Mathematical Structure of Analytic Mechanics

Thesis Submitted in partial Fulfillment for the Degree of M.SC. in Mathematics

BY:
Khalifa Abd ALmajeed Mohammed ALSadig

Supervisor:
Dr. Emad ELdeen Abdallah AbdelRahim

2016
Dedication

To my …

Father,

Mother,

Brothers,

Sisters,

Teachers,

And to lovely friends
Acknowledgements

Thanks first and last to Allah who enabled to conduct this by grace of him and denoted strength and patience.
Thanks to my supervisor Dr. EmadELdeenAbdallah Abdel Rahim for this advices and help.
Also thanks are due to all teachers in the Department of Mathematics.
Abstract

In this work, we try to set up a geometric setting for Lagrangian systems that allows to appreciate both theorems of Emmy Noether. We consistently use differential form and a geometric approach, in this research, we also discuss electrodynamics with gauge potentials as an instance of differential co-homology. Also we emphasize the role of observables with some examples and applications.
الخلاصة

لقد حاولنا في هذا العمل إقامة الوضع الهندسي لأنظمة لأجرانج، والذي يسمح بتقدير جميع مبرهنتان إيمي نوثر. وقد ضمنا في هذا العمل استخدام استمرارية التفاضل والمقارنة الهندسية، وناقشنا أيضاً الديناميكا الكهربائية مع قياس الجهد كحالة الهومولوجيا المصاحبة التفاضلية. وأكدنا أيضاً دور الملاحظات مع بعض الأمثلة والتطبيقات.
The contents

Subject ................................................................. Page
Dedication .......................................................... I
Acknowledgements .............................................. II
Abstract ........................................................... III
Abstract (Arabic) ............................................... IV
The contents ....................................................... V
Introduction ....................................................... VI

Chapter (1)

Newtonian Mechanics
Section (1-1) : Galilei space ....................................... 1
Section (1-2) : Dynamics of Newtonian systems ............ 9

Chapter (2)

Lagrangian mechanics
Section (2-1) : variational calculus and system with constraints ... 15
Section (2-2) : Lagrangian systems and Lagrangian Dynamics ...... 28
Section (2-3) : Symmetries, Noether identities and Natural Geometry. 38

Chapter (3)

Classical field theories
Section (3-1) : Maxwell’s equations and special relativity ........ 48
Section (3-2) : Electrodynamics as a gauge ...................... 61

Chapter (4)

Hamiltonian mechanics
Section (4-1) : (pre-)Symplectic manifolds .................... 69
Section (4-2) : Poisson manifolds and Hamiltonian systems .......... 77
Section (4-3) : Hamiltonian dynamics and Legendre transform ...... 85
References ......................................................... 94
**Introduction:**

It is true, of course, that physics chooses 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 be for by the mathematician.

In this research our goal is to discuss the appropriate mathematics to studying physics, roughly familiar with all classes of theoretical physics, so it is organized as follows:

Firstly, we discuss the general concepts of Newtonian mechanics, via our definition of affine space, with some remarks. Also we study the Gelilei space, and defined the equivalence class of a Galilean structure of Galilean coordinate, with some remarks, then we study the dynamics of Newtonian systems with some examples.

In chapter 2, we deal with Lagrangian mechanics, because of this, we study the basic concepts of the variational calculus and system with constraints with some observations. We illustrate the definition of Lagrangian systems and Lagrangian dynamics, the kinematical system, and the concept of a Global trajectories. We discuss the symmetries. In different descriptions, Noether identities and natural Geometry with some examples and applications.

In chapter 3, we give an explaining of Maxwell’s equations in terms of the exterior algebra. We study the Minkowski space with remarkable points. Also we discuss the electrodynamics, and investigate some of its comments on Minkowski space, and some applications in term of gauge theory.
Finally, we discuss some notions of the linear algebra, and the mathematical structures of the pre-symplectic manifolds. We study the Poisson manifolds, and Hamiltonain systems with some applications. Also, we describe the Hamiltonain dynamics and Lengendre transform with some examples and applications.
Chapter (1)
Newtonian mechanics

Section (1-1): Galilei space

In classical physics, the idea that there exists empty space should be accepted as central.

A basic postulate requires empty space to be specially 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 \((A, V)\), consisting of a set \( A \) and a \( k \)–vector space \( V \) together with an action of the abelian group \((V, +)\) underlying \( V \) on the set \( A \) 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 \( A \) is a map

\[
\rho : V \times A \rightarrow A
\]
Such that
\[ \rho(v + \omega, a) = \rho(v, \rho(\omega, a)) \text{ for all } v, \omega \in V, a \in A \]

(2) An action is called transitive, if for all \( p, q \in A \) 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 \( A \) and write \( \dim A = \dim_k V \). We also say that the affine space \( (A, V) \) is modelled over the vector space \( V \).

(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 \( A \) 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 special homogeneity. The fact that an affine space does not have a distinguished element is a nontrivial 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 \((A_1, V_1)\) and \((A_2, V_2)\) be affine spaces modelled over vector spaces \(V_1, V_2\) over the same field \(k\). A morphism \((A_1, V_1) \rightarrow (A_2, V_2)\) or affine map is a map

\[
\varphi : A_1 \rightarrow A_2
\]

for which there exists a \(k\)-linear map \(A \varphi : V_1 \rightarrow V_2\) such that

\[
\varphi (p) - \varphi (q) = A \varphi (p - q) \quad \text{for all } p, q \in A_1
\]

Remarks (1-1-4):

(1) Note that the \((V, +)\)-equivariant morphisms are those morphisms for which \(A \varphi = id V\).

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

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

(4) We will see in the appendix that affine space \(\mathbb{A}^n\) is an \(n\)-dimensional manifold. The choice of a point \(p \in A\) provides a natural global coordinate chart.
**Definitions (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 \( v, \omega \in V \) with \( H = \mathbb{R}v + \mathbb{R}_{\geq 0}\omega \). The boundary of a halfplane is the only line through the origin contained in it.

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

**Proposition (1-1-6):**

For any three-dimensional vector space, the map

\[
\{ \text{Scalar products on } V \} / \mathbb{R}_{>0} \to \text{Rotation groups}
\]

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

**Definitions (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 \( \omega \in V \) intersects \( l \) in \( v \in l \) and \( \omega = \lambda v \) with \( \lambda \geq 0 \) we define the norm on \( W \) by \( |\omega| = \lambda \).
Definitions(1-1-8):

(1) An n-dimensional Euclidean space \( E^n \) is an n-dimensional affine space \( A^n \) 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 \( \varphi \) for which the linear map \( A_\varphi \) 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 \( E^n \) is a semi-direct product of the subgroup \( V \) of translations given by the action of \( V \) and the rotation group \( O(V,b) \):

\[
Aut(E) = V \ltimes O(V,b)
\]

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.

Definitions(1-1-10):

(1) A Galilei space \( (A,V,t,\langle \cdot , \cdot \rangle) \) consists of

(i) An affine space \( A \) over a real four-dimensional vector space \( V \). The elements of \( A \) are called events or space time points.
(ii) 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

$$\text{Cont}(a) = \{ b \in \mathbb{A} \mid t(b - a) = 0 \}$$

is the subset of events simultaneous to $a \in \mathbb{A}$.

**Remark (1.1.11):**

As the morphisms of two Galilei spaces $(A_1, V_1, t_1, \langle \cdot , \cdot \rangle_1)$ and $(A_2, V_2, t_2, \langle \cdot , \cdot \rangle_2)$ we consider those affine maps

$$\varphi : A_1 \rightarrow A_2$$

which respect time differences

$$t_1(b - a) = t_2(\varphi(b) - \varphi(a))$$

for all $a, b \in A_1$,

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

**Remarks (1.1.12):**

(1) The group of automorphisms of a Galilei space is a proper subgroup of the affine group of the underlying affine space.
(2) Two Galilei spaces are isomorphic, but not canonically isomorphic. The automorphism group $\text{Aut}(G)$ of the Galilean coordinate space $G$ is called the Galilei group.

(3) We consider three classes of automorphisms in $\text{Aut}(G)$:

(i) Uniform motions with velocity $v \in \mathbb{R}^3$:

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

(ii) Special translations by $x_0 \in \mathbb{R}^3$ combined with time translations by $t_0 \in \mathbb{R}$:

$$g_2(t, x) = (t + t_0, x + x_0)$$

(iii) Rotations and reflections in space with $O \in O(2)$:

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

**Definitions (1-1-13):**

(1) Let $X$ be any set and

$$\phi : X \to G$$

be a bijection of sets. (On the right hand side, it would be formally more correct to write the set underlying $G$.) We say that $\phi$ provides a global Galilean coordinate system on $X$.

(2) We say that two Galilean coordinate systems $\phi_1, \phi_2 : X \to G$ are in relative uniform motion, if

$$\phi_1 \circ \phi_2^{-1} \in \text{Aut}(G).$$
Any global Galilean coordinate system $\phi : X \to G$ 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 $\phi$ to be an affine map:

$$x + v = \phi^{-1}(\phi(x) + v)$$

for all $x \in X, v \in \mathbb{R}^4$.

**Definitions (1-1-14):**

(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 (1-1-15):**

(1) By definition, two different inertial systems for the same Galilean structure are in uniform relative motion.

(2) There are no distinguished inertial systems.
Section (1-2): Dynamics of Newtonian systems

Definitions (1-2-1):

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

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

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

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

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

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

(3) The velocity in $t_0 \in I$ is defined as the derivative:

$$\dot{\varphi}(t_0) = \frac{d\varphi}{dt} \bigg|_{t_0}$$

(4) The acceleration in $t_0 \in I$ is defined as the second derivative:

$$\ddot{\varphi}(t_0) = \frac{d^2\varphi}{dt^2} \bigg|_{t_0}$$

(5) The graph

$$\{(t, x(t)) \mid t \in I\} \subset \mathbb{R} \times \mathbb{R}^3$$
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 be seen as a deep abstraction from many observations in nature.

We first discuss the situation in a fixed coordinate system:

**Definitions (1-2-2): [Newtonian determinism]**

1. A Newtonian trajectory of a point particle

   \[ \varphi: I \to \mathbb{R}^3 \text{, with } I = (t_0, t_1) \subset \mathbb{R} \]

   is completely determined by the initial position \( \sigma(t_0) \) and the initial velocity

   \[ \frac{d}{dt} \varphi \mid_{t=t_0} \]

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 special coordinates of \( N \) particles, the second are their velocities and the third is time, such that for all Newtonian trajectories the Newtonian equation

   \[ \frac{d^2}{dt^2} x^i = \ddot{x}^i = F^i(x, \dot{x}, t) \; , \; i = 1, \ldots, N \]
Definition (1-2-3):

Let $\mathbb{A}$ be a Galilei space, $I$ an interval of eigentime and $\varphi: I \to \mathbb{A}$ a smooth function. Then $\varphi$ 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.

Remarks (1-2-4):

(1) We immediately have the following consequences of the Newtonian principle of relativity.
(i) Invariance under time translations: the force $F$ does not depend on time $t$.
(ii) Invariance under special translations: $F$ depends only on the relative coordinates $\varphi^i - \varphi^1$.

(2) We deduce Newton’s first law: a system consisting of a single point is described in an inertial system by a uniform motion $(t, x_0 + t V_0)$. In particular, the acceleration vanishes, $\ddot{x} = 0$.

In the following we will discuss some examples:

Examples (1-2-5):

(1) The harmonic oscillator is defined by the potential

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

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

$$m \ddot{x} = -D x$$
They have the general solution \( \varphi(t) = A \cos(\omega t - \varphi_0) \) with that have to be determined from the initial conditions.

2. 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

\[
\int_{t_a}^{t_b} \varphi(t) \dot{\varphi}(t) dt
\]

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.

3. To write down Newton’s law of gravity, we consider a potential energy depending only on the distance \( r \) from the center of gravity:

\[
U(x_1, x_2, x_3) = -\frac{k}{r} \quad \text{with} \quad r = \sqrt{x_1^2 + x_2^2 + x_3^2}
\]

\[
\dot{\varphi} = -\nabla U = \frac{k}{r^3} \varphi
\]

This potential is central for the description not only of macroscopic systems like a plane 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.
**Observation (1-2-6):**

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

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

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

$$\ddot{\varphi} = - \text{grad} V(\varphi)$$

(1-1)

of second order. For any trajectory $\varphi: I \rightarrow \mathbb{R}^3$ that is a solution of (1-1) we consider the real-valued function

$$\epsilon: I \rightarrow \mathbb{R}$$

$$\epsilon(t) = \frac{1}{2} ||\dot{\varphi}(t) ||^2 + V(\varphi(t))$$

For its derivative, we find

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

where in the last step we use the equation of motion (1-1).

**Lemma (1-2-7):**

The system (1-1) is equivalent to the following system of ordinary differential equation of first order for the function $y: I \rightarrow \mathbb{R}^6$

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

The solutions of this system are called phase curves.
Observation (1-2-8):

This suggests to introduce $\mathbb{R}^6$ with coordinates $x, y, z, u_x, u_y, u_z$. This space is sometimes called the phase space $\mathbb{P}$ of the system. We equip the phase space $\mathbb{P}$ with the energy function

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

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

On phase space, we consider the ordinary differential equations

$$\frac{dx^i}{dt} = u_{x_i}, \quad \frac{du_{x_i}}{dt} = -\text{grad}_i V(x_1, x_2, x_3) \quad (1-2)$$

The solutions of (1-2) are called phase curves. They are contained in subspaces of $\mathbb{P}$ of constant value of $E$.

Observations (1-2-9):

(1) Assume that for any point $M \in \mathbb{P}$ a global solution $x_M(t)$ of (1-2) with initial conditions $M$ exists. This allows us to define a mapping

$$g_t : \mathbb{P} \rightarrow \mathbb{P} \text{ by } g_t(M) = x_M(t).$$

(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 $g_{t_1} g_{t_1} = g_{t_1 + t_2}$ and thus a smooth action

$$g : \mathbb{R} \times \mathbb{P} \rightarrow \mathbb{P}$$

$$(t, M) \rightarrow g(t, M)$$

called the phase flow.
Chapter (2)
Lagrangian mechanics

Section (2-1): Variational calculus and system with constraints

We begin this section by studying the variational calculus.

Observations (2-1-1):

(1) Our objects are trajectories, i.e. smooth functions $\phi : I \to \mathbb{R}^N$, $I = [t_1, t_2]$ on an interval with values in $\mathbb{R}^N$. We say that the trajectory is parameterized 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:

$$L_1(\phi) = \int_I ||\dot{\phi}|| \, dt$$

or its energy by

$$L_2(\phi) = \int_I ||\dot{\phi}||^2 \, dt$$

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

(2) Like the length of a trajectory, the functional we are interested in depend on the trajectory $\phi$ and its derivatives, i.e. on positions and velocities. Positions take their values in positions space, in our case
\( \mathbb{R}^N \) with Cartesian coordinates \((x^i, \ldots, x^N)\).

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

\[
J^1(\mathbb{R}^N) = \mathbb{R}^N \oplus \mathbb{R}^N
\]

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

\[
\mathbb{R}^N \hookrightarrow J^1(\mathbb{R}^N)
\]

\((x^1, ..., x^N) \mapsto (x^1, ..., x^N, 0, ..., 0)\).

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:

\[
J^2(\mathbb{R}^N) = \mathbb{R}^N \oplus \mathbb{R}^N \oplus \mathbb{R}^N
\]

with coordinates \((x^i, x^i_t, x^i_{tt})\).

We also use the abbreviation \( x^i_t = x^i_{tt} \). If we continue this way, we get a system of vector spaces for derivatives up to order \( \propto \)

\[
J^\propto (\mathbb{R}^N) \cong (\mathbb{R}^N)^{\propto + 1}
\]

with embeddings \( J^\propto \hookrightarrow J^{\propto + 1}(\mathbb{R}^N) \).

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

(3) Let us now explain in which sense the spaces we have just constructed are recipients for the derivatives of a smooth trajectory
\( \varphi : I \rightarrow \mathbb{R}^N \).

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

\[
\frac{d^n}{d\tau^n} \varphi^i
\]

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

\[
j^\alpha \varphi : I \rightarrow J^\alpha \mathbb{R}^N
\]

with components given by

\[
(j^\alpha \varphi)_{n}^{i} = \frac{d^n}{d\tau^n} \varphi^i
\]

We call this function the prolongation of the trajectory.

(4) We now formulate the variational problem we wish to solve:

Given a function like length or energy

\[
l : J^\alpha (\mathbb{R}^N) \rightarrow \mathbb{R}
\]

find all smooth trajectories

\[
\varphi : I \rightarrow \mathbb{R}^N
\]

which extremize the function

\[
L(\varphi) = \int_{I} l(j^\alpha \varphi) \, dt
\]
This problem is not yet correctly posed: if we minimize the length on all trajectories, the constant trajectories are obvious and trivial minima.

**Definition (2-1-2):**

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

\[
E(l) : (\mathbb{R}^N) \to \mathbb{R}
\]

with components

\[
E(l)^i = \frac{\partial l}{\partial x^i} - \sum_{j, \beta} \frac{\partial^2 l}{\partial x^i \partial x^j_{\beta}} x^{i}_{\beta+1}
\]

We call the operator that associates to \( l \) the function \( E(l) \) 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 \( J^1(\mathbb{R}^N) \) to smooth functions on \( J^{r+1}(\mathbb{R}^N) \). We then have

   \[
   E(l)^i = \frac{\partial l}{\partial x^i} - D \frac{\partial l}{\partial x^i_t}
   \]
(2) For any smooth trajectory, we obtain an \( \mathbb{R}^N \)-valued function on \( I \subset \mathbb{R} \) with

\[ i\text{-th component } \varphi \to E(l)^i \circ j^\alpha(\varphi) = E(l) [\varphi] \]

In terms of this function, the derivative of a variational family with respect to the variational parameter \( \varepsilon \) becomes

\[ \frac{d}{d\varepsilon} L(\varphi_{\varepsilon}) = \int_I (l) [\varphi_{\varepsilon=0}] \cdot \frac{d}{d\varepsilon} \varphi_{\varepsilon} \]

**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,

\[ \int_{t_1}^{t_2} f \cdot h \, dt = 0 \]

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 neighborhood \( U \) of \( t^* \) such that: \( f(t) > c > 0 \), for all \( t \in U \).
Using standard arguments from real analysis, we find a smooth function $h$ with support in $U$ such that $h|_{t^*} = 1$ for some neighborhood $\bar{U}$ of $t^*$ contained in $U$.

We thus find the inequalities:

$$\int_{t_0}^{t_1} f \cdot h = \int_U f \cdot h \geq \int_{\bar{U}} h \Big|_{C>0}$$

contradicting our assumption.

**Proposition (2-1-5):**

Let $L$ be a functional given on trajectories given by the smooth function

$$l : J^a \left( \mathbb{R}^N \right) \rightarrow \mathbb{R}$$

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

$$E(l) [\varphi] = 0$$

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

**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 $V : \mathbb{R}^N \rightarrow \mathbb{R}$. 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:

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

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

\[ \frac{\partial l}{\partial x^i} = -\text{grad}_i V \frac{\partial l}{\partial x^i_t} = x^i_t \]

and find for the Euler-Lagrange operator

\[ E(l)^i = -\text{grad}_i V - Dx^i_t \]

This gives us a system of \( N \) equations

\[ x^i_{tt} = -\text{grad}_i V \]

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

\[ Dx^i_t(\varphi) = \ddot{x}^i = -\text{grad}_i V(\varphi(t)) \]

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

\[ l : J(\mathbb{R}^N) \rightarrow \mathbb{R} \]

which is also called the Lagrangian function of the system.

(2) Let \( : [t_0, t_1] \rightarrow \mathbb{R} \) be a real valued function. The length of the curve:

\[ \gamma = \{(t, x) : x = \varphi(t) \text{ mit } t_0 \leq t \leq t_1 \} \subset \mathbb{R}^2 \]
is given by
\[ L(\gamma) = \sqrt{\int_{t_0}^{t_1} (1 + \dot{\varphi}^2) \, dt } \]

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

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

which reduces to \( x_{tt} = 0 \). Hence we get the differential equation \( \ddot{\varphi} = 0 \) on the trajectories which have as solutions \( \varphi(t) = ct + d \).

**Definitions (2-1-7):**

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

2. Given a Lagrangian \( l(x', x_t', t) \) and a trajectory \( \varphi : I \to \mathbb{R}^N \), we call \( L(\varphi) = \int_I l(\varphi) \, dt \) the corresponding action for the trajectory \( \varphi \).
Observations (2-1-8):

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

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

Proposition (2-1-9):

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

Proof:

For any trajectory $\phi: I \rightarrow \mathbb{R}^N$, that is a solution of the Euler-Lagrange equations, we have for a cyclic coordinate $x^i$

$$\frac{d}{dt} \frac{\partial t}{\partial x_t^i} \circ j^i \phi = \frac{\partial t}{\partial x_t^i} \circ j^i \phi = 0$$

Now we will discuss system with constrain.

Proposition (2-1-10):

Consider the auxiliary function

$$\tilde{\mathcal{I}} = \mathbb{R}^N \times \mathbb{R}^f \rightarrow \mathbb{R}$$
\[(x, \lambda) \mapsto f(x) + \sum_{\alpha=1}^{r} \lambda_{\alpha} g_{\alpha}(x)\]

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 \((x_0, \lambda_0)\) 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

\[
0 = \left. \frac{\partial \tilde{f}}{\partial x} \right|_{(x_0, \lambda_0)} = g_{\alpha}(x_0) \quad \text{and} \quad \left. \frac{\partial \tilde{f}}{\partial x} \right|_{x_0} = \sum_{\alpha=1}^{r} \lambda_{\alpha} \frac{\partial}{\partial x_{\alpha}} g_{\alpha}(x_0)
\]

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 \(f|_M\).

**Observations (2-1-11):**

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

\[
l(x, x_i) = \frac{1}{2} \sum_{i=1}^{N} (x_i^i)^2 - V(x)
\]

that is constrained by unknown forces to a sub-manifold \(M \subset \mathbb{R}^N\) 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
\[ \varphi : I \to \mathbb{R}^N \]

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

\[ \int_I l \circ j^1 \varphi . \, dt \]

We introduce Lagrangian multipliers and minimize the functional

\[ f(\epsilon, \lambda) = \left( \int_I \circ j^1 \varphi_\epsilon + \sum_{\alpha=1}^r \lambda_\alpha g_\alpha(\varphi_\epsilon) \right) \, dt \]

The derivative with respect to \( \epsilon \) yields the following additional term

\[ \sum_{\alpha=1}^r \lambda_\alpha \frac{\partial g_\alpha}{\partial x^i} \cdot \frac{d\varphi_k^i}{d\epsilon} \bigg|_{\epsilon=0} \]

so that the equation of motions for a trajectory \( \varphi \) become

\[ E(l) \circ j^1(\varphi) = \ddot{\varphi} + \text{grad} \, V(\varphi) = -\sum \lambda_\alpha \text{grad} \, g_\alpha(\varphi) \quad (2.1) \]

The right hand side describes additional forces constraining the motion to the sub-manifold \( M \). The concrete form of the forces described by the gradients of the functions \( g_\alpha \) is, in general, unknown.

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

\[ q : M \ni U \to \mathbb{R}^{N-r} \]
of the sub-manifold, 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

\[
\frac{\partial}{\partial q^\alpha} = \frac{\partial x^i}{\partial q^\alpha} \frac{\partial}{\partial x^i}, \alpha = 1, \ldots, N-r
\]

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

\[
\frac{\partial}{\partial q^\alpha} = \frac{\partial x}{\partial q^\alpha} \quad (2.2)
\]

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

\[
\varphi: I \rightarrow U \subset M.
\]

Using the embedding \( U \rightarrow \mathbb{R}^N \), we can see this also as a trajectory \( \tilde{\varphi} \) in \( \mathbb{R}^N \): The chainrule yields

\[
\frac{d\tilde{\varphi}^i}{dt} = \frac{\partial x^i}{\partial q^\alpha} \cdot \frac{d\varphi^\alpha}{dt}
\]

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 \( x_i \) of \( J^1\mathbb{R}^N \), seen as a function on jet space \( J^1M \):

\[
x^i_t = x^i_t(q^\alpha, q_t^\alpha, t) = \frac{\partial x^i}{\partial q^\alpha} q^\alpha_t \quad (2.3)
\]

and expressed in coordinates \( q^\alpha, q^\alpha_t \) on \( J^1M \). We find as an obvious consequence
\[
\frac{\partial x_t^i}{\partial q_t^i} \frac{\partial x^i}{\partial q^a} \quad (2.4)
\]

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

\[
\frac{d^2 \phi}{dt^2} = \frac{\partial^2 x_t^i}{\partial q^\beta \partial q^a} \cdot \frac{d \phi^\beta}{dt} \cdot \frac{d \phi^a}{dt} + \frac{\partial x_t^i}{\partial q^a} \cdot \frac{d^2 \phi^a}{dt^2}
\]

This yields the following expression of the coordinate function \(x_{tt}^i\) on \(J^2 \mathbb{R}^N\) as a function on \(J^2 M\), in terms of the coordinates \(q_t^a\) on \(J^2 M\):

\[
x_{tt}^i = x_{tt}^i(q^\alpha, q_t^\beta, t) = \frac{\partial x_t^i}{\partial q^a} q_{tt}^a + \frac{\partial^2 x_t^i}{\partial q^a \partial q^\beta} q_t^\beta q_t^a \quad (2.5)
\]

(4) Since we do not know the right hand side of the equation of motion (2.1), we take the scalar product \(\langle \cdot, \cdot \rangle\) in \(\mathbb{R}^N\) with each of the tangent vector (2.2) in \(T_p M\). This yields \(r\) equations

\[
\langle x_{tt}^i, \frac{\partial \phi}{\partial q^a} \rangle = -\langle \text{grad} V, \frac{\partial x}{\partial q^a} \rangle, \text{ with } \alpha = 1, \ldots, r
\]

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

\[
\langle \text{grad} V, \frac{\partial x}{\partial q^a} \rangle = \frac{\partial V}{\partial x^i} \frac{\partial x^i}{\partial q^a} = \frac{\partial V}{\partial q^a}
\]

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

**Proposition (2-1-12):**
Suppose a natural system on \( \mathbb{R}^N \) with potential \( V(x) \) is constrained by unknown forces to a sub-manifold \( M \subset \mathbb{R}^N \). Then the equations of motion on \( M \) are the Euler-Lagrange equations for a Lagrangian \( l(q, q_t) \) on \( J^1M \) obtained from the Lagrangian on \( J^1\mathbb{R}^N \) by restricting \( V \) to \( M \) and transforming the coordinates \( x_i \) as in (2.3).

**Section (2-2): Lagrangian Systems and Lagrangian Dynamics**

We begin with the concepts of Lagrangian systems.

**Definitions (2-2-1):**

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.

2. Global trajectories are global sections of \( \pi \), i.e. smooth maps \( \varphi : E \to I \) such that \( \pi \circ \varphi = \text{id}_I \). 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 \( \pi \circ \varphi = \text{id}_U \). The system of local sections forms a sheaf on \( I \).

3. A system defined by constraints on the system : \( E \to I \) is a submanifold \( i : E' \to E \) such that \( \pi' = \pi \circ i \) is a surjective submersion \( E' \to I \).

**Remarks (2-2-2):**

1. We will generalize the situation by considering a general surjective submersion \( \pi : E \to M \), where \( M \) is not necessarily an interval.
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 considered as local coordinate functions on M.

Definitions (2-2-3):

1. Let $E, M$ be smooth manifolds, $\dim M = m$ and let $\pi : E \to M$ be a surjective submersion. We say that $\pi$ 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 $\pi$, i.e. the set of smooth functions $s_U : U \to E$ such that $\pi \circ s_U = \text{id}_U$. These sets form a sheaf on M.

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

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

   and the derivative operators

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

Example (2-2-4):

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 $p_1$

$$\varphi : \mathbb{R} \to \mathbb{R} \times M$$
is given by a smooth trajectory
\[ \tilde{\phi} : \mathbb{R} \to M \]

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

\[
(t_0, x^i, x_i^t) \circ j^1_t \varphi = (t_0, x^i \tilde{\phi}(t), \left. \frac{d}{dt} \right|_{t=t_0} x_i^t \tilde{\phi}(t))
\]

Thus, all information that appears is the tangent vector

\[
\left. \frac{d\tilde{\phi}}{dt} \right|_{t=t_0}
\]

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

\[
J^1 \pi \to \mathbb{R} \times TM
\]

\[
j^1_{t_0} \varphi \to \left( t_0, \left. \frac{d\tilde{\phi}}{dt} \right|_{t=t_0} \right)
\]

**Definition (2-2-5):**

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

\[
\cdots \to J^2 \pi \overset{\pi_{2,1}}{\longrightarrow} J^1 \pi \overset{\pi_{1,0}}{\longrightarrow} E \overset{\pi}{\longrightarrow} M
\]

It comes with a family of natural projections

\[
\pi_{\infty,j} : J^\infty \pi \to J^j \pi
\]
**Definition (2-2-6):**
A real-valued function \( f : J^\infty \pi \to \mathbb{R} \) is called smooth or local, if it factorizes through a smooth function on a finite jet bundle \( J^r \pi \). I.e. there exists \( r \in \mathbb{N} \) and a smooth function \( f_r : J^r \pi \to \mathbb{R} \) such that:

\[
f = f_r \circ \pi_{\infty,r}
\]

**Remarks (2-2-7):**

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

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

\[
\pi_{k+1,k}^* : C^\infty(J^k \pi) \to C^\infty(J^{k+1} \pi)
\]

of algebras. The embedding is given by considering a function in \( C^\infty(J^k \pi) \) 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 \( C^\infty(J^\infty) \) has the structure of a filtered commutative algebra.

**Definition (2-2-8):**

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 be expressed as the integral over the pullback of a local function \( l \in \text{Loc}(E) \).
Definition (2-2-9):
A vector field on the jet bundle $J^p\pi$ is defined as a derivation on the ring $C^\infty(J^p\pi)$ of smooth functions.

Remark (2-2-10):
In local coordinates $(x^i, u_j^a)$ on $J^p\pi$, a vector field can be described by a formal series of the form

$$X = \sum_i A_i \frac{\partial}{\partial x^i} + \sum_\alpha B_\alpha^a \frac{\partial}{\partial u_j^a}$$

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

Definition (2-2-11):
The prolongation of a (local) vector field $X \in \Gamma(U)$ is the (local) vector field $\text{pr}^\infty X$ on $J^p\pi$ acting on a smooth function $f \in C^\infty(J^p\pi)$ in the point $j^\infty s(p) \in J^p\pi$ as

$$\text{pr}^\infty X_{j^\infty s(p)} f = X_p(f \circ j^\infty s)$$

Definition (2-2-12):
For a multi-index $I = (i_1, \ldots, i_n)$, we introduce the following operator acting on local functions defined on a coordinate patch:
A total differential operator is a mapping from \( \text{Loc}(E) \) 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 \( Z^I \in \text{Loc}(E) \) is a local function.

**Definitions (2-2-13):**

1. An Ehresmann connection on the fibred manifold \( \pi: E \to M \) is a smooth vector sub-bundle \( H \) of the tangent bundle \( TE \) over \( E \) such that
   \[
   TE = H \oplus V
   \]
   where the direct sum of vector bundles over \( E \) is defined fiberwise.
2. The fibers 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.

**Proposition (2-2-14):**

Let \( \pi: E \to M \) be a fibred manifold and \( J^\infty \pi \) be the corresponding jet bundle.

The following holds:

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

**Remarks (2-2-15):**

The horizontal bundle \( H \) is a vector bundle of rank \( \dim M \). The restriction \( d\pi|_H: H \to TM \) is an isomorphism.
In the following we will discuss of Lagrangian dynamics.

**Definitions (2-2-16):**

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 $l \in \Omega^{n,0}(J^\infty \pi)$, 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 $\text{vol}_M \in \Omega^n(M)$ is a volume form on $M$ and that $L \in \text{Loc}(E)$ is a local function. Then
   
   $$l = L \cdot \pi_\infty^*(\text{vol}_M) \in \Omega^{n,0}(J^\infty \pi),$$

   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 $I \subset \mathbb{R}$.

**Definition (2-2-17):**

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 $pr_E(X)$, such that:

(i) $X$ and $pr_E(X)$ agree on functions on $E$.

(ii) The Lie derivative of $pr_E(X)$ preserves the contact ideal:

$$L_{pr_E(X)}C(J^\infty \pi) \subset C(J^\infty \pi).$$

**Remarks (2-2-18):**
(1) We present expressions in local adapted coordinates. Consider a vector field on $E$,

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

(2) We then write the prolongation as

$$\text{pr}_E(X) = Z^i \frac{\partial}{\partial x^i} + Z_\alpha \frac{\partial}{\partial u_\alpha}$$

The first condition in the definition immediately yields

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

One can determine the coefficients to be

$$Z_\alpha^i = D_i(b^\alpha - u_\alpha^i) + u_\alpha^i a^i$$

**Definition (2-2-19):**

Given a Lagrangian density

$$l = l(x^i, u^a_i) dx^i \wedge \ldots \wedge dx^n \in \Omega^{n,0}(J^\infty \pi)$$

the Euler-Lagrange form $E(l) \in \Omega^{n,0}(J^\infty \pi)$ is defined as

$$E(l) = E_\alpha(l) \Theta^\alpha dx^i \wedge \ldots \wedge dx^n$$

with

$$E_\alpha(l) = \frac{\partial l}{\partial u^\alpha} - D_i \frac{\partial l}{\partial u^\alpha_i} + D_{ij} \frac{\partial l}{\partial u^\alpha_{ij}} + \ldots = (-D)_K \frac{\partial l}{\partial u^\alpha_K}$$
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 $\Omega^{n,1}(J^\infty \pi)$. Forms on $J^\infty \pi$ of type $(n, s)$ with $s \geq 1$ are $d_H$-closed, but turn out to be not even locally $d_H$-exact.

**Lemma (2-2-20):**

We have the following identities for the inner Euler operators:

1. For $\eta \in \Omega^{n-1,s}(J^\infty \pi)$, we have
   $$I(d_H \eta) = 0 .$$

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

3. For any $\omega \in \Omega^{n,s}(J^\infty \pi)$ there exists $\eta \in \Omega^{n-1}(J^\infty \pi)$ such that:
   $$\omega = I(\omega) + d_H \eta .$$

4. For any Lagrangian density $\lambda \in \Omega^{n,0}(J^\infty \pi)$, we have
   $$E(\lambda) = I(d_V \lambda)$$

This suggests to extend the bicomplex.

**Propositions (2-2-21):**

1. Let $l \in \Omega^{n,0}(J^\infty \pi)$, be any Lagrangian. We claim that then there always exists $\eta \in \Omega^{n-1,1}(J^\infty \pi)$ such that
   $$d_V l = E(l) + d_H(\eta) .$$
(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 \( \sigma \in \Omega^{n-1,0}(\mathcal{J}^\infty\pi) \) such that

\[
L_{p_{E}(X)}l = \iota_{p_{E}X}(E(l)) + d_{H}\sigma .
\]

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

\[
\left. \frac{d}{d\epsilon} \right|_{\epsilon=0} l(\mathcal{J}^\infty s_{\epsilon}) = L_{p_{E}(X)}l = \iota_{p_{E}X}(E(l)) + d_{H}\sigma
\]
Section (2-3): Symmetries, Noether identities and Natural Geometry

First we will discuss symmetries and Noether identity.

Definitions (2-3-1):

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

(2) Due to the integrability of the Cartan distribution $H$, all vector fields in $\text{vect}_H$ are symmetries.

The Lie algebra of vector fields $\text{vect}_H$ is, by definition of $\text{vect}_{\text{sym}}(\pi)$, an ideal in the Lie algebra $\text{vect}_{\text{sym}}(\pi)$ of all symmetries. We introduce the Lie algebra of non-trivial symmetries as the quotient

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

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

We will need a different description of symmetries.

Observations (2-3-2):

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

(2) Since we assumed that vector fields in $\text{vect}_{\text{sym}}$ are vertical, we can write $\phi_X$ in local coordinates as:
\[ Q^\alpha \frac{\partial y}{\partial u^\alpha} \text{ with } Q^\alpha \in \text{Loc}(E) \]

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

\[
\begin{array}{ccc}
\pi^*(\pi) & \longrightarrow & E \\
\downarrow & & \downarrow \pi \\
J^\alpha \pi & \longrightarrow & M
\end{array}
\]

We call \( \kappa(\pi) \) the space of smooth sections of the bundle \( \pi^*(\pi) \to J^\alpha \pi \) on jet space.

**Definition (2-3-3):**

An element \( \varphi_X \) is called a generating section of a symmetry \( X \) or also an evolutionary vector field.

**Remarks (2-3-4):**

(1) We denote the symmetry of the fibred manifold corresponding to a section \( \varphi \in \kappa(\pi) \) by \( E_{\varphi} \). Some authors reserve the term evolutionary vector field for this symmetry.

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

\[ \varphi = \sum_{\alpha}^m \varphi^\alpha \frac{\partial}{\partial u^\alpha} \]

with local functions \( (\varphi^1, \ldots, \varphi^m) \), then \( E_{\varphi} \) is its prolongation \( \text{pr}_E \)

\[ E_{\varphi} = \sum_{\alpha}^m D_1(\varphi^\alpha) \frac{\partial}{\partial u^\alpha} \]
The first goal of this subsection is to explore symmetries of Lagrangian systems \((\pi : E \to M, l)\).

**Remark (2-3-5):**

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 \(i = 1, \ldots, m\),

\[
pr_E(Q_E)(l) = D_K(Q_E^a) \frac{\partial l}{\partial u^a_i} = D_i j^i
\]

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

**Definition (2-3-6):**

Let \((\pi : E \to M, l)\) be a Lagrangian system. A \(m-1\)-form \(\alpha \in \Omega^{m-1,0}(J^\infty \pi)\) 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 \((j^\infty s)^* \in \Omega^{n-1}(M)\) is closed.

**Remarks (2-3-7):**

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 \(\alpha_s = \alpha \circ j^\infty(t) = \alpha(t, s(t), s_t(t), \ldots)\) on the interval \(I\) such that

\[
\frac{d}{dt} \alpha_s = 0
\]

This implies

\[
\alpha_s(t_1) - \alpha_s(t_0) = \frac{d}{dt} \alpha_s \int_{t_0}^{t_1} dt = 0
\]
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

\[
\alpha_s = \sum_{i=1}^m (\alpha_s)_i dx^i
\]

(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, \( \mathbb{A}^n \rightarrow \mathbb{A}^1 \). 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 \( j_s = \rho_s + dt \wedge j_s \)

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

\[
0 = dj = dt \wedge \left( \frac{\partial \rho_s}{\partial t} + d_{\nu} j_s \right)
\]

where \( d_{\nu} \) is the vertical (i.e. here: special) differential. Fix a certain \( n-1 \)-dimensional volume at fixed time. We then have:

\[
\frac{d}{dt} \int_V \rho_s = \int_V \frac{\partial}{\partial t} \rho_s = -\int_V d_{\nu} j_s = -\int_{\partial V} j_s
\]

where in the last step we used Stokes’ theorem. This has the following interpretation: forever solution \( s \) of the equations of motion, we find a quantity (“charge”) \( \rho_s \) that can be assigned to any spacial volume \( V \) and a “current” \( j_s \) whose flux across the boundary \( \partial V \) describes the loss or gain of this quantity. Together, they form the \( m \) components of the conserved \( m-1 \)-form.
**Definition (2-3-8):**

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

$$\int_M (j^0 s) * (FZ(G)) \, d\text{vol}_M = \int_M (j^0 s) * (Z^+(F)G) \, d\text{vol}_M$$

for all sections $s : M \to E$ and all local functions $F, G \in \text{Loc}_E$. It follows that

$$FZ(G)d\text{vol}_M = Z^+(F)Gd\text{vol}_M + d_H \zeta$$

for some $\zeta = \zeta(Z, F, G) \in \Omega^{n-1,0}(j^\infty \pi)$ 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

$$Z^+(F) = (-D)_j (Z^j F) .$$

**Theorem (2-3-9): [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 (2-3-5). 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 jetspace reads explicitly

$$\left( Q^\alpha \frac{\partial}{\partial u^\alpha_i} - j_i \right) dx^i$$
Proof:

(1) Let $s$ be a solution of the equations of motion. Since $Q$ is a variational symmetry, we have

$$0 = (\int M D_i j^i) \circ j^\infty S = \left(\int_M \frac{\partial l}{\partial u_i^\alpha} (D_i Q^\alpha)\right) \circ j^\infty S$$

Repeated integration by parts on $M$ yields by the previous comment on adjoint operators and local functions $\zeta^i$ on jet space such that

$$0 = \left(\int_M D_j \frac{\partial l}{\partial u_i^\alpha} \right) Q^\alpha \circ j^\infty S + \int_M \zeta^i \circ j^\infty S$$

Since the Lagrangian does explicitly depend on $M$, the Euler-Lagrange equations for the section $s$ take the form

$$\left( (-D)_j \frac{\partial l}{\partial u_i^\alpha} \right) \circ j^\infty S = 0$$

We thus learn that

$$0 = D_i \left( \int_M j^i \right) = 0$$

(2) If the Lagrangian depends only on first order derivatives, the divergence term is explicitly

$$D_i \left( \int_M \frac{\partial l}{\partial u_i^\alpha} Q^\alpha \right) \circ j^\infty S$$

so that the conserved quantity is given by the one-form on jet space

$$\left[ \frac{\partial l}{\partial u_i^\alpha} Q^\alpha - j^i \right] dx^i$$
**Definition (2-3-10):**

A gauge symmetry of a Lagrangian system \((\pi: E \to M, l)\) consists of a family of local functions, for \(\alpha = 1, \ldots, n\) and all multi-indices \(I\),

\[
R^\alpha_I: J^\infty \pi \to \mathbb{R}
\]

for \(\alpha = 1, \ldots, n\) and all multi-indices \(I\), such that for any local function \(\epsilon: J^\infty \pi \to \mathbb{R}\)

the evolutionary vector field \(R^\alpha_I(D_I\epsilon)\frac{\partial}{\partial \pi^\alpha}\) on \(E\) is a variational symmetry of \(l\).

**Remarks (2-3-11):**

(1) Loosely speaking, a gauge symmetry is a linear mapping from local functions on \(J^\infty \pi\) into the evolutionary vector fields on \(E\) preserving the Lagrangian. It is crucial for a gaugesymmetry that there is a symmetry for every local function.

(2) Notice that the coefficients of the vector field depend linearly on \(\epsilon\) 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 \(\epsilon\) on \(J^\infty \pi\).

**Theorem (2-3-12):** [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 \(J^\infty \pi\), the following statements are equivalent:
(i) The functions \( \{ R^\alpha \} \) define a gauge symmetry of \( l \), i.e., \( R^\alpha (D_1 \epsilon) \frac{\partial}{\partial u^\alpha} \) is a variational symmetry of \( l \) for any local function \( \epsilon: \mathcal{F}^\omega \pi \to \mathbb{R} \).

(ii) \( R^\alpha (D_1 \epsilon) E_\alpha (l) \) is a divergence for any local function \( \epsilon \).

(iii) The functions \( \{ R^\alpha \} \) define Noether identities of \( l \), i.e., \( R^\alpha (D_1 (E_\alpha (l))) \) is identically zero on the jet bundle.

Second we will discuss the Natural geometry.

**Definitions (2-3-13):**

1. Denote by \( \text{Man}_n \) the category of \( n \)-dimensional manifolds and open embeddings. Let \( \text{Fib}_n \) be the category of smooth fiber bundles over \( n \)-dimensional manifolds with morphisms differentiable maps covering morphisms of their bases in \( \text{Man}_n \).

2. A natural bundle is a factor \( \mathcal{B}: \text{Man}_n \to \text{Fib}_n \) such that for each \( M \in \text{Man}_n \), \( \mathcal{B}(M) \) is a bundle over \( M \). Moreover, \( \mathcal{B}(M') \) is the restriction of \( \mathcal{B}(M) \) for each open sub-manifold \( M' \subset M \), the map \( \mathcal{B}(M') \to \mathcal{B}(M) \) induced by \( M' \hookrightarrow M \) being the inclusion \( \mathcal{B}(M') \hookrightarrow \mathcal{B}(M) \).

**Theorem (2-3-14):** (Krupka, Palais, Terng)

For each natural bundle \( \mathcal{B} \), there exists \( l \geq 1 \) and a manifold \( \mathfrak{B} \) with a smooth \( \mathfrak{G}_{\mathcal{B}}^{(l)} \) action such that there is a factorial isomorphism

\[
\mathcal{B}(M) \cong \text{Fr}^l(M) \times_{\mathfrak{G}_{\mathcal{B}}^{(l)}} \text{B} = (\text{Fr}^l(M) \times \text{B}) / \mathfrak{G}_{\mathcal{B}}^{(l)}
\]
Conversely, each smooth $GL_n$-manifold $B$ induces, a natural bundle $\mathcal{B}$. We will call $B$ the fiber of the natural bundle $\mathcal{B}$. If the action of $GL_n$ on $B$ does not reduce to an action of the quotient $GL_{n-1}$ we say that $\mathcal{B}$ has order $l$.

**Examples (2-3-15):**

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

   $$T(M) \cong Fr(M) \times_{GL_n} \mathbb{R}^n$$

   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 $\text{Lin}(\Lambda^m(\mathbb{R}^n), \mathbb{R})$, with the obvious induced $GL_n$-action. The presentation

   $$\Omega^m(M) \cong Fr(M) \times_{GL_n} \text{Lin}(\Lambda^m(\mathbb{R}^n), \mathbb{R})$$

   is also classical. A particular case is $\Omega^0(M) \cong Fr(M) \times_{GL_n} \mathbb{R} \cong M \times \mathbb{R}$ 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.

**Definition (2-3-16):**

Let $\mathfrak{F}$ and $\mathfrak{G}$ be natural bundles. A (finite order) natural differential operator $\mathcal{D} : \mathfrak{F} \rightarrow \mathfrak{G}$ is a natural transformation (denoted by the same symbol) $\mathcal{D} : \mathfrak{F}^{(k)} \rightarrow \mathfrak{G}$, for some $k \geq 1$. We denote the space of all natural differential operators $\mathfrak{F} \rightarrow \mathfrak{G}$ by $Nat(\mathfrak{F}, \mathfrak{G})$. 

47
Examples (2-3-17):

(1) Given natural bundles $\mathcal{B}'$ and $\mathcal{B}''$ with fibers $B'$ resp. $B''$, there is an obviously defined natural bundle $\mathcal{B}' \times \mathcal{B}''$ with fiber $B' \times B''$. With this notation, the Lie bracket is a natural operator $[-,-] : T \times T \rightarrow T$ and the covariant derivative an operator $\nabla : \operatorname{Con} \times T \times T \rightarrow T$, where $T$ is the tangent space functor and $\operatorname{Con}$ the bundle of connections. The corresponding equivariant maps of fibers can be easily read off from local formulas given in Examples.

(2) The operator $\mathcal{O}_\phi : T \times \Omega^1 \rightarrow C^\infty$ from the Example above is induced by the $\operatorname{GL}_n$-equivariant map

$$O_\phi : \mathbb{R}^n \times (\mathbb{R}^n)^* \rightarrow \mathbb{R} \quad \text{given by } o_\phi(v, \alpha) = \phi(\alpha(v)).$$
Chapter (3)
Classical field theories

Section (3-1): Maxwell’s equations and Special relativity

We start explaining Maxwell’s equation on a Galilei space $\mathbb{A}$.

Observations (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 $\rho$ at time $t$ as a three-form on $\mathbb{A}_t$:

   $$\rho_t(x) \, dx^1 \wedge dx^2 \wedge dx^3 \in \Omega^3(\mathbb{A}_t)$$

   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 $\rho$ 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 $\partial V$, then electric charge can pass through its boundary $\partial V$. The amount of charge passing per time should be described by the integral over a two form
Again, this two-form will depend on time so that we introduce the current density as a 3-form: \( j(x, t) = dt \wedge j_t(x) \in \Omega^{1,2}(\mathbb{A}) \).

We then have the natural conservation law for any closed volume \( V \in \mathbb{A}_t \)

\[
\frac{d}{dt} \rho = \int_V \int_{\partial V} j_t \wedge \epsilon \Omega^2(\mathbb{A})
\]

which by Stokes’ theorem takes the form

\[
\frac{\partial}{\partial t} \int_V \rho = - \int_{\partial V} \int_{\partial V} j_t = - \int_{\partial V} d_V j_t
\]

Since this holds for all volumes \( V \subset \mathbb{A}_t \), we have the infinitesimal form of the conservation law

\[
\frac{\partial}{\partial t} \rho + d_V j_t = 0
\]

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 \( j \):

\[
j = \rho - j = \rho, dx \wedge dy \wedge dz - dt \wedge j_t \in \Omega^3(\mathbb{A})
\]

defined on the four-dimensional space \( \mathbb{A} \). We find

\[
dj = \frac{\partial}{\partial t} \rho, dt \wedge dx \wedge dy \wedge dz + dt \wedge d_V j_t = 0
\]

(4) Since the charge-current density 3-form is closed, \( dj = 0 \), it is exact by the Poincare lemma. We can find a two-form \( H \in \Omega^2(\mathbb{A}) \), the excitation 2-form such that

\[
dH = j.
\]
The fact that such a 2-form exists even globally is the content of the inhomogeneous Maxwell equation.

(5) Using the bigrading of 2-forms on $\Lambda$, we can write

$$H = dt \wedge H + D$$

with $H \in \Omega^{0,1}(\Lambda)$ and $D \in \Omega^{0,2}(\Lambda)$. 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

$$j = \rho - dt \wedge j = dH = -dt \wedge d_{\nabla}H + dt \wedge \partial_t D + d_{\nabla}D$$

are equivalent to

$$d_{\nabla}D = \rho \quad \text{and} \quad \partial_t D = d_{\nabla}H - j,$$

where the vertical derivative $d_{\nabla}$ is just the special exterior derivative.

The first equation is the Coulomb-Gauss law, the second equation the Oersted-Ampere equation. The term containing the time derivative $-\partial_t D$ of the excitation is sometimes called Maxwell’s term.
Example (3-1-2):

Gauß’ law reads in integral form for a volume $V$ in space

$$Q = \rho d^3 x = - \int _{\partial V} D \cdot d\mathbf{S}$$

The electric flux through the surface $\partial V$ 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: $\rho(\mathbf{x}, t) = f(r)\hat{e}_r$.

with

$$\hat{e}_r = \frac{1}{\sqrt{x^2 + y^2 + z^2}} (x, y, z) \in \mathcal{T}_{(x,y,z)}(\mathbb{R}^3 \setminus \{0\})$$

the radial unit vector field on $\mathbb{R}^3 \setminus \{0\}$.

Integrating over the two-sphere with center 0 and radius $r$, we find

$$q = \overline{D} \int_{S_r^2} = 4\pi r^2 f(r)$$

and thus Coulomb’s law:

$$\overline{D}(\mathbf{x}, t) = \frac{q}{4\pi r^2} \hat{e}_r$$

In the case of a static field, Coulomb’s law and the principle of superposition of charges conversely implies the first inhomogeneous Maxwell equation $d\mathbf{v} D = \rho$. 

52
Observations (3-1-3):

(1) Based on empirical evidence, we impose on the electromagnetic field \( dF \) the condition to be closed,

\[
dF = 0 .
\]

These are the homogeneous Maxwell equations.

(2) In the three-dimensional language \( F = dt \wedge E + B \), the equation

\[
0 = dF = dt \wedge dV E + dt \wedge \partial_t B + dV B
\]

is equivalent to the following two equations:

\[
dV B = 0 \quad \text{and} \quad \partial_t B = -dV E .
\]

Example (3-1-4):

Faraday’s law of induction yields a relation between the induced voltages along a loop \( \partial F \) bounding a surface \( F \)

\[
U_{\text{ind}} = \oint_{\partial F} E
\]

and the magnetic flux \( \Phi^{mag} = \overline{B} \)

through the surface which reads

\[
U_{\text{ind}} = d\oint_{\partial F} E = -\frac{d}{dt} \left( \overline{B} \right) - \oint_{F} \frac{d}{dt} \Phi^{mag} = -\frac{d}{dt} \left( \overline{B} \right)
\]

This law is the basis of the electric motor and the electrical generator. The minus sign is quite famous: it is called Lenz’ rule.
Observations (3-1-5):

(1) The following space-time relations are important:

(i) In empty space, one has the relation

\[ H = -\lambda \ast F \]

which makes sense in four dimensions. The constant \( \lambda \) is a constant of nature. Sometimes \( \ast F \) is called the dual field strength.

(ii) In axion-electrodynamics, one has two constants of nature

\[ -H = -\lambda_1 \ast F + \lambda_2 F . \]

(iii) In realistic media, the constants typically depend on frequencies or, equivalent, wavelengths, and yield quite complicated relations between \( F \) and \( H \).

(2) We summarize the Maxwell equations in vacuo:

\[ dF = 0 \text{ and } d(\ast F) = J , \]

which are the homogeneous and the inhomogeneous Maxwell equations. As a consequence we have the continuity equation

\[ dJ = d(\ast F) = 0 . \]

It is complemented by the equation for the Lorentz force

\[ f = I(F) \wedge J . \]
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:

\[ \text{div} B = 0 \text{ and } \partial_t B + \text{rot} \vec{E} = 0 \]

\[ \text{div} E = \rho \text{ and } \text{rot} B - \partial_t E = \vec{j} \]

**Observations (3-1-6):**

(1) Consider the Maxwell equations in vacuo without external sources, i.e. \( j = 0 \). We find

\[ dF = 0 \text{ and } d \ast F = 0 . \]

The last equation implies \( *F = 0 \). We thus have for the Laplace operator on differential forms

\[ \Delta F = \delta dF + d\delta F = 0 \]

so that in the vacuum without external source, \( J = 0 \), 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:

\[ \text{div} \vec{E} = 0 \text{ , } \text{rot} \vec{E} = -\frac{\partial \vec{B}}{\partial_t} \]

\[ \text{div} \vec{B} = 0 \text{ , } \text{rot} \vec{B} = \frac{\partial \vec{E}}{\partial_t} \]
Taking the rotation of the second equation yields

$$\Delta \vec{E} = \Delta \vec{E} - \text{grad} (\text{div} \vec{E}) = - \text{rot} \text{rot} \vec{E} =$$

$$= \text{rot} \frac{\partial}{\partial t} \vec{B} = \frac{\partial}{\partial t} \text{rot} \vec{B} = \frac{\partial^2 \vec{E}}{\partial t^2}$$

A dimensional analysis shows that we should restore a factor $c^2$ with the dimension of the square of a velocity. Hence we get

$$\Box \vec{E} = \left( \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta \right) \vec{E} = 0$$

with $\Box$ the so-called d’Alembert operator. Similarly, one finds

$$\Box \vec{B} = \left( \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta \right) \vec{B} = 0$$

We discuss our postulates of special relativity.

**First postulate:** space time is homogeneous

**Second postulate:** the laws of physics are of same form for all observers in relative uniform motion.

**Third postulate:** velocity of light as a limit velocity.

**Definitions (3-1-7):**

(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

$$\phi: \mathbb{M} \rightarrow (\mathbb{R}^4, \eta)$$

which induces an isometry on the difference space.

(4) The light cone in $T_p \mathbb{M}$ is the subset

$$\text{LC}_p = \{ x \in T_p \mathbb{M} | \eta_p(x, x) = 0 \}$$

**Lemma (3-1-8):**

Let $\Lambda: \mathbb{M} \rightarrow \mathbb{M}$ be a diffeomorphism that preserves the light cones. Then $\Lambda$ is an affine mapping. For the induced map on tangent space, one has

$$\eta_p(\Lambda_p x, \Lambda_p y) = a(\Lambda) \eta_p(x, y)$$

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 $\text{SO}(3, 1)$.

**Remarks (3-1-9):**

(1) Consider the bijection of vectors in $\mathbb{R}^4$ to two-dimensional hermitian matrices

$$\Phi: (x^0, x^1, x^2, x^3) \mapsto \begin{pmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{pmatrix}$$
and

\[ \det \Phi(x) \Phi(y) = -(x, y) \]

Lie group \( SL(2, \mathbb{C}) \) acts on hermitian matrices as \( H \leftrightarrow SHS^* \). This action induces an isomorphism of

\[ \text{PSL}(2, \mathbb{C}) = SL(2, \mathbb{C})/\{ \pm 1 \} \]

and the proper orthochronous Lorentz group.

(2) Using the parity transformation

\[ P = \text{diag}(+1, -1, -1, -1) \]

and time reversal

\[ T = \text{diag}(-1, +1, +1, +1) \]

we can write every element of \( L \) uniquely in the form

\[ \Lambda = P^n T^m \Lambda_0 \text{with} \, n, m \in \{0, 1\}, \Lambda_0 \in L_+^1 \]

**Definitions (3-1-10):**

(1) A non-zero vector \( x \) in the difference space for Minkowski space \( \mathbb{M} \) is called

- time like, if \( x^2 = (x, x) < 0 \)
- light like or null, if \( x^2 = (x, x) = 0 \)
- space like, 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 Minkowskispace 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.

**Lemmas (3-1-11):**

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.

**Definitions (3-1-12):**

1. A velocity unit vector \(\hat{w}\) is a future directed timelike 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 $O_{\alpha \hat{w}}$ or sometimes even with $O_{\hat{w}}$, 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 $O_{\hat{w}}$ is given by the observer-dependent time function

$$\Delta t_{\hat{w}}(x, y) = (\hat{w}, x - y).$$

**Remarks (3-1-13):**

(1) For any two observers $O_{\hat{w}}$ and $O_{\hat{v}}$ 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 $O_{\hat{w}}$ and $y$ happens before $x$ for $O_{\hat{v}}$. This fact might be called the “relativity of simultaneity.”

(2) The future $I^+(x)$ of an event $x \in \mathbb{M}$ equals

$$I^+(x) = \{ y \in \mathbb{M} | (\hat{v}, y - x) < 0 \text{ for all observers } O_{\hat{v}} \}$$

i.e. the set of events that happens after $x$ for all initial observers. This justifies the qualifier absolute future for $I^+(x)$.

**Observation (3-1-14):**

Consider two observers $O_{\hat{v}}$ and $O_{\hat{w}}$ 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 $x - y = t \in \mathbb{R}$ with some $t \in \mathbb{R}$. For the time elapsed between these two events, the observer $O_{\hat{w}}$ 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 $O_\theta$ for the events $x \in \mathbb{M}$ and $y \in \mathbb{M}$ which is

$$\Delta t_\theta (x, y) = |(\hat{v}, x - y)|$$

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:

$$a^2 = (x - y)^2 + 2(x - y, \hat{v})^2 + (x - y, \hat{v})^2(\hat{v}, \hat{v})$$

$$= (x - y)^2 + (x - y, \hat{v})^2 = -t^2 + t^2(\hat{v}, \hat{w})^2 > 0.$$

We have thus

$$x - y = -\Delta t_\varphi(x, y) \hat{v} + a$$

which implies

$$\Delta t_\varphi(x, y)^2 = |(\hat{v}, x - y)|^2 + a^2 = \Delta t_\theta(x, y)^2 + a^2$$

Hence

$$|\Delta t_\theta (x, y)| > |\Delta t_\varphi(x, y)|.$$

Thus the moving observer $O_\theta$ measures a strictly longer time interval than the observer $O_\varphi$ at rest.
Section (3-2): Electrodynamics as a gauge theory

We start our discussion with some comments on electrodynamics on Minkowski space $\mathbb{M}$.

Remarks (3-2-1):

(1) Consider Maxwell’s equations on a star-shaped region $U \subset \mathbb{M}$ on which from the homogeneous Maxwell equations $dF = 0$ and the fact that Minkowski space is contractible, we conclude that there exists a one-form $A \in \Omega^1(\mathbb{M})$ such that $dA = F$. 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

$$A' = A + d\lambda \in \Omega^1(\mathbb{M})$$

also obeys the equation $dA' = F$. 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.

Observations (3-2-2):

(1) For a geometric interpretation, we consider on the disjoint union
\[ \tilde{L} = \bigsqcup_{\alpha \in I} (U_\alpha \times \mathbb{C}) = \bigcup_{\alpha \in I} (U_\alpha \times \{\alpha\} \times \mathbb{C}) \]

the equivalence relation

\[(x, \alpha, g_{\alpha\beta} z) \sim (x, \beta, z).\]

Then the quotient

\[ L = \tilde{L} / \sim \]

is a smooth manifold and comes with a natural projection to \( M \) which provides a complex line bundle \( \pi : L \to M \).

(2) This bundle comes with a linear connection which locally is

\[ \nabla_\alpha = d + \frac{1}{i} A_\alpha \]

Locally, we can consider the two form

\[ F_\alpha = dA_\alpha \]

which is trivially closed. On twofold overlaps, we have

\[ F_\alpha - F_\beta = dA_\alpha - dA_\beta = d(\delta A)_{\alpha\beta} = d \log g_{\alpha\beta} = 0 \]

so the locally defined two-forms \( F_\alpha \) patch together into a globally defined two-form which we interpret as the electromagnetic field strength.
Remark (3-2-3):

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

$$E^i = -\frac{\partial \Phi}{\partial x^i} - \partial_t A^i \quad \text{and} \quad B^i = \text{rot} A$$

where we are using three-dimensional notation. Then the equations of motion for the Lagrangian

$$l(x, x_t, t) = \frac{m}{2} x_t^2 - q(x_t^i \Lambda_i(x, t) + \Phi(x, y))$$

are because of

$$\frac{\partial l}{\partial x^i} = -q x_t^2 \frac{\partial \Lambda^i}{\partial x^i} - q \frac{\partial \Phi}{\partial x^i}$$

given by

$$m \ddot{\phi}^i - q \frac{\partial \Lambda^i}{\partial x^j} \phi^j = -q \dot{\phi}^i \frac{\partial \Lambda^i}{\partial x^t} - q \frac{\partial \Phi}{\partial x^i} - q \partial_t A^i$$

which is just

$$m \ddot{\phi}^i = q (\dot{\phi} \wedge B)^i + q E^i$$

and thus the Lorentz force.
Remarks (3-2-4):

(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 $l(M)$, we take

$$dF = 0 \text{ and } d^* F = 0.$$ 

(We consider here for simplicity the case without external sources; external source requires relative cohomology.) It obviously has waves as classical solutions: any solution obeys $dF = 0$ and $\delta F = 0$ and is thus harmonic.

Given any solution, one can define its electric flux

$$[F] \in H^l_{dR}(M)$$

and its magnetic flux

$$[*F] \in H^{n-l}_{dR}(M)$$

(2) If one wants to introduce an action, one has to break electric magnetic duality and use one of the equations, say $dF = 0$, 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, is 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 $d^* F[A,g] = 0$. 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 $\lambda$ the symmetry $S[A+d\lambda] = S[A]$. For a general local
function, we neglect its dependence on derivatives. Differential relations between the equations of motion are then introduced by the relation \( dF[A] = 0 \).

We summarize in this subsection some important aspects of general relativity in the form of several postulates.

**Observations (3-2-5):**

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 para-compact.

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.

We discuss the postulates.

(i) Local causality.

(ii) Local conservation of energy momentum.

(iii) Field equation: Einstein’s equation read:

\[
(R_{ab}[g] - \frac{1}{2} R[g] g_{ab}) + \Lambda g_{ab} = \frac{8\pi G}{c^4} T_{ab}
\]  

(3.1)
Lemma (3-2-6):

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

$$g_p^{(1)} = \lambda(p) \cdot g_p^{(1)}$$

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 $\lambda$

$$0 = g(X + \lambda Y, X + \lambda Y) = g(X,X) + 2 \lambda g(X,Y) + \lambda^2 g(Y,Y)$$

has two roots

$$\lambda_1, \lambda_2 = \frac{g(X,Y)}{g(Y,Y)} \pm \frac{\sqrt{g(X,Y)^2 - g(X,X)g(Y,Y)}}{g(Y,Y)}$$

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.

Now suppose that $W,Z \in T_pM$ and $W + Z$ are not light-like. Then:

$$g(W,Z) = \frac{1}{2} (g(W,W) + g(Z,Z) - g(W + Z, W + Z))$$.
Each of the terms on the right hand side can be compared to either X or Y and is thus fixed.

**Proposition (3-2-7):**

The tensor $T$ is a symmetric covariantly conserved tensor: if evaluated on any field configuration $\psi$ 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 $l = L d\text{vol}_g$

$$I = \int_D l = \int_{\Phi(D)} l = \int_D \Phi^*(l)$$

and thus $Z$

$$\int_D l - \Phi^*(l) = 0$$

If the diffeomorphism $\Phi$ is generated by a vector field $X$, we have

$$L \times l = \int_D l$$
We find

$$\int_{D} L_X(l \, d \text{vol}_g) = \int_{D} \left( \frac{\partial L}{\partial \psi} - D_e \left( \frac{\partial L}{\partial \psi_e} \right) \right) L_X \psi \, d \text{vol}_g + \int_{D} T^{ab} L_X g_{ab} \, d \text{vol}_g$$

The first term vanishes due to the equations of motion. For the second term, we need the Liederivative of the metric

$$L_X g_{ab} = 2X_{(a,b)}$$

which equals the symmetrization of the covariant derivative. Thus

$$0 = T^{ab} L_X g_{ab} \, d \text{vol}_g = 2 \left( \int_{D} \varepsilon^{ab} X_a \right) \, d \text{vol}_g$$

The first term can be transformed into an integral over $\partial D$ which vanishes since the vectorfield $X$ vanishes on $\partial X$. Since this identity holds for all vector fields $X$, we have covariant conservation of the energy momentum tensor, $T^b_a$

**Remarks (3-2-8):**

(1) The equation (3.1) 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

$$S[g, \Phi] = \int_{D} \left( R[g] - 2\Lambda \right) \text{vol}_g + l(\Phi, g)$$

where $\Phi$ stands symbolically for other fields in the theory.
Chapter (4)
Hamiltonian mechanics

Section (4-1): (Pre-) symplectic manifolds

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 $(V, \omega)$ consisting of a $k$-vector space $V$ and a non-degenerate two-form $\omega \in \Lambda^2 V$, i.e. an anti-symmetric 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 $(e_i)_{i=1}^{2n}$ and its dual basis $(e^*_i)_{i=1}^{2n}$. 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 \( \varphi \in \text{GL}(V) \) of a symplectic real vector space \((V, \omega)\) that preserve \( \omega \), i.e. \( \varphi^* \omega = \omega \), is a (non-compact) Lie group \( \text{Sp}(V) \) 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.

**Definitions (4-1-3):**

Let \( M \) be a smooth manifold.

(1) A 2-form \( \omega \in \Omega^2(M) \) is called a pre-symplectic form, if it is closed, \( d\omega = 0 \), 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 \( \omega \) 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 pre-symplectic 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 \( \omega^n \) of the symplectic form \( \omega \) 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 \( \omega^n \). Since volume preserving smooth maps
have a Jacobian of determinant1, they are local diffeomorphisms. Symplectic
maps between two symplectic manifolds of the same dimension are thus local
diffeomorphisms.

**Definition (4-1-5):**

Let M be a smooth manifold of any dimension. The symplectic manifold
\((T^*M, \omega_0)\) is called the canonical phase space associated to the configuration
space manifold M.

**Remarks (4-1-6):**

(1) Since the canonical phase space comes with the canonical one-form \( \theta \), 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 compactsymplectic manifolds are quite important for mathematical
physics, but more difficult toobtain.

**Theorem (4-1-7): (Darboux’ theorem)**

Let \((M, \omega)\) be a \((2n + k)\)-dimensional presymplectic manifold with rank \( \omega = 2n \).
Then we can find for any point \( m \in M \) a neighborhood \( U \) and a local coordinate chart

\[
\psi : U \to \mathbb{R}^{2n+k}
\]

\[
\psi_{(u)}(q^1, \ldots, q^n, p_1, \ldots, p_n, \eta^1, \ldots, \eta^k)
\]

such that \( \omega|_U = \sum_{i=1}^{n} dp_i \wedge q^i \). Such coordinates are called Darboux coordinates or canonical coordinates. One can choose the covering such that the coordinate changes are canonical transformations.

**Remarks (4-1-8):**

(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 \( n > 1 \).

**Proposition (4-1-9):**

Let \( (M, \omega) \) be 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

\[
\text{df}(Y) = \omega(X_f, Y)
\]

for all local vector fields \( Y \) on \( M \), or, equivalently, \( \text{df} = \iota_{X_f} \omega \).
Definitions (4-1-10):

(1) Let $f$ be a smooth function on a symplectic manifold $(M, \omega)$. The vector field $X_f$ such that $df = \iota_{X_f} \omega$ is called the symplectic gradient of $f$.

(2) A vector field $X$ on a symplectic manifold $(M, \omega)$ 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 $\iota_X \omega = df$, 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 $(M, \omega)$ is called 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 $t$.

Proposition (4-1-11):

Let $(M, \omega)$ 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.

Lemma (4-1-12):

Let $X$ be any vector field on a symplectic vector space $(V, \omega)$. 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 $v \in V$

$$\omega(X_v, w) = dh_v(w) \text{ for all } w \in V.$$  

We differentiate both sides as a function of $v \in V$ in the direction of $u$ and find

$$\omega(DX_v(u), w) = D^2h_v(u, w).$$

Here $D^2h_v$ is the Hessian matrix of second derivatives in the point $v$. Since second derivatives are symmetric, this expression equals

$$D^2h_v(u, w) = D^2h_v(w, u) = \omega(DX_v(w), u) = -\omega(u, DX_v(w)).$$

Lemma (4-1-13):

A linear vector field is Hamiltonian, if and only if it is skew symplectic, i.e. if and only if

$$\omega(v, Aw) = -\omega(Av, w) \text{ for all } v, w \in V.$$  

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

$$DX_v(u) = \lim_{t \to 0} \frac{A(v+tu)}{t} = Au$$

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 bilinear pairing given by $\omega(A \cdot, \cdot)$, we find

$$\langle dhv, u \rangle = \frac{1}{2} (\omega(Au, v) + \omega(Av, u)) = \frac{1}{2} (-\omega(u, Av) + \omega(Av, u)) = \omega(Av, u)$$

**Lemma (4-1-14):**

Let $X$ be any vector field on a symplectic vector space $(V, \omega)$. 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 $D_{X_v}$ is skew symplectic for all $v \in V$. Then consider the function

$$h(v) = \int_0^1 \omega(X_{tv}, v) dt$$

on $V$. Then $h$ is a Hamiltonian function for $X$, since we have for all $u \in V$

$$\langle dhv, u \rangle = \int_0^1 \omega(D_{X_{tv}}(tu), v) + \omega(X_{tv}, v) dt$$

$$= \omega \int_0^1 (tD_{X_{tv}}(u), v) + \omega(X_{tv}, v) dt$$

$$= \omega \int_0^1 (tD_{X_{tv}}(v) + X_{tv}, u))dt$$

$$= \omega \left( \int_0^1 \frac{d}{dt} (X_{tv}, u) \right) = \omega(X_v, u)$$
Section (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.

Definitions (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 $·$ and $\{−, −\}$, with the following properties
   
   (i) $(P, ·)$ is an associative algebra.
   
   (ii) $(P, \{−, −\})$ is a Lie algebra.
   
   (iii) The bracket $\{·, ·\}$ provides for each $x ∈ P$ a derivation on $P$ for the associative product:
   
   $\{x, y · z\} = \{x, y\}z + y\{x, z\}$.
   
   The product $\{·, ·\}$ is also called a Poisson bracket. Morphisms of Poisson algebras are $k$-linear maps $Φ : P → P'$ that respect the two products,

   $\{ Φ(v), Φ(w)\} = Φ(\{v, w\})$

   And $Φ(v) · Φ(w) = Φ(v · w)$ for all $v, w ∈ P$.

2. A Poisson manifold is a smooth manifold $M$ with a Lie bracket $\{−, −\}$:

   $C^∞(M) × C^∞(M) → C^∞(M)$

   such that the algebra of smooth functions $(C^∞(M), \{−, −\})$ 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
$\Phi^* : C^\infty(M_0) \to C^\infty(M)$ is an isomorphism of Poisson algebras.

**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 $(M, \omega)$. We use the symplectic gradient to define the following Poisson bracket

$$\{f, g\} = -\langle dg, x_f\rangle = \omega(X_f, X_g)$$

for $f, g \in A$. From the last equation in observation (4.1.10), we find in local Darboux coordinates

$$\{f, g\} = d_g(X_f) = \frac{\partial g}{\partial P_i} \frac{\partial f}{\partial q^i} - \frac{\partial g}{\partial q^i} \frac{\partial f}{\partial P_i}$$

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 $\omega^{ij}$ 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.

$$\{ \Phi^* f, \Phi^* g \} = \Phi^* \{ f, g \} \quad \text{for all } f, g \in C^\infty(N, \mathbb{R}).$$

Observations(4-2-3):

(1) We discuss Poisson manifolds in more detail. For any smooth function $h \in C^\infty(M)$ on a Poisson manifold $M$, the map

$$\{ h, - \} : C^\infty(M) \to C^\infty(M)$$

$$f \mapsto \{ h, f \}$$

is a derivation and thus provides a global vector field $X_h \in \text{vect}(M)$ with

$$X_h(f) \equiv df(X_h) = \{ h, f \}.$$

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.

$$\frac{d\varphi}{dt} \big|_{t=t_0} = X_h(\varphi(t_0)).$$

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:

$$\frac{d}{dt} f \circ \varphi(t) = df \left( \frac{d\varphi}{dt} \right) = df(X_h) \left|_{\varphi_t = \{ h, f \}(\varphi(t))} \right..$$
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.

**Definitions (4-2-4):**

(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 \(\varphi_t\) 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 \((M, \omega)\), together with a function

\[ h : M \to \mathbb{R}, \]

We consider \(M\) with the Poisson structure induced by the symplectic structure \(\omega\).
Remarks (4-2-5):

(1) Since the Poisson bracket is anti-symmetric, \(\{h, h\} = 0\) 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 \(h\),

\[
\{h, f\} = 0 \text{ and } \{h, g\} = 0 .
\]

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\((C^\infty(M), \{\cdot, \cdot\})\).
Remarks (4-2-6):

(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.
Proposition (4-2-7):

Let \((M, \omega, h)\) 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.

Proposition (4-2-8): (Poincare’s recurrence theorem)

Let \((M, \omega, h)\) be a Hamiltonian system with the property that for any point \(p \in M\) the orbit \(O_p\) is bounded in the sense that it is contained in a subset \(D \subseteq M\) of finite volume. Then for each open subset \(U \subseteq M\), there exists an orbit that intersect the set \(U\) infinitely many times.

Remark (4-2-9):

The time evolution \(s : I \to M\) of a particle in a Hamiltonian system \((M, \omega, h)\) 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 = dimM. At time $t$, we then have a probability density $p_t \in \Omega^{2n}(M)$ with

$$\varphi_{-t}^* p_t = p_0$$

Suppose for simplicity that we have $p_t = f_t \omega^{\Lambda n}$ with $f_t : M \to \mathbb{R}$ a smooth function and $\omega^{\Lambda n}$ the Liouville volume on $M$. Then $f_t = (\varphi_{-t})^* f_0$ and thus

$$\left. \frac{d}{dt} \right|_{t=t_0} f_t = LX_h f_0 = \{h, f_0\},$$

so that the Poisson bracket also describes the evolution of probability distributions.
Section (4-3): Hamiltonian dynamics and Legendre Transform

Now we will discuss time dependent Hamiltonian dynamics.

Observations (4-3-1):

(1) We want to consider time dependent Hamiltonian systems. To this end, we consider a surjective submersion \( E \rightarrow I \) with \( I \subset \mathbb{R} \) an interval and require for each \( t \in I \) the fiber \( E_t = \pi^{-1}t \subset E \) to be a symplectic manifold. The manifold \( E \) is called the evolution space of the system.

(2) Since the evolution space \( E \) is odd-dimensional, it cannot be a symplectic manifold any longer. Rather, \( E \) is a pre-symplectic 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 pre-symplectic manifolds with higher dimensional leaves.

(3) 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 pre-symplectic formulation has been proposed in particular to deal with systems in which no natural parametrization is known, e.g. for massless relativistic particles.

Examples (4-3-2):

(1) We take a symplectic manifold \((M, \omega_0)\) and obtain a pre-symplectic manifold \( I \times M \) with pre-symplectic form \( \text{pr}_2^*\omega_0 \). This cannot be the correct pre-symplectic form, since the leaves of the pre-symplectic 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

\[ \omega_h = \text{pr}^* \omega_0 - dh \wedge dt \in \Omega^2(I \times M). \]

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 pre-symplectic manifold \( (I \times M, \omega_h) \).

(3) Consider local Darboux coordinates \((q^i, p_i)\) on \( M \). Then

\[ \omega_0 = \sum_{i=0}^n dp_i \wedge dq^i \]

and, with the summation convention understood, in local coordinates \((q^i, p_i, t)\) on \( I \times M \)

\[ \omega_h = dp_i \wedge dq^i - \frac{\partial h}{\partial p_i} dp^i \wedge dt - \frac{\partial h}{\partial q^i} dq^i \wedge dt \]

It is easy to check that the nowhere vanishing vector field \( X_h \)

\[ (X_h)|_{(p,q,t)} = \frac{\partial h}{\partial p_i} \frac{\partial}{\partial p_i} - \frac{\partial h}{\partial q^i} \frac{\partial}{\partial q^i} + \frac{\partial}{\partial t} \]

obeys

\[ \iota_{X_h} \omega_h = 0 \]

and thus is tangent to the leaves.
Definitions(4-3-3):

(1) A time dependent Hamiltonian system consists of a symplectic manifold \((M, \omega_0)\) and a smooth function

\[ h : I \times M \to \mathbb{R}, \]

called the (time-dependent) Hamiltonian function. Introduce the vector field \(\overline{X}_h \in \text{vect}(I \times M)\) in \(p \in I \times M\) as the sum

\[ \overline{X}_h = X_h + \frac{\partial}{\partial t} \]

where \(X_h\) is the symplectic gradient of \(h\) on the slice. The integral curves of \(\overline{X}_h\) describe the physical trajectories for the Hamiltonian function \(h\).

(2) The pre-symplectic manifold \((I \times \overline{M}, \omega - dh \wedge dt)\) is called evolution space of the system. The space of leaves is called the phase space of the system. It has the structure of an asymplectic 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.

Observations(4-3-4):

(1) Consider a Hamiltonian system \((M, \omega, h)\). 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 phase flow for \(h\)

\[ \varphi_t : M \to M \]
on \( M \) leads to a family of diffeomorphisms on evolution space \( E \) which parameterize the leaves

\[
t' \mapsto (\varphi_t(x), t' + t).
\]

We already know that the one-forms \( \Theta \) and \( \Theta_E \) 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 \( d\Theta = \omega \).

**Proposition (4-3-5):**

Let \( Y_1 \) and \( Y_2 \) be two curves in evolution space \( E \) that encircle the same leaves. Then we have

\[
\int_{Y_1} \Theta_E = \int_{Y_2} \Theta_E
\]

The one-form \( \Theta_E \) is called Poincare-Cartan integral invariant.

**Proof:**

Let \( \Sigma \) be the surface formed by those parts of the leaves intersecting \( Y_1 \) (and thus \( Y_2 \)) that is bounded by the curves, \( \partial \Sigma = Y_2 - Y_1 \). Such a surface is called a flux tube. Stokes’ theorem implies

\[
\int_{Y_2} \Theta_E = \int_{Y_1} \Theta_E = \int_{\partial (\Sigma)} \Theta_E = \int_{\Sigma} d\Theta_E = \int_{\Sigma} \omega_E = 0
\]

since \( \Sigma \) consists of leaves whose tangent space is by definition the kernel \( \ker \omega_E \) of the pre-symplectic form on evolution space.
Observations (4-3-6):

(1) We now specialize to curves in evolution space $M$ at constant time, $\gamma_i \in p_i^{-1}(t_i)$, to find

$$\int_{\gamma_1} \Theta = \int_{\gamma_2} \Theta$$

(2) Consider now an oriented two-chain $\sum \in M$ such that $\partial \sum = \gamma$. Stokes’ theorem then implies that

$$\int_{\sum} \omega = \int_{\sum} d\Theta = \int_{\partial \sum} \Theta = \int_{\gamma} \Theta$$

is invariant under the phase flow. We have derived this result earlier.

In the following we will discuss the Legendre Transform.

Observations (4-3-7):

(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) 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 (4-3-8):

Let $U \subset V$ be a convex subset of a real vector space. Given a convex function $f : U \to \mathbb{R}$, there 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 $f$. 

90
Remarks (4-3-9):

(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

\[
g(p) = m \frac{-1}{\alpha-1} \frac{p^\beta}{\beta} \text{ with } \frac{1}{\alpha} + \frac{1}{\beta} = 1
\]

As special cases, we find with \( \alpha = 2 \) for the function

\[
f(x) = \frac{m}{2} x^2 \text{ the Legendre transform}
\]

\[
g(p) = \frac{p^2}{2m}, \quad \text{and for } m = 1 \text{ for the function}
\]

\[
f(x) = \frac{x^\alpha}{\alpha} \text{ the Legendre transform}
\]

\[
g(p) = \frac{p^\beta}{\beta}
\]

with \( \frac{1}{\alpha} + \frac{1}{\beta} = 1 \)

(2) The definition of the Legendre transform via a maximum implies the inequality

\[
F(x, p) = xp - f(x) \leq g(p)
\]
for all values of \( x, p \) where the functions \( f \) and \( g \) are defined. A function and its Legendre transform are thus related by

\[
p x \leq f(x) + g(p).
\]

Applying this to the second example just discussed, we find the classical inequality

\[
P x \leq \frac{1}{\alpha} x^\alpha + \frac{1}{\beta} p^\beta, \text{ with } \frac{1}{\alpha} + \frac{1}{\beta} = 1
\]

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

\[
||f||_\alpha = \left( \int_M |f|^\alpha \right)^{\frac{1}{\alpha}}
\]

and similarly \( g \in L^\beta(M) \) and \( M \) a measurable space.

**Observations (4-3-10):**

(1) Given a Lagrange function \( l : I \times TM \to \mathbb{R} \), we define the Legendre transformation

\[
\Lambda : I \times TM \to I \times T^*_M
\]

for \( (t, q, q_t) \in I \times T_q M \) as the element \( (t, \Lambda_{t,q}(q_t)) \), where the element \( \Lambda_{t,q}(q_t) \in T^*_q M \) is defined by

\[
\langle \Lambda_{t,q}(q_t), \omega \rangle = \frac{d}{d\epsilon} \bigg|_{\epsilon=0} l(t, q, q_t + \epsilon \omega) = \frac{\partial l}{\partial q_t} \omega^t
\]
(2) In many cases of interest, $\Lambda$ 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 Legendre transform of $l$:

![Diagram]

This can be used as a Hamiltonian function for a time-dependent Hamiltonian system on $I \times T^*M$.

**Theorem (4-3-11):**

Let $M$ be a smooth manifold. Let $l : I \times TM \to \mathbb{R}$ be a time-dependent Lagrangian function such that the Legendre 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 Legendre transform of $l$:

$$\begin{array}{ccc}
I \times TM & \xrightarrow{\Lambda} & I \times T^*M \\
\Lambda & & \\
\downarrow & & \downarrow \\
l & & h \\
\mathbb{R} & & \mathbb{R}
\end{array}$$

This sets up a bijection of Lagrangian and Hamiltonian systems with a bijection of classical trajectories: A section $: I \to I \times M$ is a solution of the Euler Lagrange equations for $l$, if and only if the Legendre transform of its jet prolongation

$$\Lambda \circ j^1 \varphi : I \to I \times TM \to I \times T^*M$$

obeys the Hamilton equations for $h$. 
References:


B. Herbert Goldstein; Charles Poole and John Safko: Classical mechanics, Pearson/Addison Wesley, Upper Saddle River, 2002.


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


L. Some of our treatment of electrodynamics follows the review:

M. 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

N. Rolf Berndt: An introduction to symplectic geometry, American Mathematical Society, Providence, 2001