ and its dual basis $(e_i^*)_{i=1,\dots 2n}$. Then*
$$\omega = e_1^* \wedge e_2^* + e_3^* \wedge e_4^* + \ldots + e_{2n-1}^* \wedge e_{2n}^* \in \Lambda^2(\mathbb{R}^{2n})$$
is a symplectic form.
3. *Given any finite-dimensional k-vector space W, the k-vector space $V := W \oplus W^*$ has a canonical symplectic structure given by*
$$\omega((b, \beta), (c, \gamma)) := \beta(c) - \gamma(b) \ .$$
4. *The subset of linear endomorphisms $\varphi \in \mathrm{GL}(V)$ of a symplectic real vector space (V, ω) that preserve ω, i.e. $\varphi^* \omega = \omega$, is a (non-compact) Lie group $\mathrm{Sp}(V)$ of dimension $\frac{\dim V(\dim V+1)}{2}$. Its elements are also called symplectic or canonical maps.*

We now extend these notions of linear algebra to smooth manifolds.

Definition 4.1.3
Let M be a smooth manifold.

1. *A 2-form $\omega \in \Omega^2(M)$ is called a presymplectic form, if it is closed, $\mathrm{d}\omega = 0$, and of constant rank.*
2. *A non-degenerate pre-symplectic form is called a symplectic form.*
3. *A smooth manifold M together with a (pre-)symplectic form ω is called a (pre-)symplectic manifold.*

4. *A morphism $f : (M, \omega_M) \to (N, \omega_N)$ of (pre-)symplectic manifolds is a differentiable map $f : M \to N$ that preserves the (pre-)symplectic form, $f^*\omega_N = \omega_M$. Such morphisms are also called symplectomorphisms or canonical transformations.*

Remarks 4.1.4.

1. *Since for any point p of a symplectic manifold M the tangent space T_pM is a symplectic vector space, the dimension of a symplectic manifold is necessarily even. The dimension of a presymplectic manifold, in contrast, can be odd or even.*

2. *One verifies by direct computation that on a symplectic manifold M of dimension $\dim M = 2n$, the n-th power $\omega^{\wedge n}$ of the symplectic form ω is a volume form on M. It is called the Liouville volume and the induced measure on M is called the Liouville measure.*

3. *A symplectic map between two symplectic manifolds of the same dimension preserves the Liouville volume $\omega^{\wedge n}$. Since volume preserving smooth maps have a Jacobian of determinant 1, they are local diffeomorphisms. Symplectic maps between two symplectic manifolds of the same dimension are thus local diffeomorphisms.*

Examples 4.1.5.

1. *In the Euclidean space $\mathbb{R}^3$ with standard Cartesian coordinates x, y, z, the unit sphere S^2 is the embedded submanifold $\iota : S^2 \to \mathbb{R}^3$, defined the solutions of the equation $x^2 + y^2 + z^2 = 1$. The pullback $\iota^*\omega$ of the differential form $\omega = x\mathrm{d}y \wedge \mathrm{d}z + y\mathrm{d}z \wedge \mathrm{d}x + z\mathrm{d}x \wedge \mathrm{d}y$ endows S^2 with the structure of a symplectic manifold. Polar coordinates $0 < \theta < \pi$ and $0 < \varphi < 2\pi$, the form reads*

$$\iota^*\omega = \sin\theta \mathrm{d}\varphi \wedge \mathrm{d}\theta \ .$$

   More generally, any oriented two-dimensional surface equipped with an area 2-form is a symplectic manifold.

2. *For any smooth manifold M of dimension n, the cotangent bundle T^*M has a natural structure of a symplectic manifold of dimension $2n$.*

   We explain this in more detail: local coordinates $(x^i)_{i=1\ldots n}$ on M give local coordinates (x^i, p_i) for the total space of the cotangent bundle T^*M, by describing a local one-form λ_x in the point $x = (x^i)$ as a linear combination

$$\lambda_x = \sum_{i=1}^{n} p_i \mathrm{d}x^i \quad \textit{with} \quad p_i \in \mathbb{R} \ .$$

   Under changes of local coordinates, we have

$$\mathrm{d}\tilde{x}^i = \sum_{j=1}^{n} \frac{\partial \tilde{x}^i}{\partial x^j} \mathrm{d}x^j$$

   and thus

$$\tilde{p}_i = \frac{\partial x^j}{\partial \tilde{x}^i} p_j \ .$$

Consider an atlas $(U_\alpha)_{\alpha\in I}$ for a smooth manifold M together with the induced atlas $(\tilde{U})_{\alpha\in I}$ of the total space T^*M. Consider the locally defined two-form on the total space T^*M of the cotangent bundle

$$(\omega_0)_\alpha := \sum_{i=1}^n \mathrm{d}p_i \wedge \mathrm{d}x^i \in \Omega^2(\tilde{U}_\alpha) \ .$$

It is clear that this two-form is non-degenerate.

This two-form is independent of the choice of local coordinates: if $\tilde{x}^i$ are different local coordinates, we find

$$\sum_{i=1}^n \mathrm{d}\tilde{p}_i \wedge \mathrm{d}\tilde{x}^i = \sum_{i,j,k=1}^n \frac{\partial \tilde{x}^i}{\partial x^k}\frac{\partial x^j}{\partial \tilde{x}^i}\mathrm{d}p_j \wedge \mathrm{d}x^k = \omega_0 \ .$$

As a consequence, the locally defined two-forms patch together to a globally defined two-form on the cotangent bundle T^*M. One easily verifies that the two-form ω_0 is non-degenerate.

3. *To see that the two-form $\omega_0 \in \Omega^2(T^*M)$ is closed, we remark that locally on each coordinate neighborhood U_α for T^*M, we can introduce the one-form*

$$\Theta_\alpha := \sum_{i=1}^n p_i \mathrm{d}x^i \in \Omega^1(\tilde{U}_\alpha) \ .$$

Its derivative is

$$\mathrm{d}\Theta_\alpha = \sum_{i=1}^n \mathrm{d}p^i \wedge \mathrm{d}q_i = \omega_0 \ ;$$

hence the two-form ω_0 is closed and thus a symplectic form on the cotangent bundle T^*M.

4. *More invariantly, consider the bundle projection $\pi : T^*M \to M$ of the cotangent bundle and its differential*

$$\pi_* T(T^*M) \to TM \ .$$

Define a one-form on T^*M which acts at the point $w^* \in T^*M$ on $v \in T_{w^*}(T^*M)$ by

$$\theta_{w^*}(v) \ := w^*(\pi_*(v)) \ .$$

Writing this in local coordinates (x, p) for T^*M

$$v = \sum_{i=1}^n \xi^i \frac{\partial}{\partial x^i} + \eta_i \frac{\partial}{\partial p_i}$$

we find from $\pi(x, p) = x$

$$\pi_*(v) \ = \sum_{i=1}^n \xi^i \frac{\partial}{\partial x^i}$$

and with $w^* = (x^i, p_i) = \sum_{i=1}^n p_i \mathrm{d}x^i$, we get

$$\theta_{w^*}(v) = w^*(\pi_*(v)) \ = \sum_i p_i \xi^i = \theta_{(x,p)}(v) \ ,$$

so that the one-form θ is even globally defined. It is called the canonical one-form of the cotangent bundle.

5. *One can show that for any diffeomorphism $f : M \to N$, the induced map*

$$T^*f : T^*N \to T^*M$$

is symplectic.

Definition 4.1.6
Let M be a smooth manifold of any dimension. The symplectic manifold (T^*M, ω_0) is called the canonical phase space associated to the configuration space manifold M.

Remarks 4.1.7.

1. *Since the canonical phase space comes with the canonical one-form θ, the symplectic form is in this case not only closed, but even exact.*

2. *The canonical phase space is a non-compact symplectic manifold. Examples of* compact *symplectic manifolds are quite important for mathematical physics, but more difficult to obtain.*

The following result gives crucial insight into peculiarities of symplectic geometry:

Theorem 4.1.8 (Darboux' theorem)**.**
Let (M, ω) be a $(2n + k)$-dimensional presymplectic manifold with $\operatorname{rank}\omega = 2n$.
Then we can find for any point $m \in M$ a neighborhood U and a local coordinate chart

$$\psi : U \to \mathbb{R}^{2n+k}$$
$$\psi(u) = \left(q^1, ..., q^n, p_1, ..., p_n, \eta^1, ..., \eta^k\right)$$

such that $\omega\big|_U = \sum_{i=1}^n \mathrm{d}p_i \wedge \mathrm{d}q^i$. Such coordinates are called Darboux coordinates or canonical coordinates. One can choose the covering such that the coordinate changes are canonical transformations.

Proof:
We present the proof for the symplectic case only.

- Since we have to prove a local statement, we can assume without loss of generality that we consider a symplectic form ω that is defined on a neighborhood of 0 in $\mathbb{R}^{2n}$. We can also assume that at 0, the symplectic form on the tangent space $T_0\mathbb{R}^{2n}$ in 0 has been brought to the standard form

$$\omega_0 = \sum_{j=0}^{n} \mathrm{d}x_{2i-1} \wedge \mathrm{d}x_{2i} \ .$$

- We will construct a diffeomorphism φ defined in a neighborhood of 0 such that

$$\varphi^*\omega = \omega_0 \ .$$

  Using φ as a local coordinate system for the neighborhood of zero in $\mathbb{R}^{2n}$, we have then found a coordinate system in which the symplectic form is constant.

- Consider the one-parameter family of closed two-forms

$$\omega^t := \omega_0 + t(\omega - \omega_0)$$

interpolating linearly between the constant two-form ω_0 and ω. We will construct a family of diffeomorphisms φ^t such that

$$(\varphi^t)^*\omega^t = \omega_0$$

and then take $\varphi := \varphi_1$. Differentiating with respect to t, we find that the family X^t of vector fields corresponding to the family of diffeomorphisms we are looking for has to obey

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}(\varphi^t)^*\omega^t = (\varphi^t)^*(L_{X^t}\omega^t + \frac{\mathrm{d}}{\mathrm{d}t}\omega^t) \ .$$

Using Cartan's formula and the fact that ω is closed, we see that this is equivalent to finding vector fields X^t that have to obey

$$0 = (\varphi^t)^*(\mathrm{d}\iota_{X^t}\omega^t + \omega - \omega_0) \ .$$

Thus the one-parameter family of vector fields X_t we are looking for has to obey the equation

$$\mathrm{d}\iota_{X^t}\omega^t + \omega - \omega_0 \ . \qquad (*)$$

- We now find a solution to this equation. Since the two-form $\omega - \omega_0$ is closed, it is locally exact. We can thus find a one-form β such that

$$\mathrm{d}\beta = \omega - \omega_0 \ .$$

Since β is only determined up to an additive constant, we can assume that $\beta_0 = 0$.

Since $\omega_0^t = \omega_0$ for all $t \in [0,1]$, we can find a neighborhood W of 0 such that ω_y^t is non-degenerate for all $w \in W$ and $t \in [0,1]$. We can thus find a unique family of vector fields X^t such that

$$\iota_{X^t}\omega^t = -\beta \quad \text{on} \quad W \ .$$

This shows that the family X^t solves equation $(*)$.

□

Remarks 4.1.9.

1. *Darboux' theorem implies that symplectic geometry is locally trivial - in contrast to Riemannian geometry where curvature provides local invariants, and where in Riemannian coordinates the metric can be brought to a standard form in one point only. In other words, locally two symplectic manifolds are indistinguishable.*

2. *While any manifold can be endowed with a Riemannian structure, there are manifolds which cannot be endowed with a symplectic structure. For example, there is no symplectic structure on the spheres S^{2n} for $n > 1$.*

As in the case of a metric, a symplectic form allows us to relate differential forms, like the total differential of a smooth function, to vector fields. We thus have:

Proposition 4.1.10.
Let (M,ω) by a symplectic manifold and $f \in C^\infty(M,\mathbb{R})$ be a smooth function. Then there is a unique smooth vector field X_f on M such that

$$\mathrm{d}f(Y) = \omega(X_f, Y) \quad \text{for all local vector fields } Y \text{ on } M,$$

or, equivalently, $\mathrm{d}f = \iota_{X_f}\omega$.

Definition 4.1.11

1. *Let f be a smooth function on a symplectic manifold (M,ω). The vector field X_f such that $\mathrm{d}f = \iota_{X_f}\omega$ is called the symplectic gradient of f.*

2. *A vector field X on a symplectic manifold (M,ω) is called a Hamiltonian vector field, if there is a smooth function f such that $X = X_f$. Put differently, a vector field X is Hamiltonian, if there is a function f such that $\iota_X\omega = \mathrm{d}f$, i.e. if the one-form $\iota_X\omega$ is exact. The function f is called a Hamiltonian function for the vector field X.*

3. *A vector field X on a symplectic manifold (M,ω) is called locally Hamiltonian, if the one-form $\iota_X\omega$ is closed.*

Remarks 4.1.12.

1. *The kernel of map given by the symplectic gradient to the subspace* $\mathrm{HamVect}(M)$ *of Hamiltonian vector fields*

$$\begin{array}{rcl} C^\infty(M,\mathbb{R}) & \to & \mathrm{HamVect}(M)) \\ f & \mapsto & X_f \end{array}$$

   are the locally constant functions on M.

2. *Since the symplectic form ω is closed, Cartan's formula gives for the Lie derivative with respect to any vector field X*

$$L_X\omega = \ (\iota_X\mathrm{d} + \mathrm{d}\iota_X)\omega = \mathrm{d}\iota_X\omega \ .$$

   A vector field X is thus locally Hamiltonian, i.e. $\iota_X\omega$ is closed, if and only if the Lie derivative $L_X\omega = 0$ vanishes.

3. *A vector field X on a symplectic manifold (M,ω) is locally Hamiltonian, if the family $\varphi_t : M \to M$ of diffeomorphisms of M associated to the vector field is a symplectic transformation for each t.*

   Suppose that the diffeomorphisms are symplectic, i.e. $\varphi_t^*\omega = \omega$. This immediately implies $L_X\omega = 0$ and thus by the previous remark $\mathrm{d}\iota_X\omega = 0$.

4. *To compute the symplectic gradient in local Darboux coordinates (q^i, p_i) defined on $U \subset M$, we make the ansatz*

$$X_f = \sum_{i=1}^{n} \xi^i \frac{\partial}{\partial q^i} + \tilde{\xi}_i \frac{\partial}{\partial p_i} \in \mathrm{vect}(U)$$

with local coordinate functions ξ^i and $\tilde{\xi}_i$. We evaluate both sides of the defining equation on a local vector field

$$Y = \sum_{i=1}^{n} \eta^i \frac{\partial}{\partial q^i} + \tilde{\eta}_i \frac{\partial}{\partial p_i} \in \mathrm{vect}(U) \ .$$

Since we work in Darboux coordinates, we find

$$\omega(X_f, Y) = \sum_{i=1}^{n} \left(\tilde{\xi}_i \eta^i - \xi^i \tilde{\eta}_i \right)$$

which has to equal

$$\mathrm{d}f(Y) = \sum_{i=1}^{n} \frac{\partial f}{\partial q^i} \eta^i + \frac{\partial f}{\partial p_i} \tilde{\eta}_i$$

The comparison of the coefficients of η^i and $\tilde{\eta}_i$ yields for the symplectic gradient the local expression

$$X_f = \sum_{i=1}^{n} -\frac{\partial f}{\partial p_i} \frac{\partial}{\partial q^i} + \frac{\partial f}{\partial q^i} \frac{\partial}{\partial p_i} \in \mathrm{vect}(U) \ .$$

We now present an important characterization of locally Hamiltonian vector fields:

Proposition 4.1.13.
Let (M, ω) be a symplectic manifold. The flow of a vector field X consists of symplectic transformations, if and only if the vector field X is locally Hamiltonian.

In particular, the flow of a Hamiltonian vector field preserves the Liouville volume $\omega^{\wedge n}$.

This raises the question of how big the subgroup $\mathrm{Ham}(M)$ of symplectic transformations generated by Hamiltonian vector field is inside the group of $\mathrm{Symp}(M)$ of all symplectic transformations - or rather its connected component $\mathrm{Symp}_0(M)$ containing the identity. If the manifold M is closed and its first de Rham cohomology vanishes, then the two groups coincide. For general closed manifolds, $\mathrm{Ham}(M)$ is the commutator subgroup of $\mathrm{Symp}_0(M)$.

We have to show a local statement. Due to Darboux' theorem, this reduces to a statement on open subsets of symplectic vector spaces. For the proof, we thus consider the following special situation: Let (V, ω) be a symplectic vector space which we can consider as a symplectic manifold with constant symplectic form. We canonically identify the tangent space $T_pV \cong V$ for all points $p \in V$. A linear vector field is then a linear map

$$v \mapsto Av$$

with $A \in \mathrm{GL}(V)$,

Lemma 4.1.14.
Let X be any vector field on a symplectic vector space (V, ω). If X is Hamiltonian, then the linear map $DX_v : V \to V$ is skew symplectic for all points $v \in V$.

Proof:
Let X be a Hamiltonian vector field with Hamiltonian function h. By definition, we have in every point $v \in V$

$$\omega(X_v, w) = \mathrm{d}h_v(w) \quad \text{for all} \quad w \in V \ .$$

We differentiate both sides as a function of $v \in V$ in the direction of u and find

$$\omega(DX_v(u), w) \ = D^2 h_v(u, w) \ .$$

Here $D^2 h_v$ is the Hessian matrix of second derivatives in the point v. Since second derivatives are symmetric, this expression equals

$$D^2 h_v(u, w) = D^2 h_v(w, u) = \omega(DX_v(w), u) \ = -\omega(u, DX_v(w)) \ .$$

□

Lemma 4.1.15.
A linear vector field is Hamiltonian, if and only if it is skew symplectic, i.e. if and only if

$$\omega(v, Aw) = -\omega(Av, w) \quad \textit{for all} \quad v, w \in V \ .$$

Proof:

- Let X be a Hamiltonian vector field with Hamiltonian function h. In the previous lemma, we have already shown the identity

  $$\omega(DX_v(u), w) \ = -\omega(u, DX_v(w)) \ .$$

  To differentiate the linear vector field $X_v = Av$ in the direction of w, we note

  $$DX_v(u) \ = \lim_{t \to 0} \frac{A(v + tu) - Av}{t} = \ Au \ .$$

  Inserting this result, we find

  $$\omega(Au, w) = \omega(DX_v(u), w) = -\omega(u, DX_v(w)) = -\omega(u, Aw)$$

  for all $u, w \in V$ so that the endomorphism A is skew symplectic.

- Conversely, let A be skew symplectic. We introduce the function

  $$h(v) \ = \frac{1}{2}\omega(Av, v)$$

  on V. We claim that then the linear vector field $X_v = Av$ is the symplectic gradient of the function h and thus symplectic. Indeed,by the Leibniz rule for the blinear pairing given by $\omega(A\cdot, \cdot\rangle$, we find

  $$\langle \mathrm{d}h_v, u\rangle = \frac{1}{2}\left(\omega(Au, v) + \omega(Av, u)\right) = \frac{1}{2}\left(-\omega(u, Av) + \omega(Av, u)\right) = \omega(Av, u) \ .$$

□

Lemma 4.1.16.
Let X be any vector field on a symplectic vector space (V,ω). Then X is hamiltonian, if and only if the linear map $DX_v:\ V\to V$ is skew symplectic for all points $v\in V$.

Proof:

- Assume that the vector field X is Hamiltonian. Then the statement that DX_v is skew symmetric has already been shown.
- Conversely, suppose that DX_v is skew symplectic for all $v\in V$. Then consider the function

$$h(v) := \int_0^1 \mathrm{d}\ t\omega(X_{tv}, v)$$

on V. Then h is a Hamiltonian function for X, since we have for all $u\in V$

$$\begin{aligned}
\langle \mathrm{d}h_v, u\rangle &= \int_0^1 \mathrm{d}t\ \omega(DX_{tv}(tu), v) + \omega(X_{tv}, u)\\
&= \int_0^1 \mathrm{d}t\ \omega(tDX_{tv}(u), v) + \omega(X_{tv}, u)\\
&= \omega\left(\int_0^1 \mathrm{d}t\ (tDX_{tv}(v) + X_{tv}, u\right)\\
&= \omega\left(\int_0^1 \mathrm{d}t\frac{\mathrm{d}}{\mathrm{d}t}(tX_{tv}), u\right) = \omega(X_v, u)
\end{aligned}$$

□

Proof:
of the proposition. For simplicity, we restrict to linear vector fields; the general case follows by similar calculations as in the previous lemma.

- Assume that the vector field X is linear, $X_v = Av$. We have to show that then the flow consists of linear symplectic transformations φ_t, if and only if A is skew symplectic. For any pair $v, w\in V$, we compute

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=t_0}\varphi_t^*\omega(v,w) &= \frac{\mathrm{d}}{\mathrm{d}t}\omega(\varphi_t v, \varphi_t w)\\
&= \omega(\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t v, \varphi_t w) + \omega(\varphi_t v, \frac{\mathrm{d}}{\mathrm{d}t}\varphi_t w)\\
&= \omega(Av, w)\ + \omega(v, Aw)
\end{aligned}$$

where we used $\varphi_{t_0} = \mathrm{id}$. This shows the claim immediately for linear vector fields.

□

We comment on the relation between between symplectic and presymplectic vector space. We start with linear algebra.

Observation 4.1.17.
Let (V,ω) be a presymplectic vector space. Consider the linear subspace

$$\ker\,\omega = \{v \in V \mid \omega(v,-) = 0\} = \{v \in V \mid \iota_v\omega = 0\} \subset V\ .$$

On the quotient vector space $\bar{V}$, the canonical projection

$$\begin{array}{rccc} \pi: & V & \to & V/\ker\omega =: \bar{V} \\ & v & \mapsto & [v] \end{array}$$

yields a well-defined symplectic structure with

$$\bar{\omega}([v],[w]) \ = \omega(v,w)$$

where v, w are any representatives of $[v]$ and $[w]$, i.e. $\pi(v) = [v]$ and $\pi(w) = [w]$.

It is obvious to ask whether a similar procedure exists for presymplectic manifolds.

Observation 4.1.18.

- *Let (M,ω) be a presymplectic manifold. For every point $p \in M$, consider the linear subspace*

  $$\ker_p\omega \subset T_pM$$

  which gives a distribution on M, i.e. a collection of linear subspaces

  $$\ker\omega = \bigcup_{p\in M} \ker_p\omega \subset TM$$

  which, by the assumptions on a presymplectic manifold has constant rank.

- *Consider two two vector fields $V, W \in \mathrm{vect}(M)$ with values in $\ker\omega$, i.e. $\iota_V\omega = \iota_W\omega = 0$. Their Lie bracket obeys*

  $$\iota_{[V,W]}\omega = L_V\iota_W\omega - \iota_W L_V\omega = L_V\iota_W\omega - i_W(\mathrm{d}\iota_V + \iota_V\mathrm{d})\omega = 0\ ,$$

  where we used $\iota_V\omega = \iota_W\omega = 0$ and the fact that ω is closed. Thus the Lie bracket $[V,W]$ takes values in $\ker\omega$ as well. One says that the distribution $\ker\omega$ is integrable.

- *This allows us to apply a theorem of Frobenius that asserts that there is a <u>foliation</u> of M, i.e. M can be written as a disjoint union of submanifolds $\{\mathcal{L}_\alpha\}_{\alpha\in A}$, called the <u>leaves</u> of the foliation, whose tangent spaces are just the subspaces $\ker\omega$, i.e. for $p \in \mathcal{L}_\alpha$*

  $$T_p\mathcal{L}_\alpha = \ker_p\omega\ .$$

- *In general, the set of leaves $U_M = M/\ker\omega$ has no natural structure of a smooth manifold. There are, however, cases, when this is true. One sufficient condition is the existence of local <u>slices</u>: for every point $p \in M$, one can find a submanifold Σ_p that intersects every leaf at most once and whose tangent space complements $\ker\omega$ in every point $q \in \Sigma$,*

  $$T_qM = \ker_q\omega \oplus T_q\Sigma\ .$$

- *If the space of leaves U_M happens to be a manifold, it has dimension $\dim M - \dim\ker\omega$. In this case, U_M carries a symplectic structure ω_U such that the projection*

  $$\pi: \ M \to M/\ker\omega = U_M$$

  is a morphism of presymplectic manifolds, $\pi^*\omega_U = \omega$.

We will use these structures to study time-dependent mechanical systems.

4.2 Poisson manifolds and Hamiltonian systems

We use the symplectic gradient X_f that is associated to any smooth function f to endow the (commutative) algebra of smooth functions on a symplectic manifold with additional algebraic structure.

Definition 4.2.1
Let k be a field of characteristic different from two.

1. *A Poisson algebra is a k vector space P, together with two bilinear products $\cdot$ and $\{-,-\}$, with the following properties*

   - *$(P,\cdot)$ is an associative algebra.*
   - *$(P,\{-,-\})$ is a Lie algebra.*
   - *The bracket $\{\cdot,\cdot\}$ provides for each $x \in P$ a derivation on P for the associative product:*
   $$\{x, y\cdot z\} = \{x,y\}z + y\{x,z\}.$$
   The product $\{\cdot,\cdot\}$ is also called a Poisson bracket. Morphisms of Poisson algebras are k-linear maps $\Phi : P \to P'$ that respect the two products, $\{\Phi(v),\Phi(w)\} = \Phi(\{v,w\})$ and $\Phi(v)\cdot\Phi(w) = \Phi(v\cdot w)$ for all $v, w \in P$.

2. *A Poisson manifold is a smooth manifold M with a Lie bracket*
   $$\{-,-\}: \quad C^\infty(M)\times C^\infty(M) \to C^\infty(M)$$
   such that the algebra of smooth functions $(C^\infty(M),\{-,-\})$ together with the pointwise product of functions is a (commutative) Poisson algebra. A morphism of Poisson manifolds is a smooth map $\Phi : M \to M'$ such that the linear map $\Phi^* : C^\infty(M') \to C^\infty(M)$ is a morphism of Poisson algebras. (It suffices to check that it preservers the Poisson bracket.)

Examples 4.2.2.

1. *Every associative algebra $(A,\cdot)$ together with the commutator*
   $$[x,y] := x\cdot y - y\cdot x$$
   has the structure of a Poisson algebra. This Poisson bracket is trivial, if the algebra is commutative.

2. *Consider the associative commutative algebra $A = C^\infty(M)$ of smooth functions on a smooth symplectic manifold (M,ω). We use the symplectic gradient to define the following Poisson bracket*
   $$\{f,g\} := -\langle \mathrm{d}g, X_f\rangle = \omega(X_f, X_g)$$
   for $f, g \in A$. From the last equation in observation 4.1.11, we find in local Darboux coordinates
   $$\{f,g\} = \mathrm{d}g(X_f) = \frac{\partial g}{\partial p_i}\frac{\partial f}{\partial q^i} - \frac{\partial g}{\partial q^i}\frac{\partial f}{\partial p_i}\ .$$

One should verify that this really defines a Poisson bracket on $C^\infty(M,\mathbb{R})$. More generally, one has in local coordinates

$$\{f,g\} = \omega^{ij}\partial_i g \partial_j f \ ,$$

where ω^{ij} is the inverse of the symplectic form. The corresponding tensor for a general Poisson manifold need not be invertible.

3. *One can show that a diffeomorphism $\Phi : M \to N$ of symplectic manifolds is symplectic, if and only if it preserves the Poisson bracket, i.e.*

$$\{\Phi^* f, \Phi^* g\} = \Phi^*\{f,g\} \quad \textit{for all} \quad f,g \in C^\infty(N,\mathbb{R}) \ .$$

Proposition 4.2.3.
For any symplectic manifold (M,ω), the symplectic gradient provides a morphism of Lie algebras from the Lie algebra $C^\infty(M,\mathbb{R})$ of smooth functions with the Poisson structure to the Lie algebra $\mathrm{vect}(M)$ *of vector fields, i.e. for any smooth functions $f,g \in C^\infty(M,\mathbb{R})$, we have*

$$X_{\{f,g\}} = [X_f, X_g] \, .$$

As a consequence, Hamiltonian vector fields from a Lie subalgebra of the Lie algebra of all vector fields.

Proof:
Using Cartan's formulae, we get with the symplectic gradients X_f and X_g

$$\begin{aligned} \iota_{[X_f,X_g]}\omega &= L_{X_f}\iota_{X_g}\omega + \iota_{X_g}L_{X_f}\omega \\ &= \mathrm{d}\iota_{X_f}\iota_{X_g}\omega + \iota_{X_f}\mathrm{d}\iota_{X_g}\omega \\ &= \mathrm{d}(\omega(X_g,X_f)) = \ \iota_Z\omega \end{aligned}$$

where Z is the Hamiltonian vector field corresponding to the function $\omega(X_g,X_f) = \{f,g\} \in C^\infty(M,\mathbb{R})$. In the first equality, we use that $L_{X_f}\omega = 0$ for the Hamiltonian vector field X_f. In the third equality, we use $\mathrm{d}\iota_{X_g}\omega = \mathrm{d}\mathrm{d}g = 0$. Since the symplectic form ω is non-degenerate, this implies $Z = [X_f, Y_g]$.

□

Observation 4.2.4.

1. *We discuss Poisson manifolds in more detail. For any smooth function $h \in C^\infty(M)$ on a Poisson manifold M, the map*

$$\begin{aligned} \{h,-\}: \ C^\infty(M) &\to C^\infty(M) \\ f &\mapsto \{h,f\} \end{aligned}$$

is a derivation and thus provides a global vector field $X_h \in \mathrm{vect}(M)$ with

$$X_h(f) \ \equiv \mathrm{d}f(X_h) = \{h,f\} \ .$$

We can consider its integral curves; thus any function on a Poisson manifold gives a first order differential equation and thus some "dynamics".

2. *Let $\varphi : I \to M$ be an integral curve for the Hamiltonian vector field X_h, i.e.*

$$\frac{\mathrm{d}\varphi}{\mathrm{d}t}\Big|_{t=t_0} = X_h(\varphi(t_0)) \ .$$

Let $f \in C^\infty(M, \mathbb{R})$ be a smooth function on M. Then $f \circ \varphi : I \to \mathbb{R}$ is a real-valued smooth function on I. We compute its derivative:

$$\frac{\mathrm{d}}{\mathrm{d}t} f \circ \varphi(t) = \mathrm{d}f(\frac{\mathrm{d}\varphi}{\mathrm{d}t}) = \mathrm{d}f(X_h)\big|_{\varphi(t)} = \{h, f\}(\varphi(t)) \ .$$

For this equation, the short hand notation

$$\dot{f} = \{h, f\}$$

is in use.

This motivates the following definition which provides an alternative description of classical mechanical systems with time-independent configuration space in terms of a single real-valued function.

Definition 4.2.5

1. *A time-independent generalized Hamiltonian system consists of a Poisson manifold $(M, \{\cdot, \cdot\})$, together with a function*

$$h : M \to \mathbb{R} \ ,$$

called the Hamiltonian function. The manifold M is called the phase space of the system. The integral curves of the Hamiltonian vector field X_h are called the trajectories of the system. The family φ_t of diffeomorphisms associated to the Hamiltonian vector field X_h is called the phase flow of the system. The algebra of smooth functions on M is also called the algebra of observables.

2. *A time-independent Hamiltonian system consists of a symplectic manifold (M, ω), together with a function*

$$h : M \to \mathbb{R} \ ,$$

We consider M with the Poisson structure induced by the symplectic structure ω.

Example 4.2.6.
A (time-independent) natural Hamiltonian system is constructed from a smooth Riemannian manifold X and a real-valued function

$$V : X \to \mathbb{R} \ .$$

The relevant symplectic manifold M of the system is the total space of the cotangent bundle, $M = T^*X$ with the canonical symplectic structure ω_0. The metric on X gives a bilinear form $g_q : T_qX \times T_qX \to \mathbb{R}$ on each fibre T_qX of the tangent bundle and, as a consequence, to a dual bilinear form $g_q^* : T_q^*X \times T_q^*X \to \mathbb{R}$ on each fibre T_q^*X of the cotangent bundle.

The Hamiltonian function is

$$\begin{array}{rcl} h : T^*X & \to & \mathbb{R} \\ T_q^*X \ni v & \mapsto & \frac{1}{2} g_q^*(v, v) + V(q) \end{array}$$

We use the structure of Poisson manifolds and symplectic manifolds to derive some statements about Hamiltonian systems:

Remarks 4.2.7.

1. *Since the Poisson bracket is antisymmetric, $\{h,h\}=0$ holds. Thus the Hamiltonian function h itself is a conserved quantity for the Hamiltonian dynamics generated by it. More generally, a smooth function $f\in C^\infty(M,\mathbb{R})$ on a Poisson manifold is conserved, if and only if $\{h,f\}=0$. Two functions f,g on a Poisson manifold are said to Poisson-commute or to be in involution, if $\{f,g\}=0$ holds.*

2. *Suppose that two smooth functions f,g on a Poisson manifold are in involution with the Hamiltonian function h,*

$$\{h,f\}=0 \quad \textit{and} \quad \{h,g\}=0\ .$$

Then the Jacobi identity implies

$$\{h,\{f,g\}\}=-\{g,\{h,f\}\}-\{f,\{g,h\}\}=0\ .$$

Put differently, if f and g are conserved quantities, also the specific combination of their derivatives given by the Poisson bracket is a conserved quantity.

Conserved functions thus form a Lie subalgebra and even a Poisson subalgebra, since the Poisson bracket acts as a derivation:

$$\{h,f\cdot g\}=\{h,f\}\cdot g+f\cdot\{h,g\}=0\ .$$

The Poisson subalgebra of conserved quantities is just the centralizer of h in the Lie algebra $(C^\infty(M),\{\cdot,\cdot\})$.

3. *A Poisson manifold $(M,\{\cdot,\cdot\})$ of dimension $2n$ said to be completely integrable, if there are n globally defined linearly independent smooth real-valued functions $(f_i)_{i=1,\dots n}$ which Poisson commute, $\{f_i,f_j\}=0$ for all i,j. Moreover, it is required that the union of the regular level sts of the moment map*

$$\begin{array}{rcl}\mathcal{H}:M & \to & \mathbb{R}^n\\ x & \mapsto & (f_1(x),\dots f_n(x))\end{array}$$

have full measure in M.

A Hamiltonian system is called integrable, if it is integrable as a Poisson manifold and if the Hamiltonian function h is contained in the subspace spanned by the family $(f_i)_{i=1,\dots n}$.

The Arnold-Liouville theorem states in particular that then every compact connected regular level $\mathcal{H}^{-1}(c)$ of the moment mapis an n-dimensional torus L which is invariant under the phase flow φ_t and moreover carries a quasi-periodic motion. Writing $L\cong\mathbb{R}^n/\mathbb{Z}^n$ one can find angular coordinates θ such that the dynamics looks like $\varphi_t\theta=\theta+tv$. Whether the orbit closes or not depends on the arithmetic properties of the components of the so-called rotation vector v: if all components are rational, the trajectory on L is periodic; otherwise every trajectory is uniformaly distributed on L.

This is an example of the structural *theorems on mechanical systems on can prove in the Hamiltonian approach. The famous Kolmogorov-Arnold-Moser theorem investigates on how the behaviour of an integrable system changes under small perturbations. It describes precise conditions on when the invariant n-dimensional tori we have just described persist.*

We also discuss the situation in local coordinates.

Remarks 4.2.8.

1. *Suppose that M is even a symplectic manifold so that we can consider local Darboux coordinates. We can apply the equation $\dot{f} = \{h, f\}$ to the coordinate functions q^i and p_i and find for the time derivatives of the coordinate functions on a trajectory of h:*

$$\dot{q}^i = \frac{\partial h}{\partial p_i} \quad \text{and} \quad \dot{p}_i = -\frac{\partial h}{\partial q^i} \ .$$

These equations for the coordinate functions evaluated on a trajectory are called Hamilton's equations for the Hamilton function h.

2. *In the case of a natural Hamiltonian system, the Hamiltonian is in Darboux coordinates*

$$h(q,p) = \frac{1}{2} g^{ij}(q) p_i p_j + V(q^i) \ .$$

Hamilton's equations thus become

$$\dot{q}^i = g^{ij}(q) p_j \quad \text{and} \quad \dot{p}^i = -\frac{\partial V}{\partial q^i} \ .$$

This includes systems like the one-dimensional harmonic oscillator with Hamiltonian function

$$h(p,q) = \frac{1}{2m} p^2 + \frac{D}{2} q^2$$

and the Hamiltonian function

$$h(p,q) \ = \frac{1}{2m} p^2 - \frac{k}{|q|}$$

for the Kepler problem.

We can also make immediate comments on the time evolution in Hamiltonian systems: let φ_t the family of diffeomorphisms associated to the Hamiltonian vector field X_h. According to a previous proposition, this family consists of symplectomorphisms. We have thus:

Proposition 4.2.9.
Let (M, ω, h) be a Hamiltonian system. For any closed oriented surface $\Sigma \in M$ in phase space for the integral over the symplectic form is constant under the phase flow,

$$\int_{\varphi_t(\Sigma)} \omega = \int_\Sigma \varphi_t^* \omega = \int_\Sigma \omega \ .$$

One says that the sympletic form leads to an integral invariant.

The phase flow has as further invariants all exterior powers $\omega^{\wedge k}$ of the symplectic form, in particular the Liouville volume $\omega^{\wedge n}$. We need the following simple

Lemma 4.2.10.
Let M be a smooth manifold with a volume form and $g : M \to M$ be a volume preserving diffeomorphism that is mapping a measurable subset $D \subset M$ of finite volume to itself. For any point $p \in D$ and any neighborhood $U \subset D$ of this point p, we can find a point $x \in U$ which comes back to the neighborhood U in the sense that there is some $n \in \mathbb{N}$ such that $g^n x \in U$.

Proof:
The infinitely many images $U, gU, \ldots, g^n U, \ldots$ are measurable sets of the same finite non-zero volume and are all contained in the subset D of finite volume. Hence they cannot be disjoint. Thus, there are integers k and l with $k > l$ such that

$$g^k U \cap g^l U \neq \emptyset \quad .$$

We deduce that $g^{k-l} U \cap U \neq \emptyset$. □

To formulate the following famous statement, we consider a Hamiltonian system (M, ω, h) with phase flow φ_t. To simplify the notation, we assume that the phase flow φ_t is globally defined.

For any point $m \in M$, we call the set $\{\varphi_t(m)\}_{t \in \mathbb{R}}$ the orbit $\mathcal{O}_m$ of m. Put differently, the orbit of $m \in M$ are all points of M the point m will after a sufficiently long time or will reach under a sufficiently long time on its trajectory.

Proposition 4.2.11 (Poincaré's recurrence theorem)**.**
Let (M, ω, h) be a Hamiltonian system with the property that for any point $p \in M$ the orbit $\mathcal{O}_p$ is bounded in the sense that it is contained in a subset $D \subset M$ of finite volume. Then for each open subset $U \subset M$, there exists an orbit that intersect the set U infinitely many times.

Remark 4.2.12.
The recurrence theorem applies in particular to a (time-independent) natural Hamiltonian system (T^*X, ω_0, g, V) defined on $X = \mathbb{R}^n$ with a potential $V : \mathbb{R}^n \to \mathbb{R}$ with the property that $V(x) \to \infty$ for $|x| \to \infty$.

Indeed, since the Hamiltonian h itself is a conserved quantity, the phase flow preserves for any $\eta \in \mathbb{R}$ the subset

$$D_\eta := \{v \in T^*X \mid h(v) = \eta\} \ .$$

Recall that for a natural system, we have $h(q,p) = \frac{1}{2} g^*(p,p) + V(q)$; since the first term is positive definite, we need $V(q) \leq \eta$. Because of our condition on the growth of V, this implies that $q \in \mathbb{R}^n$ is contained in a sufficiently large ball, in $\mathbb{R}^n$, i.e. $|q| < R$. Then for a point q in this ball, only momenta p of modulus $\frac{1}{2} g_q(p,p) = \eta - V(q)$ are admitted. This subset is contained in subset of T^*X of finite volume.

Remark 4.2.13.
The time evolution $s : I \to M$ of a particle in a Hamiltonian system (M, ω, h) is given the integral curve of the hamiltonian vector field X_h. This implies

$$s(t) = \varphi_t(s(0))$$

where $\varphi_t : M \to M$ is the phase flow of the Hamiltonian vector field X_h on M. Put differently, we have

$$\varphi_{-t}^* \circ s(0) = s \circ \varphi_{-t}(t) \ = s(0) \ .$$

We could equally well consider the time evolution of a probability distribution on phase space M, i.e. of a normalized non-negative top form $p \in \Omega^n(M)$ with $n = \dim M$. At time t, we then have a probability density $p_t \in \Omega^n(M)$ with

$$\varphi_{-t}^* p_t = p_0 \ .$$

Suppose for simplicity that we have $p_t = f_t \omega^{\wedge n}$ with $f_t : M \to \mathbb{R}$ a smooth function and $\omega^{\wedge n}$ the Liouville volume on M. Then $f_t = (\varphi_{-t})^* f_0$ and thus

$$\frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} f_t = L_{X_h} f_0 = \{h, f_0\} \ ,$$

so that the Poisson bracket also describes the evolution of probability distributions.

4.3 Time dependent Hamiltonian dynamics

Observation 4.3.1.

- *We want to consider time dependent Hamiltonian systems. To this end, we consider a surjective submersion $\pi : E \to I$ with $I \subset \mathbb{R}$ an interval and require for each $t \in I$ the fibre $E_t := \pi^{-1} t \subset E$ to be a symplectic manifold. The manifold E is called the evolution space of the system.*

- *Since the evolution space E is odd-dimensional, it cannot be a symplectic manifold any longer. Rather, E is presymplectic manifold of rank $\dim E - 1$. It has the property that the projection of $\ker \omega$ to I is non-vanishing.*

  This has generalizations to field theory: in this case, it has been proposed to consider presymplectic manifolds with higher dimensional leaves.

- *By the slice theorem, there exists a foliation of evolution space with one-dimensional leaves. We require that the integral curves of the dynamics we are interested in parametrize these leaves. A presymplectic formulation has been proposed in particular to deal with systems in which no natural parametrization is known, e.g. for massless relativistic particles.*

Let us consider an example.

Example 4.3.2.

1. *We take a symplectic manifold (M, ω_0) and obtain a presymplectic manifold $I \times M$ with presymplectic form $pr_2^* \omega_0$. This cannot be the correct presymplectic form, since the leaves of the presymplectic manifold $(I \times M, \omega_0)$ are submanifolds of the form $I \times \{m\}$ with $m \in M$, i.e. have trivial dynamics on the space M. of leaves.*

2. *For any smooth function*

   $$h : I \times M \to \mathbb{R}$$

   we consider

   $$\omega_h = pr^* \omega_0 - \mathrm{d}h \wedge \mathrm{d}t \in \Omega^2(I \times M) \ .$$

   This form is closed as well and has rank $\dim M$. Thus any choice of function h endows $I \times M$ with the structure of a presymplectic manifold $(I \times M, \omega_h)$.

3. *Consider local Darboux coordinates (q^i, p_i) on M. Then*

   $$\omega_0 = \sum_{i=1}^{n} \mathrm{d}p_i \wedge \mathrm{d}q^i$$

and, with the summation convention understood, in local coordinates (q^i, p_i, t) *on* $I \times M$

$$\omega_h = \mathrm{d}p_i \wedge \mathrm{d}q^i - \frac{\partial h}{\partial p_i}\mathrm{d}p^i \wedge \mathrm{d}t - \frac{\partial h}{\partial q^i}\mathrm{d}q^i \wedge \mathrm{d}t \ .$$

It is easy to check that the nowhere vanishing vector field X_h

$$(X_h)\big|_{(p,q,t)} = \frac{\partial h}{\partial p_i}\big|_{(p,q,t)}\frac{\partial}{\partial q^i} - \frac{\partial h}{\partial q^i}\big|_{(p,q,t)}\frac{\partial}{\partial p_i} + \frac{\partial}{\partial t}$$

obeys

$$\iota_{X_h}\omega_h = 0$$

and thus is tangent to the leaves.

4. *Integral curves* $s : \mathbb{R} \to I \times \tilde{M}$ *of the vector field* X_h *obey*

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\tau}q^i(s(\tau)) &= \frac{\partial h}{\partial p_i} \\ \frac{\mathrm{d}}{\mathrm{d}\tau}p_i(s(\tau)) &= -\frac{\partial h}{\partial q^i} \\ \frac{\mathrm{d}}{\mathrm{d}\tau}t(s(\tau)) &= 1 \end{aligned}$$

The last equation allows us to identify the parameter τ *of the integral curve with* t*. We thus obtain Hamiltonian equations with time dependent Hamiltonian function.*

This generalizes as follow:

Definition 4.3.3

1. *A a time dependent Hamiltonian system consists of a symplectic manifold* (M, ω_0) *and a smooth function*

$$h : I \times M \to \mathbb{R} \ ,$$

called the (time-dependent) Hamiltonian function. Introduce the vector field $\tilde{X}_h \in \mathrm{vect}(I \times M)$ *in* $p \in I \times M$ *as the sum*

$$\tilde{X}_h = X_h + \frac{\partial}{\partial t}$$

where X_h *is the symplectic gradient of* h *on the slice. The integral curves of* $\tilde{X}_h$ *describe the physical trajectories for the Hamiltonian function* h*.*

2. *The presymplectic manifold* $(I \times \tilde{M}, \omega - \mathrm{d}h \wedge \mathrm{d}t)$ *is called evolution space of the system. The space of leaves is called the phase space of the system. It has the structure of a symplectic manifold. A leaf of the foliation is an unparametrized trajectory. The images of trajectories in evolution space* E *in phase space are called (Hamiltonian) trajectories.*

Observation 4.3.4.

1. *Consider a Hamiltonian system* (M, ω, h)*. Choose local Darboux coordinates* (p, q) *on* M *and the related local Darboux coordinates* (t, p, q) *on evolution space* $E := I \times M$*.*

2. *The symplectic form ω on M is exact on the coordinate neighborhood U_α of a Darboux coordinate:*

$$\omega\big|_{U_\alpha} = \sum_i \mathrm{d}p_i \wedge \mathrm{d}q^i = \mathrm{d}\left(\sum_i p_i \mathrm{d}q^i\right) = \mathrm{d}\Theta_\alpha \ .$$

Similarly, we have on evolution space $E = I \times M$ the presymplectic form ω_E which is locally

$$\begin{aligned}(\omega_E)\big|_{I\times U_\alpha} &= p_2^*\omega - p_2^*\mathrm{d}h \wedge p_1^*\mathrm{d}t \\ &= \mathrm{d}\left(\sum_i p_i \mathrm{d}q^i - h\mathrm{d}t\right) \\ &= \mathrm{d}(\Theta_E)_\alpha \ .\end{aligned}$$

The one-form $\Theta_E = \sum_i p_i \mathrm{d}q^i - h\mathrm{d}t \in \Omega^1(I\times U_\alpha)$ is a locally defined one-form on evolution space.

3. *The phase flow for h*

$$\varphi_t : \ M \to M$$

on M leads to a family of diffeomorphisms on evolution space E which parametrize the leaves

$$t' \mapsto (\varphi_t(x), t' + t) \, .$$

We already know that the one-forms Θ and Θ_E are globally defined, if M is a cotangent bundle. We will assume from now on that the symplectic form is not only closed but even exact and that a globally defined one-form $\Theta \in \Omega^1(M)$ has been chosen such that $\mathrm{d}\Theta = \omega$.

Proposition 4.3.5.
Let γ_1 and γ_2 be two curves in evolution space E that encircle the same leaves. Then we have

$$\int_{\gamma_1} \Theta_E = \int_{\gamma_2} \Theta_E.$$

The one-form Θ_E is called Poincaré-Cartan integral invariant.

Proof:
Let Σ be the the surface formed by those parts of the leaves intersecting γ_1 (and thus γ_2) that is bounded by the curves, $\partial\Sigma = \gamma_2 - \gamma_1$. Such a surface is called a flux tube. Stokes' theorem implies

$$\int_{\gamma_2} \Theta_E - \int_{\gamma_1} \Theta_E = \int_{\partial\Sigma} \Theta_E = \int_\Sigma \mathrm{d}\Theta_E = \int_\Sigma \omega_E = 0,$$

since Σ consists of leaves whose tangent space is by definition the kernel $\ker \omega_E$ of the presymplectic form on evolution space. □

Observation 4.3.6.

1. *We now specialize to curves in evolution space M at constant time, $\gamma_i \in p_1^{-1}(t_i)$, to find*

$$\int_{\gamma_1} \Theta = \int_{\gamma_2} \Theta$$

2. *We can obtain such a pair of curves from a closed curve γ in phase space $\tilde{M}$ parametrized by $s \in [0,1]$ and first lifting γ to a curve*

$$\gamma_1(s) = \big(\gamma(s), t_1\big)$$

parametrized by $s \in [0,1]$ at fixed time $t_1 \in I$ and and then applying a phase flow by a time Δt to obtain a closed curve

$$\gamma_2(s) = \big(\varphi_{\Delta t}\gamma(s), \Delta t + t_1\big)$$

in evolution space at another fixed time $t_2 = \Delta t + t_1$.

3. *Using the surface*

$$\Sigma = \Big\{ \big(\varphi_t\gamma(s), t + t_1\big),\ s \in [0,1]\,,\ t \in [0, \Delta t] \Big\}$$

formed by the leaves, we obtain by proposition 4.3.5 the equality

$$\int_\gamma \Theta = \int_{\varphi_t\gamma} \Theta$$

of integrals along curves in phase space $\tilde{M}$.

4. *Consider now an oriented two-chain $\Sigma \in M$ such that $\partial\Sigma = \gamma$. Stokes' theorem then implies that*

$$\int_\Sigma \omega = \int_\Sigma \mathrm{d}\Theta = \int_{\partial\Sigma} \Theta = \int_\gamma \Theta$$

is invariant under the phase flow. We have derived this result earlier.

4.4 The Legendre transform

We still have to understand how to relate Hamiltonian and Lagrangian mechanical natural systems with second order equations of motion. Both systems are supposed to have the same configuration space manifold M. The Lagrangian system is given by a Lagrangian function

$$l: \ \ I \times TM \to \mathbb{R}$$

and the Hamiltonian system by a function

$$h: \ \ I \times T^*M \to \mathbb{R}\ .$$

This raises the question on whether there is a function h on T^*M such that the vector field $\{h, -\}$ on T^*M has the property that the projection of its integral curves $s : I \ \to T^*M$ to M are just the physical trajectories of the mechanical system given by the Lagrangian l.

The fibres of the two vector bundles TM and T^*M over M are dual vector spaces. We are thus interested in relating functions defined over dual vector spaces.

Observation 4.4.1.

1. *Let V be a real vector space and $U \subset V$ be a convex subset, i.e. with $x, y \in U$, also all points of the form $tx + (1-t)y$ with $t \in [0,1]$ are contained in U. In other words, along with two points x, y, the subset U also contains the line segment connecting x and y.*

2. *Let $f : U \to \mathbb{R}$ be a real-valued piecewise smooth continuous convex function, i.e.*

$$f(tx + (1-t)y) \leqslant t\, f(x) + (1-t)\, f(y)$$

for all $x, y \in I$ and all $t \in [0,1]$. This implies that the matrix of second derivatives, the Hessian $\frac{\partial^2 f}{\partial x^i \partial x^j}$, is positive definite, wherever it is defined.

3. *Denote by V^* the vector space dual to V. Consider the function*

$$\begin{array}{rcl} F : V^* \times U & \to & \mathbb{R} \\ (p, x) & \mapsto & \langle p, x\rangle - f(x)\ . \end{array}$$

We fix a linear form $p \in V^*$ and assume that for this p the maximum $\max_{x\in U} F(p,x)$ exists. The fact that f is convex implies that it is assumed for a unique $x = x(p) \in U$. In case f is differentiable in x, we find $\mathrm{d}f_x = p$. This equation defines implicitly $x \in U$ as a function of $p \in V^*$.

4. *If $\dim_{\mathbb{R}} V = 1$, the situation is more specifically that U is an interval, and the implicit equation fixing x in terms of p becomes $f'(x) = p$.*

Definition 4.4.2
Let $U \subset V$ be a convex subset of a real vector space. Given a convex function $f : U \to \mathbb{R}$, the real-valued function g defined on a subset of V^* by

$$g(p) := \max_{x\in U} \langle p, x\rangle - f(x)$$

is called the Legendre transform of f.

Remarks 4.4.3.

1. *Consider as an example the function $f(x) = m\frac{x^\alpha}{\alpha}$ on $\mathbb{R}$ with $m > 0$ and $\alpha > 1$. Then for fixed $p \in \mathbb{R}$, the function*

$$F(p,x) = px - m\frac{x^\alpha}{\alpha}$$

has in x an extremum for $p = mx^{\alpha-1}$. Hence we find for the Legendre transform

$$g(p) = m^{-\frac{1}{\alpha-1}}\, \frac{p^\beta}{\beta} \quad \text{with} \quad \frac{1}{\alpha} + \frac{1}{\beta} = 1.$$

As special cases, we find with $\alpha = 2$ for the function

$$f(x) = \frac{m}{2}x^2 \quad \text{the Legendre transform} \quad g(p) = \frac{p^2}{2m}$$

and for $m = 1$ for the function

$$f(x) = \frac{x^\alpha}{\alpha} \quad \text{the Legendre transform} \quad g(p) = \frac{p^\beta}{\beta}$$

with $\frac{1}{\alpha} + \frac{1}{\beta} = 1$.

2. *The definition of the Legrendre transform via a maximum implies the inequality*

$$F(x,p) = xp - f(x) \leqslant g(p)$$

for all values of x, p where the functions f and g are defined. A function and its Legendre transform are thus related by

$$px \leqslant f(x) + g(p).$$

Applying this to the second example just discussed, we find the classical inequality

$$px \leqslant \frac{1}{\alpha}x^\alpha + \frac{1}{\beta}p^\beta \quad \textit{with } \frac{1}{\alpha} + \frac{1}{\beta} = 1.$$

This is Young's inequality. It can be used to prove Hölder's inequality

$$\|fg\|_1 \leq \|f\|_\alpha \cdot \|g\|_\beta$$

where $f \in L^\alpha(M)$ with norm

$$\|f\|_\alpha := \left(\int_M |f|^\alpha\right)^{1/\alpha}$$

and similarly $g \in L^\beta(M)$ and M a measurable space.

3. *If f is a smooth function with $f'' > 0$, then its Legendre transform g is again a convex function. It turns out that the Legendre transform is an involution.*

4. *It is instructive to specialize to a positive definite quadratic form*

$$f : \ V \to \mathbb{R}$$

with $f(q) = \frac{1}{2}A_{ij}q^iq^j$ with $A = (A_{ij})$ a matrix with real entries. The hessian A being positive definite implies in particular that the matrix A is invertible. Minimizing

$$F(p,q) = \langle p, q\rangle - f(q)$$

for a fixed p yields $p_i = A_{ij}q^j$. Since A is invertible, this can be solved for q. We find $q^i = (A^{-1})^{ij}p_i$ and thus

$$g(p) = p_i(A^{-1})^{ij}p_j - \frac{1}{2}(A^{-1})^{ij}p_ip_j = \frac{1}{2}(A^{-1})^{ij}p_ip_j \ .$$

We now use these ideas to relate Lagrangian and Hamiltonian functions.

Observation 4.4.4.

1. *Given a Lagrange function $l : I \times TM \to \mathbb{R}$, we define the Legrendre transformation*

$$\Lambda : \ I \times TM \ \to I \times T^* \ M$$

for $(t, q, q_t) \in I \times T_qM$ as the element $(t, \Lambda_{t,q}(q_t))$, where the element $\Lambda_{t,q}(q_t) \in T^*_qM$ is defined by

$$\langle \Lambda_{t,q}(q_t), w\rangle = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0} l(t, q, q_t + \epsilon w) = \frac{\partial l}{\partial q^i_t}w^i \ .$$

2. *One can then consider the so-called energy function*

$$\begin{array}{rccl} \epsilon: & I \times TM & \to & \mathbb{R} \\ & (t,q,q_t) & \mapsto & \langle \Lambda_{t,q}(q_t), v\rangle - l(t,q,q_t) \ . \end{array}$$

In the case of a natural Lagrangian mechanical system (M,g,V) with Lagrangian function

$$l(t,q,q_t) = \frac{1}{2} g_{t,q}(q_t,q_t) - V(q)$$

we have

$$\langle \Lambda_{t,q}(q_t), w\rangle = g_{t,q}(q_t,w)$$

and thus

$$\epsilon(t,q,q_t) = \frac{1}{2} g_{t,q}(q_t,q_t) + V(q) \ .$$

3. *We say that the Lagrangian is non-degenerate, if the map*

$$\Lambda_{t,-} : TM \to T^*M$$

is a local diffeomorphism, i.e. if the determinant

$$\det\left(\frac{\partial^2 l}{\partial q_t^i \partial q_t^j}\right)_{i,j=1,\dots \dim M}$$

is a non-vanishing function on TM.

4. *In many cases of interest, Λ is even a global diffeomorphism of TM to T^*M. Then one can work with the phase space T^*M instead of the kinematical space TM.*

We will now restrict to this case. Then the function

$$h: \ I \times T^*M \to \mathbb{R}$$

defined by the Legrendre transform of l:

$$\begin{array}{ccc} I \times TM & \xrightarrow{\ \Lambda\ } & I \times T^*M \\ & {\scriptstyle l}\searrow \quad \swarrow{\scriptstyle h} & \\ & \mathbb{R} & \end{array}$$

can be used as a Hamiltonian function for a time-dependent Hamiltonian system on $I \times T^*M$.

5. *In the example of a natural Lagrangian mechanical system (M,g,V) with*

$$l(t,q,q_t) = \frac{1}{2} g_{t,q}(q_t,q_t) - V(t,q) \quad .$$

Then only the quadratic form enters in the Legendre transform and we find

$$h(t,q,p) = \frac{1}{2} g^*_{t,q}(p,p) \ + V(t,q) \ .$$

where (p, q) are natural Darboux coordinates on the symplectic manifold T^*M.

In such coordinates, we have

$$h(q^1, \ldots, q^N, p_1, \ldots p_N) = \sum_{i=1}^{N} p_i q_t^i - l(t, q, q_t)$$

with the relation

$$p_i = \frac{\partial l}{\partial q_t^i}$$

defining q_t^i as a function $\tilde{q}_t^i(t, q_i, p_i)$.

As an example, we get for the Lagrangian

$$l(q, q_t) = \frac{m}{2}\left((q_t^1)^2 + (q_t^2)^2 + (q_t^3)^2\right) - V(q^1, q^2, q^3)$$

the relation

$$p_i = m q_t^i$$

and thus

$$h(q, p) \ = \frac{1}{2m}\left((p_1)^2 + (p_2)^2 + (p_3)^2\right) + V(q^1, q^2, q^3)$$

We next show that the Hamiltonian function defined by the Legendre transform encodes the same physical information.

Theorem 4.4.5.
Let M be a smooth manifold. Let

$$l : I \times TM \ \to \mathbb{R}$$

be a time-dependent Lagrangian function such that the Legrendre transform exists as a global diffeomorphism

$$\Lambda : \ I \times TM \ \to I \times T^* \ M$$

of manifolds fibred over the interval I. Let

$$h : \ I \times T^*M \to \mathbb{R}$$

be the Legrendre transform of l:

$$\begin{array}{ccc} I \times TM & \xrightarrow{\ \Lambda\ } & I \times T^*M \\ & {\scriptstyle l}\searrow \quad \swarrow{\scriptstyle h} & \\ & \mathbb{R} & \end{array}$$

This sets up a bijection of Lagrangian and Hamiltonian systems with a bijection of classical trajectories: A section $\varphi : I \to I \times M$ is a solution of the Euler Lagrange equations for l, if and only if the Legendre transform of its jet prolongation

$$\Lambda \circ \mathrm{j}^1\varphi : I \to I \times TM \to I \times T^*M$$

obeys the Hamilton equations for h.

Remark 4.4.6.
These statements are sometimes summarized in the statement that for a Lagrangian function l and a Hamiltonian function h that are related by a Legendre transform, the Euler-Lagrange equations

$$\mathrm{D}\frac{\partial l}{\partial \tilde{q}^i_t} = \frac{\partial l}{\partial q^i} \ ,$$

i.e. $n = \dim M$ ordinary differential equations of second order, are equivalent to the Hamiltonian equations

$$\dot{p}^i = -\frac{\partial h}{\partial q^i}, \quad \dot{q}^i = \frac{\partial h}{\partial p^i},$$

i.e. $2n$ ordinary differential equations of first order.

Proof:
We consider the time-dependent Hamiltonian in local coordinates

$$h = p_i \tilde{q}^i_t - l(q^i, \tilde{q}^i_t, t) : I \times T^*M \to \mathbb{R} \ ,$$

where we introduce the functions

$$\tilde{q}^i_t = q^i_t \circ \Lambda^{-1} : \ I \times T^*M \to \mathbb{R}$$

on evolution space $I \times T^*M$. (In physicists notation, this is sometimes written as $q^i_t = q^i_t(p, q, t)$.) The total derivative of the function h on $I \times T^*M$ is the one-form

$$\mathrm{d}h = \frac{\partial h}{\partial p_i}\mathrm{d}p_i + \frac{\partial h}{\partial q^i}\mathrm{d}q^i + \frac{\partial h}{\partial t}\mathrm{d}t \in \Omega^1(I \times T^*M) \tag{11}$$

Using the identity $p_i = \frac{\partial l}{\partial \tilde{q}^i_t} \circ \Lambda^{-1}$ of locally defined functions on evolution space implied by the Legendre transform, we find for our function h

$$\begin{aligned}
&\mathrm{d}h = \mathrm{d}(p_i \tilde{q}^i_t - l(q^i, q^i_t, t)) = \\
&= \tilde{q}^i_t \mathrm{d}p_i + \underbrace{\Big(p_i \frac{\partial \tilde{q}^i_t}{\partial p_j} - \frac{\partial l}{\partial q^i_t}\frac{\partial \tilde{q}^i_t}{\partial p_j} \Big)}_{=0} \mathrm{d}p_j - \frac{\partial l}{\partial q^i}\mathrm{d}q^i + \underbrace{\Big(p_i \frac{\partial \tilde{q}^i_t}{\partial q^j} - \frac{\partial l}{\partial q^i_t}\frac{\partial \tilde{q}^i_t}{\partial q^j} \Big)}_{=0} \mathrm{d}q^j - \frac{\partial l}{\partial t}\mathrm{d}t \\
&= \tilde{q}^i_t \mathrm{d}p_i - \frac{\partial l}{\partial q^i}\mathrm{d}q^i - \frac{\partial l}{\partial t}\mathrm{d}t
\end{aligned} \tag{12}$$

By comparison of the coefficients in equation (11) and (12), we find the following identities of locally defined functions on the evolution space $I \times T^*M$:

$$\tilde{q}^i_t = \frac{\partial h}{\partial p_i}, \quad \frac{\partial h}{\partial q^i} = -\frac{\partial l}{\partial q^i} \circ \Lambda^{-1}, \quad \frac{\partial h}{\partial t} = -\frac{\partial l}{\partial t} \circ \Lambda^{-1} \tag{13}$$

Consider any section $\varphi : I \to I \times M$. Then the first equation of (13) implies

$$\frac{\mathrm{d}}{\mathrm{d}t} q^i(\Lambda \circ \mathrm{j}^1\varphi) = \frac{\mathrm{d}\varphi^i}{\mathrm{d}t}(t) = q^i_t(\mathrm{j}^1\varphi(t)) \overset{(13.1)}{=} \frac{\partial h}{\partial p_i}(\Lambda \circ \mathrm{j}^1\varphi(t)) \ .$$

Suppose now that the section $\varphi : I \to I \times M$ is even a solution of the Euler-Lagrange equations for l. The Euler-Lagrange equations for the trajectory φ read

$$\frac{\partial l}{\partial q^i} \circ \mathrm{j}^1\varphi = (D\frac{\partial l}{\partial q_t^i}) \circ \mathrm{j}^1\varphi = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial l}{\partial q_t^i} \circ \mathrm{j}^1\varphi\right) \quad .$$

Using the second equation and using the identity

$$\frac{\partial l}{\partial q_t^i} \circ \Lambda^{-1} = p_i$$

from the Legrendre transform yields

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left(p_i(\Lambda \circ \mathrm{j}^1\varphi)\right) &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial l}{\partial q_t} \circ \mathrm{j}^1\varphi\right) \\
&\overset{\text{ELG}}{=} \frac{\partial l}{\partial q^i} \circ \mathrm{j}^1\varphi \\
&= \frac{\partial l}{\partial q^i} \circ \Lambda^{-1} \circ \Lambda \circ \mathrm{j}^1\varphi \\
&\overset{(13.2)}{=} -\frac{\partial h}{\partial q^i} \circ \Lambda \circ \mathrm{j}^1\varphi
\end{aligned}$$

We have thus shown that the function

$$\Lambda \circ \mathrm{j}^1\varphi : I \to I \times T^*M$$

fulfills the Hamiltonian equations for the Hamiltonian function given by the Legendre transform.

The converse of the assertion is shown analogously. □

5 Quantum mechanics

5.1 Deformations

We start by explaining the mathematical notion of a formal deformation of an associative algebra. In this section, k is a field of characteristic 0.

Definition 5.1.1

1. *The ring $k[[t]]$ of formal power series with coefficients in k is, as a vector space, the k-algebra of sequences $(a_k)_{k\in\mathbb{N}_0}$ with $a_k \in k$. We write*

$$(a_k) = a_0 + a_1 t + a_2 t^2 + \ldots = \sum_{i\in\mathbb{N}_0} a_i t^i \ .$$

It is endowed with the structure of an associative unital algebra by the addition

$$\sum_{i\in\mathbb{N}_0} a_i t^i + \sum_{i\in\mathbb{N}_0} b_i t^i = \sum_{i\in\mathbb{N}_0} (a_i + b_i) t^i$$

and the multiplication

$$\left(\sum_{i\in\mathbb{N}_0} a_i t^i\right)\left(\sum_{i\in\mathbb{N}_0} b_i t^i\right) = \sum_{i\in\mathbb{N}_0}\left(\sum_{j=0}^{i} a_j b_{i-j}\right) t^i$$

2. *For any k-vector space V, we denote by $V[[t]]$ the $k[[t]]$-module of formal power series*

$$\sum_{i\in\mathbb{N}_0} v_i t^i \quad \textit{with} \quad v_i \in V \ .$$

3. *Let A be an associative k-algebra with unit. An algebra homomorphism*

$$\epsilon : \ A \to k$$

is called an augmentation, if $\epsilon(\lambda \cdot 1_A) = \lambda$ holds for all $\lambda \in k$.

Remarks 5.1.2.

1. *The ring $k\,[[t]]$ of formal power series with coefficients in k admits the augmentation*

$$\begin{array}{rcl} \epsilon : \ k\,[[t]] & \to & k \\ \sum_{i\in\mathbb{N}_0} a_i t^i & \mapsto & a_0. \end{array}$$

It can be shown that the ring $k[[t]]$ is a local Noetherian ring with maximal ideal $tk[[t]]$.

One might think heuristically of the augmentation as an evaluation at $t = 0$ or as a limit $t \to 0$. One should, however, keep in mind that the ring $k[[t]]$ does not have any other evaluations, and that the limit does not make literally sense.

2. *An augmentation endows the ground field k with the structure of an A-bimodule by*

$$\begin{array}{rcl} A \times k & \to & k \\ (a, \lambda) & \mapsto & \epsilon(a \cdot \lambda) \\ k \times A & \to & k \\ (\lambda, a) & \mapsto & \epsilon(\lambda \cdot \alpha) \ . \end{array}$$

3. *Formal power series $\mathbb{C}[[t]]$ are the appropriate recipient for results obtained in perturbation theory: the result a_k obtained in order k is taken as the coefficient of t^k. Note that here no questions of convergence do not make sense and that it does not even make sensee to replace t by a complex number.*

Definition 5.1.3

1. *A formal deformation or deformation quantization of an associative k-algebra A is an associative $k[[t]]$-bilinear map*

$$* : \quad A[[t]] \times A[[t]] \to A[[t]]$$

such that

$$v * w = v \cdot w \bmod tA[[t]]$$

for all formal power series $u, v \in A[[t]]$. Here $v \cdot w$ is the degree preserving product on $A[[t]]$ inherited from A. The associative product on $A[[t]]$ is also called a star product.

2. *Let J_A be the group of $k[[t]]$-module automorphisms g of $A[[t]]$ such that*

$$g(u) \ = u \bmod tA[[t]] \quad \text{ for all } u \in A[[t]] \ .$$

We say that two formal deformations $*$ and $*'$ of the associative k-algebra A are equivalent, if there is an element $g \in J_A$ such that

$$g(u * v) \ = g(u) *' g(v) \ \ .$$

Remarks 5.1.4.

1. *Identifying $A \cong At^0 \subset A[[t]]$, the product of two elements $a, b \in A$ in a formal deformation is of the form*

$$a * b = ab + B_1(a, b)t + \cdots + B_n(a, b)t^n + \cdots \ .$$

We introduce the notation $B_0(a, b) := a \cdot b$. The sequence of the k-bilinear maps $B_i : A \times A \to A$ determines the product $*$, because it is $k[[t]]$-linear. Associativity amounts to the infinite set of conditions

$$\sum_{\substack{i+j=k \\ i,j \geqslant 0}} B_i \left(B_j(a, b), c\right) = \sum_{\substack{i+j=k \\ i,j \geqslant 0}} B_i \left(a, B_j(b, c)\right)$$

for all k and for all $a, b \in A$.

2. *For $g \in J_A$ and $a \in A$, we have*

$$g(a) \ = a + g_1(a)t + g_2(a)t^2 + \cdots + g_n(a)t^n + \cdots$$

for certain k-linear maps $g_i : A \to A$. These maps determine $g \in J_A$, because g is $k[[t]]$-linear.

3. *If A admits a unit element 1_A, each formal deformation admits a unit element 1_*. Any formal deformation $*$ is equivalent to a formal deformation $*'$ such that $1_* = 1_A$.*

The following lemma relates Poisson algebras and formal deformations:

Lemma 5.1.5.
Let A be an associative and commutative *k-algebra. Let $*$ be an associative (but not necessarily commutative) formal deformation of the multiplication on A. For $a, b \in A$, put*

$$\{a, b\} := B_1(a, b) - B_1(b, a) \ .$$

1. *The map $\{\cdot, \cdot\}$ is a Poisson bracket on A.*

2. *The bracket $\{\cdot, \cdot\}$ only depends on the equivalence class of $*$.*

Proof:
The map

$$\begin{array}{rccl} [\cdot, \cdot] & : A[[t]] \times A[[t]] & \to & A[[t]] \\ & (u, v) & \mapsto & \frac{1}{t}(u * v - v * u) \end{array}$$

is well-defined, since A is commutative.

As we have seen in example 4.2.2 for any associative algebra, the commutator endows the associative algebra with the structure of a Poisson algebra. Also after dividing by t, we have the structure of a Poisson algebra. The bracket $\{\cdot, \cdot\}$ is the reduction of $[\cdot, \cdot]$ modulo t and thus still a Poisson bracket.

If $g \in J$ yields an equivalence of $*$ with $*'$, we have

$$B_1(a, b) + g_1(ab) = B_1'(a, b) + g_1(a)b + ag_1(b) \quad \text{for all } a, b \in A \ .$$

Thus the difference

$$B_1(a, b) - B_1'(a, b) = g_1(a)b + ag_1(b) - g_1(ab)$$

is symmetric in a, b and does not contribute to the antisymmetric bracket $\{\cdot, \cdot\}$. □

This motivates the following

Definition 5.1.6

1. *Given a formal deformation $(A[[t]], *)$ of an associative commutative algebra A, the Poisson algebra $(A, \cdot, \{\ \})$ is called the* classical limit *of $(A[[t]], *)$.*

2. *A deformation quantization of a Poisson manifold M is a deformation quantization of the Poisson algebra $(C^\infty(M), \cdot, \{-, -\})$ in which all maps B_i are differential operators.*

Theorem 5.1.7. *(Kontsevich 1997, Fields medal 1998)*
If A is the commutative algebra of smooth functions on a smooth manifold M, then each Poisson bracket on A lifts to an associative formal deformation.

Remarks 5.1.8.

1. *In other words, the map from equivalence classes of formal deformations to Poisson structures is surjective, if A is the algebra of smooth functions on a smooth manifold M. Kontsevich even constructs a section of this map. It is canoncial up to equivalence.*

2. *The result is surprising, since one can give examples of finite-dimensional Poisson algebras whose bracket does* not *lift to formal deformations.*

3. *For $M = \mathbb{R}^2$ with Poisson bracket*

$$\{f, g\} := \frac{\partial f}{\partial x_1}\frac{\partial g}{\partial x_2} - \frac{\partial g}{\partial x_2}\frac{\partial f}{\partial x_1}$$

the deformation given by Kontsevich is

$$f * g = \sum_{n=0}^{\infty} \frac{\partial^n f}{\partial x_1^n}\frac{\partial^n g}{\partial x_2^n}\frac{t^n}{n!} \ .$$

Observation 5.1.9.

1. *The phase space of a Hamiltonian system is a Poisson manifold. The Poisson structure can be interpreted as a hint to the existence of a family of associative algebras (which in general are not commutative any longer). This idea is at the basis of all ideas of quantizations.*

2. *The Poisson bracket on a phase space is in local Darboux coordinates*

$$\{f, g\} = \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} \ .$$

Hence it has dimension $[length \times momentum]^{-1} = action^{-1}$. As a consequence, the deformation parameter t has to be a dimensionful quantity, $t = \hbar$ with the dimension of an action. We expect thus in quantum mechanics a new fundamental constant of nature, Planck's constant. It has the value of $1.054 \cdot 10^{-34} Js$.

3. *A dimensionful quantity like $\hbar$ is neither big nor small per se. Only its ratio with a quantity of the same dimension, i.e. a characteristic action of a system can be small or big.*

4. *There are no evaluation homomorphisms for a formal deformation for values of the deformation parameter other than zero. Hence we have to look for a different framework that allows to formulate convergence properties: normed algebras.*

5.2 Kinematical framework for quantum mechanics: C^*-algebras and states

Definition 5.2.1
A (unital) C^*-algebra $(A, \|\cdot\|, *)$ is a complex (unital) associative algebra A together with a norm $\|\cdot\| : A \to \mathbb{R}_{\geqslant 0}$ and a $\mathbb{C}$-antilinear map

$$\begin{array}{rccc} * : & A & \to & A \\ & a & \mapsto & a \end{array}$$

where $\mathbb{C}$-antilinear means that

$$(\lambda a + \mu b)^* = \bar{\lambda} a^* + \bar{\mu} b^* \quad \text{for all } \lambda, \mu \in \mathbb{C},\ a, b \in A$$

such that the following conditions are obeyed:

1. *$(A, \|\cdot\|)$ is a complete normed topological vector space.*
2. *The norm is submultiplicative*
$$\|a \cdot b\| \leq \|a\| \cdot \|b\| \quad \text{for all } a, b \in A\ .$$
3. *The antilinear map $*$ is an involution, $a^{**} = a$ for all $a \in A$.*
4. *The antilinear map $*$ is an algebra antihomomorphism, $(ab)^* = b^* a^*$ for all $a, b \in A$.*
5. *The antilinear map $*$ is an isometry for the norm, $\|a^*\| = \|a\|$ for all $a \in A$.*
6. *The so-called C^*-property holds,*
$$\|a^* a\| = \|a\|^2 \quad \text{for all } a \in A\ .$$

A morphism $\pi : A \to B$ of unital C^*-algebras is a unital algebra morphism for which

$$\pi(a^*) = \pi(a)^* \quad \text{for all} \quad a \in A\ .$$

Remarks 5.2.2.

1. *Properties 1. and 2. in the definition imply that any C^*-algebra is in particular a Banach algebra.*
2. *We will see later that a unital morphism $\pi : A \to B$ of unital C^*-algebras automatically obeys*
$$\|\pi(a)\| \leq \|a\|\ .$$
In particular, it is continuous.

3. *The C^*-property $\|a^*a\| = \|a\|^2$ implies that*

$$\|1\| = 1 \quad \text{and} \quad \|a^*\| = \|a\|$$

We should point out that the axioms 1.-6. are not independent. For example, axiom 5. can be deduced from axiom 2,3 and 6.

4. *Elements $a \in A$ such that $a^* = a$ are called self-adjoint. They form a real subspace, but in general not an associative subalgebra. The unit 1_A can be shown to be self-adjoint. A self-adjoint element $p \in A$ is called a projector, if it is an idempotent, i.e. if $p^2 = p$. An element is called an isometry, if $a^*a = 1$ amd unitary, if $a^*a = aa^* = 1$.*

Example 5.2.3.

- *Let $\mathcal{H}$ be a separable Hilbert space, i.e. a complete unitary vector space with a topological basis that is at most countable. The scalar product $\langle -, - \rangle$ on a Hilbert space $\mathcal{H}$ is in our conventions always linear in the first argument and antilinear in the second argument. A $\mathbb{C}$-linear map*

$$A: \ \mathcal{H} \to \mathcal{H}$$

is called bounded, if the so-called operator norm

$$\|A\| := \sup_{\|x\|=1} \|Ax\|$$

is finite. The inequality

$$\|Ax - Ax\| = \|A(x-y)\| \leq \|A\| \cdot \|x - y\|$$

implies that such $\mathbb{C}$-linear maps are continuous. The space of bounded operators on a Hilbert space is a Banach algebra $B(\mathcal{H})$.

- *From Riesz representation theorem, we deduce that for given $A \in B(\mathcal{H})$ and $v \in \mathcal{H}$ the continuous $\mathbb{C}$-linear function*

$$\langle A\cdot, v\rangle: \ \mathcal{H} \to \mathbb{C}$$

can be represented by the scalar product with a unique vector $A^*v \in \mathcal{H}$:

$$\langle A\cdot, v\rangle = \langle \cdot, A^*v\rangle \ .$$

This defines the involution $*$ on $B(\mathcal{H})$.

Assuming that axiom 5 has been proven, we check the C^* property of $B(\mathcal{H})$:

$$\begin{aligned} \|A\|^2 &= \sup_{\|x\|=1} \|Ax\|^2 = \sup_{\|x\|=1} \langle Ax, Ax\rangle = \sup_{\|x\|=1} \langle x, A^*Ax\rangle \\ &\leq \sup_{\|x\|=1} \|x\| \cdot \|A^*Ax\| = \|A^*A\| \overset{Axiom\ 2}{\leq} \|A\| \cdot \|A^*\| \\ &\overset{Axiom\ 5}{=} \|A\|^2 \ , \end{aligned}$$

where we used the definitions and the Cauchy-Schwarz identity.

- *A theorem of Gelfand and Naimark asserts that every C^*-algebra is isomorphic to a closed $*$-subalgebra of $B(\mathcal{H})$.*

Examples 5.2.4.

1. *Let X be a locally compact Hausdorff space. Denote by $C_0(X)$ the set of all continuous functions $f : X \to \mathbb{C}$ such that for any $\epsilon > 0$ there is a compact subset $K_\epsilon \subset X$ such that $|f(x)| < \epsilon$ for $x \notin K_\epsilon$. We call $C_0(X)$ the algebra of continuous functions vanishing at infinity. Obviously, if X is compact, this equals the algebra $C(X)$ of all continuous functions on X, i.e. $C_0(X) = C(X)$.*

   Since all functions $f \in C_0(X)$ are bounded, we can introduce the supremum norm

$$\|f\| := \sup_{x \in X} |f(x)| \ .$$

   Complex conjugation of the values of the function gives a natural involution,

$$f^*(x) \ := \overline{f(x)} \ .$$

   Then $(C_0(X), \|\cdot\|, *)$ is a commutative C^*-algebra. It is unital, if and only if X is compact.

2. *Let X be a smooth manifold. Consider*

$$C_0^\infty(X) := \ C^\infty(X) \cap C_0(X) \ ,$$

   the algebra of smooth functions vanishing at infinity. It is a subalgebra of $C_0(X)$ and inherits the norm and the involution $*$. It obeys all axioms of a commutative C^*-algebra, except for the fact that it is not complete. Rather, it is dense in $C_0(X)$.

The idea is that a C^*-algebra serves as an algebra of observables for a quantum mechanical system and thus provides the *kinematical* setting. In concrete cases, we might think of this algebra heuristically as a "quantization" of the Poisson algebra of functions on phase space (M, ω). The dynamics of a Hamiltonian system (N, ω, h) is specified by an element h in the Poisson algebra. Correspondingly, the dynamics of a quantum mechanical system will be specified by an element of the C^*-algebra. We postpone the discussion of this aspect of quantum mechanics and restrict ourselves to kinematical aspects.

In classical Hamiltonian mechanics, we consider the time evolution of particles or, more generally, of probability distributions. Quantum mechanics is an inherently probabilistic theory; we thus need the analogue of a probability distribution. In particular, we want to associate to any observable at least an expectation value. This motivates the following definition:

Definition 5.2.5
Let A be a unital C^*-algebra. A state ω on A is a normed positive linear functional

$$\omega : \ \ A \to \mathbb{C}.$$

Positivity for a linear form $\omega \in A^*$ means $\omega(a^*a) \geqslant 0$ for all $a \in A$ and the normalization condition is explicitly

$$\|\omega\| := \sup_{\|a\|=1} |\omega(a)| = 1 \ .$$

Remarks 5.2.6.

1. *The set of all states is convex: if ω_1 and ω_2 are states, then for all $t \in [0, 1]$ the convex linear combination*

$$\omega_t = t\omega_1 + (1 - t)\omega_2$$

   is a state.

2. *A point x of a convex subset X of a real vector space V is called extremal, if for every segment $\overline{yz} \subset X$ containing x one has either $y = x$ or $z = x$. Extremal states on a C^*-algebra are called pure states. States that are not pure states are also called mixed states.*

3. *Pure states exist: according to the Krein-Mil'man theorem every compact convex subset of a locally convex vector space equals the closed convex hull of its extremal points.*

Examples 5.2.7.

1. *The C^*-algebra $C_0(\mathbb{R})$ is not unital. Normalized states are just normalized Radon measures, i.e. normalized probability measures.*

   The pure states of $C_0(\mathbb{R})$ are given by the evaluation of the continuous function f at some point $x \in \mathbb{R}$,

   $$\delta_x(f) = f(x)$$

   and thus in bijection to points in $\mathbb{R}$. Such a measure is called a Dirac measure.

2. *Consider the C^*-algebra $B(\mathcal{H})$ of bounded operators on a Hilbert space $\mathcal{H}$. An endomorphism ρ of a Hilbert space is called positive, if $\langle \rho x, x\rangle \geq 0$ for all $x \in \mathcal{H}$. An positive endomorphism ρ of $\mathcal{H}$ is called a density matrix, if the endomorphism $\rho \circ a$ is trace class for all $a \in A$. Since this includes the case $a = \mathrm{id}_{\mathcal{H}}$, the endomorphism ρ itself is trace class and thus in particular continuous. Then*

   $$\omega_\rho(a) := \frac{\mathrm{Tr}(\rho a)}{\mathrm{Tr}(\rho)}$$

   is a state on the C^*-algebra $B(\mathcal{H})$.

   Let $\psi \in \mathcal{H}$ be a unit vector and

   $$\begin{array}{rccc} \rho: & \mathcal{H} & \to & \mathcal{H} \\ & x & \mapsto & \langle x, \psi\rangle \psi \end{array}$$

   the orthogonal projection to the one-dimensional subspace spanned by ψ. Then the state

   $$\omega_\rho(a) = \langle a\psi, \psi\rangle$$

   is a pure state.

We next endow the continuous dual of a Banach space, i.e. the $\mathbb{C}$-vector space of continuous linear functions on A, with a topology. Two topologies are of particular importance:

Remark 5.2.8.

1. *Let A be a complex Banach space and A^* its continuous dual. It can be endowed with the structure of a Banach space by the operator norm*

   $$\|\mu\| := \sup_{\|a\|=1} |\mu(a)| \ .$$

2. *There is a second important topology on the continuous dual A^*, the weak $*$ topology or topology of pointwise convergence. This topology is generated by the following collection of subsets of A^*: for $\mu_0 \in A^*$, $a \in A$ and $\delta > 0$ we set*

   $$U_\delta(\mu_0; a) := \left\{ \mu \in A^* \;\middle|\; |\mu(a) - \mu_0(a)| < \delta \right\} \ .$$

   This implies that a sequence (λ_n) with $\lambda_n \in A^*$ converges to $\lambda \in A^*$, if and only if $\lim_{n\to\infty} \lambda_n(a) = \lambda(a)$ for all $a \in A$,

Definition 5.2.9
Consider a C^* algebra A.

1. *An element $\mu \in A^*$ is called a character, if $\mu \neq 0$ and*

$$\mu(ab) = \mu(a)\mu(b) \quad \text{for all } a, b \in A \ .$$

2. *The subset $M(A) \subset A^*$ of all characters of A can be endowed with the topology induced by the weak $*$-topology on A^*. This topological space is called the Gelfand spectrum of the C^*-algebra A.*

Example 5.2.10.
Let X be a locally compact Hausdorff space and $C_0(X)$ be the C^*-algebra of complex-valued continuous functions on X. Then for every point $p \in X$, the evaluation map $f \mapsto f(p)$ is a character. The map

$$\begin{array}{rcl} X & \to & M(C_0(X)) \\ p & \mapsto & \mu_p(f) = f(p) \end{array}$$

that associates to every point p the character corresponding to evaluation in p is injective and continuous.

Proof:

- To see injectivity, consider two different points $p \neq q$. Since X is a locally compact Hausdorff space, we can find a continuous function $f \in C_0(X)$ with $f(p) \neq f(q)$. The inequality
$$\mu_p(f) = f(p) \neq f(q) = \mu_q(f)$$
implies that the characters given by evaluation at p and q respectively are different, $\mu_p \neq \mu_q$.
- Consider a convergent sequence $p_i \to p$ in X. Then for every $f \in C_0(X)$, continuity of f implies
$$\mu_{p_i}(f) = f(p_i) \to f(p) = \mu_p(f) \ ,$$
which is just pointwise convergence $\mu_{p_i} \to \mu_p$ in $(C_0(X))^*$, i.e. convergence in the weak $*$-topology.

□

Definition 5.2.11
Let A be a C^*-algebra. Any element $a \in A$ determines a complex-valued function on the Gelfand spectrum of $M(A)$:

$$\begin{array}{rcccl} \hat{a}: & M(A) & \to & \mathbb{C} \\ & \mu & \mapsto & \mu(a) \end{array}$$

The map that associates to an element of A a function on the spectrum

$$\begin{array}{rccc} \mathcal{G}: & A & \to & C_0(M(A)) \\ & a & \mapsto & \hat{a} \end{array}$$

is called the Gelfand transform.

We quote the following theorem:

Theorem 5.2.12 (Gelfand-Naimark)**.**
Let A be a commutative *C^*-algebra. Then the Gelfand transform $\mathcal{G}$ is an isometric $*$-algebra isomorphism.*

Remarks 5.2.13.

1. *A commutative C^*-algebra is thus in a natural way isomorphic to the algebra of continuous functions on a topological space that is encoded in the algebra, the Gelfand spectrum. The point of view to see C^*-algebras as generalizations of algebras of functions is the starting point for non-commutative geometry, more precisely non-commutative measure theory.*

2. *The pure states on a commutative C^*-algebra are its characters. Hence any commutative C^*-algebra is in a natural way isomorphic to the C^*-algebra of continuous functions on its pure states.*

3. *The Gelfand spectrum $M(A)$ is compact, if and only if the C^* algebra is unital.*

Definition 5.2.14

1. *A $*$-representation of a C^*-algebra A on a separable Hilbert space $\mathcal{H}$ is a $*$-preserving ring homomorphism*

$$\pi: \quad A \to B(\mathcal{H}) \ .$$

If the algebra A is unital, the map π is required to preserve the unit element. Otherwise, one requires the subspace

$$\left\{\pi(a)\xi \;\middle|\; \text{for all } \ a \in A,\, \xi \in \mathcal{H}\right\} \ \subset \mathcal{H}$$

to be dense in $\mathcal{H}$.

2. *A representation is called faithful, if π is injective. A subset $U \subset \mathcal{H}$ is called invariant under A, if*

$$\pi(A)U := \{\pi(a)u \mid a \in A, u \in U\} \subset U \ .$$

A representation is called irreducible, if the only closed vector subspaces of $\mathcal{H}$ invariant under A are $\{0\}$ and $\mathcal{H}$.

3. *A vector $\xi \in \mathcal{H}$ in a $*$-representation $\pi : A \to B(\mathcal{H})$ is called cyclic, if the subspace*

$$\{\ \pi(a)\xi \mid a \in A\ \}$$

is dense in $\mathcal{H}$. If a cyclic vector exists, the representation is called cyclic.

4. *Two representations* $\pi_1 : A \to B(\mathcal{H}_1)$ *and* $\pi_2 : A \to B(\mathcal{H}_2)$ *are called* unitarily equivalent, *if there exists a unitary operator* $U : \mathcal{H}_1 \to \mathcal{H}_2$ *such that for all* $a \in A$

$$U \circ \pi_1(a) \;=\; \pi_2(a) \circ U \;.$$

Remarks 5.2.15.

1. *The commutative* C^**-algebra* $C(X)$ *of continuous functions on an open subset* $X \subset \mathbb{R}^n$ *has a natural representation on the Hilbert space* $L^2(X, \mu)$ *of functions that are square integrable by the Lebesgues measure* μ *by multiplication. The constant function* 1 *is cyclic, because the continuous functions are dense in* $C_0(X)$*.*

2. *Every non-zero vector in an irreducible representation is cyclic.*

3. *In contrast, an arbitrary non-zero vector in a cyclic representation is not necessarily cyclic.*

Next, we construct states on an abstract C^*-algebra A from a suitable $*$-representations of A.

Definition 5.2.16
Consider a C^**-algebra* A *and a* $*$*-representation* $\pi : A \to B(\mathcal{H})$*.*

1. *For every non-zero vector* $v \in \mathcal{H} \setminus \{0\}$ *the function*

$$\omega_v(a) := \frac{\langle \pi(a)v, v\rangle}{\langle v, v\rangle} \quad a \in A$$

is a state on A*. Such a state is called a* vector state*; it is also called a* ray state*, because non-zero vectors in the same one-dimensional subspace of* $\mathcal{H}$ *give the same state.*

2. *More generally, for every density matrix* ρ *on* $\mathcal{H}$ *the function*

$$\omega_\rho(a) = \mathrm{Tr}\, \rho\pi(a)/\mathrm{Tr}\rho \quad a \in A$$

is a state on A*. Such states are called* normal states*.*

Our goal is to show that *all* states on a C^*-algebra can be obtained this way.

Lemma 5.2.17.
Let ω *be a state on a* C^**-algebra* A *with unit. Then we have:*

1. *The map*

$$\begin{array}{rcl} A \times A & \to & \mathbb{C} \\ (a, b) & \mapsto & \omega(a^* \cdot b) \end{array}$$

is a positive semi-definite, Hermitian sesquilinear form.

2. *The Cauchy-Schwarz inequality holds:*

$$|\omega(b^*a)|^2 \leq \omega(a^*a) \cdot \omega(b^*b) \;\; \textit{for all} \;\; a, b \in A\,.$$

3. $\omega(a^*) = \overline{\omega(a)}$ *for all* $a \in A$.

4. $|\omega(a)|^2 \leq \omega(a^*a)$ *for all* $a \in A$.

5. $\omega(1) = \|\omega\| = 1$.

Proof:
The sesquilinearity and positive semi-definiteness of the bilinear form follows immediately from the definitions. To see that the form is hermitian, consider for $a, b \in A$ and any $z \in \mathbb{C}$ the element $c := az + b \in A$ and compute

$$\begin{aligned} 0 &\leq \omega(c^*c) \\ &= |z|^2\omega(a^*a) + \overline{z}\omega(a^*b) + z\omega(b^*a) \ + \omega(b^*b) \ . \end{aligned} \tag{14}$$

Thus the imaginary part of $\overline{z}\omega(a^*b) + z\omega(b^*a)$ has to vanish for all $z \in \mathbb{C}$. For $z = 1$, we obtain

$$\mathrm{Im}\,\omega(a^*b) \ = -\mathrm{Im}\,\omega(b^*a)$$

and from $z = \mathrm{i}$, we obtain from equation (14)

$$\mathrm{Re}\omega(a^*b) \ = \mathrm{Re}\omega(b^*a)$$

and thus $\omega(a^*b) = \overline{\omega(b^*a)}$. This shows that the form is hermitian.

Setting $z = -\frac{\omega(a^*b)}{\omega(a^*a)}$ in equation (14) yields

$$0 \leq \frac{|\omega(a^*b)|^2}{\omega(a^*a)} - \frac{|\omega(a^*b)|^2}{\omega(a^*a)} - \frac{|\omega(a^*b)|^2}{\omega(a^*a)} + \omega(b^*b)$$

and thus the Cauchy-Schwarz inequality.

Since A has a unit which is self-adjoint and since the form given by ω is hermitian, we have

$$\omega(a^*) = \omega(a^*1) \ = \overline{\omega(1^*a)} = \overline{\omega(a)} \ .$$

To show assertion 4, we use the Cauchy-Schwarz identity and $\omega(1) \leq 1$ to compute

$$|\omega(a)|^2 = |\omega(1^*a)|^2 \leq \omega(1^*1) \cdot \omega(a^*a) = \omega(1) \cdot \omega(a^*a) \leq \omega(a^*a) \ .$$

Using $\omega(1) = \omega(1^*1) \ \geq 0$, we compute

$$|\omega(a)|^2 \leq \omega(1^*1) \cdot \omega(a^*a) \leq \omega(1) \cdot \|\omega\| \cdot \|a^*a\| = \omega(1) \cdot \|a\|^2 \ .$$

We thus have

$$1 = \|\omega\|^2 \leq \sup_{a \in A \setminus \{0\}} \frac{|\omega(a)|^2}{\|a\|^2} \leq \omega(1)$$

and hence $\omega(1) = 1$. □

Lemma 5.2.18.
Let ω *be a state on a* C^**-algebra* A*. Then the following holds:*

1. $\omega(a^*a) = 0$ *if and only if* $\omega(ba) = 0$ *for all* $b \in A$.

2. *We have*
$$\omega(b^*a^*ab) \leq \|a^*a\| \cdot \omega(b^*b) \quad \text{for all } a, b \in A \ .$$

Proof:

1. Suppose $\omega(a^*a) = 0$. Then the Cauchy-Schwarz identity implies
$$|\omega(ba)|^2 \leq \omega(a^*a)\omega(b^*b) = 0$$
for all $b \in A$. The converse direction is obvious.

2. If $\omega(b^*b) = 0$, then according to the first assertion, we have $\omega(cb) = 0$ for all $c \in A$, in particular for $c = b^*a^*a$, so that the inequality holds. We may thus assume $\omega(b^*b) \neq 0$ and put
$$\rho(c) := \frac{\omega(b^*cb)}{\omega(b^*b)} \ .$$
Clearly, ρ is a positive linear functional with $\|\rho\| = \rho(1) = 1$ and thus a state. From 5. in lemma 5.2.17, we have $\rho(a^*a) \leq \|a^*a\|$.

□

Definition 5.2.19
Let A be a C^*-algebra and ω a state on A. Elements of A are also called observables. For $a \in A$, we call
$$\langle a \rangle_\omega := \omega(a)$$
the expectation value of the observable a in the state ω and
$$\Delta_\omega(a)^2 := \langle (a - \langle a \rangle)^2 \rangle_\omega = \langle a^2 \rangle_\omega - \langle a \rangle_\omega^2 = \omega(a^2) - \omega(a)^2 \ .$$
the variance of a in the state ω.

We are now able to derive a lower bound on the product of the variances. It implies that there is no state in which two observables with canonical commutation relations have both variance zero.

Proposition 5.2.20 (Heisenberg's uncertainty relations)**.**
Given a C^*-algebra A and self-adjoint elements $a, b \in A$ and a state ω of A, the following relation for the variance holds:
$$\Delta_\omega(a)\Delta_\omega(b) \geqslant \frac{1}{2}\big|\omega([a, b])\big| \ .$$

In the special case when we consider two element $p, q \in A$ with commutator $[p, q] = i\hbar 1$ for some real positive number $\hbar$, i.e. when p and q obey canonical commutation relations, we find
$$\Delta_\omega(q)\Delta_\omega(p) \geqslant \frac{\hbar}{2} \ .$$

Proof:

We decompose the product ab into the sum of the anti-commutator $\{a,b\} := ab + ba$ and the commutator:

$$ab = \frac{1}{2}(ab+ba) \ + \frac{1}{2}(ab-ba) \ = \frac{1}{2}\{a,b\} + \frac{1}{2}\,[a,b]\,.$$

If a and b are self-adjoint, the anti-commutator is self-adjoint, the commutator is anti self-adjoint:

$$\begin{aligned}\{a,b\} &= ab+ba = a^*b^* + b^*a^* = (ba)^* + (ab)^* = \{a,b\}^* \\ [a,b] &= -\,[b,a]^* \ .\end{aligned}$$

By lemma 5.2.17.3, the state ω yields a real value on the anticommutator and a purely imaginary value on the commutator. Thus

$$\big|\omega(ab)\big|^2 = \frac{1}{4}\big|\omega(\{a,b\}) + \omega([a,b])\big|^2 = \frac{1}{4}\omega(\{a,b\})^2 + \frac{1}{4}\big|\omega([a,b])\big|^2.$$

We find an upper bound for the left hand side using the Cauchy-Schwarz identity for ω:

$$\big|\omega(ab)\big|^2 \leq \omega(b^2)\omega(a^2),$$

Altogether we find

$$\omega(a^2)\omega(b^2) \geqslant \frac{1}{4}\big|\omega([a,b])\big|^2.$$

Now set $\tilde{a} := a - \omega(a)1$ and $\tilde{b} := b - \omega(b)1$ and note that

$$\Big[\tilde{a},\tilde{b}\Big] = [a,b] \ \ ;$$

we thus deduce from the inequality for $\tilde{a}$ and $\tilde{b}$ the inequality

$$\Delta_\omega(a)^2\Delta_\omega(b)^2 \geqslant \frac{1}{4}|\omega([a,b])|^2.$$

□

Theorem 5.2.21 (Gelfand-Naimark-Segal)**.**

1. *Let ω be a state on a unital C^*-algebra A. Then there is a $*$-representation π_ω of A on some Hilbert space $(\mathcal{H}_\omega, \langle\cdot,\cdot\rangle_\omega)$ with a cyclic vector $\Omega_\omega \in \mathcal{H}_\omega$ such that*

   $$\omega(a) = \langle \pi_\omega(a)\Omega_\omega, \Omega_\omega\rangle_\omega$$

   holds for all $a \in A$. This representation is unique up to unitary equivalence.

2. *The representation π_ω is irreducible, if and only if ω is a pure state.*

Proof:
The idea is to pass to a quotient of A make the degenerate hermitian product $(a,b) \mapsto \omega(a^*b)$ on A non-degenerate. By assertion 1 of lemma 5.2.18, the null space

$$N_\omega := \{a \in A \mid \omega(a^*a) = 0\}$$

is closed. By assertion 2 in the same lemma, it is a left ideal of A. Therefore, the pairing

$$\begin{array}{rcl} A/N_\omega \times A/N_\omega & \to & \mathbb{C} \\ ([a],[b]) & \mapsto & \omega(b^*a) \end{array}$$

is a well-defined Hermitian scalar product $\langle \cdot, \cdot \rangle_\omega$ and endows A/N_ω with the structure of a pre-Hilbert space. Denote by $\mathcal{H}_\omega$ the completion of the pre-Hilbert space.

Then the map

$$\pi_\omega : \ A \to B(A/N_\omega)$$

with

$$\pi_\omega(a).[b] := \ [ab]$$

is well-defined and satisfies

$$\|\pi_\omega(a) \cdot [b]\|^2 = \omega(b^*a^*ab) \ \leq \|a^*a\| \cdot \omega(b^*b) = \|a\|^2 \|[b]\|^2$$

so that $\|\pi_\omega(a)\| \leq \|a\|$ and $\|\pi_\omega\| \leq 1$. The map π_ω thus extends to a representation

$$\pi_\omega : \ A \to B(\mathcal{H}_\omega)$$

This gives a $*$-representation of A with cyclic vector $\Omega_\omega := [1] \in \mathcal{H}_\omega$, since

$$\pi_\omega(A).\Omega_\omega = A/N_\omega \subset \mathcal{H}_\omega$$

is dense in $\mathcal{H}$. The state ω can be expressed using this representation:

$$\omega(a) \ = \omega(1^*a1) = \langle [a1], [1] \rangle_\omega = \langle \pi_\omega(a).\Omega_\omega, \Omega_\omega \rangle_\omega$$

□

The importance of the Gelfand-Naimark theorem is that it allows to consider the Hilbert space as a derived concept in quantum mechanics and the algebra of observables as the fundamental object.

Another consequence is the following

Remark 5.2.22.
Let A be a C^*-algebra and $S(A)$ the set of states for A. The direct sum representation

$$\oplus_{\omega \in S(A)} \pi_\omega : A \to L\big(\oplus_{\omega \in S(A)} \mathcal{H}_\omega \big)$$

is called the universal representation of A. It can be shown to be faithful. Hence any C^* algebra has a faithful representation and can be identified with a subalgebra of the C^*-algebra of bounded operators on a separable Hilbert space.

Example 5.2.23.
For $A = C(X)$ with X a topological space and with a state ω_μ given by a probability measure μ on X as

$$\omega_\mu(f) \ = \int_X f \mathrm{d}\mu \ ,$$

the representation space of the GNS representations is $L^2(X, \mu)$.

5.3 Composite systems and Bell's inequality

Observation 5.3.1.

1. *The composition of two Hamiltonian mechanical systems (M_1, ω_1, h_1) and (M_2, ω_2, h_2) is described by the smooth product manifold $M_1 \times M_2$ with symplectic form $p_1^*\omega_1 + p_2^*\omega_2$. If the Hamiltonian $p_1^*h_1 + p_2^*h_2$ is chosen, the system is said to be non-interacting.*

2. *In quantum mechanics, the situation is as follows: given two C^*-algebra A_1 and A_2, there are many norms on the algebraic tensor product $A_1 \otimes A_2$ that yield the structure of a C^*-algebra on the corresponding completion. However, given two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, the algebraic tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$ has a unique completion.*

We choose to work with the following tensor product:

Definition 5.3.2
Let $(A, \|\cdot\|_A, *)$ and $(B, \|\cdot\|_B, *)$ two C^*-algebras. The projective C^*-norm $\|\cdot\|_\pi$ on the algebraic tensor product $A \otimes B$ is defined by

$$\|c\|_\pi := \inf \Big\{ \sum_{j=1}^{n} \|a_j\|_A \cdot \|b_j\|_B \;\Big|\; c = \sum_{j=1}^{n} a_j \otimes b_j \Big\} .$$

The completion of $A \otimes B$ with respect to this C^*-norm is called the projective tensor product and is denoted by $A \otimes_\pi B$.

Remark 5.3.3.
The projective tensor product has the following universal property: let A, B and C be C^*-algebras and let $\varphi : A \to C$ and $\psi : B \to C$ be $*$-morphisms such that $[\varphi(a), \psi(b)] = 0$ for all $a \in A$ and all $b \in B$. Then there exists a unique $*$-morphism $\chi : A \otimes B \to C$ such that

$$\chi(a \otimes b) = \varphi(a)\psi(b) \qquad \text{for all} \quad a \in A, b \in B .$$

Observation 5.3.4.
We now study states on the tensor product $A \otimes_\pi B$. Given two linear functionals $\mu : A \to \mathbb{C}$ and $\nu : B \to \mathbb{C}$, we get a linear functional $\mu \otimes \nu : A \otimes B \to \mathbb{C}$ with $\mu \otimes \nu(a \otimes b) = \mu(a) \cdot \nu(b)$. In the projective norm, we have

$$\|\mu \otimes \nu\|_\pi = \|\mu\|_A \cdot \|\nu\|_B .$$

Finally, one checks that $\mu \otimes \nu$ is positive, if μ and ν are positive.

Definition 5.3.5
Let A and B be C^*-algebras and let μ be a state on A and ν a state on B. The unique extension of $\mu \otimes \nu$ to the projective tensor product $A \otimes_\pi B$ is called a product state.

Let us now assume that the C^*-algebras are unital. Given a state τ on $A \otimes_\pi B$, we get states on the factors by

$$\tau^A(a) := \tau(a \otimes 1_B) \qquad \text{and} \qquad \tau^B(b) := \tau(1_A \otimes b) .$$

If τ is a product state, $\tau = \mu \otimes \nu$, then $\tau^A = \mu$ and $\tau^B = \nu$. Put differently, a measurement in the product state τ of an observable $a \otimes b$ simply results in the product of the measurements in the states μ and ν. For a general state on $A \otimes B$, this is not necessarily true.

Definition 5.3.6

1. *A state τ on $A \otimes_\pi B$ is called correlated, if there exists $a \in A$ and $b \in B$ such that $\tau(a \otimes b) \neq \tau^A(a) \cdot \tau^B(b)$.*

2. *A state τ on $A \otimes_\pi B$ is called decomposable, if it is the pointwise limit of convex combinations of product states. (The set of decomposable states is thus the weak-$*$ closure of the convex hull of the product states.)*

3. *A state τ on $A \otimes_\pi B$ is called entangled, if it is not decomposable.*

For the proof of the following proposition, we refer to the literature:

Proposition 5.3.7.
Let A and B be C^*-algebras with unit. If A or B are abelian, then all states on the projective tensor product $A \otimes_\pi B$ are decomposable.

Thus identifying entangled states in a quantum mechanical system amounts to establishing that indeed non-commutative algebras have to be used in quantum mechanics. In this way, one establishes that quantum mechanics is fundamentally different from classical theories.

A different aspect of entangled states is as follows: suppose we spatially separate the two subsystems. If we had only product states, this would mean that the expectation values of the two systems are fixed in a completely independent way. Having only decomposable states would amount to having only a *classical* probability superpose in the form of a convex linear combination whose coefficients could be interpreted as probabilities. As we will see in an example below, experiments tell us that entangled states exist and that quantum mechanics can thus not be reduced to such "classical" superpositions.

Lemma 5.3.8 (Bell's inequality)**.**
Let A and B be C^*-algebras and let τ be a decomposable state on $A \otimes_\pi B$. Then for all self-adjoint elements $a, a' \in A$ and $b, b' \in B$ of norm smaller or equal than 1, we have

$$|\tau(a \otimes (b - b'))| + |\tau(a' \otimes (b + b'))| \leq 2$$

Proof:
For a product state $\tau = \mu \otimes \nu$, we have

$$\tau(a \otimes (b - b')) = \mu(a)\nu(b) - \mu(a)\nu(b')$$
$$= \mu(a) \cdot \nu(b) \cdot (1 \pm \mu(a') \cdot \nu(b')) - \mu(a) \cdot \nu(b')(1 \pm \mu(a') \cdot \nu(b))$$

Since by assumption $|\mu(a)|$, $|\mu(a')|$, $|\nu(b)|$ and $|\nu(b')| \leq 1$, we have

$$\begin{aligned} |\tau(a \otimes (b - b'))| &\leq |1 \pm \mu(a') \cdot \nu(b')| + |1 \pm \mu(a') \cdot \nu(b)| \\ &= 1 \pm \mu(a') \cdot \nu(b') + 1 \pm \mu(a') \cdot \nu(b) \\ &= 2 \pm \tau(a' \otimes (b + b')) \end{aligned}$$

Hence the inequality holds for product states. If τ is a convex combination of product states,

$$\tau = \sum_{j=1}^{n} \lambda_j \mu_j \otimes \nu_j$$

with $\lambda_i \in [0,1]$ and $\sum_{j=1}^n \lambda_i = 1$ we obtain

$$\begin{array}{c} |\tau(a \otimes (b-b'))| + |\tau(a' \otimes (b+b'))| \\ \leq \sum_{j=1}^n \lambda_j \left\{ (\mu_j \otimes \nu_j (a \otimes (b-b')) + (\mu_j \otimes \nu_j)(a' \otimes (b+b')) \right\} \leq 2 \end{array}$$

For pointwise limits of convex combinations, the inequality holds by continuity. □

The following example is quite important:

Example 5.3.9.
Let $A = B = M(2 \times 2, \mathbb{C})$ be complex matrix algebras. Denote by (e_1, e_2) the ordered standard basis of $\mathbb{C}^2$. On $A \otimes B$, we consider the Bell state, i.e. the pure state with vector

$$\Omega := \frac{1}{\sqrt{2}} (e_1 \otimes e_1 - e_2 \otimes e_2) \ .$$

Such a state can be realized experimentally e.g. in terms of spin degrees of freedom of the decay of a spin zero system into two particles, e.g. of the η meson in the rare decay $\eta \to \mu^+ + \mu^-$. In praxis, one rather uses proton-proton scattering at low energies.

To see that the Bell state is entangled, consider the four observables

$$a := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad a' := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad b := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad b' := \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$$

We compute

$$\begin{array}{rcl} \tau(a \otimes (b-b') & = & \sqrt{2}\tau(a \otimes a) = \sqrt{2}\langle (a \otimes a)\Omega, \Omega \rangle \\ & = & \sqrt{2}\langle \Omega, \Omega \rangle = \sqrt{2} \end{array}$$

and similarly $\tau(a' \otimes (b+b')) = \sqrt{2}$ and hence

$$|\tau(a \otimes (b-b')| + |\tau(a' \otimes (b+b'))| = 2\sqrt{2} > 2 \ .$$

Thus the state τ violates Bell's inequality and is therefore entangled. This entanglement has been established experimentally with very high accuracy.

5.4 Dynamics of quantum mechanical systems

We need a final piece of mathematical theory to discuss concrete quantum mechanical systems, which allows us to deal with specific operators that represent physical observables. One of these operators will be the Hamiltonian which describes the dynamics of the system.

We need a few measure-theoretic preliminaries:

Definition 5.4.1

1. *Let X be a set. A sigma-algebra is collection $\mathcal{X}$ of subsets of X with the properties:*

- *The collection $\mathcal{X}$ is not empty.*
- *The collection $\mathcal{X}$ is closed under complements: if $U \subset X$ is in $\mathcal{X}$, then also $X \setminus U \in \mathcal{X}$.*
- *The collection $\mathcal{X}$ is closed under countable unions: if each element of the family $(U_i)_{i\in\mathbb{N}}$ is in $\mathcal{X}$, then $\cup_{i\in\mathbb{N}} U_i \in \mathcal{X}$.*

2. *Let now X be a topological space. A Borel set is a subset of X that can be obtained from open sets in X (or, equivalently, from closed sets in X) through the operations of countable unions, countable intersections, and relative complements.*

   The collection of all Borel sets on X forms a σ-algebra, the Borel σ-algebra of the topological space which is the smallest σ-algebra containing all open sets (or, equivalently, all closed sets).

Definition 5.4.2
Let A be a unital C^*-algebra. Denote by $\mathcal{B}(\mathbb{R})$ the sigma-algebra of Borel subsets of $\mathbb{R}$ with its standard topology. A (normalized) projector-valued measure on $\mathbb{R}$ with values in A is a mapping

$$P : \mathcal{B}(\mathbb{R}) \to A$$

satisfying the following axioms:

1. *The measure is projector valued: For every Borel subset $E \subset \mathbb{R}$, $P(E)$ is an orthogonal projector, i.e.*
$$P(E) = P(E)^2 \quad \text{and} \quad P(E) = P(E)^* \ .$$

2. *The measure is normalized: One has*
$$P(\emptyset) = 0 \quad \text{and} \quad P(\mathbb{R}) = 1_A \ .$$

3. *Additivity: for every countable disjoint union $E = \sqcup_{n=1}^{\infty} E_n$ of Borel subsets with pairwise empty intersection, one has*
$$P(E) = \lim_{n\to\infty} \sum_{i=1}^{n} P(E_i) \ .$$

Remarks 5.4.3.
1. *To every projector-valued measure P on $\mathbb{R}$, we associate the function*
$$\begin{array}{ccc} \mathbb{R} & \to & A \\ \lambda & \mapsto & P((-\infty, \lambda)) \end{array}$$
which, by abuse of notation, we denote by P as well. This function only takes orthogonal projectors as values and has the properties:
$$\lim_{\lambda\to-\infty} P(\lambda) = 0 \qquad \begin{array}{c} P(\lambda)P(\mu) = P(\min\{\lambda,\mu\}) \\ \text{and} \qquad \lim_{\lambda\to\infty} P(\lambda) = 1_A \\ \lim_{\mu\to\lambda^-} P(\mu) = P(\lambda) \end{array}$$
It is therefore called a projector valued resolution of the identity.

2. *Suppose, we are given a state ω on A and a projector-valued measure. Then the map*

$$\begin{array}{rcl} \mathcal{B}(\mathbb{R}) & \stackrel{P}{\to} & A \stackrel{\omega}{\to} \mathbb{R} \\ E & \mapsto & \omega(P(E)) \end{array}$$

defines a probability measure on $\mathbb{R}$.

In the special case when A is the algebra $B(\mathcal{H})$ of bounded operators on a separable Hilbert space $\mathcal{H}$, for any $v \in \mathcal{H} \setminus \{0\}$ the distribution function

$$P_v(\lambda) = \langle P(\lambda)v, v\rangle$$

defines a bounded measure on $\mathbb{R}$. It is a probability measure, if $|v| = 1$. More generally, for any pair of vectors $v, w \in \mathcal{H} \setminus \{0\}$, the function $\langle P(\lambda)v, w\rangle$ defines a complex measure on $\mathbb{R}$.

In quantum mechanics, *unbounded* linear operators on separable Hilbert spaces $\mathcal{H}$ are omnipresent. Their range of definition is only a dense subspace of $\mathcal{H}$.

Definition 5.4.4
Let $\mathcal{H}$ be a separable Hilbert space and let A be a linear operator in $\mathcal{H}$ with domain a linear subspace $D(A) \subset \mathcal{H}$.

1. *An operator is called closed, if its graph*

$$\Gamma(A) := \{(v, Av) \in \mathcal{H} \times \mathcal{H} \quad \textit{with} \quad v \in D(A)\}$$

is a closed subspace of $\mathcal{H} \times \mathcal{H}$.

2. *If the domain is dense, $\overline{D(A)} = \mathcal{H}$, the domain $D(A^*)$ of the adjoint operator consists of those $v \in \mathcal{H}$ such that there is $w \in \mathcal{H}$ with the property that*

$$\langle Ax, v\rangle = \langle x, w\rangle \quad \textit{for all} \;\; x \in \mathcal{H} \; .$$

We set set $w = A^*v$ for $v \in D(A)$.

3. *An operator with dense domain is called self-adjoint, if $A = A^*$. Equivalently, A is self-adjoint, if and only if A is symmetric and $\overline{D(A) = D(A^*)}$.*

4. *A symmetric operator is called essentially self-adjoint, if its closure $\overline{A} := A^{**}$ is self-adjoint.*

If A is bounded, it is an element of the C^*-algebra $B(\mathcal{H})$, and self-adjointness amounts to $A^* = A$. One should be aware of the fact that the C^*-algebra $B(\mathcal{H})$ does not contain unbounded operators. Still, unbounded self-adjoint operators yield interesting structures inside the C^*-algebra $B(\mathcal{H})$, as we will see.

We introduce the notion of the spectrum of an operator:

Definition 5.4.5

1. *Let A be a linear operator on a separable Hilbert space $\mathcal{H}$. (This operator is not necessarily bounded.) The spectrum of A is defined as*

$$\mathrm{Spec}(A) = \left\{ \lambda \in \mathbb{C} \;\middle|\; A - \lambda \mathrm{id}_{\mathcal{H}} \quad \textit{has no} \text{ bounded } \textit{inverse} \right\}$$

and the resolvent set is defined as

$$r(A) = \left\{ \lambda \in \mathbb{C} \;\middle|\; A - \lambda \mathrm{id}_{\mathcal{H}} \quad \textit{has a} \text{ bounded } \textit{inverse} \right\} \ ;$$

thus $r(A) = \mathbb{C} \setminus \mathrm{Spec}(A)$.

2. *Let a be an element of an abstract $\mathbb{C}^*$ algebra A with unit. We call*

$$r_A(a) := \{\lambda \in \mathbb{C} \; | \lambda 1_A - a \in A^\times\}$$

the resolvent set of a and

$$\mathrm{Spec}_A(a) = \mathbb{C} \setminus r_A(a)$$

the spectrum of a.

3. *The non-negative real number*

$$\rho_A(a) := \ \sup\{|\lambda| \mid \lambda \in \mathrm{Spec}_A(a)\}$$

is called the spectral radius of a.

Remarks 5.4.6.

1. *The definitions agree for elements of the C^*-algebra $B(\mathcal{H})$ of* bounded *operators on a separable Hilbert space $\mathcal{H}$.*

2. *We have*

$$\rho_A(a) = \lim_{n\to\infty} \|a^n\|^{\frac{1}{n}} = \inf_{n\in\mathbb{N}} \|a^n\|^{\frac{1}{n}} \leq \|a\| \ .$$

An element $a \in A$ is called normal, if $aa^* = a^*a$. (Any self adjoint element is obviously normal.) If a is normal, we have $\rho_A(a) = \|a\|$.

3. *The equality in 2. implies that unit-preserving $*$-morphisms $\pi : A \to B$ obey $\|\pi(a)\| \leq \|a\|$.*

First, note that if $a \in A^\times$ has a multiplictive inverse, then $\pi(a)^{-1} = \pi(a^{-1})$, since π is unit-preserving. Thus

$$\lambda 1_B - \pi(a) \ = \pi(\lambda 1_A - a)$$

so that $\lambda \in \mathbb{C}$ is in the resolvent of $\pi(a)$, if it is in the resolvent of a. Thus $r_A(a) \subset r_B(\pi(a))$ and thus $\mathrm{Spec}_B(\pi(a)) \subset \mathrm{Spec}_A(a)$. Thus

$$\rho_B(\pi(a)) \leq \rho_A(a) \ .$$

Since a^*a and $\pi(a)^*\pi(a)$ are self-adjoint, the C^*-property yields the estimate

$$\|\pi(a)\|^2 = \|\pi(a)^*\pi(a)\| = \rho_B(\pi(a)^*\pi(a)) = \rho_B(\pi(a^*a)) \leq \rho_A(a^*a) = \|a\|^2 \ .$$

4. *For any eigenvalue λ of an operator A on the Hilbert space $\mathcal{H}$, the operator $A - \lambda \cdot 1$ has a non-trivial kernel. Thus the eigenvalue λ is an element of the spectrum* $\mathrm{spec}A$*. The collection of eigenvalues is also called the point spectrum. If $\mathcal{H}$ is infinite-dimensional, the point spectrum is, in general, a proper subset of the spectrum.*

   As an example, consider the so-called bilateral shift T on the Hilbert space $l^2(\mathbb{Z})$ of sequences of complex numbers indexed by the integers which is defined by

$$T(\cdots, a_{-1}, \hat{a}_0, a_1, \cdots) = (\cdots, \hat{a}_{-1}, a_0, a_1, \cdots)$$

   where the hat denotes the zero-th position. Since T is a unitary operator, its spectrum is contained in the unit circle in $\mathbb{C}$.

   Direct calculation shows T has no eigenvalues, but every $\lambda \in \mathbb{C}$ with $|\lambda| = 1$ is in the spectrum. This can be seen by considering the family of unit vectors

$$x_n := \frac{1}{\sqrt{n}}(\ldots, 0, 1, \lambda^{-1}, \lambda^{-2}, \ldots, \lambda^{1-n}, 0, \ldots)$$

   for which we have

$$\|Tx_n - \lambda x_n\| = \sqrt{\frac{2}{n}} \to 0 \ .$$

   This immediately implies that $T - \lambda 1$ cannot have a bounded *inverse.*

We are now ready to formulate von Neumann's spectral theorem for self-adjoint operators on a separable Hilbert space.

Theorem 5.4.7.
For every self-adjoint operator A on a separable Hilbert space $\mathcal{H}$, there exists a unique projector-valued resolution $P_A(\lambda)$ of the identity satisfying the following properties:

- *(Spectral decomposition)*
  The range of definition of A can be characterized as

$$D(A) = \{v \in \mathcal{H} : \int_{-\infty}^{\infty} \lambda^2 \mathrm{d}\langle P_A(\lambda)v, v\rangle < \infty\}$$

  and for every $v \in D(A)$

$$Av = \int_{-\infty}^{\infty} \lambda \mathrm{d}P_A(\lambda)v$$

  defined as a limit of Riemann-Stieltjes sums. The support of the spectral measure $\mathrm{d}P_A(\lambda)$ equals the spectrum $\mathrm{Spec}(A)$ of A.

- *(Functional calculus)*
  For every continuous function f on $\mathbb{R}$, $f(A)$ is a linear operator on $\mathcal{H}$ with dense domain

$$D(f(A)) = \{v \in \mathcal{H} : \int_{-\infty}^{\infty} |f(\lambda)|^2 \mathrm{d}\langle P_A(\lambda)v, v\rangle < \infty\}$$

  and for every $v \in D(f(A))$

$$f(A)v = \int_{-\infty}^{\infty} f(\lambda)\mathrm{d}P_A(\lambda)v$$

- *One has* $f(A)^* = \overline{f}(A)$ *with* $\overline{f}$ *the complex conjugate function. The operator* $f(A)$ *is bounded, iff the continuous function* f *is bounded on the spectrum* $\mathrm{Spec}(A)$*. We thus have a map defined on bounded continuous functions on the spectrum*

$$\begin{array}{rcl} C_{\mathrm{bd}}(\mathrm{Spec}(A)) & \to & A \\ f & \mapsto & f(A) \end{array}$$

which is an injection of the commutative C^**-algebra* $C_{\mathrm{bd}}(\mathrm{Spec}(A))$ *into* $B(\mathcal{H})$*.*

In a sense, an unbounded self-adjoint operator A on a Hilbert space encodes information on such commutative C^*-subalgebras.

Let now A be an abstract C^*-algebra. Using a faithful representation of A, one deduces an analogues theorem for a self-adjoint element a of an abstract C^*-algebra A. Again, there is a projector-valued measure $P_a(\lambda)$ on $\mathrm{Spec}(a)$ such that for any continuous function f on $\mathrm{Spec}(a)$, one has

$$f(a) = \int_{\mathrm{Spec}(a)} f(\lambda) \mathrm{d}P_a(\lambda) \ .$$

In particular,

$$a = \int_{\mathrm{Spec}(a)} \lambda \mathrm{d}P_a(\lambda) \ .$$

For any state ω, one obtains

$$\omega(a) = \int_{\mathrm{Spec}(a)} \lambda \omega(\mathrm{d}P_a(\lambda))$$

and thus a probability measure $\omega(\mathrm{d}P_a(\lambda))$ on $\mathrm{Spec}(a)$.

In this case, the essence of the so-called continuous functional calculus can be summarized by

Proposition 5.4.8.
Let A *be a* C^**-algebra with unit. Let* a *be normal. Then there is a unique* $*$*-morphism*

$$\begin{array}{rcl} C(\mathrm{Spec}_A(a)) & \to & A \\ a & \mapsto & f(a) \end{array}$$

which has the standard meaning if f *is a polynomial function and for which the following statements hold:*

1. $\|f(a)\| = \|f\|_{C(\mathrm{Spec}_A(a))}$ *for all* $f \in C(\mathrm{Spec}_A(a))$*. On the right hand side, we have the supremum norm.*
2. *If* B *is another* C^**-algebra with unit and* $\pi : A \to B$ *a unit preserving* $*$*-morphism, then* $\pi(f(a)) = f(\pi(a))$ *for all* $f \in C(\mathrm{Spec}_A(a))$*.*
3. $\mathrm{Spec}_A(f(a)) = f(\mathrm{Spec}_A(a))$ *for all* $f \in C(\mathrm{Spec}_A(a))$*.*

Proof:
For any polynomial function $P \in \mathbb{C}[X]$, the operator $P(a)$ is normal and hence

$$\begin{array}{rcl} \|P(a)\| = \rho_A(P(a)) & = & \sup\{|\mu| \mid \mu \in \mathrm{Spec}_A(P(a))\} = \sup\{\, |P(\lambda)| \mid \lambda \in \mathrm{Spec}_A(a)\} \\ & = & \|P\|_{C_{\mathrm{Spec}_A(a)}} \end{array}$$

Thus the map $P \mapsto P(a)$ extends uniquely to a linear map from the closure of the algebra of polynomials on $\mathrm{Spec}_A(a)$ with the supremum norm to A. Since the polynomials form an algebra with a unit which contains complex conjugates and is separating norms, the Stone-Weierstraß theorem asserts that the closure of the algebra of all polynomials is the algebra $C(\mathrm{Spec}_A(a))$ of all continuous functions. By continuity, this is a $*$-morphism and preserves the norm.

The second assertion holds for polynomials and follows for continuous f since π is continuous.

For the third assertion, let $\lambda \in \mathrm{Spec}_A(a)$. Choose polynomials $P_n \to f$ in $C(\mathrm{Spec}_A(a))$. Then $P_n(\lambda) \in \mathrm{Spec}_A(P_n(a))$, i.e. $P_n(a) - P_n(\lambda) \cdot 1 \notin A^\times$. Since the complement of $A^\times$ is closed, we have in the limit $f(a) - f(\lambda) \cdot 1 \notin A^\times$. This shows $f(\mathrm{Spec}_A(a)) \subset \mathrm{Spec}_A(f(a))$.

Conversely, let $\mu \notin f(\mathrm{Spec}_A(a))$. Then we have for the function $g := (f - \mu)^{-1} \in C(\mathrm{Spec}_A(a))$. Then

$$g(a)\,(f(a) - \mu \cdot 1_A) = (f(a) - \mu \cdot 1_A)g(a) \ = 1$$

implies that $f(a) - \mu \cdot 1_A \in A^\times$, thus $\mu \notin \mathrm{Spec}_A(f(a))$. □

We are now ready to formulate some axioms of a quantum mechanical system:

Definition 5.4.9

1. *The first datum of a quantum mechanical system is kinematical and consists of a C^*-algebra A. The self-adjoint elements of A are called observables. The possible outcomes of the measurements are given by the spectrum* $\mathrm{Spec}(a) \subset \mathbb{R}$. *Note that the set of possible outcomes of measurements are independent of the state.*

2. *The state of a quantum mechanical system can be described by a state of its C^*-algebra A. Given a quantum mechanical system A in a state ω, the result of a measurement of an observable $a \in A$ cannot be predicted. Rather, quantum mechanics is a probabilistic theory. On the possible outcomes of the measurements, the spectrum* $\mathrm{Spec}(a) \subset \mathbb{R}$, *quantum mechanics predicts the probability measure*

$$\mu_{a,\rho} = \omega(\mathrm{d}P_a(\lambda))$$

for the outcome. In particular, the expectation value for an observable a in a state ω equals

$$\langle a \rangle_\omega = \int_{\mathrm{Spec}(a)} \lambda \omega(\mathrm{d}P_a(\lambda)) = \omega(a)$$

and the variance equals

$$\Delta_\omega(a) = \langle (a - \langle a \rangle)^2 \rangle_\omega = \langle a^2 \rangle_\omega - (\langle a \rangle_\omega)^2 = \omega(a^2) - \omega(a)^2 \ .$$

3. *The measurement process influences the state: if the value $a_0 \in \mathrm{Spec}_A(a)$ for an observable a has been measured, the system is in a state where the observable a has variance zero. This is also called the collapse of the wave function.*

4. *One says that a finite set of observables $\vec{A} := \{A_1, A_2, \ldots, A_n\}$ can be measured simultaneously, if they commute pairwise. In this case, there is a projector valued measure P on $\mathbb{R}^n$ given by*

$$P_{\vec{A}}(E_1 \times E_2 \times \ldots \times E_n) = P_{A_1}(E_1) \cdot P_{A_2}(E_2) \cdot \ldots \cdot P_{A_n}(E_n)$$

Then there can exist vectors in a $*$-representation and thus states which are simultaneous eigenstates so that the predictions with variance zero are possible for all observables simultaneously.

5. *Let $I \subset \mathbb{R}$ be an interval. The dynamics of the quantum mechanical system with observable algebra A is given by a map*

$$\alpha : \ I \to A$$

that is strongly continuous and takes its values in the unitary elements of A. We also write $\alpha_t := \alpha(t)$.

6. *A quantum mechanical system is said to have time independent dynamics if the dynamics is defined on all of $\mathbb{R}$ and if the map α is a strongly continuous one-parameter group of unitary elements on A. (Here we restrict ourselves implicitly on a time-independent dynamics.)*

7. *We may take two different point of views:*

   - *In the Heisenberg picture, the observables evolve in time according to*

$$a(t) = \alpha_t^{-1} a \alpha_t \ .$$

     The idea is thus to implement time evolution by inner automorphisms of A. States, in contrast, do not depends on time.
     This induces a time-dependent family of probability measures for each pair (a, ρ) with a an observable and ρ a state on A. Using the GNS construction, we can find a Hilbert space $\mathcal{H}_\rho$ and a cyclic vector $\Omega \in \mathcal{H}_\Omega$ such that

$$\rho(a(t) = \langle a(t)\Omega, \Omega \rangle$$

     is the time-dependent expectation value.

   - *In the Schrödinger picture, observables are time independent, but the state evolves in time.*

8. *Consider a quantum mechanical system (A, α_t) with time-independent dynamics. If we realize the C^*-algebra A as a subalgebra of $B(\mathcal{H})$, there is an "infinitesimal description" of the unitary one-parameter subgroup*

$$\begin{array}{ccc} \mathbb{R} & \to & B(\mathcal{H}) \\ t & \mapsto & \alpha_t \end{array}$$

   in terms of an unbounded self-adjoint operator H on $\mathcal{H}$:

$$\alpha_t = \exp(\frac{\mathrm{i}}{\hbar} H t)$$

   The self-adjoint operator H is called the Hamiltonian or Hamilton-operator. In the Heisenberg picture, any observable $a \in A$ obeys the Heisenberg equation of motion

$$\frac{\mathrm{d}}{\mathrm{d}t} a(t) = \frac{\mathrm{i}}{\hbar} [H, a] \ .$$

This equation should be compared with the equation for the time evolution of functions f on phase space in Hamiltonian systems, which are the observables in classical mechanics:

$$\frac{\mathrm{d}}{\mathrm{d}t} f = \{h, f\} \ .$$

We can see heuristically the commutator as a quantization of the Poisson bracket and the Hamiltonian H as the quantum counter part of the Hamiltonian function in classical Hamiltonian dynamics.

9. *The relation between the Heisenberg picture and the Schrödinger picture is as follows: Consider the unitarily transformed operator*

$$\tilde{a}(t) := \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}Ht} \cdot a(t) \cdot \mathrm{e}^{\frac{\mathrm{i}}{\hbar}Ht} \ .$$

It does not depend on time:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\, \tilde{a}(t) = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}Ht} \left[-\frac{\mathrm{i}}{\hbar} H, a(t) \right] \mathrm{e}^{\frac{\mathrm{i}}{\hbar}Ht} \\ + \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}Ht} \left[\frac{\mathrm{i}}{\hbar} H, a(t) \right] \mathrm{e}^{\frac{\mathrm{i}}{\hbar}Ht} = 0 \end{aligned}$$

We consider the expectation value in the state ρ at the time t and find

$$\begin{aligned} \rho(a(t)) &= \langle a(t)\psi_0, \psi_0 \rangle = \langle U(t)a(t)U^{-1}(t)U(t)\psi_0, U(t)\psi_0 \rangle \\ &= \langle \tilde{a}\psi(t), \psi(t) \rangle \ . \end{aligned}$$

In the last step, we introduce the time-dependent vector $\psi(t) := U(t)\psi_0 \ \in \mathcal{H}$.

10. *This way, we obtain the Schrödinger picture which we summarize as follows:*

    - *Observables are time-independent.*
    - *(Pure) states are represented as vectors in a separable Hilbert space $\mathcal{H}$. They are time-dependent. They obey the (time-dependent) Schrödinger equation:*

$$\mathrm{i}\hbar \frac{\mathrm{d}}{\mathrm{d}t}\, \psi = H\, \psi.$$

The Schrödinger equation is taken in many books on quantum mechanics as a starting point. One first argues, following de Broglie, that quantum-mechanically, particles have also a wave-like nature. This can be based on experiments like the double-slit experiment which show interferometric patterns. The subtle point is that the wave function does not describe probabilities, but rather its absolute square does. As remarked first by Debye, any decent wave needs a wave equation; this equation was found – actually as a classical limit of a relativistic equation, the Klein-Gordon equation – by Schrödinger.

11. *The eigenvector equation for the Hamiltonian H*

$$H\, \psi = \mathrm{E}\, \psi$$

is also called time-independent Schrödinger equation. Indeed, every eigenvector ψ_E of H

$$H\,\psi_E = E\psi_E$$

provides a solution of the Schrödinger equation:

$$\psi(t) = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}Et}\cdot\psi_E\ ,$$

since

$$\mathrm{i}\hbar\,\dot{\psi}(t) = \mathrm{i}\hbar\,\frac{-\mathrm{i}}{\hbar}\,\mathrm{E}\,\psi_E = H\,\psi_E\ .$$

The (generalized) eigenvectors of a self-adjoint operator being complete, one can describe a general state as a superposition of eigenstates and discuss time evolution componentwise.

5.5 Quantization

Next, we would like to obtain interesting examples of quantum mechanical systems (A, H). One path leading to them is to deduce them from a classical mechanical Hamiltonian system (M, ω, h). A scheme affording this is called "quantization". We formulate a naive version of quantization which one might call "Dirac's wish list". To this end, we are realizing the C^*-algebra explicitly.

Definition 5.5.1

1. *Consider a unital C^*-algebra $A \subset B(\mathcal{H})$. We endow A with a 1-parameter family of Lie brackets*
$$\frac{1}{\mathrm{i}\hbar}\,[a,b] =: \{a,b\}_\hbar$$
parametrized by $\hbar \in \mathbb{R}^\times$. These Lie brackets preserve the real subspace $\{x \,|\, x = x^*\} \subset A$ of self-adjoint elements and endow it with the structure of a real Lie algebra.

2. *Let $(A_{\text{class}}, \cdot, \{\cdot,\cdot\})$ be a commutative unital real Poisson algebra. A strong quantization of A is a C^*-algebra $A \subset B(\mathcal{H})$, together with a homomorphism of real Lie algebras*
$$Q:\ (A_{\text{class}}, \{\cdot,\cdot\}) \to (A, \{\cdot,\cdot\}_\hbar)$$
for some real *value of $\hbar$ such that*
   (a) *Q preserves the unit, $Q(1) = 1$.*
   (b) *The Lie-algebra representation of A_{class} on $\mathcal{H}$ induced by $Q(A) \subset B(\mathcal{H})$ is irreducible.*

Remarks 5.5.2.

1. *The advantage of having a quantization at hand is that our intuition coming from classical physics gives us a direct identification of the important observables in A.*

2. *We only require compatibility of Q with the Lie algebra structure on A, not with the commutative product.*

3. *It can be helpful to admit unbounded operators with dense domain on the Hilbert space $\mathcal{H}$ in the image of Q. In this case, the image of A is required to be contained in the real vector subspace of symmetric operators on $\mathcal{H}$.*

   For example, consider a strong quantization of the Poisson algebra of smooth functions on a symplectic manifold M. Let $f \in C^\infty(M)$ be a smooth function whose symplectic gradient X_f is not a completely integrable vector field on M. Then the full quantization of f can be shown to take as its value an unbounded operator.

Even when one admits *unbounded* operators, strong quantizations need not exist. It is known that the cotangent bundle T^*S^2 of the 2-sphere does not admit a strong quantization while the cotangent T^*T^2 of the 2-torus admits a strong quantization.

Our goal is now to show that a strong quantization does not exist for the simplest case, the Poisson algebra of smooth functions on the two-dimensional symplectic vector space $T^*\mathbb{R}$.

Observation 5.5.3.

1. *Consider a symplectic vector space (V, ω). Fix a Darboux basis of V; then the coordinates (p_i, q^i) of elements $v \in V$ define linear functions which are in particular elements in $A_{\text{class}} = C^\infty(V)$. A strong quantization*

$$Q : \ (A_{\text{class}}, \{\cdot,\cdot\}) \to (A, \{\cdot,\cdot\}_\hbar)$$

   therefore provides (essentially) self-adjoint operators on $\mathcal{H}$

$$Q^i := Q(q^i), \quad P_i := Q(p_i) \ .$$

   The axioms of strong quantization require them to obey Heisenberg's commutation relations or canonical commutation relations :

$$\left[Q^i, Q^j\right] = 0, \quad [P_i, P_j] = 0 \quad \left[P_j, Q^k\right] = \mathrm{i}\hbar\, \delta^k_j \ .$$

2. *It is a helpful strategy to first investigate representations of the subalgebra of A spanned by Q^i and P_j; more precisely, of the subalgebra of* bounded operators *generated by them. When $\mathcal{H}$ is finite-dimensional, we can take the trace of the last relation and use the fact that the trace is cyclic to find*

$$0 = \mathrm{tr}\left[P_j, Q^k\right] = \mathrm{i}\hbar\, \delta^k_j \mathrm{trid}_{\mathcal{H}} = \mathrm{i}\hbar\, \delta^k_j \dim_{\mathbb{C}} \mathcal{H}$$

   This implies $\dim_{\mathbb{C}} \mathcal{H} = 0$; thus there are no strong quantizations on a finite-dimensional *Hilbert space $\mathcal{H}$.*

3. *To avoid the use of unbounded operators, we consider the n-parameter groups generated by Q^i and P_j respectively:*

$$U(\alpha) = \mathrm{e}^{\mathrm{i}\alpha^j P_j} \qquad \text{and} \qquad V(\beta) = \mathrm{e}^{\mathrm{i}\beta_j Q^j}$$

   with $\alpha, \beta \in \mathbb{R}^n$. The Weyl algebra is the closed subalgebra of $B(\mathcal{H})$ generated by

$$W(\alpha, \beta) := U(\alpha)V(\beta) \ .$$

   For such operators Weyl's rules read

$$W(\alpha, \beta)W(\alpha', \beta') = \mathrm{e}^{-\mathrm{i}\omega((\alpha,\beta),(\alpha',\beta')))/2} W(\alpha + \alpha', \beta + \beta')$$

   with ω the symplectic form on V. They lead to exponentiated form of the canonical commutation relations.

We now note that if (V,ω) is a symplectic vector space, its dual (V^*,ω) is a Lie-subalgebra of the classical observable algebra $A_{\text{class}} = C^\infty(V)$. As a first step to investigate quantizations, we study maps associating to elements $v \in V^*$ elements in a C^*-algebra A that obey the Weyl algebra. Since V^* is a symplectic vector space itself, we simplify notation by denoting in the next definition the symplectic vector space in question by V. The results are to be applied, however, to linear coordinate functions.

Definition 5.5.4
Let (V,ω) be a symplectic vector space. A Weyl system of V consists of a unital C^*-algebra A and a map $W: V \to A$ such that for all $v, w \in V$ we have

1. $W(0) = 1_A$
2. $W(-v) = W(v)^*$
3. $W(v)W(w) = \mathrm{e}^{-\mathrm{i}\omega(v,w)/2}W(v+w)$.

We do not endow the symplectic vector space with a topology and thus do not require the map W to be continuous.

Observation 5.5.5.
There is a direct construction of Weyl systems: let $\mathcal{H} = L^2(V,\mathbb{C})$ be the Hilbert space of square-integrable functions on V with respect to the counting measure with Hermitian product

$$\langle F, G\rangle := \sum_{v\in V} F(v)\overline{G(v)}\ .$$

We define

$$\left(W(v)F\right)(w) := \mathrm{e}^{\mathrm{i}\omega(v,w)/2}F(v+w)\quad .$$

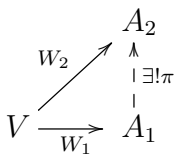
Denote by $\mathrm{CCR}(V,\omega)$ the C^*-algebra of $B(\mathcal{H})$ generated by $W(v)$ for $v \in V$. Then the pair $(\mathrm{CCR}(V,\omega), W)$ forms a Weyl system for the symplectic vector space (V,ω).

We need to impose a minimality requirement on the C^*-algebra A.

Definition 5.5.6
A Weyl system (A,W) of a symplectic vector space (V,ω) is called a CCR representation of (V,ω), if A is generated as a C^*-algebra by the elements $W(v)$ with $v \in V$.

Here CCR stands for canonical commutation relations. The proof of the following theorem is not too hard:

Theorem 5.5.7.
Let (V,ω) be a symplectic vector space and let (A_1,W_1) and (A_2,W_2) be two CCR representations of (V,ω). Then there exists a unique $*$-isomorphism $\pi: A_1 \to A_2$ such that the diagram



commutes.

Corollary 5.5.8.
There is a unique functor

$$\mathrm{CCR}: \ \mathit{SympVect} \to C^*\text{-}\mathit{Alg}$$

from the category of symplectic vector spaces to the category of C^*-algebras with injective unit-preserving $*$-morphisms. (Observe that symplectic maps are automatically injective.)

In the case of a symplectic endomorphism $S : V \to V$, the induced $*$-morphism $\mathrm{CCR}(S)$ is also called a Bogoliubov transformation.

Proof:
Suppose, we have a symplectic map

$$S: \ (V_1, \omega) \to (V_2, \omega_2)$$

and a CCR representation $\mathrm{CCR}(V_2, \omega_2), W_2$ of the second symplectic vector space. One immediately sees that $(\mathrm{CCR}(V_2, \omega_2), W_2 \circ S)$ is a Weyl system of (V_1, ω_1). The theorem thus yields a morphism $\mathrm{CCR}(S)$ such that the diagram

$$\begin{array}{ccc} V_1 & \xrightarrow{\quad S \quad} & V_2 \\ {\scriptstyle W_1}\downarrow & & \downarrow{\scriptstyle W_2} \\ \mathrm{CCR}(V_1, \omega_1) & \xrightarrow[\mathrm{CCR}(S)]{} & \mathrm{CCR}(V_2, \omega_2) \end{array}$$

commutes. By uniqueness, we conclude $\mathrm{CCR}(\mathrm{id}_V) = 1$ and $\mathrm{CCR}(S_2 \circ S_1) = \mathrm{CCR}(S_2) \circ \mathrm{CCR}(S_1)$. □

There is an important result about the representations of Weyl systems.

Observation 5.5.9.

1. *Let (V, ω) be a finite-dimensional symplectic vector space of dimension $2n$ and $U \subset V$ a Lagrangian subspace of V, i.e. a linear subspace of dimension n such that $\omega\big|_U = 0$. We can find a complementary Lagrangian subspace $U' \subset V$ such that $U \oplus U' = V$ and such that*

   $$\omega\big|_{U \times U'} : U \times U' \to \mathbb{R}$$

   is a non-degenerate bilinear pairing. Denote by $P: \ X \to U$ the projection to U.

2. *Endow the finite-dimensional real vector space U with the Lebesgue measure and consider the separable Hilbert space $\mathcal{H} = L^2(U, \mathbb{C})$. For $v \in V$ and $f \in \mathcal{H}$ define an element $v.f \in \mathcal{H}$ by*

   $$(v.f)(x) \ := \ \mathrm{e}^{\mathrm{i}\omega(v,x)} f(x - Pv) \ .$$

   Explicitly, elements of U act as translations and, infinitesimally, as derivative operators. Elements of U' act by multiplication with a phase factor:

   $$\begin{aligned} [U(\alpha)f](x) &:= f(x - \alpha) \quad \text{for } \alpha \in U \\ [V(\beta)f](x) &:= e^{i\omega(\beta,x)} f(x) \quad \text{for } \beta \in U' \ . \end{aligned}$$

3. *This defines a representation of the Weyl algebra* $\mathrm{CCR}(V,\omega)$ *on* $\mathcal{H}$*:*

$$\begin{aligned}\big(v_1.(v_2.f)\big)(x) &= f(x-Pv_1-Pv_2)\mathrm{e}^{\mathrm{i}\omega(v_1,x-v_2)}\mathrm{e}^{\mathrm{i}\omega(v_2,x)}\\ &= \mathrm{e}^{\mathrm{i}\omega(v_1+v_2,x)}f(x-Pv_1-Pv_2)\mathrm{e}^{-\mathrm{i}\omega(v_1,v_2)}\\ &= \mathrm{e}^{-\mathrm{i}\omega(v_1,v_2)}\left((v_1+v_2).f\right)(x)\end{aligned}$$

Definition 5.5.10

1. *The representation of* $\mathrm{CCR}(V,\omega)$ *we have just defined on* $\mathcal{H}=L^2(U)$ *is called the* Schrödinger representation*.*
2. *More generally, given any finite or infinite-dimensional Hilbert space* X*, we obtain an analogous representation of the Weyl algebra on the infinite-dimensional Hilbert space* $\mathcal{H}_X=\mathrm{L}^2(\mathbb{R}^n,X)$*.*

These are indeed all representations of the Weyl algebra $\mathrm{CCR}(V,\omega)$.

Theorem 5.5.11 (Stone - von Neumann)**.**
Let (V,ω) *be a finite-dimensional symplectic vector space.*

1. *For any* $*$*-representation of the Weyl algebra* $\mathrm{CCR}(V,\omega)$ *on a separable Hilbert space* $\mathcal{H}$*, there exists a separable Hilbert space* V *and a unitary operator*

$$U:\ \mathcal{H}\to\mathcal{H}_X$$

providing an isomorphism of $*$*-representations to the Schrödinger representation on* $\mathcal{H}_X$*. The Hilbert space* X *is also called the* internal Hilbert space*.*

2. *The representation of the Weyl algebra on* $\mathcal{H}_V$ *is irreducible, iff* $\dim_{\mathbb{C}}V=1$*.*

Observation 5.5.12.

1. *Usually, a slightly different formulation of the Schrödinger representation is used. Let* (V,ω) *be a finite-dimensional symplectic vector space,* $\dim_{\mathbb{R}}V=2n$*. Choose a Darboux basis* (b_i,b^i)*. Introduce the Lagrangian subspace* $U:=\mathrm{span}(b_i)_{i=1,\dots,n}$*. We think of this subspace as "positions" of the system and denote the coordinate functions by* $(q^i)_{i=1,\dots,n}$*. Thus* $q^i\in U^*$*.*

   We then define $U':=\mathrm{span}(b^i)_{i=1,\dots n}$ *and think of this subspace as momenta. The corresponding linear coordinate functions are denoted by* $(p_i)_{i=1,\dots,n}$*. Thus* $p_i\in(U')^*$*.*

   We then use the symplectic form $\omega:V\times V\to\mathbb{R}$ *to identify* $U'\cong U^*$ *and thereby* $(U')^\perp\cong U$*.*

2. *We now introduce the Hilbert space* $\mathcal{H}=L^2(U,\mathbb{C})$ *of square integrable functions of positions. The CCR representation we want to construct is for the Lie algebra* V^* *of linear functions in the coordinates.*

   We associate to the coordinate function $q^i\in U^*$ *the operator* Q^i *with*

$$(Q^i\psi)(q)=q^i\psi(q)$$

which for the Weyl algebra means for $\alpha\in U^*$

$$U(\alpha)\psi(q)=\mathrm{e}^{\mathrm{i}\langle\alpha,q\rangle}\psi(q)\ .$$

3. *A function p_i should be represented by a derivative operator P_i:*

$$(P_i\psi)(q) = \frac{\partial}{\mathrm{i}\partial q^i}\psi(q)$$

which for the Weyl algebra means for $\beta \in (U')^* \cong U$

$$V(\beta)\psi(q) = \psi(q+\beta) \ .$$

We call this representation the Schrödinger representation in position space. A generalization including an internal Hilbert space V is obvious.

4. *For quantizations of the classical phase space $T^*\mathbb{R}$, we can thus without loss of generality require that the strong quantization map is a map*

$$Q: \quad C^\infty(T^*\mathbb{R}) \to B(\mathcal{H}_V)$$

for some finite-dimensional Hilbert space V and restricts to the Schrödinger representation in position space for the Lie subalgebra of $C^\infty(T^*\mathbb{R})$ that is spanned by linear coordinate functions.

Lemma 5.5.13.
For any strong quantization $Q : C^\infty(T^*\mathbb{R}) \to B(\mathcal{H}_V)$ for with these properties, we have the quadratic identities

$$\begin{aligned} Q(q^2) &= (Q(q))^2 \\ Q(p^2) &= (Q(p))^2 \\ Q(qp) &= \frac{1}{2}\left(Q(q)Q(p) + Q(p)Q(q)\right). \end{aligned}$$

Proof:

- We introduce the shorthand $\hat{f} := Q(f)$ for any smooth function $f \in C^\infty(T^*\mathbb{R})$. We first note that for any smooth real function f with $\{f, q\} = 0$, $\hat{f}$ has to be of the form

$$\hat{f}\psi(q) = A(q)\psi(q) \ .$$

with $A(q)$ a hermitian operator on V. (To derive this in the simplest way, one should impose on the quantization the additional requirement that $\{Q(f), Q(g)\} = 0$ implies that the spectral projections of $\hat{f}$ and $\hat{q}$ commute.)

- In particular, $\{q^2, q\} = 0$ implies $\widehat{q^2} = A(q)$. From $\{p, q^2\} = -2q$, we deduce $\frac{1}{\mathrm{i}\hbar}\left[\hat{p}, \widehat{q^2}\right] = -2\hat{q}$ which amounts in the Schrödinger representation on $\mathcal{H}_V$ to

$$\frac{1}{\mathrm{i}\hbar}\left[\frac{\hbar}{\mathrm{i}}\partial_q, A(q)\right]\psi(q) = -A'(q)\psi(q) \stackrel{!}{=} -2q\psi(q)$$

which implies $A'(q) = 2q$ and thus

$$\widehat{q^2} = \hat{q}^2 - 2e_-$$

with e_- a hermitian endomorphism of V. Analogously, we obtain $\widehat{p^2} = \hat{p}^2 + 2e_+$ with a hermitian endomorphism e_+ of V.

- From the Poisson bracket $4pq = \{q^2, p^2\}$, we deduce

$$\widehat{qp} = \frac{1}{\mathrm{i}\hbar}\left[\widehat{q^2}, \widehat{p^2}\right] = \frac{1}{\mathrm{i}\hbar}\left[\hat{q}^2, \hat{p}^2\right] - \frac{1}{\mathrm{i}\hbar}\left[e_-, e_+\right] =$$
$$= \frac{1}{2}(\hat{q}\hat{p} + \hat{p}\hat{q}) + h \quad \text{with} \quad h = \frac{1}{\mathrm{i}\hbar}\left[e_+, e_-\right]$$

- Further calculations furnish the relations

$$\frac{1}{\mathrm{i}\hbar}\left[e_+, e_-\right] = h, \quad \frac{1}{\imath\hbar}\left[h, e_\pm\right] = \pm 2 e_\pm$$

  which are just the commutation relations for three generators of the non-compact real Lie algebra $\mathrm{sl}(2, \mathbb{R})$. This Lie algebra is known *not* to have finite-dimensional representations where e_+, e_- and h are hermitian.

□

Definition 5.5.14
Consider the symplectic vector space $V = T^*\mathbb{R}$. and the Poisson algebra $C^\infty(T^*\mathbb{R})$. Denote by p and q the linear coordinate functions given by the canonical Darboux basis. Obviously $p, q \in C^\infty(T^*\mathbb{R})$.

The Poisson algebra $C^\infty(T^*\mathbb{R})$ contains the following Lie subalgebras:

- *The Lie subalgebra*

$$\mathcal{F}_{(1)} := \mathrm{span}(1, p, q) \subset C^\infty(T^*\mathbb{R})$$

  of polynomials in p, q of order at most one. It has a basis $1, p, q$ and commutation relations

$$\{p, q\} = 1 \quad \{p, p\} = \{q, q\} = \{1, p\} = \{1, q\} = 0 \ .$$

  A three-dimensional real Lie-algebra isomorphic to $\mathcal{F}_{(1)}$ is called a Heisenberg Lie algebra.

- *More generally, we denote for $n \in \mathbb{N} \cup \{\infty\}$ by $\mathcal{F}_{(n)}$ the Lie subalgebra generated by polynomials in p and q of total degree at most n. Similarly, we denote for $m, n \in \mathbb{N} \cup \{\infty\}$ by $\mathcal{F}_{(m,n)}$ the Lie subalgebra generated by polynomials of degree at most m in p and degree at most n in q.*

Theorem 5.5.15 (Groenewold, van Hove)**.**
Consider the symplectic manifold $T^*\mathbb{R}$ with the Poisson algebra $C^\infty(T^*\mathbb{R})$ and its Lie subalgebra $\mathcal{F}_{(1)}$. Let X be a finite-dimensional Hilbert space. Then the Schrödinger representation in position space

$$Q : \mathcal{F}_{(1)} \to B(\mathcal{H}_X)$$

cannot be extended to any Lie subalgebra of $C^\infty(T^*\mathbb{R})$ containing polynomials in q and p which are in both variables of degree strictly bigger than two.

In particular, there is no strong quantization

$$Q : C^\infty(T^*\mathbb{R}) \to (B(\mathcal{H}), \{\cdot, \cdot\}_\hbar) \ .$$

Proof:

- We will show that in general, one has for real polynomials f, g in one variable the identities

$$\widehat{f(q)} = f(\hat{q}), \quad \widehat{f(p)} = f(\hat{p})$$

and

$$\widehat{g(q)p} = \frac{1}{2}\left(g(\hat{q})\hat{p} + \hat{p}g(\hat{q})\right) \qquad \text{and} \quad \widehat{g(g)q} = \frac{1}{2}\left(g(\hat{p})\hat{q} + \hat{q}g(\hat{p})\right) \ .$$

We only need these identities for $f(x) = x^3$ and $g(x) = x^3$ and restrict our proof to these cases.

- Applying this result to the identity

$$\frac{1}{9}\{q^3, p^3\} = \frac{1}{3}\{q^2p, p^2q\}$$

in the Poisson algebra $C^\infty(T^*\mathbb{R})$, we find that a strong quantization Q applied to the left hand side yields

$$\frac{1}{9\mathrm{i}\hbar}\left[\hat{q}^3, \hat{p}^3\right] = \hat{q}^2\hat{p}^2 - 2\mathrm{i}\hbar\hat{q}\hat{p} - \frac{2}{3}\hbar^2 1,$$

while for the right hand side we obtain

$$= \hat{q}^2\hat{p}^2 - 2\mathrm{i}\hbar\hat{q}\hat{p} - \frac{1}{3}\hbar^2 1 \ .$$

The two results differ, hence we have obtained a contradiction and no strong quantization exists.

- It remains to prove the assertion in the first item. Since the two operators $\hat{q}^3$ and $\widehat{q^3}$ commute with $\hat{q}$, we conclude that their difference equals a hermitian multiplication operator $A(q)$,

$$\widehat{q^3} - \hat{q}^3 = A(q) \ .$$

The identity

$$\left[\widehat{q^3}, \hat{p}\right] = \mathrm{i}\hbar\widehat{\{q^3, p\}} = 3\mathrm{i}\hbar\widehat{q^2} = 3\mathrm{i}\hbar\hat{q}^2 = \left[\hat{q}^3, \hat{p}\right]$$

implies that $A(q) = A$ is constant,

$$\widehat{q^3} = \hat{q}^3 + A \ .$$

To determine the constant, we compute

$$\widehat{q^3} = \frac{1}{3}\widehat{\{q^3, qp\}} = \frac{1}{3\mathrm{i}\hbar}\left[\widehat{q^3}, \widehat{qp}\right] = \frac{1}{3\mathrm{i}\hbar}\left[\hat{q}^3 + A, \frac{1}{2}(\hat{q}\hat{p} + \hat{p}\hat{q})\right] = \hat{q}^3 \ ,$$

hence $A = 0$. The identity $\hat{p}^3 = \widehat{p^3}$ is derived analogously.

- A straightforward calculation shows

$$\widehat{q^2p} = \frac{1}{6}\{\widehat{q^3, p^2}\} = \frac{1}{6\mathrm{i}\hbar}\left[\widehat{q^3}, \widehat{p^2}\right] == \frac{1}{6\mathrm{i}\hbar}\left[\hat{q}^3, \hat{p}^2\right] = \frac{1}{2}\left(\hat{q}^2\hat{p} + \hat{p}\hat{q}^2\right) \ .$$

  The expression for $\widehat{p^2q}$ is derived analogously.

□

Remarks 5.5.16.

- *The Schrödinger representation can be extended to the to the Lie subalgebra* $\mathcal{F}_{(2)} = \mathrm{span}(1, p, q, p^2, q^2, pq)$ *of* $C^\infty(T^*\mathbb{R})$ *and to the Lie subalgebra* $\mathcal{F}_{(\infty,1)} = \mathrm{span}(q^i, pq^i)_{i=0,1,\ldots}$. *These two Lie subalgebras are the only maximal Lie subalgebras of* $\mathcal{F}_\infty$ *containing the Heisenberg Lie algebra* $\mathcal{F}_{(1)}$.

- *Already the Lie subalgebra*

$$\mathcal{F}_{(1)} = \mathrm{span}(1, p, q) \subset C^\infty(T^*\mathbb{R})$$

  contains sufficiently many variables to separate all points of the phase space $T^*\mathbb{R}$*: given two configurations in phase space, i.e. two points* $x_1, x_2 \in T^*\mathbb{R}$*, we can find an element* $f \in \mathcal{F}_{(1)}$ *such that* $f(x_1) \neq f(x_2)$*. In this sense, we still have enough classical observables in* $\mathcal{F}_{(1)}$ *to distinguish all distinct configurations. In a sense, asking for all observables to be quantized, is asking for too much.*

We have thus to weaken our notion of quantization. We still want to keep enough information to be able to separate points in the classical phase space: otherwise, we would really quantize a *different* classical system. This leads to the following definition:

Definition 5.5.17
An imperfect quantization of a Poisson algebra $(A_{\mathrm{class}}, \cdot, \{\cdot,\cdot\})$ *is an injective linear map*

$$Q : A_{\mathrm{class}} \to A$$

into a unital C^**-algebra* $A \subset B(\mathcal{H})$ *of bounded operators on a separable Hilbert space* $\mathcal{H}$ *such that*

1. *The unit of the Poisson algebra is mapped to the unit of the* C^**-algebra.*
2. *The representation of the subalgebra generated by* $Q(A_{\mathrm{class}})$ *on the Hilbert space* $\mathcal{H}$ *is irreducible.*
3. *There exists a Lie subalgebra* A'_{class} *of* A_{class} *such that the polynomials in* A'_{class} *are dense in* A_{class} *and that the restriction* $Q|_{A'_{\mathrm{class}}}$ *is an injective morphism of Lie algebras.*

Remarks 5.5.18.

1. *Imperfect quantizations exist for most physical systems.*
2. *For polynomials in* A'_{class} *of higher order, the quantization condition only holds in the weakened form*

$$Q(\{f, g\}) = \frac{1}{\mathrm{i}\hbar}[Q(f), Q(g)] + O(\hbar^2) \ .$$

5.6 Symmetries in quantum mechanics

Observation 5.6.1.

- *Consider a quantum mechanical system described by a C^*-algebra A. We have seen that for any state $\omega \in S(A)$, we get a probability measure $\mu_{a,\omega}$ for any self-adjoint element $a \in A$ which describes the probability distribution of possible outcomes of experiments. It is thus natural to define a symmetry of a quantum mechanical system as pair of mappings: $\alpha : A \to A$ mapping observables to observables and $\overline{\alpha} : S(A) \to S(A)$ of states such that*

$$\mu_{S(a),\overline{S}(\omega)} = \mu_{a,\omega} \ .$$

  One can show that this implies that $\overline{\alpha}$ is an affine mapping.

- *Let us assume that $\overline{\alpha}$ is also invertible. Then it maps pure states on pure states, since the set of states is the convex hull of the subset of pure states.*

- *Now assume that we have realized the algebra A as a subalgebra of bounded operators on a Hilbert space $\mathcal{H}$. Suppose that the state ω is represented by a density matrix P, i.e. that we have a state ω_P such that $\omega_P(a) = \mathrm{tr}_{\mathcal{H}} Pa$ for all $a \in A$. From*

$$\overline{\alpha}(\omega_P)(\alpha x) = \omega_P(x) = \mathrm{tr}_{\mathcal{H}} Px = \mathrm{tr}_{\mathcal{H}} \alpha(P)\alpha(x) \quad \text{for all } x \in A$$

  we conclude that the state $\overline{\alpha}$ has density matrix $\alpha(P)$,

$$\overline{\alpha}(\omega_P) = \omega_{\alpha(P)} \ .$$

  This implies that for all $P' \in A$, we have

$$\mathrm{tr}_{\mathcal{H}} \alpha(P)\alpha(P') = \overline{\alpha}(\omega_P)(\alpha(P')) = \omega_P(P') \ = \ \mathrm{tr}_{\mathcal{H}} PP' \ .$$

  Pure states correspond to projections of rank one. Let $P, P' \in A$ be such rank one projections to one-dimensional subspaces spanned by the unit vectors ψ and ψ', respectively. We then have from $P_{\psi'}(\psi) = \langle \psi', \psi \rangle \psi$

$$\mathrm{tr}_{\mathcal{H}} PP' = \langle \psi, P_{\psi'} \psi \rangle = \langle \psi, \psi' \rangle \cdot \langle \psi', \psi \rangle = |\langle \psi', \psi \rangle|^2 \ .$$

- *Denote by $\mathbb{P}\mathcal{H}$ the set of one-dimensional subspaces (which are also called rays) of the separable Hilbert space $\mathcal{H}$. This is actually an infinite-dimensional manifold modelled over an separable Hilbert space. Let L_1, L_2 be points in $\mathbb{P}\mathcal{H}$ which are represented by unit vectors $\psi_1, \psi_2 \in \mathcal{H}$; consider the function $p : \mathbb{P}\mathcal{H} \times \mathbb{P}\mathcal{H} \to \mathbb{R}$*

$$p(L_1, L_2) := |\langle \psi', \psi \rangle|^2 \ .$$

We thus define

Definition 5.6.2

1. *We denote by $\mathrm{Aut}_{qm}(\mathbb{P}\mathcal{H})$ the group of smooth maps $\mathbb{P}g : \mathbb{P}\mathcal{H} \to \mathbb{P}\mathcal{H}$ that preserve the function p, i.e.*

$$p(\mathbb{P}gL_1, \mathbb{P}g\ L_2) = p(L_1, L_2) \quad \text{for all lines} \quad L_1, L_2 \text{ in } \mathcal{H} \ .$$

2. *Let (A, H) be a quantum mechanical system that is represented on a Hilbert space $\mathcal{H}$. A symmetry of $(A, H, \mathcal{H})$ is a map $g : \mathcal{H} \to \mathcal{H}$ with the following properties:*

   (a) *The map g induces a map $\mathbb{P}g \in \mathrm{Aut}_{qm}(\mathbb{P}\mathcal{H})$.*

   (b) *The map g respects A: $ga = ag$ for all $a \in A$.*

   (c) *The map g commutes with the one-parameter group generated by the Hamiltonian H, i.e. $[U_H(t), g] = 0$ for all $t \in \mathbb{R}$.*

Observation 5.6.3.

1. *A real linear map $S : \mathcal{H} \to \mathcal{H}$ is called antiunitary, if it is $\mathbb{C}$-antilinear and*

$$\langle Sv_1, Sv_2\rangle = \overline{\langle v_1, v_2\rangle} \quad \text{for all} \quad v_1, v_2 \in \mathcal{H} \ .$$

2. *Let $G(\mathcal{H})$ be the group of all unitary and antiunitary operators on $\mathcal{H}$. This is a Banach Lie group with two connected components, consisting of unitary and antiunitary operators respectively. Obviously, both unitary and antiunitary operators induce maps in $\mathrm{Aut}_{qm}(\mathbb{P}\mathcal{H})$. We have thus a group homomorphism*

$$G(\mathcal{H}) \ \to \mathrm{Aut}_{qm}(\mathbb{P}\mathcal{H})$$

   whose kernel consists of the subgroup $M_{\mathbb{T}} \subset G(\mathcal{H})$ of multiplication operators by unit norm scalars $\alpha \in \mathbb{C}$:

$$\begin{array}{rccc} m_\alpha : \mathcal{H} & \to & \mathcal{H} \\ v & \mapsto & \alpha v \ . \end{array}$$

   We get the following short exact sequence of groups:

$$1 \to M_{\mathbb{T}} \to G(\mathcal{H}) \to \mathrm{Aut}_{qm}(\mathbb{P}\mathcal{H}) \ .$$

   Note that the subgroup $M_{\mathbb{T}}$ is not central, since conjugation by antiunitary operators acts on $M_{\mathbb{T}}$ by complex conjugation: if g is antiunitary, we get from the antilinearity of g that $g \circ m_\alpha(v) = g(\alpha v) = \overline{\alpha} g(v) = m_{\overline{\alpha}} \circ g(v)$ and thus $g \circ m_\alpha = m_{\overline{\alpha}} \circ g$.

Proposition 5.6.4 (Wigner)**.**
The homomorphism $G(\mathcal{H}) \ \to \mathrm{Aut}_{qm}(\mathbb{P}\mathcal{H})$ is surjective.

Remarks 5.6.5.

1. *Put differently, any symmetry of a quantum mechanical system (A, H) represented on a Hilbert space $\mathcal{H}$ can be described in terms of either a unitary or an antiunitary operator on $\mathcal{H}$.*

2. *A conceptual proof of this proposition can be based on the fact that elements of $\mathrm{Aut}_{qm}(\mathbb{P}\mathcal{H})$ are just the isometries of the projective space $\mathbb{P}\mathcal{H}$ with the Fubini-Study metric. For details, we refer to Dan Freed: On Wigner's theorem,* `arXiv:1112.2133v1`.

3. *An important symmetry described by an anti-unitary operator is time reversal. For details, in particular including systems with spin and Kramers degeneracy as a consequence, we refer to section 4.4 the book by Sakurai.*

Remarks 5.6.6.

1. *For a general C^*-algebra A, a symmetry is described by a Jordan endomorphism which subsumes the notion of a morphism and an anti-morphism. A Jordan morphism $\varphi : A \to B$ is a $\mathbb{C}$-linear map which preserves the star and the anticommutator,*

$$\varphi(a^*) = \varphi(a)^* \quad \text{and} \quad \varphi(\{a, b\}) = \{\varphi(a), \varphi(b)\} \quad \text{for all } a, b \in A \ .$$

2. *While it is true that every Jordan automorphism of the C^*-algebra $B(\mathcal{H})$ is either an automorphism or an antiautomorphism, this does not hold true any longer for an arbitrary C^*-algebra. Rather the following proposition holds: let A be a C^*-algebra, $\mathcal{H}$ a separable Hilbert space and $\varphi : A \to B(\mathcal{H})$ a Jordan morphism. Let B be the C^*-subalgebra of $B(\mathcal{H})$ generated by the image $\varphi(A)$. Then there exists a projector E on $\mathcal{H}$ commuting with all elements of B such that*

$$a \mapsto \varphi(a)E$$

is a morphism of C^*-algebras and

$$a \mapsto \varphi(a)(1 - E)$$

is an antimorphism.

In particular, if $\mathcal{H}$ is an irreducible representation of B, then φ is either a morphism or an antimorphism.

Observation 5.6.7.

1. *So far, we have been considering a single symmetry operation. Suppose we have a whole set G of them, i.e. we have a subset*

$$G \hookrightarrow \mathrm{Aut}_{qm}(\mathbb{P}\mathcal{H}) \ .$$

Without loss of generality, we can replace the set G by the subgroup of $\mathrm{Aut}_{qm}(\mathbb{P}\mathcal{H})$ it generates.

For simplicity, we now assume that all symmetries are unitary. We might prefer to work directly with a unitary representation of G on $\mathcal{H}$ and thus look for a lift

$$\begin{array}{ccc} & & \mathrm{U}(\mathcal{H}) \\ & \overset{\hat{\rho}}{\nearrow} & \big\downarrow \\ G & \xrightarrow[\rho]{} & \mathrm{Aut}_{qm}(\mathbb{P}\mathcal{H}) \end{array}$$

In general, such a lift does not exist. Rather, if we just choose a representative $\hat{\rho}(g)$ in $\mathrm{U}(\mathcal{H})$ of each $\rho(g)$, we can only conclude that for each pair of elements $g_1, g_2 \in G$, there is an element $\xi(g_1, g_2) \in \mathbb{T}$ such that

$$\hat{\rho}(g_1)\hat{\rho}(g_2) = \xi(g_1, g_2)\hat{\rho}(g_1 \cdot g_2) \ .$$

One says that $\hat{\rho}$ only provides a ray representation, a "representation up to a phase" or a projective representation of the group G on $\mathcal{H}$.

2. *We derive an important constraint on the phases: from the associativity of composition of endomorphisms, we have for any triple* $g_1, g_2, g_3 \in G$

$$\hat{\rho}(g_1)[\hat{\rho}(g_2)\hat{\rho}(g_3)] = [\hat{\rho}(g_1)[\hat{\rho}(g_2)]\hat{\rho}(g_3)$$

We compute the left hand side:

$$\hat{\rho}(g_1)[\hat{\rho}(g_2)\hat{\rho}(g_3)] = \xi(g_2, g_3)\hat{\rho}(g_1)\hat{\rho}(g_2g_3) = \xi(g_2, g_3)\xi(g_1, g_2 \cdot g_3)\hat{\rho}(g_1g_2g_3)$$

and the right hand side

$$[\hat{\rho}(g_1)[\hat{\rho}(g_2)]\hat{\rho}(g_3) = \xi(g_1, g_2)\hat{\rho}(g_1g_2)\hat{\rho}(g_3) = \xi(g_1, g_2)\xi(g_1 \cdot g_2, g_3)\hat{\rho}(g_1g_2g_3) \ .$$

By comparison, we find that ξ *has to obey the so-called cocycle relation*

$$\xi(g_2, g_3)\xi(g_1, g_2 \cdot g_3) = \xi(g_1, g_2)\xi(g_1 \cdot g_2, g_3) \quad \text{for all } g_1, g_2, g_3 \in G \ .$$

3. *As an example, consider the following action of the group* $G = \mathbb{Z}_2 \times \mathbb{Z}_2$ *with generators* g_1, g_2 *and relations* $(g_1)^2 = (g_2)^2 = e$ *and* $g_1g_2 = g_2g_1$. *Take* $\mathcal{H} = \mathbb{C}^2$ *the two-dimensional Hilbert space and* $\mathbb{P}\mathcal{H}$ *its (complex) projective space. An action of* G *on* $\mathbb{P}\mathcal{H}$ *is defined by the automorphisms*

$$\rho(g_1) = \mathbb{P}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad \rho(g_2) = \mathbb{P}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \ .$$

For the obvious lift to $\mathcal{H}$,

$$\hat{\rho}(g_1) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad \hat{\rho}(g_2) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \ ,$$

we find

$$\hat{\rho}(g_1) \cdot \hat{\rho}(g_2) \ = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{and} \quad \hat{\rho}(g_2) \cdot \hat{\rho}(g_1) \ = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \ ,$$

so that we do not reproduce the relation $g_1g_2 = g_2g_1$.

4. *The function* $\xi : G \times \ G \to \mathbb{T}$ *can be modified by using elements of* $\mathbb{T}$ *to redefine the lift*

$$\tilde{\rho}(g) := \eta(g)\hat{\rho}(g)$$

where $\eta : G \to \mathbb{T}$ *is some function. Then the modified function reads:*

$$\tilde{\xi}(g_1, g_2) = \eta(g_1)\eta(g_2)\eta(g_1g_2)^{-1}\xi(g_1, g_2) \ .$$

The projective representation of G *can be reduced to an ordinary representation of* G, *if for the corresponding function* $\xi : G \times G \to \mathbb{T}$ *there exists a function* $\eta : G \to \mathbb{T}$ *such that* $\tilde{\xi}(g_1, g_2) = 1$ *for all* $g_1, g_2 \in G$.

Such functions need not exist; they do not exist for all projective representations of the group $G = \mathbb{Z}_2 \times \mathbb{Z}_2$, *e.g. not for the projective representation presented above. The question on whether they exist or not is answered by the mathematical theory of group cohomology: they exist, if and only if the second group cohomology* $H^2_{group}(G, \mathbb{T})$ *of* G *with values in* $\mathbb{T}$ *vanishes.*

5. *An important group having finite-dimensional unitary* projective representations *is the rotation group* $\mathrm{SO}(3)$*. To discuss this in more detail, we note that there is a surjection of Lie groups*

$$1 \to \{\pm 1\} \to \mathrm{SU}(2) \stackrel{\pi}{\to} \mathrm{SO}(3) \to 1$$

with kernel a cyclic group of order two which turns out to be the center of the Lie group $\mathrm{SU}(2)$*.*

This group homomorphism can be constructed as follows: denote by H *the real linear subspace of hermitian traceless* 2×2*-matrices. This vector space is isomorphic to* $\mathbb{R}^3$ *by*

$$\begin{array}{rcl} \Phi : \mathbb{R}^3 & \to & H \\ (x_1, x_2, x_3) & \mapsto & \begin{pmatrix} x_3 & x_1 + \mathrm{i}x_2 \\ x_1 + \mathrm{i}x_2 & -x_3 \end{pmatrix} \end{array}$$

We have $\det \Phi(x) = -x_3^2 - x_1^2 - x_2^2 = -\|x\|^2$*. For any* $g \in \mathrm{SU}(2)$*, we find* $\mathrm{tr} g\Phi(x)g^{-1} = \mathrm{tr}\Phi(x) = 0$ *and* $\det g\Phi(x)g^{-1} = \det \Phi(x)$*. Thus there is a linear map* $\tilde{g} : \mathbb{R}^3 \to \mathbb{R}^3$ *such that* $g\Phi(x)g^{-1} = \Phi(\tilde{g}x)$*. Obviously, the linear map* $\tilde{g}$ *preserves the norm and thus the scalar product of* $\mathbb{R}^3$*. It can be shown to have determinant one and is thus an element in* $\mathrm{SO}(3)$*. Closer inspection shows that* $\pi : g \to \tilde{g}$ *is a surjective group homomorphism with kernel* $\pm 1 \in \mathrm{SU}(2)$*.*

6. *In general, let* $\pi : \ G \to H$ *be a group homomorphism,* $\mathcal{H}$ *be a finite-dimensional or infinite-dimensional Hilbert space and*

$$\rho : \ H \to \mathrm{U}(\mathcal{H})$$

a unitary representation of the group H *on* $\mathcal{H}$*. Then*

$$\pi^*\rho := \rho \circ \pi : \ G \stackrel{\pi}{\to} H \stackrel{\rho}{\to} \mathrm{U}(\mathcal{H})$$

is a unitary representation of G*, called the* pullback *of the representation* ρ *along the group homomorphism* π*. It is easy to check that projective representations can be pulled back to representations that are, in general, projective as well.*

7. *We now quote the following general statement: if* G *is a compact, connected and simply connected Lie group, then the group cohomology* $H^2_{group}(G, \mathbb{T})$ *vanishes, i.e. there are no projective representations. This applies in particular to the Lie group* $\mathrm{SU}(2)$*.*

We conclude that any projective representation of the rotation group $\mathrm{SO}(3)$ *can be pulled back to an* ordinary *representation of* $\mathrm{SU}(2)$*. For this reason, one usually prefers to work in quantum mechanics with the Lie group* $\mathrm{SU}(2)$*; it then suffices to work with ordinary representations of* $\mathrm{SU}(2)$ *only.*

Explicitly, the irreducible representations of $\mathrm{SU}(2)$ *with half-integer spin are only projective representations of* $\mathrm{SO}(3)$ *while the irreducible representations of* $\mathrm{SU}(2)$ *with integer spin are genuine representations of* $\mathrm{SO}(3)$*.*

There is a different way to cope with projective representations. For simplicity, we restrict ourselves for the moment to a finite group G. The idea is to take care of the phases $\xi(g_1, g_2) \in \mathbb{T}$ by incorporating them in an algebraic structure related to the group G.

Definition 5.6.8

1. *Let k be a field and G a group. The group algebra $k[G]$ is the associative k-algebra defined on the k-vector space freely generated by the set G with distinguished basis $(b_g)_{g\in G}$ and multiplication derived from the group law in G:*

$$b_{g_1} \cdot b_{g_2} := b_{g_1 g_2} \ .$$

It is an associative algebra with unit b_e, where e is the neutral element of G.

2. *Given a function*

$$\xi : \ G \times G \to \mathbb{T}$$

such that

$$\xi(g_1 g_2, g_3) \cdot \xi(g_1, g_2) = \xi(g_1, g_2 g_3) \cdot \xi(g_2, g_3) \quad \text{for all} \quad g_1, g_2, g_3 \in G$$

we can define on the vector space freely generated by G the multiplication

$$b_{g_1} \cdot b_{g_2} := \xi(g_1, g_2) b_{g_1 g_2} \ .$$

We obtain an associative algebra $k_\xi[G]$ with unit b_e. It is called the twisted group algebra, of the group G twisted by the cocycle ξ.

Remarks 5.6.9.

1. *Two twisted group algebras $k_{\xi_1}[G]$ and $k_{\xi_2}[G]$ are isomorphic as associative algebras, if and only if there is a function $\eta : \ G \to \mathbb{T}$ such that $\xi_2(g_1, g_2) \ = \eta(g_1)\eta(g_2)\eta(g_1 g_2)^{-1}\xi_1(g_1, g_2)$ for all $g_1, g_2 \in G$. The isomorphism is given by*

$$\begin{aligned} k_{\xi_1}[G] \ &\to \ k_{\xi_2}[G] \\ b_g \ &\mapsto \ \eta(g) b_g \ . \end{aligned}$$

Such an isomorphism exists, if and only if the group cohomology classes are equal, $[\xi_1] = [\xi_2] \in H^2_{group}(G, \mathbb{C}^\times)$.

2. *Thus projective k-linear representations of a group G with given cocycle ξ are just ordinary representations of the corresponding twisted group algebra $k_\xi[G]$.*

We are thus lead to consider *algebras* as symmetry structures.

Observation 5.6.10.

1. *Let A be a C^*-algebra which we think of as an symmetries. For example, A could be the complex group algebra $\mathbb{C}[G]$ of a finite group G with involution $(b_g)^* = b_{g^{-1}}$.*

   Recall that a representation of A on a separable Hilbert space $\mathcal{H}$ is a $*$-morphism $A \to B(\mathcal{H})$.

2. *The structure of an associative algebra does not suffice to discuss all aspects of symmetries one wants to have in quantum mechanics. For example, we would like to make the statement that a state, e.g. the vacuum Ω, is invariant under the action of the symmetry algebra. For the action of a group G, this implies*

$$(b_g).\Omega = g.\Omega = \Omega \quad \text{for all } g \in G \ .$$

We therefore consider the following linear functional on the group algebra $K[G]$:

$$\begin{array}{rccl} \epsilon: & \mathbb{C}[G] & \to & \mathbb{C} \\ & b_g & \mapsto & 1 \quad \textit{for all } g \in G \end{array}$$

This linear function is actually an algebra morphism.

We thus require our symmetry algebra to have a distinguished algebra morphism $\epsilon: A \to \mathbb{C}$ which defines a distinguished one-dimensional representation, the trivial representation. Vectors transforming in the trivial representation,

$$h.v = \epsilon(h)v \quad \textit{for all} \quad h \in H \ ,$$

are called invariant vectors.

3. *Suppose now we have two quantum mechanical systems carrying each a representation of the same symmetry algebra A. We thus have star-morphisms $\rho_1 : A \to B(\mathcal{H}_1)$ and $\rho_2 : A \to B(\mathcal{H}_2)$. We want to be able to couple this system to a composite system with Hilbert space $\mathcal{H}_1 \hat{\otimes} \mathcal{H}_2$ which should carry again a representation of A. By the universal property from remark 5.3.3, we get a morphism of $*$-algebras*

$$\rho_1 \otimes \rho_2 : A \otimes_\pi A \to \mathcal{B}(\mathcal{H}_1) \otimes_\pi \mathcal{B}(\mathcal{H}_2) \cong B(\mathcal{H}_1 \hat{\otimes} \mathcal{H}_2) \ .$$

We want, however, rather a representation of A on the Hilbert space of the composite system. This can be afforded by the additional datum of an $*$-morphism of algebras

$$\Delta: \ A \to A \otimes_\pi A \ .$$

On the composite system, we then have the representation

$$\rho := \Delta^*(\rho_1 \otimes \rho_2): \ A \to \mathcal{B}(\mathcal{H}_1) \otimes_\pi \mathcal{B}(\mathcal{H}_2) \ .$$

4. *In the case of a group algebra, we have the coproduct*

$$\Delta(b_g) = b_g \otimes b_g \ .$$

Suppose, we have eigenvectors $v_1 \in \mathcal{H}_1$ and $v_2 \in \mathcal{H}_2$,

$$\rho_1(g)v_1 = \lambda_1 v_1 \quad \textit{und} \quad \rho_2(g)v_2 = \lambda_2 v_2 \ .$$

We then find in the composite system

$$\rho(g)v_1 \otimes v_2 = g.v_1 \otimes g.v_2 = (\lambda_1 \cdot \lambda_2) \ v_1 \otimes v_2 \ .$$

One says that the symmetry $g \in G$ leads to a multiplicative quantum number. Examples of such symmetries in quantum mechanics include parity and charge conjugation.

We are thus lead to consider the following algebraic structure to describe symmetries of quantum mechanical systems:

Definition 5.6.11
Let k be a field.

1. *A k-vector space A, together with a k-linear map*

$$m: \ A \otimes_k A \to A$$

and a linear map $\eta: \ k \to A$ is called an (associative, unital) algebra, if the following conditions hold

$$m \circ (m \otimes_k \mathrm{id}_A) = m \circ (\mathrm{id}_A \otimes_k m) \quad \textit{[associativity]}$$
$$m \circ (\eta \otimes_k \mathrm{id}_A) = \mathrm{id}_A = m \circ (\mathrm{id}_A \otimes_k \eta) \quad \textit{[unitality]}$$

The map η is called a unit map and $\eta(1_k) =: 1_A \in A$ is the unit element of A. The map m is called the product.

2. *A k-vector space C, together with a k-linear map*

$$\Delta: \ C \to C \otimes_k C$$

and a linear form $\epsilon: \ C \to k$ is called a (coassociative, counital) coalgebra, if the following conditions hold

$$(\Delta \otimes_k \mathrm{id}_C) \circ \Delta = (\mathrm{id}_C \otimes_k \Delta) \circ \Delta \quad \textit{[coassociativity]}$$
$$(\epsilon \otimes_k \mathrm{id}_C) \circ \Delta = \mathrm{id}_C \ = (\mathrm{id}_C \otimes_k \epsilon) \circ \Delta \quad \textit{[counitality]}$$

The linear form ϵ is called the counit and Δ is called the coproduct.

3. *A k-vector space A that is endowed with the structure of both a coalgebra (A, Δ, ϵ) and of an algebra (A, m, η) such that Δ and ϵ are morphisms of algebras (or, equivalently such that m and η are morphisms of coalgebras) is called a bialgebra.*

We are thus lead to consider bialgebras as symmetry structures of quantum mechanical systems.

Observation 5.6.12.
Sometimes, one finds more structure. In the case of groups, we have the inverse which gives rise to a linear map

$$\begin{array}{rcl} S: k[G] & \to & k[G] \\ b_g & \mapsto & b_{g^{-1}} \end{array}$$

It is clear that S is a linear anti-algebramorphism, $S(a \cdot b) = S(b) \cdot S(a)$ of the bialgebra $k[G]$. It is also an anti-coalgebra morphism.

This leads to the following further definition:

Definition 5.6.13
Let $(A, m, \eta, \Delta, \epsilon)$ be a bialgebra. A map $S: A \to A$ is called an antipode, if

$$m \circ (S \otimes_k \mathrm{id}_A) \circ \Delta = 1_A \epsilon = m \circ (\mathrm{id}_A \otimes_k S) \circ \Delta$$

holds.

If an antipode exists, it is unique. A bialgebra that admits an antipode is called a Hopf algebra.

Remarks 5.6.14.

1. *We explain the importance of the antipode: if the k-vector space V carries a representation of a Hopf algebra A,*
$$\rho : \quad A \to \mathrm{End}(V) \ ,$$
then the dual of V carries a representation of A by
$$\rho^\vee : \ A \stackrel{S}{\to} A \stackrel{\rho^*}{\to} \mathrm{End}(V^*) \ .$$
Indeed, we have for all $a, b \in A$:
$$\rho^\vee(a \cdot b) = \rho^*(S(a \cdot b)) = \rho^*(S(b) \cdot S(a)) = \rho^*(S(a)) \circ \rho^*(S(b)) = \rho^\vee(a) \circ \rho^\vee(b)$$

2. *The other axioms of the antipode imply that the two natural maps, the evaluation*
$$\begin{array}{rcl} V^* \otimes_k V & \to & k \\ (\beta, v) & \mapsto & \beta(v) \end{array}$$
and, for $\dim_k V < \infty$, the coevaluation
$$\begin{array}{rcl} k & \to & \mathrm{End}(V) \cong V \otimes V^* \\ \lambda & \mapsto & \lambda \mathrm{id}_V \end{array}$$
are morphisms of representations of A.

3. *A relation between the star and the antipode is always fulfilled: the endomorphism $*S^{-1}* = (*S*)^{-1}$ satisfies all conditions for the antipode and thus by the uniqueness of the antipode has to be equal to the antipode. Thus,*
$$S * S* = \ \mathrm{id}_A \ .$$

4. *The universal enveloping algebra $U(\mathfrak{g})$ of any Lie-algebra $\mathfrak{g}$ is another example of a Hopf algebra. On elements $x \in \mathfrak{g} \subset U(\mathfrak{g})$, the coproduct is defined*
$$\Delta(x) = \ x \otimes 1 + 1 \otimes x \quad \text{and} \quad \epsilon(x) = 0 \ .$$
The antipode is defined by $S(1) = 1$ and $S(x) = -x$ for all $x \in \mathfrak{g}$.
Suppose, we consider again a composite system and have eigenvectors $v_1 \in \mathcal{H}_1$ and $v_2 \in \mathcal{H}_2$ for some $x \in \mathfrak{g}$:
$$\rho_1(x)v_1 = \lambda_1 v_1 \quad \text{and} \quad \rho_2(x)v_2 = \lambda_2 v_2 \ .$$
We then find in the composite system
$$\rho(x)v_1 \otimes v_2 = x.v_1 \otimes 1.v_2 + 1.v_1 \otimes x.v_2 = (\lambda_1 + \lambda_2) \ v_1 \otimes v_2 \ .$$
One says that the symmetry $x \in \mathfrak{g}$ leads to an additive quantum number. An example of an additive quantum number is provided by angular momentum.

5. *We do not discuss the rather important point of compatibility between symmetries and statistics. A first example of such symmetries are supersymmetries. It is a rather important discovery that in quantum field theories in low dimensions, there are more general statistics and the ones of bosons and fermions, and that these statistics are compatible with symmetry structures more general than groups and Lie algebras, so-called quantum groups. For a review, refer to G. Mack and V. Schomerus: "A short introduction to quantum symmetry", J. Geom. and Phys. 11 (1993) 361-366.*

5.7 Examples of quantum mechanical systems

In this section, we discuss several important examples: we start with one-particle systems: the step potential, the harmonic oscillator and the hydrogen atom. More detailed discussions of these problems can be found in any book on quantum mechanics.

Example 5.7.1 (Step potential)**.**

1. *The time-independent Schrödinger equation*

$$H\psi = -\frac{\hbar^2}{2m}\psi'' + V(x) = E\psi$$

in $L^2(\mathbb{R})$ is particularly simple to treat if the potential is piecewise constant.

2. *If V is constant, $V(x) = V_0$ we get as solutions plane waves*

$$\psi_k(x) = A(k)\mathrm{e}^{\mathrm{i}kx} + B(k)\ \mathrm{e}^{-\mathrm{i}kx} \quad \textit{with} \quad k = \sqrt{2m(E-V_0)}/\hbar\ .$$

3. *Even if V jumps at a point $a \in \mathbb{R}$, one can show that the solution of the ordinary differential equation has to be continuous with continuous derivative: for a behaviour $\psi(x) \sim \Theta(x-a)$ would imply $\psi''(x) \sim \delta'(x-a)$ and a behaviour $\psi'(x) \sim \Theta(x-a)$ would imply $\psi''(x) \sim \delta(x-a)$.*

4. *Now consider the potential $V(x) = V_0\Theta(x)$. We find for $E > V_=$ as solutions with*

$$k_< = \sqrt{2mE}/\hbar \quad \textit{and} \quad k_> = \sqrt{2m(E-V_0)}/\hbar$$

the wave functions

$$\begin{aligned} \psi(x) &= \mathrm{e}^{\mathrm{i}k_< x} + R\ \mathrm{e}^{-\mathrm{i}k_< x} \quad \textit{for} \quad x < 0 \\ \psi(x) &= T\mathrm{e}^{\mathrm{i}k_> x} \quad \textit{for} \quad x > 0\ . \end{aligned}$$

Continuity of ψ and ψ' at $x = 0$ implies

$$R = \frac{k_< - k_>}{k_< + k_>} \quad \textit{and} \quad T = \frac{2k_>}{k_< + k_>}\ .$$

We discuss this in terms of the probability currents and their relation to the incoming current j_{inc}. There is a reflected current which turns out to be $j_{\mathrm{refl}} = |R|^2 j_{\mathrm{inc}}$. This is a new phenomenon of quantum mechanics. Classically, the particle would move on to the right at a reduced speed, and there would be no reflection.

For the transmitted current, we find $j_{\mathrm{refl}} = \frac{q}{k}|T|^2 j_{\mathrm{inc}}$.

5. *If the energy E is smaller than the potential V_0, we find with*

$$\kappa := \sqrt{2m(V_0 - E)}/\hbar$$

a solution

$$\begin{aligned} \psi(x) = \ & \mathrm{e}^{\mathrm{i}k_< x} + R\ \mathrm{e}^{-\mathrm{i}k_< x} \quad \textit{for} \quad x < 0 \\ \psi(x) = \ & T\mathrm{e}^{-\kappa x} \quad \textit{for} \quad x > 0\ . \end{aligned}$$

and reflection and transmission coefficients

$$R = \frac{k_< - \mathrm{i}\kappa}{k_< + \mathrm{i}\kappa} \quad \textit{and} \quad T = \frac{2k}{k_< + \mathrm{i}\kappa}\ .$$

Thus the particle can penetrate with an exponentially decaying probability into the region $x > 0$ that is classically forbidden, but there is no probability flow in this direction.

6. *If we just have a potential well, $V(x) = V_0\Theta(a - |x|)$ with $a > 0$, a quantum mechanical particle even with energy $E < V_0$ has a certain probability to pass through the well to the very right, i.e. to $x > a$. This is called the tunnel effect and is of utmost practical importance. (We just mention that even with an energy $E > V_0$, a quantum mechanical particle also has a certain probability to get reflected at the barrier at $-a$.)*

Example 5.7.2 (Harmonic oscillator)**.**

1. *We discuss the harmonic oscillator in the Schrödinger picture using the Schrödinger representation. The classical phase space is the cotangent space $T^*\mathbb{R}$ with Hamilton function*

$$h(t, q, p) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 \quad .$$

Here $\omega > 0$ is a real parameter that characterizes the oscillator. The imperfect quantization in the Schrödinger representation yields on the Hilbert space $\mathcal{H} = \mathrm{L}^2(\mathbb{R})$ the densely defined linear operators

$$Q(x) = x \qquad Q(p) = \frac{\hbar}{\mathrm{i}}\frac{\mathrm{d}}{\mathrm{d}x} \qquad \textit{and} \quad Q(p^2) = -\hbar^2\frac{\mathrm{d}^2}{\mathrm{d}x^2}\ .$$

We thus obtain the Hamilton operator

$$H = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{m}{2}\omega^2 x^2 = \frac{\hbar\omega}{2}\left(-x_0^2\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{x^2}{x_0^2}\right)$$

where we have introduced the parameter

$$x_0 := \sqrt{\frac{\hbar}{m\omega}} \in \mathbb{R}_{>0}$$

which has the dimension of length.

2. *It is helpful to introduce the operator*

$$a := \frac{1}{\sqrt{2}} \left(x_0 \frac{\mathrm{d}}{\mathrm{d}x} + \frac{x}{x_0} \right)$$

together with its adjoint

$$a^* = \frac{1}{\sqrt{2}} \left(-x_0 \frac{\mathrm{d}}{\mathrm{d}x} + \frac{x}{x_0} \right) .$$

Heisenberg's commutation relations imply the commutation relation

$$[a, a^*] = 1 .$$

The Hamiltonian can now be expressed as

$$H = \hbar\omega \left(a^* a + \frac{1}{2} \right) .$$

Any operator of the form a^*a is positive definite, since

$$\langle a^* a v, v \rangle = \langle av, av \rangle \geq 0 .$$

Thus the spectrum of H is bounded below by $\frac{\hbar\omega}{2}$.

3. *The infimum of the spectrum is an eigenvalue: the eigenvalue equation*

$$H \, \psi_0 = \frac{1}{2} \, \hbar\omega \, \psi_0$$

is equivalent to the equation $a\psi_0 = 0$ which in the Schrödinger representation gives the following ordinary differential equation

$$x_0 \, \frac{\mathrm{d}\psi_0}{\mathrm{d}x} = -\frac{x}{x_0} \, \psi_0 .$$

It has as an L^2-normalized solution the Gaussian function

$$\psi_0(x) = \left(\frac{1}{\pi^{1/4} \sqrt{x_0}} \right) \exp \left(-\frac{x^2}{2x_0^2} \right) .$$

4. *To discuss the spectrum further, we introduce the number operator $n := a^* \, a$. From the commutation relations of a and a^*, we deduce the commutation relations*

$$[n, a^*] = a^* \qquad \textit{and} \qquad [n, a] = -a .$$

Suppose that ψ_ν is an eigenvector of n to the eigenvalue ν. Then the commutation relations imply the relations

$$\begin{aligned} n(a^*\psi_\nu) &= a^* n\psi_\nu + a^*\psi_\nu = (\nu+1)a^*\psi_\nu \\ n(a\psi) &= an\psi - a\psi = (\nu-1)a\psi, \end{aligned}$$

so that the two vectors $a^*\psi$ and $a\psi$, if they are non-zero, are eigenvectors of n as well and hence eigenvectors of the Hamiltonian H. This explains the name raising operator or creation operator for a^* and lowering operator or annihilation operator for a.

One now concludes that all eigenvectors of the Hamiltonian H are multiples of vectors of the form $(a^*)^n\psi_0$ with $\in \mathbb{Z}_{\geqslant 0}$. We compute them explicitly:

$$\psi_n(x) = \mathcal{N}_n \left(\frac{x}{x_0} + \frac{\mathrm{d}}{\mathrm{d}x}\right)^n \psi_0 = \mathcal{N}_n \mathrm{H}_n\left(\frac{x}{x_0}\right) \exp\left(-\frac{x^2}{2x_0^2}\right)$$

where $\mathcal{N}_n$ is a normalization factor and H_n is the n-th Hermite polynomial which is a polynomial of order n.

5. *The Hermite-polynomials are defined by*

$$\mathrm{H}_n(x) = (-1)^n \mathrm{e}^{x^2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \mathrm{e}^{-x^2} \ .$$

They are easy to memorize with their generating function

$$\sum_{n=0}^{\infty} \frac{1}{n!} t^n \mathrm{H}_n(x) = \mathrm{e}^{-t^2+2tx} \ .$$

6. *We next show that the functions $x^n \exp(-x^2)$ are dense in $L^2(\mathbb{R})$. Indeed, assume that there exists $f \in L^2(\mathbb{R})$ such that*

$$\int_{-\infty}^{\infty} f(x) x^n \exp(-x^2) = 0 \quad \text{for all } n \in \mathbb{N}_0 \ .$$

The integral

$$F(z) := \int_{-\infty}^{\infty} f(x) \mathrm{e}^{\mathrm{i}xz - x^2} \mathrm{d}x$$

is absolutely convergent for all $z \in \mathbb{C}$ and thus defines an entire function on $\mathbb{C}$. For its derivatives, we find

$$F^{(n)}(0) = \mathrm{i}^n \int_{-\infty}^{\infty} f(x) x^n \mathrm{e}^{\mathrm{i}xz - x^2} \ .$$

By our assumption on f, we find $F^{(n)}(0) = 0$. Since F is an entire function, this implies $F(z) = 0$ for all $z \in \mathbb{C}$. But this means that the Fourier transform of the function $f(x)\mathrm{e}^{-x^2}$ vanishes and thus the function itself. This implies that the linear subspace spanned by the eigenfunctions of the harmonic oscillator is dense in $L^2(\mathbb{R})$.

7. *This in turn implies that the spectrum of H is a pure point spectrum so that*
$$\mathrm{spec}(H) = \left\{ \hbar\omega \left(n + \frac{1}{2} \right) \mid n \in \mathbb{Z}_{\geqslant 0} \right\} .$$

8. *We comment on the underlying algebraic structure: the operators a, a^* and K span a Lie algebra isomorphic to the Heisenberg algebra $\mathfrak{g}$; in particular $[a, a^*] = K$. Denote by $\mathfrak{b}_+$ the Lie subalgebra spanned by K and a. It has a one-dimensional representation $\mathbb{C}v$ with*
$$av = 0 \quad \text{and} \quad Kv = v \ .$$
The dense subspace of $L^2(\mathbb{R})$ spanned by the eigenfunctions ψ_n is isomorphic, as a representation of the Heisenberg algebra $\mathfrak{g}$ to the induced representation
$$\mathrm{Ind}_{\mathrm{U}(\mathfrak{b}_+)}^{\mathrm{U}(\mathfrak{g})} \mathbb{C}v = \mathrm{U}(\mathfrak{g}) \otimes_{\mathrm{U}(\mathfrak{b}_+)} \mathbb{C}v \ .$$
This representation is the simplest example of a Fock space.

Example 5.7.3 (Hydrogen atom)**.**

1. *For a central potential $V(\vec{x}) = V(r)$ with $r := |\vec{x}|$ with classical Hamiltonian function*
$$h(t, q, p) = \frac{1}{2m} p^2 + V(r)$$
we introduce the generators
$$\mathrm{L}_i = \epsilon_{ijk} x_j \mathrm{P}_k$$
for angular momentum. From the canonical commutation relations of x and P, we find the commutation relations for the operators
$$[\mathrm{L}_i, \mathrm{L}_j] = i\hbar\, \epsilon_{ijk} \mathrm{L}_k \ .$$
We obtain a representation of the real Lie algebra $\mathrm{su}(2)$ by anti-self-adjoint operators iL_i.

These operators have commutation relations

$$[\mathrm{L}_i, p_j] = i\hbar\, \epsilon_{ijk} p_k$$
$$[\mathrm{L}_i, x_j] = i\hbar\, \epsilon_{ijk} x_k$$

with the operators x_j arising as the quantization of space coordinates and the momenta p_j. One computes that $\mathrm{L}^2 = \mathrm{L}_x^2 + \mathrm{L}_y^2 + \mathrm{L}_z^2$ is the quadratic Casimir operator in $\mathrm{U}(\mathrm{su}(2))$.

2. *We will decompose eigenfunctions into a radial and an angular part; for the latter, we need the Hilbert space $\mathcal{H} = L^2(S^2, \mathbb{C})$ of square integrable functions on the two-dimensional sphere S^2. The rotation group* $\mathrm{SO}(3)$ *acts on the unit sphere S^2 by rotations and thus on the Hilbert space $\mathcal{H} = L^2(S^2, \mathbb{C})$.*

   It is a standard problem of Harmonic analysis to determine how this representation decomposes into irreducible representations of $\mathrm{SO}(3)$*:*

$$L^2(S^2, \mathbb{C}) \cong \oplus_{l=0}^{\infty} V_l$$

   where V_l is the irreducible representation of $\mathrm{SO}(3)$ *of dimension $l+1$ ("of spin l") in which the quadratic Casimir operator acts with eigenvalue $\hbar^2 l(l+1)$.*

   One concludes that a Hilbert space basis of $L^2(S^2, \mathbb{C})$ is given by the countable set of functions

$$Y_{l,m} \quad \text{with} \quad l \in \{0, 1, 2, ...\}, \quad m \in \{-l, -l+1, ..., l-1, l\}$$

   with eigenvalues

$$\mathrm{L}^2 Y_{l,m} = \hbar^2 l(l+1) Y_{l,m}$$
$$\mathrm{L}_3 Y_{l,m} = \hbar m Y_{l,m}$$

   The functions $Y_{l,m}$ are called spherical harmonics. The non-negative integer l is called angular quantum number or orbital quantum number, the integer m is called magnetic quantum number.

3. *To investigate the kinetic part of the Hamiltonian, we rewrite*

$$p^2 = \frac{\mathrm{L}^2}{r^2} + p_r$$

   where we have introduced the differential operator

$$p_r := \frac{\hbar}{\mathrm{i}} \left(\partial_r + \frac{1}{r} \right)$$

for radial momentum. To solve the Schrödinger equation

$$\left[\frac{p^2}{2m} + V(r)\right] \psi(r,\theta,\phi) = E\, \psi(r,\theta,\phi)$$

in radial coordinates (r,θ,ϕ)*, we make the ansatz*

$$\psi(r,\theta,\phi) = R(r)Y_{l,m}(\theta,\phi) \ ,$$

which yields for the function $u(r) = rR(r)$ *the ordinary differential equation*

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{\hbar^2 l(l+1)}{2mr^2} + V(r)\right] u(r) = Eu(r) \ .$$

It suggests to introduce the effective potential

$$V_{\mathit{eff}} := \frac{\hbar^2 l(l+1)}{2mr^2} + V(r) \ .$$

In the special case of the Coulomb potential $V(r) = -\frac{e^2}{r}$ *, we find eigenvalues for the Hamiltonian* H*:*

$$E_n = -\frac{me^4}{2\hbar^2}\,\frac{z^2}{n^2} = \frac{E_1}{n^2} \quad n = 1,2,3,\ldots$$

The integer n *is called principal quantum number. The range of the orbital quantum number* l *is restricted for given value* n *of the principal quantum number to* $l \in \{0,1,...,n-1\}$*.*

4. *We thus find the following eigenvectors:*

Principal quantum nr.	*Angular quantum nr.*	*Magnetic quantum nr.*	*Energy*	*# states*
$n=1$	$l=0$	$m=0$	E_1	*1 (ground state)*
$n=2$	$l=0$	$m=0$	$E_1/4$	
	$l=1$	$m \in \{\pm 1, 0\}$	$E_1/4$	*4*
$\vdots$	$\vdots$			
n	$l=0,\ldots n-1$		E_1/n^2	n^2

We finally present three comments:

(a) *The energy eigenvalue does not depend on the quantum numbers* l *and* m*, yielding a degeneracy of the eigenstates. This is specific for the Coulomb potential and can be explained by the fact that for this potential the so-called Runge-Lenz vector provides three more quantities*

commuting with the Hamiltonian (but not with angular momentum). For a discussion, we refer to Chapter 3.5.3 of the book by Takhtajan. In the real hydrogen atom, this degeneracy is lifted by relativistic effects.

(b) *The spectrum has also a continuous part for* $\mathrm{E} \geqslant 0$*; the corresponding generalized eigenstates are called* scattering states.

(c) *When one measures radiative transitions in atomic spectra, one measures energy differences. These are of the form* $\Delta E = |E_1| \left(\frac{1}{n^2} - \frac{1}{m^2}\right)$*, with integers* n *and* m*, in very good agreement with spectral experiments.*

Remark 5.7.4.
There are two more main players in non-relativistic quantum mechanics:

1. *The particle with spin non-relativistic quantum mechanics.*
2. *The photon.*

Discuss the particle with spin on a spin^c*-manifold. Pauli exclusion principle.*

Remarks 5.7.5.
One of the achievements of quantum mechanics is the explanation of the stability of atoms. Let us comment on stability.

1. *We distinguish between* dynamical *and* energetic stability*. Dynamical stability means that the time evolution is well-defined for all times, independently of the initial conditions. Energetic stability means that the total energy of a system cannot become arbitrarily negative.*

   Both notions of stability are an issue in classical mechanics. In quantum mechanics, energetic stability implies dynamical stability: any symmetric bounded *operator has a self-adjoint extension, the Friedrichs extension. In the case of the Hamiltonian, this self-adjoint extension generates a dynamics for all times.*

   Hence the energetic stability of a single atom, also called stability of the first kind*, is quite striking structural feature of quantum mechanics.*

2. Stability of matter *or* stability of the second kind *concerns macroscopic matter which is supposed to consist of very many atoms. Stability of the first kind states that the total energy of the system is bounded from below. Stability of the second kind asks how the lower bound depends on the size of the system.*

Take first the volume of a system as an example of a quantity depending on the size, measured in terms of the number N of particles. The volume $V(N)$ of a system should grow with its number N of components; otherwise, more particles would take less space per particle and the volume of matter would not be extensive.

3. *In the same spirit, the total energy of a system should be extensive. Imagine we have N electrons and M nuclei with atomic numbers $\underline{Z} := (Z_1, \ldots, Z_M)$. Then the infimum $E_{N,M}(\underline{Z})$ of the energy of such a system should be bounded below in the form*

$$E_{N,M}(\underline{Z}) \geq \Xi(\max_{1\leq i\leq M} Z_i)(N+M) \ .$$

The constant Ξ depends on Planck's constant, the mass and the charge of the electron, the maximum of the charges Z_i of the nuclei, but not on the masses of the nuclei.

It is important to consider particles with spin obeying the Pauli exclusion principle and thus wave functions for the electrons

$$\psi : \mathbb{R}^{3N} \to \mathbb{C}^2$$

that are totally antisymmetric and L^2-normalizable. More precisely, the following expression has to be estimated, where we write the wave function for the electron as a function of $\underline{X} \in \mathbb{R}^{3N}$ and parameters $\underline{\sigma}$ for the spin. We consider the kinetic energy

$$T_\psi = \frac{\hbar^2}{2m} \sum_{i=1}^{N} \sum_{\underline{\sigma}} \int_{\mathbb{R}^{3N}} |\nabla_{\mathbf{x}_i} \psi(\underline{X}, \underline{\sigma})|^2 d\underline{X}$$

and the potential energy *as a function of $\underline{X} \in \mathbb{R}^{3M}$:*

$$V_\psi(\underline{R}) = \sum_{\underline{\sigma}} \int_{\mathbb{R}^{3N}} V_{\mathrm{C}}(\underline{X}, \underline{R}) |\psi(\underline{X}, \underline{\sigma})|^2 d\underline{X},$$

where we have introduced the Coulomb potential:

$$V_{\mathrm{C}}(\underline{X}, \underline{R}) = -\sum_{i=1}^{N} \sum_{j=1}^{M} \frac{Z}{|\mathbf{x}_i - \mathbf{R}_j|} + \sum_{1\leq i<j\leq N} \frac{1}{|\mathbf{x}_i - \mathbf{x}_j|} + \sum_{1\leq i<j\leq M} \frac{Z^2}{|\mathbf{R}_i - \mathbf{R}_j|}.$$

Then the expression to be estimated reads:

$$E_{N,M}(\underline{Z}) = \inf\{T_\psi + e^2 V_\psi(\underline{R}) \quad \textit{with} \quad \underline{X} \in \mathbb{R}^{3M}\}$$

for all wave functions ψ that are fermionic and normalized in a Sobolev subspace of the L^2-space.

For information on how to address this problem of analysis, we refer to the book:
Elliott H. Lieb and Robert Seiringer: Stability of Matter in Quantum Mechanics, Cambridge University Press, 2010.

Remark 5.7.6.
Minimal coupling and the importance of the potential. Discuss the Aharonov Bohm effect.

Remark 5.7.7.
Comment on quantization of the radiation field.

5.8 Quantum statistical mechanics and KMS states

Observation 5.8.1.
We first discuss the notion of a ground state. For a system with a fixed number of particles, a ground state should be a state of lowest energy. By shifting the energies such that the Hamiltonian is positive definite, $H \geq 0$, this leads to the following definition.

Definition 5.8.2
Let (A, α_t) be a time-independent quantum mechanical system with a positive definite self-adjoint Hamiltonian H. A ground state is a state $\omega \in S(A)$ such that $\overline{\alpha}_t(\omega) = \omega$.

Remarks 5.8.3.

1. *We do not impose invariance under translations and rotations (this would be inappropriate for the description of crystals) or under Galilei transformations.*

2. *Suppose a quantum mechanical system admits ground states that are* not *invariant under the full symmetry group of the theory. Then one says that the symmetry is spontaneously broken.*

Observation 5.8.4.

1. *We make the same assumptions for the Hamiltonian H: it is a self-adjoint, positive operator on a separable Hilbert space $\mathcal{H}$. We extend the time evolution from real time parameters to time parameters in the closed complex upper half-plane $\overline{\mathbb{H}}_+ := \{z \in \mathbb{C} \,|\, \mathrm{Im}(z) \geq 0\}$. We find a map*

$$\begin{array}{rcl} \overline{\mathbb{H}} & \to & B(\mathcal{H}) \\ z & \mapsto & \mathrm{e}^{\mathrm{i}zH} \end{array}$$

The operators in the image are bounded with norm $\|\mathrm{e}^{\mathrm{i}zH}\| \leq 1$. The function on the upper half plane is analytic, since

$$\mathrm{e}^{\mathrm{i}zH} = \mathrm{e}^{\mathrm{i}(z-z_0)H}\mathrm{e}^{\mathrm{i}z_0H} = \sum_{n=0}^{\infty} \mathrm{i}^n (z-z_0)^n H^n \mathrm{e}^{\mathrm{i}z_0H}$$

is a power series of bounded operators that is absolutely convergent for $|z - z_0| \leq \mathrm{Im}(z_0)$.

2. *Now suppose that ω is a ground state. Then for any pair of observables $a, b \in A$ the function*

$$F_{ab} : z \mapsto \langle \Omega_\omega, \pi_\omega(a)\mathrm{e}^{\mathrm{i}zH}\pi_\omega(b)\Omega_\omega \rangle$$

is a bounded analytic function on $\overline{\mathbb{H}}_+$. Note that here we are considering correlators in the state ω with a possibly complex difference of time.

One can even show:

Theorem 5.8.5.
A state ω for the quantum mechanical system (A, H) is a ground state, if and only if for each pair $a, b \in A$ the functions

$$F_{ab}(t) \ = \omega(a\alpha_t b \alpha_t^{-1})$$

are bounded on $\overline{\mathbb{H}}_+$ and analytic on $\mathbb{H}_+$.

For the proof, we refer to proposition 5.3.19 of the book by Bratteli and Robinson.

Observation 5.8.6.

1. *Let $(A, H, \mathcal{H})$ be a time-independent quantum mechanical system represented on a separable Hilbert space $\mathcal{H}$. If $\mathrm{e}^{-\beta H}$ is a trace class operator for all $\beta > 0$, we can consider the so-called Gibbs' state for any $\beta > 0$*

$$\omega_\beta(a) \ = Z^{-1}\mathrm{tr}\mathrm{e}^{-\beta H}a \quad \text{with} \quad Z := \mathrm{tr}\mathrm{e}^{-\beta H} \ .$$

Writing $\beta = \frac{1}{kT}$, the parameter T has the interpretation of a temperature.

2. *The basic problem in defining equilibrium states in* infinitely extended *systems is the fact that these states are typically* not *normal states of the vacuum representation. We have thus to formulate a condition on the state without assuming that it is represented by a density matrix.*

Definition 5.8.7
A state ω is said to be a KMS-state *(for Kubo-Martin-Schwinger) at temperature T or the inverse temperature $\beta = \frac{1}{kT}$, if for each pair $a, b \in A$ there the function*

$$F_{a,b}(t) := \omega(a\alpha_t(b))$$

is bounded continuous on the closed strip

$$\overline{\mathbb{H}}_\beta := \{z \in \mathbb{C} \,|\, 0 \leq \mathrm{Im}z \leq \beta\}$$

and analytic on the open strip

$$\mathbb{H}_\beta := \{z \in \mathbb{C} \,|\, 0 < \mathrm{Im}z < \beta\}$$

and one has

$$F_{a,}(t + \mathrm{i}\beta) \;=\; \omega(\alpha_t(b)a) \;.$$

Remarks 5.8.8.

1. *Suppose that we are dealing with a quantum mechanical system with time development specified by a Hamilton operator H such that e^{-tH} is trace class for all $t > 0$. Then the Gibbs state obeys the KMS condition.*

2. *One can compute the KMS states for an ideal gas and finds a two-parameter family labelled by temperature and chemical potential. This provides an inherent justification of the chemical potential without recourse to the grand canonical ensemble.*

5.9 Perturbation theory

Possibly mention WKB as another approximation scheme.

5.10 Path integral methods

Mention correlators. Then path integral and Wiener measure.

6 A glimpse to quantum field theory

We start with a very naive idea on what a quantum field is. In the simplest case, for a configuration space of the forms $p_1 : M \times \mathbb{R} \to M$, a classical field is a function on space time. In quantum mechanics, numbers are replaced by operators. So instead of ordinary functions, we study operator-valued functions on space time. It turns out that we also have to deal with something like a δ-function, we thus deal with operator-valued distributions, and they should be mathematical formulations of quantum fields. An ordinary Schwartz distribution assigns a number to each test function, by definition. An operator-valued distribution should assign an operator, which may be unbounded, to each test function. There is a precise mathematical definition of this notion of an operator-valued distribution, and we can further axiomatize a physical idea of what a quantum field should be. Such an axiomatization is known as a set of Wightman axioms.

We work on a Minkowski space R^4. The scalar product of $x = (x_0, x_1, x_2, x_3)$ and $y = (y_0, y_1, y_2, y_3)$ is $x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3$. The linear maps $\mathbb{R}^4 \to R^4$ preserving the scalar product are the Lorentz transformations. The restricted Lorentz group is the subgroup such that $\Lambda^{00} > 0$ and $\det \Lambda = 1$. The semi-direct product with translations is called the Poincaré group; the subgroup of the Poincaré group of elements $x \mapsto \Lambda x + b$ such that Λ is in the restricted Lorentz group is called the restricted Poincaré group.

The universal cover of the restricted Poincaré group is naturally identified with

$$\{(A, a) | A \in \mathrm{SL}(2, \mathbb{C}), a \in \mathbb{R}^4\} .$$

We say that two regions O_1, O_2 of Minkowski space are spacelike separated, if for any $x = (x_0, x_1, x_2, x_3) \in O_1$ and $y = (y_0, y_1, y_2, y_3) \in O_2$, we have $(x_0 - y_0)^2 - (x_1 - y_1)^2 - (x_2 - y_2)^2 - (x_3 - y_3)^2 < 0$.

We are now ready to state the Wightman axioms.

1. We have closed operators $\phi_1(f), \phi_2(f), \ldots \phi_n(f)$ on a Hilbert space H for each smooth function f on $\mathbb{R}^4$ with compact support.

2. There exists a dense subspace $D \subset H$ that is contained in the domains of all $\phi_i(f)$, $\phi^*(f)$ for all $f \in C_C^\infty(\mathbb{R}^4)$. For all such f, we have $\phi(f)D \subset D$ and $\phi_i^*(f)D \subset D$. For all $v, s \in D$, the function $f \mapsto \langle v, \phi(f)w \rangle$ is a Schwartz distribution.

3. We require that H carries a unitary representation U of the universal cover of the restricted Poincaré group and of an n-dimensional representation of

$\mathrm{SL}(2,\mathbb{C})$ such that

$$\begin{array}{lcl} U(A,a)D & = & D \\ U(A,a)\phi_i(f)U(A,a)^* & = & \sum_j S(A^{-1})_{ij}\phi_j(f(A,a)), \text{ with } f(A,a) = f(\Lambda(A)^{-1}(x-a)) \end{array}$$

Here the last identity is to be understood as identities on U.

4. If the supports of two smooth functions f and g are compact and spacelike separated, then we $[\phi_i(f), \phi_j(g)] = 0$ and $[\phi_i(f), \phi_j^*(g)] = 0$

5. There is a non-zero vector $\Omega \in D$, the vacuum vector such that

   - $U(A,a)\Omega = \Omega$ for all elements of the restricted Lorentz group
   - the spectrum of the four-parameter unitary group $U(I,a), a \in \mathbb{R}^4$, is in the positive cone

   $$\{(x_0,x_1,x_2,x_3) | x_0 > 0, x_0^2 - x_1^2 - x_2^2 - x_3^2 \geq 0\}$$

   - The subspace generated by finitely many applications of $\phi_i(f)$ and $\phi_j^*(g)$ on Ω is dense in H.

   Moreover, we require the subspace of such vectors to be one-dimensional.

Distributions and unbounded operators cause various technical difficulties in a rigorous treatment. Bounded linear operators are much more convenient for algebraic handling, and we seek for a mathematical framework using only bounded linear operators. There has been such a framework pursued by Araki, Haag and Kastler, and such an approach is called algebraic quantum field theory today. The basic reference is Haag's book and we now explain its basic ideas.

Suppose we have a family of operator-valued distributions $\{\phi\}$ subject to the Wightman axioms. For an operator valued distribution ϕ and a smooth test function $f \in C_c^\infty(\mathbb{R}^4)$ with support in a bounded region O the expression $\phi(f)$ gives an unbounded operator operator which we assume to be self-adjoint.

In quantum mechanics, observables are represented as (possibly unbounded) self-adjoint operators. We regard $\phi(f)$ as an observable in the spacetime region O. We have a family of such quantum fields and many test functions with supports contained in O, so we have many unbounded operators for each O. Applying spectral projections of these unbounded operators, we can consider the von Neumann algebra $A(O)$ generated by these projections. In this way, we obtain a family of von Neumann algebras $\{A(O)\}$ on the same Hilbert space H parameterized by bounded space time regions O. We axiomatize properties of this family.

We list the axioms for space time being Minkowski space $\mathbb{R}^4$. In this case, it is enough to consider as bounded regions double cones of the form $(x + V_+) \cap (y + V_-)$, where $x, y \in \mathbb{R}^4$ and we introduce the forward and backward cones

$$V_\pm = \{z = (z_0, z_1, z_2, z_3) \in \mathbb{R}^4 | z_0^2 - z_1^2 - z_2^2 - z_3^2 >, \pm z_0 > 0\} \ .$$

1. (Isotony) For a larger double cone $O_2 \supset O_1$, we have more test functions, and operators: $A(O_1) \subset A(O_2)$.

2. (Locality) Suppose two double cones O_1 and O_2 are spacelike separated so that no interaction between them is possible, even at the speed of light. Then the observables in the two regions should commute, leading to the requirement that the elements in $A(O_1)$ and $A(O_2)$ commute.

3. (Poincaré Covariance) The (universal cover of the) natural symmetry group of Minkowski space, the restricted Poincaré group, is required to be unitarily represented on H such that $A(gO) = U_g A(O) U_g^*$.

4. (Vacuum) There is a distinguished unit vector $\Omega \in H$, the vacuum vector such that $U_g \Omega = \Omega$ for all elements g in the restricted Poincaré group.

5. (Irreducibility) We require that $\cup_O A(O)\Omega$ is dense in H.

6. (Spectrum Condition) If we restrict the representation U to the translation subgroup, its spectrum is contained in the closure of V_+.

It is clear that the above set of axioms is very similar to the Wightman axioms. The assignment of $0 \mapsto A(O)$ is called a net (of von Neumann algebras). It is very hard to construct an example satisfying the above axioms. In the 4-dimensional Minkowski space, we have only one example known and this is known under the name of free fields.

A Differentiable Manifolds

A.1 Definition of differentiable manifolds

Manifolds are topological spaces which locally "look like" an open subset of $\mathbb{R}^n$. Examples of manifolds are abundant in geometry and mathematical physics. They are most easily visualized in the case of two-dimensions: a sphere or a doughnut locally look like $\mathbb{R}^2$. Indeed, we describe the earth in an atlas as follows:

- We have a collection of charts which are pieces of $\mathbb{R}^2$ on which some part of the earth is represented. In practice, of course, with simplifications, but this is inessential for our purposes. What matters is that every point on the earth together with a small neighborhood of it appears in at least one chart.
- Close to the boundary of each chart, we have a prescription on how to glue together the charts. In fact, the points close to the boundary of any chart have to appear in at least one more chart.

We formalize this in the following definition:

Definition A.1.1

(i) Let n be a natural number. An n-dimensional topological manifold M is a Hausdorff topological space with a countable basis of open sets which is locally homeomorphic to $\mathbb{R}^n$. We sometimes write M^n to make the dimension explicit or $\dim M = n$.

In more detail, this means: for every point $p \in M$ there exists an open neighborhood $p \in U \subset M$ and a homeomorphism $x_U : U \to V$ to some open subset $V \subset \mathbb{R}^n$.

(ii) A chart of an n-dimensional topological manifold is a homeomorphism $x : U \to V$ from an open subset of $U \subset M$ to an open subset of $V \subset \mathbb{R}^n$.

(iii) An atlas is a family I of charts $x_\alpha : U_\alpha \to \mathbb{R}^n$ such that the map

$$\sqcup_{\alpha\in I} x_\alpha^{-1} : \sqcup_{\alpha\in I} U_\alpha \to M$$

given by the homeomorphisms x_α^{-1} is a surjection.

(iv) An atlas is called differentiable, if for all pairs x, y of coordinate charts the map

$$y \circ x^{-1} : \quad \mathbb{R}^n \to \mathbb{R}^n$$

is a diffeomorphism everywhere where it is defined.

(v) A topological manifold, together with the choice of a differentiable atlas, is called a differentiable manifold or, synonymously a smooth manifold.

Remarks A.1.2.

1. *Recall that a Hausdorff space is a topological space M in which for any two disjoint points $p, q \in M$ there exist disjoint open sets U, V in M with $p \in U$ and $q \in V$.*

   The following example shows that it is necessary to impose the Hausdorff property, even though we require the topological space underlying a manifold to look locally like $\mathbb{R}^n$. Take two copies of the real axis $\mathbb{R}$ with coordinates y and y' respectively. Then identify all points with $y = y'$, if $y \neq 0$ to obtain a topological space that is not Hausdorff, because there are no disjoint open neighbourhoods of the two distinct points $y = 0$ and $y' = 0$. Still, each point is contained in a coordinate neighbourhood homeomorphic to an open interval in $\mathbb{R}$.

2. *The requirement of M to have a countable basis implies that M is paracompact: for every atlas (U_α), there exists a locally finite atlas $(V_\beta)_{\beta \in I}$ with every open set V_β contained in some open set U_α. One says that the atlas $(V_\beta)_{\beta \in I}$ is subordinate to the atlas (U_α). An atlas $(V_\beta)_{\beta \in I}$ is locally finite, if for every point $p \in M$, there exists an open neighborhood which intersects only a finite number of the sets V_β.*

   For paracompact spaces, one can find a <u>partition of unity</u> subordinate to the locally finite cover $(V_\beta)_{\beta \in I}$: this is a set of smooth real-valued functions $(g_\beta)_{\beta \in I}$ such that

   (a) The functions are bounded: $0 \leq g_\beta \leq 1$ in M for each $\beta \in I$.

   (b) The support of g_β is contained in V_β.

   (c) $\sum_{\beta \in I} g_\beta(p) = 1$ for each p. (Notice that this sum is finite for any fixed $p \in M$.)

   A partition of unity allows to write a smooth function $f \in C^\infty(M)$ as a locally finite sum

$$f = f \cdot 1 = \sum_\beta (f \cdot g_\beta)$$

   of smooth functions $f \cdot g_\beta$ with support contained in the coordinate chart V_β. This is essential for setting up an integration theory on manifolds.

Examples A.1.3.

1. *An affine space $\mathbb{A}$ modelled over $\mathbb{R}^n$ is an n-dimensional manifold. Any point $p \in \mathbb{A}$ determines via $q = p + v \mapsto v \in \mathbb{R}^n$ a globally defined coordinate chart.*

2. *Any discrete topological space is a zero-dimensional manifold.*

3. *The circle S^1 is a one-dimensional manifold. More generally, the n-sphere in an $n+1$-dimensional Euclidean vector space $(V, \langle \cdot, \cdot \rangle)$*

$$S^n := \{x \in V | \langle x, x \rangle = 1\}$$

   is an n-dimensional (compact) manifold.

4. *A torus is a two-dimensional (compact) manifold. The Klein bottle is a two-dimensional (compact) manifold.*

5. *More genrally let $L \subset \mathbb{R}^n$ be a lattice, i.e. L is the integral span of a basis of $\mathbb{R}^n$. A lattice is a discrete subgroup of the additive group of $\mathbb{R}^n$, and the quotient $\mathbb{R}^n / L$ is an n-dimensional (compact) manifold, the n-dimensional torus T^n.*

Definition A.1.4

(i) Let M, N be differentiable manifolds of dimensions m, n respectively. Let G be open in M. A continuous map

$$f : M \to N$$

is called differentiable or smooth , if for all charts the map

$$y \circ f \circ x^{-1} : \mathbb{R}^m \to \mathbb{R}^n$$

is differentiable wherever it is defined.

(ii) Differentiable maps are the morphisms in the category of differentiable or smooth manifolds. This category is also called the category of smooth manifolds. The isomorphisms are called diffeomorphisms. (Note that at this point, we do not know how to differentiate diffeomorphisms!)

(iii) For any open subset $U \subset M$ of a smooth manifold, we consider the real-valued differentiable functions $C^\infty(U, \mathbb{R}) =: \mathcal{F}(U)$. Under addition and multiplication of their values, they form an $\mathbb{R}$-algebra. For any inclusion $U \subset V$ of open subsets we have a restriction

$$\mathrm{res}^V_U : \quad \mathcal{F}^\infty(V, \mathbb{R}) \to \mathcal{F}^\infty(U, \mathbb{R})$$

which is a morphism of $\mathbb{R}$-algebras. Obviously, $\mathrm{res}^U_U = \mathrm{id}$ and for $U \subset V \subset W$ we have

$$\mathrm{res}^W_U = \mathrm{res}^V_U \circ \mathrm{res}^W_V \ .$$

These axioms are expressed by saying that we have a presheaf of $\mathbb{R}$-algebras on M.

(iv) This presheaf has two more properties. Suppose that $(U_\alpha)_{\alpha \in I}$ is an open covering of an open subset $U \subset M$.

- *If for two functions $f, g \in \mathcal{F}(U)$ the restrictions to all open subsets coincide, $\mathrm{res}^U_{U_\alpha}(f) = \mathrm{res}^U_{U_\alpha}(g)$ for all $\alpha \in I$, then the functions coincide, $f = g$. This just expresses the fact that a function is uniquely determined once we know it locally everywhere.*
- *Given a collection of functions $f_\alpha \in \mathcal{F}(U_\alpha)$ such that the restrictions agree on all twofold intersections:*

  $$\mathrm{res}^{U_\alpha}_{U_\alpha \cap U_\beta}(f_\alpha) = \mathrm{res}^{U_\beta}_{U_\alpha \cap U_\beta}(f_\beta)$$

  for all pairs $\alpha, \beta \in I$, there is a function $f \in \mathcal{F}(U)$ defined on all of U such that $f_\alpha = \mathrm{res}^U_{U_\alpha}(f)$. This expresses the fact that we can globally define functions by patching together locally defined functions, if they coincide on all overlaps.

This is expressed by saing that we have a sheaf of $\mathbb{R}$-algebras on M.

(v) For the sheaf of smooth functions on a smooth manifold, all restriction morphisms are surjective. One says that the sheaf is flabby.

(vi) *Let $U \subset M$ be open and let for $p \in U$ be $\mathcal{F}_p^0(U)$ the ideal of functions that vanish in some neighborhood of p. The quotient algebra*

$$\mathcal{F}_p(U) := \mathcal{F}(U)\big/\mathcal{F}_p^0(U)$$

is called the algebra of germs of differentiable functions on U in p.

Lemma A.1.5.
Given an inclusion $U \subset V$ of open subsets containing p, the restriction

$$\mathcal{F}(V) \to \mathcal{F}(U)$$

induces a canonical isomorphism

$$\mathcal{F}_p(V) \tilde{\to} \mathcal{F}_p(U).$$

Proof:
Injectivity of is immediate by the definition of the ideal that is divided out. Surjectivity is implied by the fact that any smooth function $\varphi \in \mathcal{F}(U)$ can be continued to a *smooth* function on the larger open set V. □

For any smooth function φ defined on a neighborhood V of p determines an element $[\varphi] \in \mathcal{F}_p$. The value of φ at p determines a map

$$\begin{array}{ccc} \mathcal{F}_p & \to & \mathbb{R} \\ [\varphi] & \mapsto & \varphi(p) \ . \end{array}$$

The germ of a function contains, however, much more information than the value of a function, but its whole local behaviour. In particular, all derivatives can be obtained.

One should be aware that this lemma does not hold, if one works with analytic rather than differentiable functions. In this case, one has to consider a direct limit over the directed system of open subsets containing the point p.

A.2 Tangent vectors and differentiation

Differentiation requires only a local knowledge of the function. We should thus define differentiation on germs of functions. We consider all operations on germs of functions that obey the algebraic rules for differentiation: linearity and the Leibniz rule.

Definition A.2.1

(i) *A tangent vector v at M in the point p is a derivation*

$$v : \mathcal{F}_p M \to \mathbb{R}$$

on the algebra $\mathcal{F}_p M$ of germs of smooth functions at the point p, i.e. an $\mathbb{R}$-linear map for which the Leibniz rule

$$v(fg) = v(f)g(p) + f(p)v(g)$$

holds for all $f, g \in \mathcal{F}_p(M)$.

(ii) *The set of all derivations in a point p has a natural structure of an $\mathbb{R}$-vector space, the* tangent space *T_pM.*

Remarks A.2.2.

(a) *Examples for tangent vectors of $\mathbb{R}^n$ are provided by partial derivatives:*

$$\frac{\partial}{\partial x^i}\Big|_p : f \mapsto \frac{\partial f}{\partial x^i}(p) \ .$$

Differentiation is linear and obeys the product rule.

(b) *For a general manifold M, choose a coordinate chart x around a point $p \in M$. Then the partial derivative for the germ in the local coordinate x defines for all $i = 1, \ldots, n$ a tangent vector:*

$$\begin{aligned} \tfrac{\partial}{\partial x^i}\Big|_p : \quad \mathcal{F}_pM &\to \mathbb{R}, \\ f &\mapsto \tfrac{\partial}{\partial x^i}(f \circ x^{-1})\Big|_{x(p)} \end{aligned}$$

Lemma A.2.3.
Let M be a smooth n-dimensional manifold.

1. *For any local coordinate x, the family $\left\{\frac{\partial}{\partial x^i}|_p\right\}_{i=1,\ldots,n}$ is a basis of T_pM.*
2. *This implies in particular $\dim_{\mathbb{R}} T_pM = \dim M$. Any tangent vector $v \in T_pM$ can thus be written uniquely as a linear combination*

$$v = \sum_{i=1}^{n} \alpha^i \ \frac{\partial}{\partial x^i}\Big|_p \ \textit{with} \quad \alpha^i \in \mathbb{R} \ .$$

We claim that $\alpha^i = v(x^i)$, i.e. the coefficients are obtained by evaluating the derivation on the germ of the coordinate function x^i.

Proof:

1. We first consider the following situation: fix an open ball U around $0 \in \mathbb{R}^n$ and a function $\varphi \in C^\infty(U, \mathbb{R})$ with $\varphi(0) = 0$. For any $i = 1, \ldots n$ consider the function

$$\psi_i(u) := \int_0^1 \frac{\partial}{\partial x^i}\varphi(tu)\ \mathrm{d}t \quad \in C^\infty(U, \mathbb{R}).$$

Obviously,

$$\psi_i(0) = \frac{\partial}{\partial u^i}\Big|_0 \varphi$$

and the chain rule shows

$$\varphi(u) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\varphi(tu) = \sum_{i=1}^{n} u^i \psi_i(u).$$

2. Given a local coordinate x around p with $x(p) = 0$ we use the preceding remark to deduce that for any function f with $f(p) = 0$, the function $f \circ x^{-1} : \mathbb{R}^n \to \mathbb{R}$ can be written as

$$f \circ x^{-1}(u) = \sum_{i=1}^{n} u^i \psi_i(u)$$

and thus

$$f = \sum_{i=1}^{n} x^i \psi_i \circ x\big|_{x^{-1}(U)} \ .$$

We conclude

$$v(f) = v(f|_{x^{-1}(U)}) = \sum_{i=1}^{n} v(x^i)\psi_i(0) = \sum_{i=1}^{n} v(x^i) \frac{\partial}{\partial x^i}\Big|_p f \ .$$

3. For a general function f, we apply the argument to the function $\tilde{f}(q) := f(q) - f(p)$.

□

Lemma A.2.4.
If x and y are local coordinate charts around $p \in M$, we have

$$\frac{\partial}{\partial y^i}\Big|_p = \sum_{j=1}^{n} \alpha_i^j \frac{\partial}{\partial x^j}\Big|_p$$

with α the square matrix given by the Jacobian matrix for function describing the change of coordinates $x \circ y^{-1} : \mathbb{R}^n \to \mathbb{R}^n$,

$$\alpha_i^k = \frac{\partial x^k}{\partial y^i} \ .$$

Proof:
The existence of the coefficients α follows from the fact that we have two bases of tangent space T_pM. Applying both sides to the function on $U \subset M$ given by the k-th coordinate

$$x^k : U \to \mathbb{R}$$

gives the concrete form of the coefficients α_i^k. □

Example A.2.5.
Let $(\mathbb{A}, V)$ be an affine space modelled over a finite-dimensional $\mathbb{R}$-vector space V. Given a point $p \in \mathbb{A}$, choose an affine line $\{p + tv,\, t \in \mathbb{R}\}$. Differentiation in the direction of the affine line,

$$\frac{\partial}{\partial v} : \quad f \mapsto \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} f(p + tv) \quad \textit{for } v \in V,$$

gives a derivation on $\mathcal{F}_p\mathbb{A}$. This provides a canonical isomorphism of the tangent space $T_p\mathbb{A}$ and the difference vector space V:

$$\begin{array}{ccc} V & \tilde{\to} & T_p\mathbb{A} \\ v & \mapsto & \frac{\partial}{\partial v} \ . \end{array}$$

Definition A.2.6
Let $f : M \to N$ be a differentiable mapping of smooth manifold.

1. *Functions can be pulled back: if $\varphi : \ U \to \mathbb{R}$ is a function defined on an open subset $U \subset N$, then $\varphi \circ f : f^{-1}(U) \to \mathbb{R}$ is a real-valued smooth function defined on the subset $f^{-1}(U) \subset M$ which is open since f is continuous.*

2. *This pull back is compatible with the quotients leading to germs. We thus get for any point $p \in M$ a ring homomorphism on the ring of germs of smooth functions*

$$\begin{array}{rcl} f^* : \mathcal{F}_{f(p)}N & \to & \mathcal{F}_pM \\ \varphi & \mapsto & \varphi \circ f \end{array}$$

   called the pullback of germs of functions.

3. *This, in turn, gives rise to a linear map on tangent vectors:*

$$f_{*p} : \ T_pM \to T_{f(p)}N$$

   by

$$(f_{*p}v)(\varphi) = v(f^*\varphi) = v(\varphi \circ f) \quad \text{for } v \in T_pM, \quad \text{and} \quad \varphi \in \mathcal{F}_{f(p)}N.$$

   For this linear mapping, different notations are in use,

$$f'_p = \mathrm{D}f|_p = T_pf \ .$$

Remarks A.2.7.
(a) One shows that

$$\mathrm{id}_{*p} = \mathrm{id}_{T_pM}$$

and that the chain rule holds:

$$(g \circ f)_{*p} = g_{*f(p)} \circ f_{*p}.$$

(b) If x is a chart around p and y a chart around $f(p)$, then f_{*p} is given with respect to the bases $\left\{ \frac{\partial}{\partial x^i}\big|_p \right\}$ and $\left\{ \frac{\partial}{\partial y^j}\big|_{f(p)} \right\}$ by the Jacobian matrix of $y \circ f \circ x^{-1}$ in the point $x(p)$.

Tangent vectors can also be thought as possible values for "velocities". Let us make this more explicit:

Remark A.2.8.
To be able to discuss velocities, we consider a trajectory parametrized by an interval $I \subset \mathbb{R}$ for an arbitrary smooth manifold M: this is a smooth map

$$\varphi : I \to M$$

Since I is a subset of an affine space, we can identify the derivative with a linear map defined on the difference vector space $\mathbb{R}$

$$\mathrm{D}\varphi : \ \mathbb{R} \cong T_tI \to T_{\varphi(t)}M \ .$$

Such a linear map can be described by its value on $1 \in \mathbb{R}$*, i.e. by*

$$\dot{\varphi}(t) := \mathrm{D}\varphi(1) \ \in T_{\varphi(t)}M \ ,$$

which is just a tangent vector.

We still have to convince ourselves that all tangent vectors arise as velocities. This will be shown later, when we discuss ordinary differential equations in this language.

We finally single out subclasses of differentiable maps:

Definition A.2.9

(i) A differentiable map $f : \ M \to N$ *is called an* immersion*, if the linear map*

$$f_{*p} : \ T_pM \to T_{f(p)}N$$

is injective for all $p \in M$*. It is called a* submersion*, if the linear maps are surjective for all* $p \in M$*.*

(ii) An embedding *is an immersion that is a homeomorphism of* M *to a subspace of* N*. (Easy exercise: find an immersion that is* not *an embedding.)*

(iii) Let N *be a smooth manifold. If the inclusion* $M \subseteq N$ *of a subset is an embedding, then* M *is called a* submanifold *of* N*.*

(iv) Let $f : M \to N$ *be a differentiable map. Then* $p \in M$ *is called a* regular point *of* f *if the linear map* f_{*p} *is surjective; otherwise it is called a* critical point *of* f*.*

(v) A point $q \in N$ *is called a* regular value *of* f*, if all* $p \in f^{-1}(q)$ *are regular points of* f*. Otherwise, it is called a* critical value*.*

The proof of the following statement uses the theorem of implicit functions. We leave it as an exercise to the reader:

Lemma A.2.10.
Let M *be a differentiable manifold of dimension* n *and* N *be a differentiable manifold of dimension* k *with* $k \leq n$*. The preimage of a regular value under a differentiable map*

$$f : M^n \to N^k$$

is a submanifold of dimension $n - k$*.*

A.3 Fibre bundles and Lie groups

Definition A.3.1
The Cartesian product *$M \times N$ of two differentiable manifolds M, N of dimension n and k respectively, is endowed with the structure of a differentiable manifold by products of coordinate charts.*

One has dim $M \times N = n + k$. The projections $\mathrm{pr}_1, \mathrm{pr}_2$

$$\begin{array}{ccc} & M \times N & \\ \swarrow & & \searrow \\ M & & N \end{array}$$

are surjective submersions: in local coordinates, the projections are the projections $\mathbb{R}^{n+k} \to \mathbb{R}^n$ and $\mathbb{R}^{n+k} \to \mathbb{R}^k$.

We next discuss the case of a manifold that is only locally a product. To this end, we single out one surjective submersion:

Definition A.3.2

(i) A surjective submersion $\pi : P \to M$ is called a differentiable fibration or fibre bundle with total space P, base space M and fibre N, if for each point $p \in M$ there exists a neighborhood U and a diffeomorphism

$$x : \pi^{-1}(U) \tilde{\to} U \times N$$

such that the diagram

$$\begin{array}{ccc} \pi^{-1}(U) & \xrightarrow{\sim} & U \times N \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle p_1} \\ U & = & U \end{array}$$

commutes. The diffeomorphism x is called a local trivialization or bundle chart. The submanifold $\pi^{-1}(q) \subset P$ is called the fibre over $q \in M$. The manifold M is called the base of the fibre bundle and the manifold P the total space of the fibre bundle.

(ii) A bundle is called trivial, if there exists global bundle chart $P \tilde{\to} M \times N$.

(iii) Let $P \overset{\pi}{\to} M$ and $P' \overset{\pi}{\to} M'$ be fibre bundles. A morphism of fibre bundles is a diffeomorphism $f : P \to P'$ such that $\pi(p_1) = \pi(p_2)$ implies $\pi'(f(p_1)) = \pi'(f(p_2))$. Put differently, there is a smooth map $\overline{f} : M \to M'$ such that the diagram

$$\begin{array}{ccc} P & \xrightarrow{f} & P' \\ \downarrow{\scriptstyle\pi} & & \downarrow{\scriptstyle\pi'} \\ M & \xrightarrow[\overline{f}]{} & M' \end{array}$$

commutes. A trivial bundle is then isomorphic to the Cartesian product.

(iv) Let U be an open subset of M. A local section on U of a bundle $\pi : P \to M$ is a differentiable mapping

$$s : U \to P$$

such that $\pi \circ s = \mathrm{id}_U$. The sets $\mathcal{F}(U)$ of sections on all open subsets $U \subset M$ forms a sheaf. The elements of $\mathcal{F}(M)$ are called global sections.

The following lemma shows why surjective submersions are particularly important:

Lemma A.3.3.
A surjection $\pi : M \to N$ of smooth manifolds has local smooth *sections, iff π is a surjective submersion.*

Proof:
If π is a surjective submersion, the existence of local sections is a consequence of the theorem of implicitly defined functions. If π admits a smooth section, the surjectivity of $\mathrm{d}\pi$ follows by differentiating the identity $\pi \circ s = \mathrm{id}_U$ to find $\mathrm{d}\pi \circ \mathrm{d}s = \mathrm{id}$. □

Remark A.3.4.
If we describe a mechanical system of N mass points moving in a Galilei space $\mathbb{A}$ in terms of world lines parametrized by an eigentime in an interval I, we start with the fibre bundle $\pi : I \times \mathbb{A}^n \to I$ given by projection on the first factor.

Trajectories are then sections of π. If we impose constraints – e.g. to describe a solid body, we impose that the spacial distances between the mass points are constant in time –, we restrict ourselves to a submanifold $\tilde{M} \subset I \times \mathbb{A}^n$ which still provides a fibre bundle over I. If the constraints depend on time, we still obtain a submanifold $\tilde{M}$ of $I \times \mathbb{A}^n$, together with a surjective submersion to I. The manifold $\tilde{M}$ is then called the (extended) configuration space of the system.

We finally need manifolds with additional structure.

Definition A.3.5
An n-dimensional smooth manifold G with a group structure such that the difference map

$$\begin{aligned} \tilde{\mu} :\ & G \times G \to G \\ & (g, h) \mapsto g h^{-1} \ , \end{aligned}$$

is smooth, is called a Lie group.

Remarks A.3.6.

1. *Note that this implies that both the inverse $g \mapsto g^{-1}$ and the multiplication $(g_1, g_2) \mapsto g_1 \cdot g_2$ are smooth maps.*

2. *Examples:*

   - *The general linear group $\mathrm{GL}(n, \mathbb{R})$ is an n^2-dimensional Lie group. It is non-abelian for $n > 1$ and non-compact.*
   - *The specialorthogonal group $\mathrm{SO}(n)$, defined as the structure group of a Euclidian vector space of dimension n is an $\frac{n(n-1)}{2}$-dimensional Lie group. It is non-abelian for $n > 2$ and compact for all n.*
   - *Let V be a finite-dimensional real vector space with a non-degenerate symmetric bilinear form $\langle \cdot, \cdot \rangle$ of signature (p, q). The group $\mathrm{O}(p, q)$ of linear transformations preserving this bilinear form,*

     $$\mathrm{O}(p, q) := \{\varphi \in \mathrm{GL}(V) \,|\, \langle \varphi v, \varphi w \rangle = \langle v, w \rangle \quad \text{for all } v, w \in V\}$$

     is an $\frac{n(n-1)}{2}$-dimensional Lie group. It is non-abelian for $n > 2$ and non-compact as soon as $pq \neq 0$. The case $(p, q) = (n-1, 1)$ is of particular importance for relativity. The corresponding non-compact Lie group is called the n-dimensional Lorentz group.

- *The unitary group* $\mathrm{U}(n)$*, defined as the structure group of a unitary vector space of dimension* n *is an* n^2*-dimensional compact* real *Lie group. It is non-abelian for* $n > 2$*.*
- *The unitary group* $\mathrm{SU}(n)$*, defined as the structure group of a unitary vector space of dimension* n *consisting of maps of determinant* 1 *is an* $n^2 - 1$*-dimensional compact real Lie group. It is non-abelian for* $n > 2$*.*
- *The Galilei group is a ten-dimensional non-compact Lie group.*

Observation A.3.7.
Let M *be a differentiable manifold of dimension* n*.*

- *The set* $\bigcup_p T_pM$ *of all tangent vectors can be endowed with the structure of a differentiable manifold as follows. For a local coordinate chart defined on* $U \subset M$

$$x \ : U \to \mathbb{R}^n$$

we consider the local bijection, called the associated bundle chart,

$$\bar{x} : \bigcup_{p \in U} T_pM \to \mathbb{R}^n \times \mathbb{R}^n$$

defined by

$$\bar{x}(v) = \big(x(p), v(x^1), ..., v(x^n)\big) \quad \text{for } v \in T_pM$$

One deduces from Lemma A.2.4 the following identity for a change of local coordinates $\overline{x}, \overline{y}$*:*

$$\bar{x} \circ \bar{y}^{-1}(a, b) = \left(x \circ y^{-1}(a), \sum_{i=1}^{n} \frac{\partial y^i}{\partial x^i} b^i, ... \right),$$

so that we have differentiable maps which endow

$$TM := \bigcup_{p \in M} T_pM$$

with the structure of a smooth manifold of dimension $2n$*. It should be noted that even if the manifold* M *is compact, the total space of the manifold* TM *is non-compact. (To see this, consider one subspace* T_pM *and a find a sequence in this real vector space without accumulation point.)*

- *In local coordinates* $\bar{x}$ *and* x*, the map*

$$\pi : \quad TM \to M \quad \text{with} \quad \pi(v) = p \ \text{ for } \ v \in T_pM$$

is the projection on the first n *components and thus a surjective submersion. This way,* TM *becomes a bundle over* M *with fibre an* n*-dimensional vector space. This bundle is called the* tangent bundle *of the smooth manifold* M*.*

- *Any smooth map* $f : M \to N$ *gives rise to a map*

$$\begin{array}{rcl} Tf = Df : TM & \to & TN \\ (v, p) & \mapsto & f_{*p}(v) \in T_{f(p)}N \end{array}$$

This map is differentiable and we have a commuting diagram

$$\begin{array}{ccc} TM & \xrightarrow{Tf} & TN \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\pi} \\ M & \xrightarrow[f]{} & N \end{array}$$

(This is, of course, just the statement that for $v \in T_mM$, we have $Tf(v) \in T_{f(m)}N$.) One says that Tf covers f. Thus T is a (covariant) functor in the category of (smooth) manifolds.

- *An important aspect of bundles is the fact that they can be pulled back : given a smooth map of manifolds $f : M \to N$ and a smooth bundle $\pi : E \to N$, one defines*

$$f^*E := M \times_N E \cong \{(m, e) | m \in M, e \in E \text{ such that } f(m) = \pi(e)\}$$

The bundle projection $f^*E \to M$ is induced from the projection on the first factor. In other words, the fibre of f^*E in the point $m \in M$ is the fibre of E in the point $f(m) \in N$. We then have a map such that the following diagram commutes:

$$\begin{array}{ccc} M \times_N E & \xrightarrow{f_*} & E \\ \downarrow & & \downarrow{\scriptstyle\pi} \\ M & \xrightarrow{f} & N \end{array}$$

A.4 Vector fields and Lie algebras

Definition A.4.1
Let M be a differentiable manifold. A differentiable vector field on M is a global section X of the tangent bundle. A local vector field on an open subset $U \subset M$ is a local section of the tangent bundle.

In other words, given a vector field X, we have for every point $p \in M$ a vector $X_p \in T_pM$ such that for all smooth functions $f \in C^\infty(M)$ the function

$$\begin{array}{rccc} X(f): & M & \to & \mathbb{R} \\ & p & \mapsto & X_p(f) \end{array}$$

is smooth.

Remarks A.4.2.
Let M be an n-dimensional smooth manifold.

(a) The set $\mathrm{vect}(M)$ *of all vector fields is an infinite-dimensional real vector space. One can define a scalar multiplication of smooth real-valued functions and vector fields. Thus* $\mathrm{vect}(M)$ *is a module over the algebra $C^\infty(M, \mathbb{R})$ of smooth functions.*

(b) For a local coordinate chart x with domain $U \subseteq M$, we define n local vector fields on U:

$$\frac{\partial}{\partial x^i}(p) = \frac{\partial}{\partial x^i}\Big|_p \quad \text{with} \quad i = 1, 2, \ldots n$$

They form a basis of the $C^\infty(U)$-module $\mathrm{vect}(U)$*. Local vector fields thus form a free $C^\infty(U)$-module; one says that vector fields form a locally free $C^\infty(M)$-module. (Globally, this statement is not necessarily true: the tangent bundle is only trivial, if the manifold is parallelizable which holds e.g. for Lie groups.)*

(c) For the action of vector fields on smooth functions $f, g \in C^\infty(M)$, the Leibniz rule holds,

$$X(fg) = X(f)g + fX(g).$$

Vector fields are thus globally defined $\mathbb{R}$-linear first order differential operators on the space $C^\infty(M)$ of smooth functions.

Definition A.4.3
A Lie algebra over a commutative ring R is an R-module L together with an antisymmetric R-bilinear mapping, called the Lie bracket or commutator

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

such that the Jacobi identity

$$[x,[y,z]] + [z,[x,y]] + [y,[z,x]] = 0$$

holds for all $x, y, z \in L$.

Vector fields over any smooth manifold M have the structure of an (infinite-dimensional) real Lie algebra:

Observation A.4.4.
(a) Given two vector fields $X, Y \in \mathrm{Vect}(M)$, we consider the following map on germs of functions:

$$\begin{aligned} X_pY :&\quad \mathcal{F}_pM \to \mathbb{R} \\ \textit{with}&\quad (X_pY)(f) := X_p(Y(f)) \end{aligned}$$

Because of the identity

$$\begin{aligned} X_pY(fg) &= X_p(Y(f)g + fY(g)) \\ &= X_pY(f)g + Y_p(f)X_p(g) + X_p(f)Y_p(g) + fX_pY(g) \end{aligned}$$

the map X_pY is not a derivation on germs of smooth functions and thus does not give rise to a vector field. However, the map

$$[X,Y]_p := X_pY - Y_pX : \mathcal{F}_pM \to \mathbb{R}$$

is a derivation and thus provides a vector field. One checks that this way $\mathrm{Vect}(M)$ *becomes a Lie algebra over $\mathbb{R}$. One should be aware that one does* not *obtain a Lie algebra over the ring $C^\infty(M)$, because we find for any germ φ of a function:*

$$\begin{aligned} [X, fY]\,\varphi &= X(fY\varphi) - fY(X\varphi) \\ &= X(f)Y(\varphi) + fXY(\varphi) - fYX(\varphi) \\ &= f\,[X,Y]\,\varphi + X(f)Y(\varphi) \\ \textit{hence}\quad [X, fY] &= f\,[X,Y] + X(f)Y\ . \end{aligned}$$

Similarly, we find $[fX, Y] = f\,[X,Y] - Y(f)X$.

(b) Given any local coordinate chart, the local basis fields $\frac{\partial}{\partial x^i}$ commute, since the action of partial derivatives on smooth functions commutes.

(c) In local coordinates, we write two vector fields X and Y as

$$X = \sum_{i=1}^{n} \xi^i \frac{\partial}{\partial x^i}$$
$$Y = \sum_{i=1}^{n} \eta^i \frac{\partial}{\partial x^i} \ .$$

We recall that $\xi^i = X(x^i)$ is given by the action of the vector field X on the i-th coordinate function x^i. We compute for the Lie bracket, using observation A.4.4 (a) and (b):

$$\begin{aligned}
[X,Y] &= \textstyle\sum_{i,j} \left[\xi^i \frac{\partial}{\partial x^i}, \eta^j \frac{\partial}{\partial x^j}\right] \\
&= \xi^i \tfrac{\partial}{\partial x^i}(\eta^j) \tfrac{\partial}{\partial x^j} + \eta^j [\xi^i \tfrac{\partial}{\partial x^i}, \tfrac{\partial}{\partial x^j}] \\
&= \xi^i \tfrac{\partial \eta^j}{\partial x^i} \tfrac{\partial}{\partial x^j} + \eta^j \xi^i [\tfrac{\partial}{\partial x^i}, \tfrac{\partial}{\partial x^j}] - \eta^j \tfrac{\partial \xi^i}{\partial x^j} \tfrac{\partial}{\partial x^i} \\
&= \textstyle\sum_{k,l} \left(\eta^l \frac{\partial \xi^k}{\partial x^l} - \xi^l \frac{\partial \eta^k}{\partial x^l}\right) \frac{\partial}{\partial x^k}
\end{aligned}$$

If the manifold carries the additional structure of a Lie group G, we can single out a subclass of vector fields:

Definition A.4.5
Let G be a smooth Lie group.

1. *For any element $g \in G$, we call the smooth map*

$$\begin{aligned} L_g : G &\to G \\ h &\mapsto gh \end{aligned}$$

defines a left action of G on itself, the <u>left translation</u> by g. This defines a smooth left action of G on itself. Similarly for any $g \in G$, the map $R_g : h \mapsto h \cdot g$ defines a right action, right translation.

2. *A global vector field $V \in \Gamma(TG)$ is called <u>left invariant</u>, if $(L_g)_* V = V$ for all $g \in G$. Right invariant vector fields are defined analogously.*

3. *One can verify that left invariant and right invariant vector fields form a Lie subalgebra of the Lie algebra of all vector fields, called the <u>Lie algebra</u> $\mathrm{Lie}(G)$ of the Lie group G. In particular, the tangent bundle of a Lie group is parallelizable.*

   Left translation acts transitively and freely on the Lie group G. Hence, a left invariant vector field is determined by its value in the neutral element $e \in G$ in T_eG As a vector space, the Lie algebra $\mathrm{Lie}(G)$ can be identified with the tangent space T_eG. In particular, it is finite-dimensional of dimension $\dim G$.

Remark A.4.6.

1. *Given a vector field V on M, we call a trajectory $\varphi: I \to M$ such that*

$$\frac{\mathrm{d}\varphi}{\mathrm{d}t}_{t=t_0} = V(\varphi(t_0))$$

an integral curve of the vector field V. In local coordinates $x: M \supset U \to \mathbb{R}^n$, the vector field reads

$$V = \sum_{i=1}^{n} V^i \frac{\partial}{\partial x^i}$$

and the trajectory can be described by a function with values in $\mathbb{R}^n$:

$$\xi = x \circ \varphi : \mathbb{R} \to \mathbb{R}^n$$

Then the condition that the trajectory is an integral curve amounts to the system of first order ordinary differential equations

$$\frac{\mathrm{d}\xi^i}{\mathrm{d}t}(t) = V^i(\xi(t))$$

This justifies the point of view that integral curves of vector fields are a geometric expression for solutions of a system of ordinary differential equations of first order.

2. *Standard theorems assert the existence of local solutions of ordinary differential equations. As a consequence, for each point $q \in M$, there is an open neighborhood U and $\epsilon > 0$ such that the vector field X defines a family $\phi_t : U \to M$ of diffeomorphisms for all $|t| < \epsilon$, obtained by taking each point $p \in U$ a parameter distance t along the integral curves of X. In fact, the ϕ_t form a local one-parameter local group of diffeomorphisms, since $\phi_{t+s} = \phi_t \circ \phi_s = \phi_s \circ \phi_t$ for $|t| < \epsilon$, $|s| < \epsilon$ and $|s + t| < \epsilon$.*

A.5 Differential forms and the de Rham complex

Definition A.5.1
Let k be a field of characteristic different from two and V be a k-vector space.

(i) A multilinear map $f : V^p = V \times \ldots \times V \to k$ is called alternating, if $f(v_1, ..., v_p) = 0$, as soon as $v_i = v_j$ for a pair $i \neq j$.

(ii) The vector space of alternating p-forms on V will be denoted by $\Lambda^p V$. We have $\Lambda^1 V = V^*$, the dual vector space, and we set $\Lambda^0 V := k$.

Lemma A.5.2.
(i) Given p one-forms $\eta^1, ..., \eta^p \in \Lambda^1 V = V^*$ we consider the multilinear map

$$\begin{array}{rccl} \eta^1 \wedge ... \wedge \eta^p : & V^{\times p} & \to & k \\ & (v_1, ..., v_p) & \mapsto & \det(\eta^i(v_j)) \end{array}$$

This defines an alternating p-form, $\eta^1 \wedge \ldots \wedge \eta^p \in \Lambda^p V$.

(ii) If $\{b_1, \dots b_n\}$ *is a basis of* V *and* $\{b^1, \dots b^n\}$ *the corresponding dual basis of* V^**, then*

$$\left\{ b^{i_1} \wedge ... \wedge b^{i_p} \,\middle|\, 1 \leqslant i_1 < i_2 < ... < i_p \leqslant n \right\}$$

is a basis of $\Lambda^p V$*.*

(iii) For a finite-dimensional vector space V *of dimension* n*, the dimension of the space of* p *forms is* $\dim \Lambda^p V = \binom{n}{p}$*. In particular, there are no* p*-forms for* $p > n$*.*

Definition A.5.3
On the graded vector space

$$\Lambda^\bullet V := \bigoplus_{p=0}^{n} \Lambda^p V$$

we define the product on the basis vectors:

$$\begin{aligned}&(b^{i_1} \wedge \ldots \wedge b^{i_k}) \wedge (b^{j_1} \wedge \ldots \wedge b^{j_l}) \\ &:= b^{i_1} \wedge \ldots \wedge b^{i_k} \wedge b^{j_1} \wedge \ldots \wedge b^{j_l}\end{aligned}$$

This product is called exterior product, indexexterior product of differential forms wedge product or Grassmann product.

Remarks A.5.4.
(a) The wedge product is associative.

(b) The wedge product is graded commutative: for $\omega \in \Lambda^i V$ *and* $\omega' \in \Lambda^j V$*, we have*

$$\omega \wedge \omega' = (-1)^{ij} \omega' \wedge \omega$$

(c) The algebra structure induced by the wedge product is independent of the choice of basis $\{b_i\}$ *of* V*. More invariantly, we define the exterior algebra* $\Lambda^\bullet(V)$ *as the quotient of the tensor algebra* $T^\bullet V^* = \oplus_{p=0}^{\infty} (V^*)^{\otimes n}$ *of the dual vector space modulo a two-sided ideal*

$$\Lambda^\bullet(V) = T^\bullet V^* / \langle \omega \otimes \omega' + \omega' \otimes \omega \rangle$$

generated by all pairs $\omega, \omega' \in V^*$*.*

(d) Alternating forms can be pulled back along a linear map $f : V \to W$*: we define a linear map*

$$\Lambda^p f : \Lambda^p W \to \Lambda^p V$$

by

$$(\Lambda^p f(\omega))\,(v_1, ..., v_p) = \omega(f v_1, ..., f v_p)$$

for $v_i \in V$ *and* $\omega \in \Lambda^p W$*. For* $p = 1$*, this is just the dual map* $f^* : W^* \to V^*$*.*

For $p = n$*, we find* $\dim_k \Lambda^n V = \begin{pmatrix} n \\ n \end{pmatrix} = 1$*. For an endomorphism* $f : V \to V$ *and* $\Lambda^n f$ *is an endomorphism of a one-dimensional vector space which we canonically identify with a scalar. From the axiomatic definition of the determinant as a normalized alternating multilinear form on square matrices, one easily derives that the scalar* $\Lambda^n f$ *equals the determinant.*

(e) *Given another linear map $g : W \to U$, we find*

$$\Lambda^p(g \circ f) = \Lambda^p(f) \circ \Lambda^p(g).$$

We also use the short hand notation f^* instead of $\Lambda^p(f)$. This way, we have for each p a contravariant functor

$$\Lambda^p : \mathrm{vect}(k) \to \mathrm{vect}(k) \ .$$

We now apply this construction of linear algebra to the tangent bundle fibrewise:

Definition A.5.5

(i) *Let M be a smooth n-dimensional manifold. The set*

$$\Lambda^p TM := \bigcup_{x \in M} \Lambda^p T_x M$$

can be endowed with the structure of a smooth manifold as in observation A.3.7. The evident projection to M turns this into a fibre bundle with fibre a vector space of dimension $\binom{\dim M}{p}$.

(ii) *In particular, for $p = n$, we get a line bundle, the <u>determinant line bundle</u>.*

(iii) *The local sections of the bundle $\Lambda^p T^*M$ on an open subset $U \subset M$ are called (local) <u>differential forms</u> $\Omega^p(U)$. They form a $C^\infty(U)$-module. The collection $\Omega^p(U)$ for all open subsets $U \subset M$ forms again a sheaf. The elements of $\Omega^p(M)$ are called global differential forms.*

(iv) *The exterior product of differential forms is defined by applying the wedge product fiberwise and endows for every open subset $U \subset M$ the vector space $\Omega^\bullet(U) := \oplus_{p=1} \Omega^p(U)$ with the structure of an infinite-dimensional graded commutative algebra*

$$\wedge : \ \ \Omega^k(U) \times \Omega^l(U) \to \Omega^{k+l}(U) \ .$$

We consider the situation also in local coordinates on the differentiable manifold M:

Remarks A.5.6.

(a) *Let $x : M \supset U \to \mathbb{R}^n$ be a coordinate chart of a smooth manifold M of dimension n. We have seen that $\left\{ \frac{\partial}{\partial x^i}\Big|_p \right\}_{i=1,\ldots,n}$ is a basis of the tangent space T_pM. We denote the dual basis of the cotangent space T_p^*M by $\left\{ \mathrm{d}x^i\big|_p \right\}_{i=1\ldots n}$.*

Recall that $\frac{\partial}{\partial x^i}$ is a local section in the tangent bundle TM, i.e. a local vector field. Similarly, $\mathrm{d}x^i$ is a local section in the cotangent bundle T^*M, i.e. a local one-form.

(b) *Correspondingly, local p-forms can be written as*

$$\omega(x) = \sum_{1 \leqslant i_1 < \ldots < i_p \leqslant n} \omega_{i_1 \ldots i_p}(x) \mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_p}$$

with smooth coefficient functions $\omega_{i_1 \ldots i_k}(x)$ on $U \subset M$. This implies that a local p-form defined on the domain of definition of the coordinate chart x is determined by its values on the local vector fields $\frac{\partial}{\partial x^i}$. The $C^\infty(U)$-module $\Omega^p(U)$ is thus locally free.

(c) Under change of local coordinates, we deduce from

$$\frac{\partial}{\partial x^i} = \frac{\partial y^j}{\partial x^i}\frac{\partial}{\partial y^j}$$

that

$$\mathrm{d}x^i = \sum_j \frac{\partial x^i}{\partial y^j}\mathrm{d}y^j$$

As a consequence, we obtain for the special case of an n-form on an n-dimensional manifold

$$\mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^n = \det\left(\frac{\partial x^i}{\partial y^j}\right)\mathrm{d}y^1 \wedge \ldots \wedge \mathrm{d}y^n$$

which justifies the name determinant line bundle for the highest non-vanishing exterior power $\Lambda^n TM$ of the cotangent bundle.

In several contexts, it is helpful to have antisymmetrized expressions in derivatives: for example, the curl of a vector field on $\mathbb{R}^3$ is given by $(\mathrm{rot}v)_i := \epsilon_{ijk}\partial_j v_k$ (with ϵ totally antisymmetric in the indices) and the electromagnetic field strength is given in terms of the vector potential as $F_{\mu\nu} := \partial_\nu A_\mu - \partial_\mu A_\nu$. This leads us to the following

Definition A.5.7
For any open subset U of a smooth manifold M, we define a map

$$\mathrm{d}: \quad \Omega^p(U) \to \Omega^{p+1}(U)$$

by defining $\mathrm{d}\omega$ for $\omega \in \Omega^p(U)$ on a zero-form, i.e. a smooth function f as

$$\langle \mathrm{d}f, X\rangle = X(f) \qquad \text{for all vector fields } X$$

and acting on the p-form

$$\omega = \omega_{i_1\ldots i_p}\mathrm{d}x^{i_1} \wedge \ldots \mathrm{d}x^{i_p}$$

as

$$\mathrm{d}\omega = \partial_{i_0}\omega_{i_1\ldots i_p}\mathrm{d}x^{i_0} \wedge \mathrm{d}x^{i_1} \wedge \ldots \mathrm{d}x^{i_p} = \mathrm{d}\omega_{i_1\ldots i_p} \wedge \mathrm{d}x^{i_0} \wedge \mathrm{d}x^{i_1} \wedge \ldots \mathrm{d}x^{i_p}\ .$$

Remarks A.5.8.

1. *In the special case of smooth functions, we find $\mathrm{d}f = \sum_{i=1}^n \partial_i f \mathrm{d}x^i$ which is the total differential of classical calculus.*

2. *One checks that this definition does not depend on the choice of local coordinates. Indeed, suppose (y^i) is another set of local coordinates and*

$$\omega = \omega_{i'_1\ldots i'_p}\mathrm{d}y^{i'_1} \wedge \ldots \mathrm{d}y^{i'_p}\ ,$$

then we have

$$\omega_{i'_1\ldots i'_p} = \frac{\partial x^{i_1}}{\partial y^{i'_1}} \cdot \frac{\partial x^{i_2}}{\partial y^{i'_2}} \cdots \frac{\partial x^{i_p}}{\partial y^{i'_p}}\omega_{i_1\ldots i_p}$$

and thus in these coordinates

$$\begin{aligned}
\mathrm{d}\omega &= \mathrm{d}\omega_{i'_1\dots i'_p}\mathrm{d}y^{i'_1}\wedge\dots\mathrm{d}y^{i'_p}\\
&= \mathrm{d}\left(\frac{\partial x^{i_1}}{\partial y^{i'_1}}\cdot\frac{\partial x^{i_2}}{\partial y^{i'_2}}\cdots\frac{\partial x^{i_p}}{\partial y^{i'_p}}\omega_{i_1\dots i_p}\right)\mathrm{d}y^{i'_1}\wedge\dots\mathrm{d}y^{i'_p}\\
&= \frac{\partial x^{i_1}}{\partial y^{i'_1}}\cdot\frac{\partial x^{i_2}}{\partial y^{i'_2}}\cdots\frac{\partial x^{i_p}}{\partial y^{i'_p}}\mathrm{d}\omega_{i_1\dots i_p}\mathrm{d}y^{i'_1}\wedge\dots\mathrm{d}y^{i'_p}\\
&\quad+\frac{\partial^2 x^{i_1}}{\partial y^{i'_1}\partial y^{j'}}\cdot\frac{\partial x^{i_2}}{\partial y^{i'_2}}\cdots\frac{\partial x^{i_p}}{\partial y^{i'_p}}\omega_{i_1\dots i_p}\mathrm{d}y^{j'}\wedge\mathrm{d}y^{i'_1}\wedge\dots\mathrm{d}y^{i'_p}\\
&= \mathrm{d}\omega_{i_1\dots i_p}\wedge\mathrm{d}x^{i_0}\wedge\mathrm{d}x^{i_1}\wedge\dots\mathrm{d}x^{i_p}\ .
\end{aligned}$$

3. *On $p+1$ (local) vector fields $v_1,...,v_{p+1}$, the exterior derivative $\mathrm{d}\omega$ acts as*

$$\begin{aligned}
&\mathrm{d}\,\omega(v_1,...,v_{p+1}) =\\
&=\sum_{i=1}^{p+1}(-1)^i v_i\omega(v_1,...,\hat{v}_i,...,v_{p+1}) +\\
&+\sum_{i<j}(-1)^{i+j}\omega\big(\,[v_i,v_j]\,,v_1,...,\hat{v}_i,\hat{v}_j,...,v_{p+1}\big)
\end{aligned}$$

Here the hat denotes expressions to be left out.

It is remarkable that, due to their antisymmetry, p-forms can be differentiated without any additional data like a connection, just using the manifold structure.

We list properties of the exterior derivative:

Remarks A.5.9.

(a) The map d *is* $\mathbb{R}$*-linear.*

(b) If $\omega\in\Omega^p(M)$*, then we have for all differential forms* η

$$\mathrm{d}(\omega\wedge\eta)=\mathrm{d}\omega\wedge\eta+(-1)^p\omega\wedge\mathrm{d}\eta\ .$$

The exterior derivative is thus a derivation of degree $+1$*.*

(c) For any smooth map $f:M\to N$*, we have as a consequence of the chain rule:*

$$\mathrm{d}f^*\omega=f^*\mathrm{d}\omega.$$

(d) The antisymmetry of forms, combined with the symmetry of second derivatives implies $\mathrm{d}^2\omega=0$ *for all* $\omega\in\Omega^\bullet M$*.*

There is a second notion of a derivative that also uses the manifold structure and does not need any additional data like a connection, the Lie derivative. Morally, it generalizes partial derivatives by allowing instead of directions given by coordinate axes directions given by arbitrary smooth vector fields.

Definition A.5.10

Let M be a smooth manifold and X be a smooth vector field on M. Recall from Remark

A.2.8 the local one-parameter group ϕ_t of diffeomorphisms given by X. Since these maps are invertible, we can push for $q := \phi_{-t}(p)$ vector fields (or even more general tensor fields)

$$\Phi_t := \phi_{t*} : T_qM \to T_pM$$

and pull back differential forms

$$\Phi_t := \phi_{-t}^* : \Omega^k(M)_q \to \Omega^k(M)_p$$

The Lie derivative with respect to the vector field X is defined for $\mathcal{O}$ being either a differential form or a vector field as

$$L_X\mathcal{O}|_p := \lim_{t\to 0} \frac{1}{t}(\mathcal{O}_p - (\Phi_t\mathcal{O})|_p) \,.$$

Remarks A.5.11.
From the properties of Φ_t, it follows:

1. *The Lie derivative L_XV of a vector field V is a vector field; the Lie derivative $L_X\omega$ of a k-form ω is a k-form. The Lie derivative has thus degree 0.*

2. *On smooth functions f, the Lie derivative with respect to the vector field X amounts to the action of the derivations, $L_Xf = X(f)$. On vector fields, this operation reduces to the Lie bracket of vector fields, $L_XY = [X, Y]$.*

3. *The Lie derivative L_X is linear and preserves operations from tensor calculus like contractions.*

4. *There is a Leibniz rule for the tensor product of tensor fields and forms:*

$$L_X(S \otimes T) = L_XS \otimes T + S \otimes L_XT \ .$$

5. *For a one-form $\omega = \omega_i \mathrm{d}x^i$, one finds for the Lie derivative the one-form $L_X\omega = (L_X\omega)_i \mathrm{d}x^i$ with coefficient functions*

$$(L_X\omega)_i = \frac{\partial \omega_i}{\partial x^j} X^j + \omega_j (\frac{\partial X^j}{\partial x^i}) \ .$$

   For a k-form, the last term is replaced by k terms of similar form with insertions of derivatives of the vector field X at any position.

6. *For any two vector fields X and Y, one finds*

$$L_X \circ L_Y \ - L_Y \circ L_X \ = \ L_{[X,Y]} \ .$$

   Applied to vector fields, this is just the Jacobi identity for the Lie bracket of vector fields.

7. *One has*

$$\mathrm{d}(L_X\omega) = L_X(\mathrm{d}\omega)$$

Observation A.5.12.
Let M be a smooth manifold and X be a smooth vector field on M.

1. *We define a linear map, called the* interior product *or* contraction

$$\iota_X : \Omega^{k+1}(M) \to \Omega^k(M)$$

by

$$(\iota_X\omega)(X_1, X_2, \ldots, X_k) = \omega(X, X_1, X_2, \ldots, X_k)$$

for any k-tuple of local vector fields $X_1, \ldots, X_k$.

2. *ι_X is $\mathbb{R}$-linear and an derivation of degree -1,*

$$\iota_X(\omega \wedge \eta) = \iota_X\omega \wedge \eta + (-1)^k \omega \wedge (\iota_X\eta)$$

for $\omega \in \Omega^k(M)$.

3. *One has*

$$\iota_X \circ \iota_Y \;=\; -\iota_Y \circ \iota_X$$

for all vector fields X, Y.

4. Cartan's magic formula *relates the Lie derivative to the interior product and the exterior differential* d*:*

$$L_X\omega = \iota_X(\mathrm{d}\omega) + \mathrm{d}(\iota_X\omega) = \; (\iota_X + \mathrm{d})^2\omega$$

and the relation

$$L_X \circ \iota_Y \; - \iota_Y L_X \;=\; \iota_{[X,Y]} \; .$$

5. *For any smooth function f, one has the relation*

$$L_{fX} = f L_X\omega + \mathrm{d}f \wedge \iota_X\omega \; .$$

The equality $\mathrm{d}^2 = 0$ for the exterior derivative of differential forms leads us to the following

Definition A.5.13
Let M be a smooth manifold of dimension n and $U \subset M$ open.

1. *A differential form $\omega \in \Omega^p(U)$ such that $\mathrm{d}\omega = 0$ is called* closed*.*

2. *A differential form $\omega \in \Omega^p(U)$ is called* exact *, if there exists a differential form $\eta \in \Omega^{p-1}(U)$ such that $\omega = \mathrm{d}\eta$.*

3. *Because of $\mathrm{d}^2\omega = 0$ for all ω, exact differential forms are closed. Hence we consider the quotient space. The real vector space for global differential forms*

$$\mathrm{H}^p_{dR}(M) := \ker\big(\mathrm{d} : \Omega^p M \to \Omega^{p+1} M\big) \Big/ \mathrm{im}\big(\mathrm{d} : \Omega^{p-1} M \to \Omega^p M\big)$$

is called the p-th de Rham cohomology *group of M. For convenience, we have adopted here the convention $\Omega^{n+1}(M) = 0, \;\; \Omega^{-1}(M) = 0$.*

4. *Given a smooth map $f : M \to N$, the pull back $f^* : \Omega^p(N) \to \Omega^p(M)$ of differential forms commutes with the exterior derivative and thus gives rise to a pull back*

$$f^* : \mathrm{H}^p_{dR}(N) \to \mathrm{H}^p_{dR}(M)$$

of de Rham cohomology groups. It should be emphasized that the error is reverted: this is why de Rham cohomology is called co-homology. *In fact, we have functoriality: for $M \xrightarrow{f} N \xrightarrow{g} P$, we get $(g \circ f)^* = f^* \circ g^*$ and we have $\mathrm{id}_M)^* = \mathrm{id}_{H^k_{dR}}(M)$.*

In particular, diffeomorphic smooth manifolds have isomorphic de Rham cohomology groups.

The field strength F in classical electrodynamics is a closed two-form, but not necessarily exact. In quantum systems, its representative $[F] \in H^2_{dR}(M)$ takes its values in a distinguished lattice in the vector space $H^2_{dR}(M)$. This is the mathematical statement of the quantization of electric flux.

Lemma A.5.14.

1. *For any smooth manifold M, the number of connected components of M equals $\dim_{\mathbb{R}} \mathrm{H}^0_{dR}(M)$.*

2. *We have $\mathrm{H}^k_{dR}(\mathbb{R}^n) = 0$ for all $k \geqslant 1$.*

3. *A smooth manifold M is said to be <u>contractible</u> to a point $p_0 \in M$, if there exists a smooth map*

$$\mathrm{H} : \quad M \times [0,1] \to M \ ,$$

called a (smooth) <u>homotopy</u>, such that

$$\mathrm{H}(p,0) = p \quad \textit{and} \quad \mathrm{H}(p,1) = p_0 \quad \textit{for all} \quad p \in M \ .$$

4. *<u>Poincaré lemma</u> asserts that any closed p-form ω with $p > 0$ on a contractible manifold M is exact.*

The Poincaré lemma encodes many analytic statements; e.g. for $p = 1$, it asserts that on any contractible subset $U \subset \mathbb{R}^n$, the only obstruction for the system of n partial differential equations

$$\frac{\partial f}{\partial x^i} = g_i(x) \quad \text{for } i = 1, \dots n$$

to have a solution f for given functions $g_i(x)$ is the fact that the equations

$$\frac{\partial g_j}{\partial x^i} = \frac{\partial g_i}{\partial x^j}$$

hold for all pairs i, j.

Example A.5.15.
Consider on the punctured plane $\mathbb{R}^2 \setminus \{0\}$ with standard Cartesian coordinates x, y the smooth one-form

$$\omega := \frac{x}{x^2+y^2}\mathrm{d}y - \frac{y}{x^2+y^2}\mathrm{d}x \quad \in \Omega^1(\mathbb{R}^2 \setminus \{0\}) \ .$$

A simple computation shows that ω is closed:

$$\mathrm{d}\omega = \frac{\partial}{\partial x}\left(\frac{x}{x^2+y^2}\right)\mathrm{d}x\wedge\mathrm{d}y - \frac{\partial}{\partial y}\left(\frac{y}{x^2+y^2}\right)\mathrm{d}y\wedge\mathrm{d}x = 0\ .$$

In radial coordinates (r,φ), using that

$$x = r\cos\varphi \quad \textit{and} \quad y = r\sin\varphi$$

implies

$$\mathrm{d}x = \mathrm{d}r\cos\varphi - r\sin\varphi\mathrm{d}\varphi \quad \textit{and} \quad \mathrm{d}y = \mathrm{d}r\sin\varphi + r\cos\varphi\mathrm{d}\varphi\ ,$$

we find

$$\begin{aligned}\omega &= \frac{1}{r^2}\left(r\cos\varphi(\mathrm{d}r\sin\varphi + r\cos\varphi\mathrm{d}\varphi) - r\sin\varphi(\cos\varphi\mathrm{d}r - r\sin\varphi\mathrm{d}\varphi)\right)\\ &= \mathrm{d}\varphi.\end{aligned}$$

The angular coordinate φ is, however, only defined on the complex plane minus a half axis. If ω would be the exterior derivative of a function $f\in C^\infty(\mathbb{R}^2\setminus\{0\})$, i.e. $\omega = \mathrm{d}f$, we would necessarily have $f = \varphi + const.$ on $\mathbb{R}^2\setminus\{half\ axis\}$. Such an f cannot exist on all of $\mathbb{R}^2\setminus\{0\}$; hence

$$\mathrm{H}^1_{dR}(\mathbb{R}^2\setminus\{0\}) \neq 0\ .$$

We conclude in particular that the punctured plane $\mathbb{R}^2\setminus\{0\}$ is not diffeomorphic to the plane $\mathbb{R}^2$.

We can also interpret this result in terms of differential equations: the system of differential equations:

$$\frac{\partial f}{\partial y}(x,y) = \frac{x}{x^2+y^2} \quad \textit{and} \quad \frac{\partial f}{\partial x}(x,y) = -\frac{y}{x^2+y^2}\mathrm{d}x$$

does not have any solution on $\mathbb{R}^2\setminus\{0\}$.

Definition A.5.16

1. *Let M be a differentiable manifold of dimension n. A nowhere vanishing n-form, i.e. $\omega(p)\neq 0$ for all $p\in M$, is called a <u>volume form</u> on M.*
2. *A manifold that admits a volume form is called <u>orientable</u>.*
3. *An <u>orientation</u> of M is an equivalence class of volume forms that differ by a positive function, i.e. $\omega'\sim\omega$, if and only if there is $f\in C^\infty(M,\mathbb{R}_{>0})$ such that $\omega' = f\omega$.*

Proposition A.5.17.
A smooth manifold M is orientable if and only if there exists an atlas $(x_i)_{i\in I}$ such that the determinant of the Jacobian of all diffeomorphisms $x_i\circ x_j^{-1}$ is positive.

Facts A.5.18.

(a) *An n-form can be integrated over any oriented n-dimensional manifold M with boundary: there is a linear map*

$$\int : \Omega^n(M) \to \overline{\mathbb{R}}$$
$$\omega \mapsto \int_M \omega$$

This is shown by first choosing a partition of unity on M that is subordinate to a coordinate system of M. One then defines local contributions to the integral in local coordinates. Since under the change of local coordinates both the n-form and the measure transform with powers of the determinant of the Jacobian, one can show that this yields a well-defined expression.

(b) *If $\phi : M \to N$ is a diffeomorphism of smooth manifolds of dimension n and if $\omega \in \Omega^n(N)$, then*

$$\int_M \phi^*\omega = \int_N \omega \ .$$

In the proof, again the fact enters that all quantities transform with powers of the Jacobian.

(c) *Let M be a smooth n-dimensional manifold with boundary and let $i : \partial M \to M$ be the embedding of the boundary. For any $(n-1)$-form $\omega \in \Omega^{n-1}(M)$ with compact support, Stokes' theorem asserts the following equality:*

$$\int_{\partial M} i^*\omega = \int_M \mathrm{d}\omega \ .$$

One should note that $\imath^*\omega$ is just the restriction of the $(n-1)$-form to the boundary. Stokes' theorem is thus frequently informally written as $\int_{\partial M} \omega = \int_M \mathrm{d}\omega$. Its proof is a generalization of the fact that the integral of a derivative is determined by its boundary values, i.e. of the fundamental theorem of calculus. As we will see, it contains the classical integral formulas of vector calculus.

For proofs and details, we refer to the books by Boot and Tu and by Madsen and Tornehave.

A.6 Riemannian manifolds and the Hodge dual

We start again with linear algebra: for any finite dimensional k-vector space V, the dual vector space V^* has the same dimension as V. The two vector spaces V and V^* are thus isomorphic, but there is no canonical isomorphism. However, a non-degenerate symmetric bilinear form

$$g : V \times V \to k$$

provides an isomorphism

$$\begin{array}{rccl} \iota_V^g : V & \tilde{\to} & V^* \\ v & \mapsto & (g(v,\cdot) : w \mapsto g(v,w)) \end{array}$$

To get a geometric notion derived from this notion of linear algebra, we endow the tangent spaces in a smooth way with non-degenerate symmetric bilinear forms:

Definition A.6.1

1. *A* Riemannian manifold *(M, g) is a smooth manifold M for which the tangent space T_pM for each point $p \in M$ has been endowed in a smooth way with the structure of a Euclidean vector space.*

   More explicitly, for each $p \in M$, there is a positive symmetric definite bilinear form

   $$g_p : \; T_pM \times T_pM \to \mathbb{R} \; ,$$

   called the metric *of M, such that for all local coordinate charts x of M the locally defined functions*

   $$g_{ij} : \; U \to \mathbb{R}$$
   $$g_{ij}(p) := g_p \left(\frac{\partial}{\partial x^i}\Big|_p , \frac{\partial}{\partial x^j}\Big|_p \right)$$

   are smooth for all $1 \leq i, j \leq \dim M$.

2. *More generally, we also consider a smooth family of non-degenerate symmetric bilinear forms*

   $$g_p : \; T_pM \times T_pM \to \mathbb{R}$$

   which can have any signature s, but whose signature is (by continuity) independent on $p \in M$ and obtain the notion of a pseudo-Riemannian manifold*. For the case of signature $(1, n-1)$, the manifold (M, g) is called a* Lorentzian manifold*.*

3. *For any open subset $U \subset M$ of a (pseudo-)Riemannian manifold, a canonical isomorphism of bundles between the tangent bundle TM and the cotangent bundle is obtained fibrewise:*

   $$\iota^g_{T_pM} : T_pM \; \to T^*_pM \quad .$$

   This gives a bijection between local smooth one-forms (which are just local smooth sections of the cotangent bundle) and local smooth vector fields (which are just local smooth sections of the tangent bundle):

   $$\iota^g_U : \mathrm{vect}(U) \xrightarrow{\sim} \Omega^1(U) \quad .$$

An important aspect of Riemannian manifolds is the fact that we can define lengths of curves on them (and also angles between intersecting curves). We can also define volumes:

Lemma A.6.2.
Let (M, g) be an oriented *(pseudo-)Riemannian manifold of dimension n. If $(\mathrm{d}x^1, \mathrm{d}x^2, \ldots \mathrm{d}x^n)$ is an oriented local basis of the cotangent bundle, then the n-forms*

$$\omega_g(U) := \sqrt{|\det g_{ij}|}\mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \ldots \wedge \mathrm{d}x^n \in \Omega^n(U)$$

patch together to a globally defined n-form $\omega_g \in \Omega^n(M)$ which is nowhere vanishing. It is called the normalized volume form *for the metric g.*

If V is a finite-dimensional vector space of dimension n, the two vector spaces $\Lambda^p V$ and $\Lambda^{n-p} V$ have the same dimension,

$$\dim \Lambda^p V = \begin{pmatrix} n \\ p \end{pmatrix} = \begin{pmatrix} n \\ n-p \end{pmatrix} = \dim \Lambda^{n-p} V \ ,$$

and are thus isomorphic. Again, without additional structure on the vector space, there is no canonical isomorphism.

Lemma A.6.3.

1. *Let V be an orientable finite-dimensional $\mathbb{R}$-vector space of dimension n. Assume further that there is a non-degenerate symmetric bilinear form g on V. Fix an orthonormal basis $(b_1, b_2, \ldots, b_n)$ on V. We extend the bilinear form g on V to a non-degenerate symmetric bilinear form g on $\Lambda^p V$ for all p by setting for $\omega, \omega' \in \Lambda^p\ V$*

$$\langle \omega, \omega' \rangle := \sum_{i_1, i_2, \ldots i_p} \omega(b_{i_1}, \ldots, b_{i_p}) \omega'(b_{i_1}, \ldots, b_{i_p}) g(b_{i_1}, b_{i_1}) \ldots g(b_{i_p}, b_{i_p})$$

2. *If the vector space is moreover oriented, there exists a unique isomorphism of vector spaces*

$$*_g : \quad \Lambda^p V \to \Lambda^{n-p} V$$

such that for all $\alpha, \beta \in \wedge^p V$ the equality

$$\alpha \wedge *_g \beta = g(\alpha, \beta) \omega_g$$

holds. The isomorphism $*$ is called the Hodge operator.

Remarks A.6.4.

1. *For V an oriented Euclidean vector space with oriented orthonormal basis $(e_1, \ldots e_n)$, the Hodge operator acts as*

$$*(e_1 \wedge e_2 \wedge \ldots \wedge e_p) = e_{p+1} \wedge e_{p+2} \wedge \ldots \wedge e_n \ .$$

2. *In particular, for $\mathbb{R}^3$ with the standard orientation and the standard Euclidean structure, one finds*

$$*\mathrm{d}x = \mathrm{d}y \wedge \mathrm{d}z \quad * \mathrm{d}y = \mathrm{d}z \wedge \mathrm{d}x \quad * \mathrm{d}z = \mathrm{d}x \wedge \mathrm{d}y$$

Indeed, the three equations on $*\mathrm{d}x$

$$\begin{array}{rcl} \mathrm{d}x \wedge (*\mathrm{d}x) & = & g(\mathrm{d}x, \mathrm{d}x) \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z = \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z \\ \mathrm{d}y \wedge (*\mathrm{d}x) & = & g(\mathrm{d}y, \mathrm{d}x) \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z = 0 \\ \mathrm{d}z \wedge (*\mathrm{d}x) & = & g(\mathrm{d}z, \mathrm{d}x) \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z = 0 \end{array}$$

have the unique solution $\mathrm{d}y \wedge \mathrm{d}z$.

3. *This allows us to formulate the cross product of vectors in $\mathbb{R}^3$, endowed with the standard structure as an oriented Euclidean vector space:*

$$\begin{array}{rccl} \wedge : & \mathbb{R}^3 \times \mathbb{R}^3 & \to & \mathbb{R}^3 \\ & (v, v') & \mapsto & \iota^{-1} * (\iota(v) \wedge \iota(v')) \end{array}$$

Indeed, we find

$$\frac{\partial}{\partial x} \wedge \frac{\partial}{\partial y} = \iota^{-1} * (\mathrm{d}x \wedge \mathrm{d}y) = \iota^{-1}(\mathrm{d}z) = \frac{\partial}{\partial z} \ .$$

The important lesson is that the cross product is really about 1-forms and produces in a natural way a 2-form, which can be identified with a 1-form just because we are working with a three-dimensional oriented vector space. In particular, angular moment is naturally a two-form. The way it is expressed as a vector changes, if one changes the orientation of $\mathbb{R}^3$*. This is meant by the expression that angular moment is an axial vector.*

4. *Another important example is* $\mathbb{R}^4$ *with coordinates* (t, x, y, z) *and a metric of signature* $(-, +, +, +)$*. We find, e.g. for the one-form* $\mathrm{d}t$ *the equations*

$$\begin{aligned} \mathrm{d}t \wedge (*\mathrm{d}t) &= g(\mathrm{d}t, \mathrm{d}t)\mathrm{d}t \wedge \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z = -\mathrm{d}t \wedge \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z \\ \mathrm{d}x \wedge (*\mathrm{d}t) &= g(\mathrm{d}x, \mathrm{d}t)\mathrm{d}t \wedge \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z = 0 \\ \mathrm{d}y \wedge (*\mathrm{d}t) &= g(\mathrm{d}y, \mathrm{d}t)\mathrm{d}t \wedge \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z = 0 \\ \mathrm{d}z \wedge (*\mathrm{d}t) &= g(\mathrm{d}z, \mathrm{d}t)\mathrm{d}t \wedge \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z = 0 \end{aligned}$$

have the unique solution $-\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z$*. Similarly, one finds*

$$*\mathrm{d}t = -\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z \quad *\mathrm{d}x = -\mathrm{d}t \wedge \mathrm{d}y \wedge \mathrm{d}z \quad *\mathrm{d}y = -\mathrm{d}t \wedge \mathrm{d}z \wedge \mathrm{d}x \quad *\mathrm{d}z = -\mathrm{d}t \wedge \mathrm{d}x \wedge \mathrm{d}y \ .$$

For 2-forms, we find from the only non-zero scalar product

$$(\mathrm{d}t \wedge \mathrm{d}x) \wedge *(\mathrm{d}t \wedge \mathrm{d}x) = \ -\mathrm{d}t \wedge \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z$$

the relation

$$*(\mathrm{d}t \wedge \mathrm{d}x) = -\mathrm{d}y \wedge \mathrm{d}z$$

and similarly

$$\begin{array}{lll} *(\mathrm{d}t \wedge \mathrm{d}x) = -\mathrm{d}y \wedge \mathrm{d}z & *(\mathrm{d}t \wedge \mathrm{d}y) = -\mathrm{d}z \wedge \mathrm{d}x & *(\mathrm{d}t \wedge \mathrm{d}z) = -\mathrm{d}x \wedge \mathrm{d}y \\ *(\mathrm{d}x \wedge \mathrm{d}y) = \mathrm{d}t \wedge \mathrm{d}z & *(\mathrm{d}y \wedge \mathrm{d}z) = \mathrm{d}t \wedge \mathrm{d}x & *(\mathrm{d}z \wedge \mathrm{d}x) = \mathrm{d}t \wedge \mathrm{d}y \end{array}$$

5. *If* g *is a non-degenerate bilinear symmetric form on a vector space* V*, then we have for the component functions*

$$(*\omega)_{i_1 i_2 \ldots i_{n-p}} = \frac{1}{p!} \omega^{j_1 \ldots j_p} \sqrt{|\det g|} \epsilon_{j_1 \ldots j_p i_1 \ldots i_{n-p}}$$

where the indices of ω *are raised and lowered with* g *and its inverse and where* ϵ *is totally antisymmetric in the indices with normalization* $\epsilon_{1,2,\ldots n} = 1$*.*

Proposition A.6.5.
Let V *be a pseudoeuclidean vector space of signature* s *and dimension* n*. Then the Hodge operator has the following properties:*

(i) $*1 = \omega_g, \quad *\omega_g = (-1)^s$

(ii) For $\alpha \in \Lambda^p V, \ \ \beta \in \Lambda^{n-p} V$*, we have*

$$g(\alpha, *\beta) = (-1)^{p(n-p)} g(*\alpha, \beta)$$

(iii) On $\Lambda^p V$, we find for the square of the Hodge operator

$$(*)^2 = (-1)^{p(n-p)+s}\,\mathrm{id}_{\Lambda^p V}$$

Observation A.6.6.
Let (M, g) be an oriented (pseudo-)Riemannian manifold.

1. *In this case, we use the canonical volume form ω_g on the manifold and extend the Hodge operator fibrewise to a smooth map*

   $$\Omega^p(U) \to \Omega^{n-p}(U)$$

   for any open subset $U \subset M$.

2. *In this way, we obtain an L^2-norm on a suitable subspace of $\Omega^p(M)$:*

   $$\langle \omega, \omega' \rangle := \int_M \omega \wedge *\omega' \qquad \text{for} \quad \omega, \omega' \in \Omega^p(M).$$

3. *In particular, we can define the adjoint*

   $$\delta = \delta_g : \Omega^p(U) \to \Omega^{p-1}(U)$$

   of the exterior derivative d*, called the* <u>*codifferential*</u> *by the equation*

   $$\langle \omega, \delta\omega' \rangle = \langle \mathrm{d}\omega, \omega' \rangle \ .$$

   The codifferential depends on the metric chosen on M, since the Hodge operator depends on the metric. On p-forms, it reads explicitly

   $$\delta = (-1)^p *^{-1} \mathrm{d} * \ .$$

   Indeed, for $\omega \in \Omega^{p-1}(M)$ and $\omega' \in \Omega^p(M)$ we deduce from the fact that the exterior differential d *is an odd derivation*

   $$\mathrm{d}(\omega \wedge *\omega') = \mathrm{d}\omega \wedge *\omega' \ + (-1)^{p-1}\omega \wedge \mathrm{d} * \omega'$$

   that

   $$\begin{aligned} (-1)^p \langle \omega, *^{-1}\mathrm{d} * \omega' \rangle &= (-1)^p \textstyle\int_M \omega \wedge *(*^{-1}\mathrm{d} * \omega' \rangle = (-1)^p \int_M \omega \wedge \mathrm{d} * \omega' \rangle \\ &= -\textstyle\int_M \mathrm{d}(\omega \wedge *\omega') + \int_M \mathrm{d}\omega \wedge *\omega' \\ &= \langle \mathrm{d}\omega, \omega' \rangle \ . \end{aligned}$$

   One checks

   $$\delta^2 = *^{-1} \circ \mathrm{d} \circ \mathrm{d} \circ * = 0 \ .$$

4. *The* <u>*Laplace operator*</u> *on differential forms is defined by*

   $$\Delta := (\delta + \mathrm{d})^2 = \delta \circ \mathrm{d} + \mathrm{d} \circ \delta \, .$$

Since the codifferential δ depends on the metric g, the Laplace operator depends on the metric g as well. It extends the usual Laplace operator on smoot functions. The Laplace operator on forms is symmetric

$$\langle \Delta\omega, \omega' \rangle = \langle \omega, \Delta\omega' \rangle$$

and positive definite

$$\langle \Delta\omega, \omega \rangle \geq 0\,.$$

The elements in its kernel are called harmonic forms. (Also the notion of a harmonic form depends on the choice of a metric.) The vector space of harmonic forms can be shown to be naturally isomorphic to de Rham cohomology. The Hodge operator induces an isomorphism of harmonic forms

$$* : H^p_\Delta(M) \to H^{n-p}_\Delta(M)$$

which implements Poincaré duality of de Rham cohomology.

We finally comment on the relation of the operations we just introduced to the classical operations of gradient, curl and divergence in three-dimensional vector calculus. To this end, we consider the special case of M being $\mathbb{R}^3$, endowed with the standard scalar product and the standard orientation.

Observation A.6.7.

1. *For a smooth function $f \in \Omega^0(U) = C^\infty(U, \mathbb{R})$, the derivative $\mathrm{d}f : TU \to T\mathbb{R} \cong \mathbb{R}$ is a one-form. The vector field corresponding to $\mathrm{d}f$ is called the gradient of the function $f \in C^\infty(U)$:*

$$\mathrm{grad}(f) := \iota^{-1}(\mathrm{d}f) \in \mathrm{vect}(U) \quad .$$

In contrast to the differential $\mathrm{d}f$, the gradient $\mathrm{grad}(f)$ depends on the metric.

2. *The exterior derivative of a one-form $\omega = A\mathrm{d}x + B\mathrm{d}y + C\mathrm{d}z \in \Omega^1(\mathbb{R}^3)$, is a 2-form*

$$\mathrm{d}\omega = (\frac{\partial C}{\partial y} - \frac{\partial B}{\partial z})\mathrm{d}y \wedge \mathrm{d}z + (\frac{\partial A}{\partial z} - \frac{\partial C}{\partial x})\mathrm{d}z \wedge \mathrm{d}x + (\frac{\partial B}{\partial x} - \frac{\partial A}{\partial y})\mathrm{d}x \wedge \mathrm{d}y$$

whose coefficient functions we recognize as the components of the curl.

We therefore define the curl operator for vector fields defined on open subsets U of $\mathbb{R}^3$ with the standard orientation as

$$\mathrm{curl} : \mathrm{vect}(U) \to \mathrm{vect}(U)$$

by

$$\mathrm{vect}(U) \xrightarrow[\iota]{} \Omega^1(U) \xrightarrow[\mathrm{d}]{} \Omega^2(U) \xrightarrow[*]{} \Omega^1(U) \xrightarrow[\iota^{-1}]{} \mathrm{vect}(U)$$

Reversing the orientation of $\mathbb{R}^3$ replaces the curl by its negative. In the slang of vector calculus, this is sometimes summarized by saying that the curl is an axial vector, rather than an ordinary polar vector.

3. *Finally, for a 1-form* $\omega = A\mathrm{d}x + B\mathrm{d}y + C\mathrm{d}z \in \Omega^1(U)$ *defined on an open subset* U *of* $\mathbb{R}^3$, *we find*

$$\begin{aligned} *\mathrm{d} * \omega &= *\mathrm{d}\left(A\mathrm{d}y \wedge \mathrm{d}z + B\mathrm{d}z \wedge \mathrm{d}x + C\mathrm{d}x \wedge \mathrm{d}y\right) \\ &= *\left(\tfrac{\partial A}{\partial x} + \tfrac{\partial B}{\partial y} + \tfrac{\partial C}{\partial z}\right)\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z = \tfrac{\partial A}{\partial x} + \tfrac{\partial B}{\partial y} + \tfrac{\partial C}{\partial z}\ . \end{aligned}$$

We therefore define the standard <u>*divergence operator*</u> *on vector fields defined on open subsets* $U \subset \mathbb{R}^3$ *as*

$$\mathrm{div}(v) : \mathrm{vect}(U) \to C^\infty(U, \mathbb{R})$$

by

$$\mathrm{vect}(U) \xrightarrow[\iota]{} \Omega^1(U) \xrightarrow[*]{} \Omega^2(U) \xrightarrow[\mathrm{d}]{} \Omega^3(U) \xrightarrow[*]{} C^\infty(U, \mathbb{R})$$

4. *The relation* $\mathrm{d}^2 = 0$ *now implies the two classical identities*

$$\mathrm{curl} \circ \mathrm{grad} f = 0 \quad \textit{and} \quad \mathrm{div} \circ \mathrm{curl}\ V = 0$$

for f *a smooth function and* V *a smooth vector field on* $U \subset \mathbb{R}^3$.

Index