QUANTUM FIELD THEORY COURSE
VERSION 03

IVAN MIRKOVIĆ

Contents

Part 0. Introduction 4
  0.1. Usefulness of QFT 4
  0.2. Classical Mechanics 6
  0.3. Quantum Mechanics 7
  0.4. Quantum Field Theory 8

Part I. Classical Mechanics 9
  0. Intro 9
  0.1. Mechanical systems 9
  0.2. Lagrangian and Hamiltonian formulation 9
  1. Lagrangian approach 10
    1.1. Manifolds 10
    1.2. Differentiation 11
    1.3. Calculus of Variations on an interval 11
    1.4. Lagrangian reformulation of Newton’s equation 13
  2. Hamiltonian approach 14
    2.1. Metrics: linear algebra 14
    2.2. The passage to the cotangent vector bundle via the kinetic energy metric 15
    2.3. Legendre transform 15
    2.4. The canonical momentum 17
    2.5. Hamiltonian reformulation of Newton’s equation 17
    2.6. Poisson structures 19

Date: Spring 2000.
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.7.</td>
<td>Canonical formalism</td>
<td>19</td>
</tr>
<tr>
<td>2.8.</td>
<td>Summary</td>
<td>20</td>
</tr>
<tr>
<td>2.9.</td>
<td>Extra: Symplectic structures</td>
<td>20</td>
</tr>
<tr>
<td>3.</td>
<td>Example: Equivalence of three actions for a <strong>Free particle</strong></td>
<td>22</td>
</tr>
<tr>
<td>3.1.</td>
<td>(A) Newtonian setting: the kinetic energy action</td>
<td>22</td>
</tr>
<tr>
<td>3.2.</td>
<td>(B) Relativistic setting: the spacetime trajectory length action</td>
<td>23</td>
</tr>
<tr>
<td>3.3.</td>
<td>(C) Relativistic setting: kinetic action with a worldline metric</td>
<td>24</td>
</tr>
<tr>
<td>3.4.</td>
<td>The philosophy of adding an auxiliary field to the action</td>
<td>27</td>
</tr>
</tbody>
</table>

**Part II. Quantum Mechanics**

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.</td>
<td>Intro</td>
<td>28</td>
</tr>
<tr>
<td>0.1.</td>
<td>Dirac axioms</td>
<td>30</td>
</tr>
<tr>
<td>1.</td>
<td>Principles of Quantum Mechanics</td>
<td>31</td>
</tr>
<tr>
<td>1.1.</td>
<td>Amplitudes</td>
<td>31</td>
</tr>
<tr>
<td>1.2.</td>
<td>Hilbert space setting (the Hamiltonian view)</td>
<td>31</td>
</tr>
<tr>
<td>1.3.</td>
<td>Feynman integrals (the Lagrangian view)</td>
<td>32</td>
</tr>
<tr>
<td>2.</td>
<td>Amplitudes</td>
<td>34</td>
</tr>
<tr>
<td>2.1.</td>
<td>Shooting electrons through holes leads to: probability, complex numbers, uncertainty principle and the dual nature of particles</td>
<td>34</td>
</tr>
<tr>
<td>2.2.</td>
<td>The amplitude principles</td>
<td>35</td>
</tr>
<tr>
<td>2.3.</td>
<td>Uncertainty principle: a quantitative version</td>
<td>36</td>
</tr>
<tr>
<td>3.</td>
<td>Operators: the Hamiltonian approach</td>
<td>38</td>
</tr>
<tr>
<td>3.1.</td>
<td>Deformation quantization: a mathematical notion of quantization</td>
<td>38</td>
</tr>
<tr>
<td>3.2.</td>
<td>Stern-Gerlach filtering apparata leads to Hilbert space structure on the space of states</td>
<td>40</td>
</tr>
<tr>
<td>3.3.</td>
<td>Time evolution of a quantum system is given by the Hamiltonian operator $\hat{H}$</td>
<td>42</td>
</tr>
<tr>
<td>3.4.</td>
<td>Example: a single electron in a vector space</td>
<td>43</td>
</tr>
<tr>
<td>3.5.</td>
<td>Example: Harmonic oscillator</td>
<td>46</td>
</tr>
<tr>
<td>4.</td>
<td>Path integrals: the Lagrangian approach</td>
<td>50</td>
</tr>
<tr>
<td>4.1.</td>
<td>Matrix coefficients of $\hat{f}$ for the normal ordering quantization</td>
<td>51</td>
</tr>
<tr>
<td>4.2.</td>
<td>Localization of the amplitude on the time interval</td>
<td>53</td>
</tr>
<tr>
<td>5.</td>
<td>Example: <strong>Relativistic free particle</strong></td>
<td>56</td>
</tr>
</tbody>
</table>
5.1. Amplitude calculation
5.2. Gauge fixing (= Equivariant integration)
5.3. Gauge fixing of the metric $g$ and length as a Teichmüller parameter (leftover parameter)
5.4. Amplitudes

Part III. (Quantum) Field Theory

0. Intro

1. Classical Field Theory: Calculus of Variations
   1.1. Variation $\frac{\delta S}{\delta x}$ and pointwise variation $\frac{\delta S}{\delta x(t)}$ of the action
   1.2. Borcherds Formalism

2. Perturbative expansion of Feynman integrals by Feynman graphs

Appendix A. Geometry

A.1. Anomaly
A.2. Spinors
Part 0. Introduction

Text.

- R.P. Feynman, *Quantum Mechanics*, volume 4 of *Feynman’s lectures on Physics.*

0.1. **Usefulness of QFT.** Quantum Field Theory is a framework for thinking about the microscopic structure of the world – the nature and behavior of elementary particles. *Our goal* is not the frontiers of physics but the *usefulness of the QFT ideas in mathematics.* Importing ideas from QFT to mathematics was the dominant trend in recent past. However we will only learn the basics of QFT and leave the mathematical applications for some other opportunity.

During a period of superficial relations between math and physics, physicists thought deeply about some specific problems and (unobserved) developed a number of ideas that later proved useful in mathematics.

0.1.1. **Some applications to mathematics.**

- Low dimensional topology
  - (1) Witten’s Chern-Simons invariants of, 3d manifolds,
  - (2) Seiberg-Witten invariants of 4d-manifolds.
- Algebraic Geometry: Mirror symmetry.
- Representation theory: Vertex algebras.
- Computing: Quantum computing.

0.1.2. **Two fundamental methods.** There are two basic formalisms used in physics. We will notice how they propagate from Classical to Quantum Mechanics and Quantum Field Theory (in the example of strings):

(1) **Lagrangian approach**

Feynman integrals calculate the probability $\langle b, a \rangle$ for the system to pass from the state $a$ to the state $b$ as the the sum of contributions from all possible histories:

$$\int \text{all ways } x \text{ to pass from } a \text{ to } b \quad \text{probability that } x \text{ will happen}$$

(2) **Hamiltonian approach**

Deformation quantization of algebras of operators

The relation of the two approaches is:

*Feynman integrals are matrix coefficients of operators.*
Mathematics has difficulties with either method, but the situation with (2) has recently improved due to Kontsevich. In (1) there is no mathematical understanding what the Feynman integrals should really mean\(^1\). The measures are not known, and if they were the integrals would be likely to diverge, and there are claims that whatever we do our expectations for the precise meaning of Feynman integrals are self-contradictory. One way is to imagine that these integrals are only the visible part of some structure finer then just calculating a number. This is a complete million \$ mystery.

0.1.3. The ability of to write down some “functions” via the Lagrangian formalism. Physicists can write down some functions on sets \(S\) that are mysterious to mathematicians. Such \(S\) is usually the moduli of something (roughly the set of isomorphism classes of something), such as: all smooth manifolds of a certain dimension, all elliptic curves, all Calabi-Yau manifolds of a certain dimension.

The usefulness of such functions:

- They may distinguish different points in \(S\).
  For instance, if \(S\) is the set of isomorphism classes of knots in \(\mathbb{R}^3\), if we know a function \(Z\) on \(S\) then \(Z(K_1) \neq Z(K_2)\) implies that the knots \(K_1\) and \(K_2\) are different. This applies to genetics since the basic structure of different DNA is that they are knotted in a different way.
  We call such functions invariants because one often thinks of a knot \(K\) in presence of some additional data (a choice of a projection to a plane), then “invariance” means that \(Z\) does not depend on these auxiliary choice.
- Detecting “dualities”.
  For instance on the moduli of Calabi-Yau manifolds of dimension \(n\), one can define two invariants \(Z_A\) and \(Z_B\) (called the A and B models). It has been observed that for each Calabi-Yau \(M\) there seems to exist another Calabi-Yau \(N\) such that \(Z_A(M) = Z_B(N)\). Then \(M\) and \(N\) are said to be mirror partners. This was spectacular because (i) features about \(M\) that are very deep and difficult to understand are often easy to read off from some simple features of \(N\), (ii) it points out an undreamed of relation between familiar geometric objects.
- Construction of interesting functions.

0.1.4. String duality before string theory: Langlands duality. The Langlands duality between reductive algebraic groups, \(G \leftrightarrow \hat{G}\), is a deep mystery in representation theory, number theory and algebraic geometry. With the development of string theory, it appears now that this is a particular manifestation of a string duality ideas conjectured by physicists.

In the remainder we survey the notes.

\(^1\)except in some simplest cases: the Gaussian integrals
0.2. Classical Mechanics. The evolution of a mechanical system is viewed as a path $x : \mathbb{R} \to \mathcal{C}$ in the space $\mathcal{C}$ of all possible configurations (states) of the system. This path satisfies Newton’s equation of motion $a = F/m$, which is a second order differential equation

$$\dddot{x} = F(x, \dot{x}, t).$$

For instance, if the force is conservative and independent of velocity and time, this becomes

$$\dddot{x} = \frac{dV}{dx}$$

for the potential $V(x)$. In particular, the situation is completely deterministic: if one knows position and velocity at one moment then there is only one possible evolution.

The two main approaches to the study of Newton’s equation of motion are the Lagrangian and Hamiltonian formulation. These are two geometric ways to think of our differential equations that uncover more symmetries of the situation.

0.2.1. Lagrangian approach to Classical Mechanics. It is also referred to as the path approach (since the main heroes are the paths of possible evolutions of the system) or the Calculus of Variations (for its mathematical underpinning). The main idea is to view Newton’s equation as a technical manifestation of the principle

**physical system evolves so that a certain quantity $S[x]$ is (locally) minimal among all possible evolutions $x$.**

Here one thinks of the curve $x$ in $\mathcal{C}$ in terms of the velocity curve $(x, \dot{x})$ in the tangent bundle $T\mathcal{C}$ to the configuration space $\mathcal{C}$. Newton’s differential equation $\dddot{x} = F(x, \dot{x}, t)$, i.e., the expression $F$, is here encoded as a function $L$ on $T\mathcal{C} \times \mathbb{R}$, called the Lagrangian of the situation. $L(x, \dot{x}, t)$ is typically (?) the difference of the kinetic and potential energy $L = T - V$.

Now the action $S[x]$ is the time integral of the Lagrangian:

$$S[x] \overset{\text{def}}{=} \int_{t_1}^{t_f} dt \ L(x(t), \dot{x}(t), t).$$

**Related mathematical ideas.** Critical points of functions are studied in Morse Theory (global structure of spaces), and Stationary Approximation Method (oscillating integrals $\int e^{iS(x)} \ dx$).

0.2.2. Hamiltonian approach to Classical Mechanics. It is also referred to as the Canonical formalism, her the word “canonical” indicates the relation with the cotangent bundle (as in “canonical transforms”). We think of it as a next step after the Lagrangian approach which lives in the tangent bundle $T\mathcal{C}$ of the configuration space. The idea is that one replaces $T\mathcal{C}$ with the cotangent bundle $T^{\ast}\mathcal{C}$ which has an additional geometric structure, the Poisson structure.
The passage from $T\mathcal{C}$ to $T^*\mathcal{C}$ is based on the interpretation of the kinetic energy as a metric $g_T$ on $\mathcal{C}$ (i.e., on the vector bundle $T\mathcal{C}$), and on the ideas of the Legendre transform (from functions on $T\mathcal{C}$ to functions on $T^*\mathcal{C}$ and the Legendre map (the vertical differential of a function on $T\mathcal{C}$ is viewed as a map from $T\mathcal{C}$ to $T^*\mathcal{C}$).

The coordinates on $T^*\mathcal{C}$ are traditionally denoted $(q,p)$ with $q = x$ the position in $\mathcal{C}$ and $p$ the momentum. One replaces the velocity curve $(x, \dot{x})$ in the tangent bundle $T\mathcal{C}$, by the momentum curve $(q,p) = (x,p)$ in the cotangent bundle $T^*\mathcal{C}$, using the identification given by the kinetic energy metric $g_T$. The Lagrangian function $L$ on $T\mathcal{C}$ is replaced by the Hamiltonian function $H$ on $T^*\mathcal{C}$, here $H = \mathcal{L}(L)$ is the Legendre transform of $L$. $H$ is typically (?) the total energy of the system $H = T + V$, i.e., kinetic + potential energy.

Now Poisson structure on $T^*\mathcal{C}$ associates to the function $H$ a vector field $\tilde{H} = \{H, -\}$ on $T^*\mathcal{C}$, and Newton’s equation of motion gets reformulated into:

The evolution of any observable $f \in C^\infty(T^*\mathcal{C})$ along the momentum curve $(Q(t), p(t))$ is governed by the Hamiltonian vector field: $\frac{df}{dt} = \{H, f\} = \tilde{H}(f)$.

So the system is governed by its energy distribution on $T^*\mathcal{C}$. If one applies this to the coordinate function $q_i, p_i$ we get Hamiltonian equations for the evolution of the momentum curve

$$\dot{q} = -H_p \quad \text{and} \quad \dot{p} = H_q.$$

**Related mathematical ideas.** The basic geometric feature used in the Hamiltonian approach to classical mechanics is the Poisson geometry, or more specially the Symplectic Geometry. The study of the Hamiltonian equations lead to the theory of Integrable systems (more formally: “Completely Integrable Systems”). This refers to particularly nice systems of PDEs that are not interesting for applied mathematicians since very few PDEs are completely integrable, but are a staple of physics and mathematics since the PDEs in these worlds are often completely integrable as a reflection of the beautiful organization of the world with many hidden structures and interrelations.

0.3. **Quantum Mechanics.**
### 0.4. Quantum Field Theory

#### Overview:

<table>
<thead>
<tr>
<th>Theory</th>
<th>Features</th>
<th>Lagrangian formulation</th>
<th>Hamiltonian formulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical Mechanics</td>
<td>$\mathbb{R}^x \rightarrow \mathcal{C}$</td>
<td>$\frac{\partial}{\partial t_i} [t_i, t_f] (x, \dot{x}, t) \rightarrow C(T \mathcal{C}, L \in C^\infty(T \mathcal{C})$</td>
<td>$\mathbb{R}^{(q,p)} \rightarrow T^* \mathcal{C}$</td>
</tr>
<tr>
<td>Deterministic Mechanics</td>
<td>$\dot{x} = F(x, \dot{x}, t)$</td>
<td>$S \in C^\infty(\text{Map}[t_i, t_f], \mathcal{C})$</td>
<td>$H = \mathcal{L}(L) \in C^\infty(T^* \mathcal{C})$</td>
</tr>
<tr>
<td></td>
<td>Stochastic</td>
<td>$d_x S = 0$ i.e., $L_x = \frac{d}{dt} L \dot{x}$</td>
<td>$\frac{d}{dt} f = {H, f} = H(f)$</td>
</tr>
<tr>
<td>Quantum Mechanics</td>
<td>$\text{Amplitude} = \text{path integral}$</td>
<td>$\text{Amplitude} = \text{matrix coefficient}$</td>
<td></td>
</tr>
</tbody>
</table>
Part I. Classical Mechanics

The classical mechanics is governed by Newton’s first law $ma = F$ which is mathematically a second order ordinary differential equation. There are two traditional approaches which geometrize the study of this Newton equation, the Lagrange and the Hamiltonian approaches. Both carry over to quantum mechanics, QFT and therefore also to strings.

0. Intro

0.1. Mechanical systems.

0.1.1. **Evolution of a system as a path in the configuration space.** Our basic space $S$ will be an affine space over certain $\mathbb{R}^p$ (we will usually say that $p = 3$). Its meaning is that these are possible positions of one object. The configuration space for $n$ objects (i.e., $n$-tuples of points in $S$) is then $C = S^n$. The *evolution* of a system of $n$ points is then a *path in the configuration space*, i.e., a map $x : \mathbb{R} \rightarrow C$ from the *time line* to the configuration space. It is an $n$-tuple $x = (x_1, ..., x_n)$ with $x_i$ mapping $\mathbb{R}$ to $S$.

0.1.2. **The nature of physical laws in classical mechanics.** The evolution is governed by the *equation of motion*, which we also call *Newton’s equation*. It is of the form

\[ \text{acceleration} = \text{force/mass}, \text{ i.e., } \ddot{x} = F(x, \dot{x}, t). \]

Notice that asking for the equation to be of this form is essentially the same as asking that

If one knows $x$ and $\dot{x}$ at $t_0$, then $x(t)$ is determined at any time $t$.

In particular, this encodes the *Deterministic Nature* of the classical mechanics.

A specific system is now described by a specific force, i.e., a specific differential equation of the form $\ddot{x} = F(x, \dot{x}, t)$. The basic example (globally conservative force) is when the force is a gradient vector field of a potential function

\[ m_i \ddot{x}_i = \frac{dV}{dx_i} \]

for the potential energy function $V = V(x)$ on $C$.

0.2. Lagrangian and Hamiltonian formulation. These are two geometric ways to think of our differential equations that uncover more symmetries of the situation.

The *Lagrangian approach* is to study the velocity curve $\dot{x}$ in the tangent bundle $TC$ to the configuration space $C$, rather then $x$ itself. The mathematical formalism here is based on the idea that the system evolves in a way which minimizes certain quantity.
The *Hamiltonian formulation* replaces $TC$ with the cotangent bundle $T^*C$, and the velocity curve with the momentum curve. The advantage is that any cotangent bundle has an additional geometric structure, the Poisson structure.

1. Lagrangian approach

1.0.1. *Physical Laws as criticality equations.* The main idea here is that Newton’s equation of motion means that the evolution $x$ of the system develops in such a way that a certain quantity, the *action* $S[x]$ of $x$, is minimal possible.

In order to construct the action $S[x]$, one starts with a function $L$ on the tangent bundle $TC$, called the Lagrangian function (or just “Lagrangian”). This is (typically?) the difference between the kinetic and potential energy

$$L \overset{\text{def}}{=} T - V.$$ 

Then the action $S[x]$ is a time integral of the Lagrangian function $L(x, \dot{x})$.

Geometrically, the physical setting is now described by a function $L$ on the tangent bundle $TC$. So, the physics happens on $TC$, and therefore the time evolution is naturally thought of as a velocity curve $(x, \dot{x})$ in $TC^2$. So, the reformulation of the Newton equation for $x$ in terms of the velocity curve, happens to be the criticality equation, of the “action functional” $S[x]$ on the space of all velocity curves $(x, \dot{x})$ in $TC$ (i.e., all curves $x$ in $C$).

1.0.2. *Functionals.* A function on a space of functions is called a functional and one denotes it by $S[x]$ rather then $S(x)$, just to remember that $x$ itself is a function. The differential of a functional $S$ with respect to a function $x$ is called variation and denoted

$$\frac{\delta S}{\delta x} \overset{\text{def}}{=} d_x S.$$ 

1.0.3. *Calculus of variations.* This is the underlying mathematical machinery, i.e., the calculation of the criticality equation (the *Euler-Lagrange equation*)

$$\frac{\delta S}{\delta x} = 0.$$ 

So, we will calculate the variation $\frac{\delta S}{\delta x}$ of the action functional given as a time integral of a Lagrangian function.

Mathematically, this is the differentiation of a function $S$ on an infinite dimensional manifold, so we review the the differential calculus on such manifolds.

1.1. *Manifolds.* The basic assumption of mathematical physics is that the configuration space is a manifold. So we review the basic manifold notions.

---

2This means that we now replace $x$ by a curve $(x, v)$ in $TC$ and therefore we view the 2nd order Newton equation $\ddot{x} = F(x, y, t)$ in $x$ as a first order system: $\dot{x} = v, \dot{v} = F(x, y, t)$ on $TC$, i.e., a flow on $TC$. 
1.1.1. **Definition.**

1.1.2. **Smooth maps.**

1.1.3. **(Co)tangent spaces.**

1.1.4. **Vector bundles.** Examples: $TM$ and $T^*M$.

1.1.5. **The differential** $d_aF = F'(a)$.

1.2. **Differentiation.** The criticality equation for the action functional $S[x]$ asks that the differential $d_xS$ vanishes. Since $x$ varies through an infinite-dimensional manifold of smooth functions, we need the calculus is infinite dimensional setting.

1.2.1. **Normed vector spaces.**

1.2.2. **Derivatives of maps between normed vector spaces.**

1.2.3. **Affine spaces over vector spaces.**

1.2.4. **Comparison of notions of a differential.**

1.3. **Calculus of Variations on an interval.** We are interested in the critical points of a functional $S[x]$ on the space of all paths $x : \mathbb{R} \to \mathcal{C}$ in a space $\mathcal{C}$. We will call $S[x]$ the *action* of the path $x$. The action functional will be a “time” integral

$$S[x] \overset{\text{def}}{=} \int_{t_1}^{t_2} L(x(t), \dot{x}(t), t) \, dt,$$

of a function $L(x,v,t)$ which is a (possibly time dependent) function of the position $x$ in $\mathcal{C}$, and velocity $v \in T_x\mathcal{C}$. So the Lagrangian $L(x,v,t)$ is a function on $T\mathcal{C} \times \mathbb{R}$.

Below we calculate the criticality equation (Euler-Lagrange equation) of this situation.

We will revisit this in the section III.1 where the maps are allowed to have a higher dimensional source.

1.3.1. **Paths.** Precisely, we are interested in the space $C^\infty(I,M)$ of smooth paths from an interval $I = [\tau_i, \tau_f]$, to a Riemannian manifold $M$. Here we consider the action of the form

$$S[x] = \int_I d\tau \, L(x, \dot{x}, \tau).$$

For simplicity we will assume $M$ to be a vector space. The parameter $\tau$ may be time (but need not). The dot above a letter denotes the differentiation with respect to $\tau$. 
1.3.2. **Theorem.** The variation (differential) of the action $S[x] = \int_I d\tau \, L(x(\tau), \dot{x}(\tau), \tau)$, in the direction of $u \in T_x(C^\infty(I, M)) = C^\infty(I, M)$, is

$$\frac{\delta S}{\delta x} u = [L_\dot{x}(x, \dot{x}, \tau) u(\tau)]_{\tau_f}^{\tau_i} + \int_I d\tau \, [L_x(x, \dot{x}, \tau) - \frac{d}{d\tau} L_\dot{x}(x, \dot{x}, \tau)] \cdot u$$

$$= \Delta(L_\dot{x}) u + \int_I d\tau \, (L_x - \frac{d}{d\tau} L_\dot{x}) \cdot u.$$ (b) The boundary term does not appear in the following cases:

- (i) If we consider paths with fixed choices of end points.
- (ii) If $x$ is a loop $x: S^1 \to M$ and $L = L(x, \dot{x})$ does not depend on time.

(c) When the boundary term does not appear, the EL-equation (criticality equation) is

$$L_x(x, \dot{x}, \tau) = \frac{d}{d\tau} L_\dot{x}(x, \dot{x}, \tau), \text{ i.e., } L_x = \frac{d}{dt} L_\dot{x}.$$ **Proof.** (a) In general a tangent vector $u \in T_x(C^\infty(I, M))$ is the section of $TM$ along $x$, i.e., a section of the pullback $x^* TM$. When $M$ is a vector space this is the same as $u \in C^\infty(I, M)$. The variation $\frac{\delta S}{\delta x}$ of $S$ at $x$ and with respect to $u$, means the differential

$$\frac{\delta S}{\delta x} u = (d_x S) u \overset{\text{def}}{=} \frac{d}{d\varepsilon} S(x + \varepsilon u) = \int_I d\tau \frac{d}{d\varepsilon}|_{\varepsilon=0} L(x + u, \dot{x} + \dot{u}, \tau)$$

$$= \int_I d\tau \, L_x(x, \dot{x}, \tau) u + L_\dot{x}(x, \dot{x}, \tau) \dot{u}.$$ It remains to use integration by parts for the "dynamic" ingredient $\dot{u}$.

(b) In the case (i) one has $u = 0$ on $\partial I$. In (ii), everything is periodic with respect to $\tau_f - \tau_i$.

1.3.3. **Pointwise variation** $\frac{\delta S}{\delta x(t)}$. Now we consider the role of the variation with respect to the value of $x$ at $t$. So, let $t \in I - \partial I$ and $v \in T_{x(t)} M$. We are calculating the variation of $S[x]$ when one moves $x$ by a distributional path $u = \delta_t(\tau) \cdot v$ concentrated at the point $t \in I$, so

$$\frac{\delta S}{\delta x(t)} v \overset{\text{def}}{=} \delta_t(\tau) \cdot v = (d_x S) \delta_t(\tau) \cdot v.$$ **Corollary.** The pointwise variation is the integral kernel for the variation operator $\frac{\delta S}{\delta x}$

$$\frac{\delta S}{\delta x(t)} = [L_x - \frac{d}{d\tau} L](x(t), \dot{x}(t), t), \quad t \in (t_i, t_f).$$

**Proof.** When $t$ is not an end point the boundary term vanishes and $(d_x S) \delta_t(\tau) v$ is

$$\int_I d\tau \, [L_\dot{x}(x(\tau), \dot{x}(\tau), \tau) - \frac{d}{d\tau} L_\dot{x}(x(\tau), \dot{x}(\tau), \tau)] \cdot \delta_t(\tau) v = [L_x - \frac{d}{d\tau} L](x(t), \dot{x}(t), t) \cdot v.$$
In terms of component functions and the Einstein summation convention this is
\[ \frac{\delta S}{\delta x(t)} v = L_{x^\mu}(x(t), \dot{x}(t), t) v^\mu - \frac{d}{d\tau} L_{\dot{x}^\mu}(x(t), \dot{x}(t), t) v^\mu. \]

1.3.4. Remarks. (a) The variation with respect to \( x \) consists of two parts, due to the appearance of \( x \) as a non-dynamical variable (i.e., the appearance of the field \( x \) itself without derivatives), and as a dynamical variable (i.e., the appearance of the derivative \( \dot{x} \) of the field). The non-dynamical term of the variation has no time derivative, while the dynamical term has time derivative with a minus (since the dynamic variable causes integration by parts).

(b) In general, i.e., when we have to allow the boundary term, EL-equation consists of two parts: \( L_x = \frac{d}{dt} L_{\dot{x}} \) and \( L_{\dot{x}}(x(\tau), \dot{x}(\tau), \tau) = 0 \) at the ends \( \tau \in \partial I = \{ \tau_i, \tau_f \} \). (For the first claim we use all \( u \)'s supported in \( (\tau_i, \tau_f) \) and the continuity at the ends. Then for the second we just vary the values of \( u \) at the ends.)

1.3.5. The case \( S(x) = \int_{t_1}^{t_2} L(x(t)) \, dt \).

1.4. Lagrangian reformulation of Newton's equation.

1.4.1. Newton's equation as a flow on \( TC \). Instead of a curve \( x \) into \( C \) one considers the corresponding velocity curve \( (x, \dot{x}) \) into the tangent bundle \( TC \). The equation on \( TC \) is of the 1st order, so it gives a flow on \( TC \).

1.4.2. Lemma. In the standard case of the conservative force, i.e., \( m\ddot{x} = -\frac{dV}{dx} \), one can indeed reformulate Newton's equation as the Euler-Lagrange equation.

Proof. The corresponding Lagrangian function \( L \in C^\infty(TC) \), is
\[ L(x(t), \dot{x}(t)) = T(\dot{x}(t)) - V(x(t)) = \frac{1}{2}m\dot{x}(t)^2 - V((x(t)). \]

The action is the time integral of the Lagrangian
\[ S[x] \overset{\text{def}}{=} \int_{\tau_i}^{\tau_f} dt \, L(x(t), \dot{x}(t), t) = \int_{\tau_i}^{\tau_f} dt \, \frac{1}{2}m\dot{x}(t)^2 - V((x(t)). \]

Since, \( L_x = -\frac{dV}{dx} \) and \( L_{\dot{x}} = m\ddot{x} \), the criticality equation \( L_x = \frac{d}{dt} L_{\dot{x}} \) is indeed
\[ -\frac{dV}{dx} = -\frac{d}{dt} m\ddot{x}, \quad \text{i.e., } m\ddot{x} = \frac{dV}{dx}. \]

1.4.3. Terminology in the presence of a Lagrangian function \( L \) on \( TC \).
1.4.4. Examples of Lagrangians.

- (a) Mass falling to Earth: \( L = \frac{1}{2}m\dot{x}^2 - mgh \).
- (b) Harmonic oscillator: \( L = \frac{1}{2}(m\ddot{x}^2 - mx^2) \).

2. Hamiltonian approach

Also called the “canonical formalism”. Here “canonical” indicates the cotangent bundle setting (as in “canonical transforms”).

Newton’s equation of motion has last been seen in the Lagrangian formalism as the Euler-Lagrange (criticality) equation for the velocity curve in the tangent bundle to the configuration space. We will now reformulate Newton’s equation in terms of the momentum curve in the symplectic variety \( T^*C \) (the “phase space”). This “Hamiltonian” equation says that the evolution of the momentum curve is given by the Hamiltonian vector \( \tilde{H} \) field on \( T^*C \), i.e., the vector field associated to the Hamiltonian function \( H \) via the symplectic structure on \( T^*C \).

Technically, the passage from \( TC \) to \( T^*C \) uses the ideas of the Legendre map and the Legendre transform, and a natural metric \( g_T \) on \( TC \) given by the kinetic energy \( T \). canonical metric. The function \( H \) is the phase space incarnation of the Lagrangian function \( L \) on the tangent bundle \( TC \), explicitly, \( H = \mathcal{L}(L) \) is the Legendre transform of \( L \). In the standard case when \( L \) is the difference of the kinetic and potential energy, \( H \) is the total energy \( H = T + V \) of the system.

The canonical momentum \( p \) is a \( T^*C \) version of the momentum \( m\dot{x} \in TC \). It is defined as the Legendre map of the Lagrangian \( L \) (i.e., the differential of \( L \)).

The kinetic energy metric \( g_T \) is used to move the velocity curve \( (x(t), \dot{x}(t)) \) in \( TC \) to the momentum curve \( (q, p(t)) \) in \( T^*C \).

In terms of the coordinates \( (q, p) \) = (position, momentum) on \( T^*C \), the Euler-Lagrange equation \( L_x = \frac{\partial}{\partial \dot{x}} L_{\dot{x}} \) is first reformulated as the Hamiltonian equation for the evolution of the momentum curve

\[
\dot{p} = -H_q \quad \text{and} \quad \dot{q} = H_p,
\]

and this is elegantly written in terms of the Poisson structure on \( T^*C \) as

\[
\dot{p} = \{H, q\} \quad \text{and} \quad \dot{q} = \{H, p\}.
\]

This gives a uniform formula for the evolution of any observable \( f \in \mathcal{C}^\infty(T^*C) \) along the momentum curve:

\[
\dot{f} = \{H, f\} = \tilde{H}f.
\]

The classical usefulness of the Hamiltonian formulation was the large group of symmetries (symmetries of Newton’s equation are given by the Galilean group).

2.1.1. *Inner products and quadratic forms.*

*Lemma.* (a) An inner product \((-,-)\) on a vector space \(V\) is the same as a positive non-degenerate quadratic form \(q\) on \(V\).

(b) It gives an isomorphism of vector spaces \(\iota: V \xrightarrow{\cong} V^*\), \(\iota: v \mapsto (v,-)\), which takes any orthonormal bases \(v_i\) to the dual bases \(p_i\). Then \(q = \sum p_i^2\).

2.1.2. *Metrics on vector bundles.* A metric \(g\) on a vector bundle \(V/X\) gives an isomorphism \(\iota_g: V \xrightarrow{\cong} V^*\).

2.1.3. *Metrics on a manifolds.* A Riemannian manifold \((M,g)\) is a manifold \(M\) with a metric \(g\) on \(M\), i.e., on the vector bundle \(TM\). On a Riemannian manifold \((M,g)\) one has the notions:

- (a) length of curves,
- (b) distance, i.e., a metric space structure,
- (c) geodesics.

2.2. *The passage to the cotangent vector bundle via the kinetic energy metric.* We will think of kinetic energy as a metric on the configuration space \(C\), i.e., on its tangent bundle \(TC\).

2.2.1. *Metric and kinetic energy.* The formula for the kinetic energy \(T = \sum m_i (\dot{x}_i)^2 / 2\) uses the standard metric on the configuration space \(C \cong \mathbb{R}^{np}\).

However, one can also think of the quadratic form \(T\) on the tangent bundle \(TC\) as a natural metric \(g_T\) on \(C\)

\[
g_T(x, \dot{x}) \overset{\text{def}}{=} (\dot{x}, \dot{x})_T \overset{\text{def}}{=} \sum m_i (\dot{x}_i)^2, \quad g_T(x; \dot{x}, \dot{y}) \overset{\text{def}}{=} \sum m_i \dot{x}_i \cdot \dot{y}_i.
\]

In this way, mass influences the geometry of the space.

The corresponding isomorphism of vector bundles \(\iota_{g_T}: TC \xrightarrow{\cong} T^*C\) is given on \(T\_xC\) by

\[
\dot{x} \mapsto (\dot{x}, -)_T = \left( \sum m_i \dot{x}_i, - \right),
\]

for the standard metric on \(C = \mathbb{R}^{3n}\).

2.2.2. *Momentum curve \((q,p): \mathbb{R} \rightarrow T^*C\).* This is the image of the velocity curve under the identification \(g_T: TC \xrightarrow{\cong} T^*C\), given by the kinetic energy metric \(g_T\) on \(C\). So, \(t \mapsto (q(t), p(t))\) where \(q(t) = x(t)\) is the position, and \(p(t)\) is the canonical momentum

\[
p(t) = \left( \sum m_i \dot{x}_i(t), - \right) = \frac{\partial L}{\partial \dot{x}}(x(t), \dot{x}(t), t).
\]

2.3. Legendre transform.
2.3.1. The differential of \( f \in C^\infty(V) \), viewed as a Legendre map \( f' : V \to V^* \). For a function \( f \) on a manifold \( M \), its differential \( f' \) is a section of \( T^*M \). However, if \( M \) is a vector space \( V \) we can view it as a “Legendre map” \( f' : V \to V^* \). (For \( f \in C^\infty(V) \), the differential at \( a \), \( f'(a) \) \( = \) \( d_a f \) \( = \) \( T_a(V) \to T_{f(a)} \mathbb{R} \) is a linear function on \( T_a v \cong V \), hence \( f'(a) \in V^* \).)

This will generalize the map of vector spaces \( V \cong V^* \) that comes from an bilinear form \((-,-)\) on \( V \) (think of it as a quadratic form on \( V \)).

If \( V \) is a vector bundle over a manifold \( M \) then a function \( f \) on \( V \) defines a fiberwise Legendre map \( \partial f : V \to V^* \) is calculated fiber by fiber, i.e., for \( m \in M \), this is the partial derivative in the vertical direction \( \partial f \mid_{V_m} = f|_{V_m} \). The differential at \( a \), \( f'(a) \) \( = \) \( d_a f \) \( = \) \( T_a(V) \to T_{f(a)} \mathbb{R} \) is linear in \( f \) \( \mid_{V_m} \). Actually, since \( V = \mathbb{R} \) it is obvious from the graph.

Examples. Observe that \( f \to f' \) is linear in \( f \). The map \( f' \) associated to a linear function \( f = p \in V^* \) is a constant function \( f' = p \) from \( V \) to \( V^* \). The map associated to a constant function is \( c' = 0 \).

Lemma. (a) For a positive definite quadratic form \( q \), map \( (\frac{q}{2})' : V \to V^* \) is the isomorphism of vector spaces given by the corresponding inner product \((-,-)\).

(b) If \( f \) is a strictly convex function then \( f' \) is a bijection.

Proof. (a) \( q = \sum p_i^2 \) for some basis \( p_i \) of \( V^* \) and then \( (u,v) = \sum p_i(u) \cdot p_i(v) \). Now, \( (\frac{q}{2})'(a) = \sum p_i(a) \cdot p_i = (a,-) \).

(b) is a generalization of (a). When \( V = \mathbb{R} \) it is obvious from the graph.

2.3.2. Legendre transform \( \mathcal{L} \) of functions on \( V \) to functions on \( V^* \). Actually, \( \mathcal{L}f \) is defined only when the Legendre map \( f' : V \to V^* \) is a bijection, since the definition uses the inverse of the Legendre map:

\[
(\mathcal{L}f)(p) \overset{\text{def}}{=} \langle v, p \rangle - f(v) \mid_{v = (f')^{-1}p} = \langle v, (f')^{-1}p \rangle - f((f')^{-1}p), \quad p \in V^*.
\]

The geometric meaning is in

Lemma. Suppose that \( f \) is strictly convex. Then \( \mathcal{L}(f) \) is defined and \( \mathcal{L}(f)(p) \) is the maximal height of the graph of the linear function \( p \) on \( V \) above the graph of the convex function \( f \).

Proof. The height at \( v \in V \) is \( h(v) = \langle v, p \rangle - f(v) \). The height has a unique critical point since \( 0 = d_v h = p - f'(v) \) means that \( v = (f')^{-1}p \). Actually, since \( f \) is strictly convex, so is the height function \( h \) and therefore \( h \) has a (unique) maximum.

2.3.3. Lemma. (Properties.) (a) For a positive definite quadratic form \( q \) on \( V \),

\[
\mathcal{L}(\frac{q}{2}) = \frac{Q}{2},
\]
where the positive definite quadratic form $Q$ on $V^*$ is obtained by transporting $q$ to $Q$ by the isomorphism $\iota_q : V \rightarrow V^*$ defined by $q$.

(b) $\mathcal{L}(f + c) = \mathcal{L}f - c$.

(c) $\mathcal{L}(cf) (p) = (\mathcal{L}f)(c^{-1}p)$.

(d) $\mathcal{L}(f + q) (p) = (\mathcal{L}f)(p - q)$ for $q \in V^*$.

(e) If $f'$ is a bijection then so is $g'$ for $g = \mathcal{L}f$.

(f) $\mathcal{L}(\mathcal{L}f) = f$.

Proof. (a) $\iota_q = (\frac{q \cdot v}{2})'$, or equivalently, $\iota_q(v) = (v, -)$ for the corresponding inner product.

2.3.4. Question. For $f$ convex and even, when is the Fourier transform of $e^{-f/2}$ equal to $e^{-\mathcal{L}f/2}$?

2.4. **The canonical momentum.** A choice of a Lagrangian function $L$ on $TC$ defines a notion of a canonical momentum

$$p = \frac{\partial L}{\partial \dot{x}}.$$ 

So, the canonical momentum $p$ is the variation (change) of the Lagrangian function corresponding to an infinitesimal change of velocity – our dynamical variable. (one also says that $\frac{\partial L}{\partial \dot{x}}$ is the momentum conjugate to the position function $x$.

Formally, the canonical momentum is the Legendre map (see 2.3.1)

$$TC \xrightarrow{p} T^*C$$

for the Lagrange function $L$ on $TC$. (The partial derivative $\frac{\partial L}{\partial \dot{x}}(x, \dot{x})$ is the differential $d_L(L|T_xC)$ of a function on a vector space $T_xC$, so it is a point of $T^*_xC$.)

2.4.1. **Comparison with the naive momentum (when the potential energy does not depend on velocity !)** In this case the canonical momentum is the change of the kinetic energy corresponding to an infinitesimal change of velocity:

$$p = \frac{dL}{d\dot{x}} = \frac{dT}{d\dot{x}} - \frac{dV}{d\dot{x}} = \frac{dT}{d\dot{x}} - \sum m_i(\dot{x}_i, \ddot{x}_i)/2 = (\sum m_i \dot{x}_i, -).$$

Here the canonical momentum (defined from the Lagrangian), coincides with the naive momentum up to the interpretation as a covector. Observe also that in this case, the canonical momentum map $TC \xrightarrow{p} T^*C$, is precisely the isomorphism given by the kinetic energy metric $g_T$ on $TC$.

Lastly, we will denote by “$T$ on $T_x^*C$”, the kinetic energy function transplanted to a function on $T_x^*C$ by the above isomorphism.

2.5. **Hamiltonian reformulation of Newton’s equation.**
2.5.1. *Hamiltonian function* $H$ on $T^*\mathcal{C}$ as the Laplace transform $\mathcal{L}L$ of the Lagrangian function $L$ on $T\mathcal{C}$. In order to define the Hamiltonian function on $T^*\mathcal{C}$ by

$$H \overset{\text{def}}{=} \mathcal{L}L,$$

we need to assume that the restrictions $L|T_x\mathcal{C}$ to any tangent space has invertible Legendre map. In practice $L|T_x\mathcal{C}$ will be strictly convex:

$$L|T_x\mathcal{C} = \text{a positive definite quadratic form } T \text{ plus a constant } -V(x).$$

2.5.2. *Lemma.* If $L$ is the difference between the kinetic and potential energy then $H$ is the total energy

*Proof.* Let $x \in \mathcal{C}$ and $p \in T^*_x\mathcal{C}$. Observe that $q = 2T$ is a positive definite quadratic form on $T_x\mathcal{C}$, and denote by $Q$ the positive definite quadratic form on $T^*_x\mathcal{C}$, obtained by transporting $q$ via the isomorphism $V \mapsto V^*$ given by $q$ then

$$H(x,p) = \mathcal{L}(L)(x,p) \overset{\text{def}}{=} \mathcal{L}(L|T_x\mathcal{C})(p) = \mathcal{L}(\frac{q}{2} - V(x))(p) = \frac{Q}{2} + V(x) = \text{“}T\text{” on } T^*_x\mathcal{C},$$

where we denote by “$T$” = “$T$ on $T^*\mathcal{C}$” the kinetic energy function transplanted to a function on $T^*\mathcal{C}$ by the same isomorphism (recall that $\dot{x} \mapsto (\sum m_i\ddot{x}_i, -)$).

2.5.3. *Hamiltonian equations for the momentum curve.*

*Lemma.* In terms of the momentum curve $\mu(t) \overset{\text{def}}{=} (q(t), p(t)) : \mathbb{R} \to T^*\mathcal{C}$, the Euler-Lagrange equation for the velocity curve becomes

$$\dot{p}(t) = -\partial_1 H(q(t), p(t), t) \quad \text{and} \quad \dot{q}(t) = \partial_2 H(q(t), p(t), t).$$

Less precisely,

$$\dot{p} = -H_q \quad \text{and} \quad \dot{q} = H_p.$$  

[A terrible crime has been committed in the last line since the symbols $p$ and $q$ have two meanings. On the RHS $p$ is a certain function of time, a component of the momentum curve, on the RHS $p$ is a coordinate on the phase space $T^*\mathcal{C}$.]

*Proof.* [We will only consider the case when $L = T(\dot{x}) - V(x)$.] While the evolution curve $x : \mathbb{R} \to \mathcal{C}$ satisfies Newton’s equation, the velocity curve $(x, \dot{x}) : \mathbb{R} \to T\mathcal{C}$ satisfies Euler-Lagrange equation and one definition equation

$$\frac{\partial L}{\partial x}(x(t), \dot{x}(t), t) = \frac{d}{dt} \frac{\partial L}{\partial \dot{x}}(x(t), \dot{x}(t), t), \quad \text{and} \quad \dot{x}(t) = \frac{dx}{dt}.$$

By the definition of the momentum curve $p(t) \overset{\text{def}}{=} \partial_2 L(x(t), \dot{x}(t))$ so EL equation is reformulated as $\dot{p}(t) = \frac{d}{dt} \partial_2 L(x(t), \dot{x}(t)) = \partial_1 L(x(t), \dot{x}(t))$. Now we move that to $T^*\mathcal{C}$

$$\dot{p}(t) = \partial_1 L(x(t), \dot{x}(t)) = \frac{\partial}{\partial x} (T(\dot{x}) - V(x)) \big|_{x=x(t), \dot{x} = \dot{x}(t)} = -\frac{\partial}{\partial x} V(x) \big|_{x=x(t), \dot{x} = \dot{x}(t)}$$
\[
= -\frac{\partial}{\partial x} \left( \left\langle T(p) + V(x) \right\rangle \right) \bigg|_{x=x(t), \ p=p(t)} = -\frac{\partial}{\partial x} H(x(t), p(t)).
\]

Also,
\[
\partial_2 H(q(t), p(t)) = \partial_2 \left( \left\langle T(p) - V(x) \right\rangle \right) \bigg|_{x=x(t), \ p=p(t)} = d_{p(t)} \left\langle T \right\rangle = \dot{x}(t).
\]

The last equality is the computation of the Lagrangian map for the function \( \left\langle T \right\rangle \big|_{x=x(t), \ p=p(t)} \) at the point \( p(t) \).

Now recall that we start with a positive definite quadratic form \( T \) and that \( \left\langle T \right\rangle \) is its Lagrangian transform, i.e., “the dual” positive definite quadratic form on the dual vector space. It remains to observe that the Legendre maps for dual quadratic forms are mutually inverse.

2.6. **Poisson structures.** We will interpret the Hamiltonian equations above in terms of the Poisson structure, a standard geometric structure on the cotangent bundle.

2.6.1. **Lie algebras.** (a) Commutator in an associative algebra. (b) Vector fields. (c) Lie groups.

2.6.2. **Poisson structures.** (a) On a commutative algebra \( A \). (b) On a manifold \( M \).

(c) The canonical Poisson structure on \( T^*M \). If \( M = \mathbb{R}^n \) then
\[
\{f, g\} \overset{\text{def}}{=} \sum_i f_{p_i} g_{q_i} - f_{q_i} g_{p_i},
\]

hence
\[
\{p_i, -\} = \partial_{q_i}, \quad \{q_i, -\} = -\partial_{p_i} \quad \text{and} \quad \{p_i, q_j\} = \delta_{ij}.
\]

(d) Function \( f \) on \( M \) defines a vector field \( \tilde{f} \overset{\text{def}}{=} \{f, -\} \) on \( M \).

2.7. **Canonical formalism.** Using the canonical Poisson structure on \( T^*C \), Hamiltonian equations are rewritten as
\[
\dot{p}(t) = \{H, p\}(\mu(t), t) \quad \text{and} \quad \dot{q}(t) = \{H, q\}(\mu(t), t),
\]

i.e.,
\[
\dot{p} = \{H, p\} \quad \text{and} \quad \dot{q} = \{H, q\}.
\]

2.7.1. **Observables.** We call \( T^*C \) the phase space (its principal property is the Poisson structure). We call functions on the phase space observables since \( \phi \in C^\infty(T^*C) \) can be observed any time \( t \), i.e., it defines a function \( \phi(t) \overset{\text{def}}{=} \phi(\mu(t)) \) of time. The evolution of \( \phi(t) \) is always given by the Hamiltonian
\[
\dot{\phi}(\mu(t), t) = \{H, \phi\}(\mu(t), t), \quad \text{i.e.,} \quad \dot{\phi} = \{H, \phi\}.
\]

(This is true for the coordinate functions \( p_i, q_i \) hence for all polynomials in them etc.)

This is the final formulation of the canonical formalism.
2.7.2. **Hamiltonian vector field.** This is the vector field $\tilde{H}$ on $T^*\mathcal{C}$ associated to the Hamiltonian function. In terms of this vector field Hamiltonian equations become

$$\dot{\mu}(t) = \tilde{H}(\mu(t)),$$

i.e., the solutions are integral curves of the flow defined by the Hamiltonian vector field.

2.7.3. **Harmonic oscillator example.** There $L = T(\dot{x}) - V(x) = \frac{m}{2}\dot{x}^2 - \frac{k}{2}x^2$, hence

$$H = T(\dot{x}) + V(x) = \frac{1}{2m}\dot{p}^2 + \frac{k}{2}q^2$$

and

$$\tilde{H}(f) = \{H, f\} = \{\frac{1}{2m}\dot{p}^2 - \frac{k}{2}q^2, f\} = \frac{1}{2m}2p\{p, f\} + \frac{k}{2}2q\{q, f\} = \frac{1}{m}p\partial_q f - kq\partial_p f.$$ 

So, $\tilde{H} = \frac{1}{m}p\partial_q - kq\partial_p$. If $k = m = 1$ then $\tilde{H} = p\partial_q - q\partial_p$ and the integral curves are circles. In general they are the ellipses $\frac{1}{m}p^2 + kq^2 = c, \ c \geq 0$.

2.8. **Summary.** We proved equivalence of several differential equations. The bulk of the time went into defining the framework for the quantities involved - manifolds, tangent and cotangent bundles. This is equivalent to specifying how these quantities transform under the change of coordinates which is the classical language that is standard in physics.

If our system of points has some constraints, its configuration space (all possible positions) will be a manifold $M$ (say a submanifold of $\mathcal{C} = \mathbb{S}^n$). Still, the Lagrangian formalism applies with $L$ a function on $TM$.

The original Newton’s law $F = ma$ looks differently in different coordinate systems, i.e., it is not invariant under the transformations of space other then the Galilean group (translations, rotations and ?). The Hamiltonian equation (canonical formalism) is invariant under a large class of transformations of $T^*\mathcal{C}$ (rather then only $\mathcal{C}$) – the ones that preserve the Poisson bracket (canonical transformations or symplectomorphisms). Reducing the problem to the simplest form by a choice of a symplectic coordinate change is said to be the strongest method in classical dynamics (Arnold).

2.9. **Extra: Symplectic structures.**

2.9.1. **Differential forms.** A $p$-form $\omega$ on a manifold $M$ is a smooth family $\omega_m, m \in M$, with each $\omega_p$ an alternating multi-liner form on $T_mM$ in $p$-variables. Let $\mathcal{A}^p(M)$ be the vector space of $p$-forms on $M$, then $\mathcal{A}^0(M) = C^\infty(M)$.

2.9.2. **De Rham differential.** De Rham differential $d = d_p : \mathcal{A}^p(M) \to \mathcal{A}^{p+1}(M)$ satisfies $d^2 = 0$. For $p = 0$, $d_0 = df$ is the usual differential etc. We say that a $p$-form $\omega$ is closed if it is killed by $d_p$. 
2.9.3. *Symplectic structures.* A symplectic form on $M$ is a closed 2-form $\omega$ such that each bilinear form $\omega_m$ on $T_mM$ is non-degenerate.

A symplectic form $\omega$ on $M$ defines a Poisson structure $\{-,-\}$ on $M$ and is in turn determined by it. So a symplectic structure is a special case of a Poisson structure.

2.9.4. *Cotangent bundles.* On $T^*M$ one has a canonical 1-form $\eta$ which in local coordinates can be written as $\eta = \sum p_i \, dq_i$. Then $\omega = d\eta = \sum dp_i \, dq_i$ is a canonical symplectic structure that gives the Poisson structure above.
3. Example: Equivalence of three actions for a Free particle

In the simplest possible example of a free particle we will see how three actions are reasonably equivalent.

- (A) In Newtonian setting (i) the obvious action is the kinetic energy and (ii) the evolution of the particle is naturally parameterized by time.
- (B) In a relativistic length formulation we consider (i) the action will be the length of the trajectory in spacetime and (ii) we add a (physically meaningless) quantity, a choice of a parameterization of the trajectory since the time is now not divorced from space so, there is no canonical parameterization.
- (C) The relativistic kinetic formulation has (i) action quadratic in $\dot{x}$, (ii) another auxiliary field – a metric $g$ on the worldline $I$.

The Newtonian kinetic energy action is physically natural. The nice mathematical features of the length action are that it is familiar to mathematicians (geodesics) and makes sense if we increase the worldsheet dimension (volume action). For physics, it passes the obvious classical test: the solution of the equation of motion is the same as in the non-relativistic setting (as it should be since a free particle is inertial).

The unpleasant feature of the length action is that the quantization is difficult since $\dot{x}$ appears under a square root (Lagrangian is really the square root of the kinetic Lagrangian!), and this makes quantization difficult. The “relativistic kinetic action” resolves the last problem (it is quadratic in $\dot{x}$), and also allows for the massless particles.

Its formula is hard to justify directly, however it passes a classical-quantum test of comparison with the length action: the solutions of the equation of motion are the same and the new action reduces to the length action at the criticality of the auxiliary field.

3.1. (A) Newtonian setting: the kinetic energy action. In general, the Lagrangian is the difference of kinetic and potential energy. The meaning of “free particle” is that the potential is absent, so

$$L = T - V = \frac{m}{2} \dot{x}^2,$$

and the action is a time integral of the Lagrangian

$$S[x] = \int_{t_i}^{t_f} dt L(\dot{x}) = \int_{t_i}^{t_f} dt \frac{m}{2} \dot{x}^2.$$

Therefore, the EL-equation $L_x = \frac{d}{dt} L_{\dot{x}}$ now says

$$0 = L_x = \frac{d}{dt} \frac{m}{2} \dot{x}^2 = m \ddot{x}, \quad \text{i.e.,} \quad \ddot{x} = 0.$$

So, acceleration is zero, velocity is constant: $\dot{x} = v$, and the solutions are the lines

$$x = x_0 + tv.$$
3.2. (B) Relativistic setting: the spacetime trajectory length action.

3.2.1. Relativistic setting. For a relativistic setting one needs a formulation in which time is not separated from space. So the evolution of a particle should be viewed as a curve \( C \) in the spacetime \( M \).

For calculations one still needs a parameterization and now we can not parameterize by time anymore. So, the price we pay is the introduction of an additional datum, a choice of a parameterization \( x \) of \( C \) by an interval \( I \). The choice of parameterization is physically meaningless since the action does not depend on it (the Lagrangian does). Physicists call this invariance under reparameterizations, these new symmetries (reparameterizations) precisely eat up the new degree of freedom.

3.2.2. Length action. The action will be the length action \( S[C] = l(C) \) for the spacetime trajectory \( C \). The idea is that the most natural evolution is the one that minimizes the trajectory length. Notice that on the level of Lagrangians this is the square root of the above kinetic energy action!

This length action is familiar mathematically (solutions are geodesics), but it is more complicated then the kinetic energy action because of the square root.

The action of the trajectory \( C \), calculated in terms of a parameterization \( x \) is

\[
S[C] \overset{\text{def}}{=} S[x] \overset{\text{def}}{=} m \int_C dl = m \int_I d\tau \sqrt{-\dot{x}^2} = \int_I d\tau L(\dot{x})
\]

where \( dl = d\tau \sqrt{-\dot{x}^2} \) is the length of a piece of the curve \( C \) and the Lagrangian is

\[
L(\dot{x}) = m\sqrt{-\dot{x}^2}.
\]

3.2.3. The role of the mass factor in action. Removing the mass factor from the action would not affect the EL-equation. However, the next lemma shows that this would fix the size of the canonical momentum to be 1.

3.2.4. Lemma. (a) The canonical momentum is\(^3\)

\[
p = \frac{m\dot{x}}{\sqrt{-\dot{x}^2}}.
\]

(b) The EL-equation (equation of motion) is

\[
\frac{d}{dt} \left( \frac{m\dot{x}}{\sqrt{-\dot{x}^2}} \right) = 0.
\]

\(^3\)Here, velocity with lower indices means that we view it as a cotangent vector. This and the mass factor are standard, a nonstandard feature is the denominator. It comes from taking the square root of the kinetic energy in the Lagrangian.
This means that \( \dot{x} = c \) is a constant vector which is necessarily a unit vector. So, all solutions are of the form

\[
\dot{x} = \phi(\tau) \cdot c,
\]

where \( c \) is a unit vector, and \( \phi \) is any function \( \phi : I \to \mathbb{R}_{\geq 0} \) (\( \phi \) is the length of \( \dot{x} \)).

3.3. (C) Relativistic setting: kinetic action with a worldline metric. We will improve the relativistic length action so that

- (i) it is quadratic in \( \dot{x} \) (a la the kinetic energy action) and
- (ii) it allows for massless particles.
- (iii) In some sense it is equivalent to the length action.

A quadratic dependence in \( \dot{x} \) is much simpler and preferable for path integrals. The price will be the introduction of another auxiliary field – a choice of a Riemannian metric \( g \) on the worldline \( I \). \( ^4 \) The action is

\[
S[x, g] \overset{\text{def}}{=} -\frac{1}{2} \int_I dl \ (\dot{x})_{G,g}^2 - m^2.
\]

\( ^4 \)In the preceding construction, on the worldline we used the pull-back metric \( g = x^*G \) coming from a Lorentzian metric \( G \) on the spacetime, but we effectively used \(-x^*G\) which was Riemannian because of the light cone assumption on the motion.
We will assume that \( \mathbf{m} \neq \mathbf{0} \) throughout, and revisit the massless case at the end (3.3.4). I do not know how to motivate this formula,\(^5\) but we will check that it has the properties (i-iii). For the notation see

3.3.1. The coordinate notation. The metrics \( g \) and \( g^* = g^{-1} \) are given by the functions

\[
g_{\tau\tau} \overset{\text{def}}{=} g(\partial_\tau, \partial_\tau) > 0 \quad \text{and} \quad g^{\tau\tau} \overset{\text{def}}{=} g^{-1}(d\tau, d\tau) = \frac{1}{g_{\tau\tau}}
\]
on \( I \). We also encoded \( g \) in terms of the \( \textit{einbein} \)

\[
e(\overset{}{\text{def}} = \sqrt{g_{\tau\tau}} = \text{length of } \partial_\tau.
\]

By \( dl \), we denote the length of a piece of the interval \( I \), i.e., \( dl = e(\tau) \cdot d\tau \).

The square speed \( -(\dot{x})^2 \overset{\text{def}}{=} -\dot{x} \cdot \dot{x}, \) is the measure of size of the velocity \( \dot{x}(\tau) \in \text{Hom}_\mathbb{R}[T_{\tau}(I), T_{x(\tau)}M] = T^*_\tau(I) \otimes T_{x(\tau)}M \), which is given by the inner product \( (g^*)_\tau \otimes g_{x(\tau)} \), so by ??,

\[
(\dot{x})^2_{G,g} = \frac{\partial x}{\partial \tau} \cdot \frac{\partial x}{\partial \tau} \cdot g^{\tau\tau}.
\]

So, in coordinates the action is

\[
S[x,g] = -\frac{1}{2} \int_I d\tau \sqrt{g_{\tau\tau}} \ (g^{\tau\tau}\dot{x}^2 - m^2) = \int_I d\tau \ L(\dot{x}, g),
\]

for the Lagrangian

\[
L(\dot{x}, g) = -\frac{1}{2} \ [(g^{\tau\tau})^\frac{1}{2} \dot{x}^2 - (g^{\tau\tau})^{-\frac{1}{2}} m^2].
\]

Notice that \( x \) is a dynamical variable and \( g \) is not.

3.3.2. Variations with respect to the inverse metric \( g^{-1} \) and the path \( x \). We have two fields so we will get two criticality equations, the vanishing of the variations (differentials) with respect to \( g^{-1} \) and \( x, \)\(^6\)

Lemma. (a) The criticality equation \( \frac{\delta S}{\delta g^{\tau\tau}} = 0 \) with respect to the inverse metric, is

\[
g^{\tau\tau} = \frac{m^2}{-\dot{x}^2}, \quad \text{i.e.,} \quad g_{\tau\tau} = \frac{1}{m^2} \ (-\dot{x}^2), \quad \text{i.e.,} \quad g = \frac{x^*(G)}{m^2}.
\]

It determines the metric \( g \) in terms of \( x \) – up to a mass factor, \( g \) is the \( x \)-pull-back of the metric \(-G\) on the spacetime.

(b) The criticality equation with respect to the path \( x \): \( \frac{\delta S}{\delta g^{\tau\tau}} = 0 \), is the requirement that \( (g^{\tau\tau})^\frac{1}{2} \dot{x} \) is a constant vector, i.e., \( \tau \)-velocity is a multiple of some vector \( c \)

\[
\dot{x} = g^{\tau\tau} \frac{1}{2} c.
\]

\(^5\)In particular, how do length and mass have the same units?\(^6\)One uses \( g^{-1} \) so that the variation \( \frac{\delta S}{\delta g^{\tau\tau}} \) is proportional to \( g \).
Proof. (a) Since the metric $g^{-1}$ is a non-dynamical field
\[
\frac{\delta S}{\delta g^{\tau\tau}} = L_{g^{\tau\tau}} = -\frac{1}{2} \left[ \frac{1}{2} (g^{\tau\tau})^{-\frac{1}{2}} \dot{x}^2 - \left( -\frac{1}{2}\right) (g^{\tau\tau})^{-\frac{3}{2}} m^2 \right] = -\frac{1}{4} (g^{\tau\tau})^{-\frac{3}{2}} \left[ g^{\tau\tau} \dot{x}^2 + m^2 \right].
\]
(b) Since $x$ is a purely dynamical field
\[
\frac{\delta S}{\delta \dot{x}} = \frac{\partial}{\partial \tau} L_{\dot{x}, g} = -\frac{1}{2} \frac{\partial}{\partial \tau} \frac{\partial}{\partial \dot{x}} \left[ (g^{\tau\tau})^{\frac{1}{2}} \dot{x}^2 - (g^{\tau\tau})^{-\frac{3}{2}} m^2 \right] = -\frac{1}{2} \frac{\partial}{\partial \tau} \left[ (g^{\tau\tau})^{\frac{1}{2}} \dot{x} \dot{x} \right].
\]

3.3.3. The relation to the length action. We will now check that the passage from the length action to $S[x, g]$ is an example of the philosophy 3.4.1 of adding an auxiliary field to the action, while keeping the physics essentially equivalent.

Lemma. (a) (Reduction at criticality.) At the criticality for the metric $g^{-1}$, the new action reduces to the length action
\[
S[x, g] \big|_{g^{-1}} = \frac{m^* c}{m^*} = \int_I d\tau \ m \sqrt{-\dot{x}^2} = S_{\text{length}}[x].
\]
(b) (Classical equivalence.) The criticality equations of the new action $S[g, x]$ in both the metric $g$ and the path $x$, together amount to the criticality in $x$ of the length action and a formula for $g$ in terms of $x$ ($g$ is the $x$-pullback of the space time metric $g = -x^* G$).
(c) The symmetries of the action $S[g, x]$ are the same as for the length action $S[x].$

Proof. (a) At the criticality equation for the inverse metric $g^{\tau\tau} = \frac{m^2}{v^2},$
\[
S[x, g] \big|_{g^{\tau\tau} = \frac{m^2}{\dot{x}^2}} = -\frac{1}{2} \int_I d\tau \ (g^{\tau\tau})^{\frac{1}{2}} \dot{x}^2 - (g^{\tau\tau})^{-\frac{3}{2}} m^2 \big|_{g^{\tau\tau} = \frac{m^2}{\dot{x}^2}}
\]
\[
= -\frac{1}{2} \int_I d\tau \ \frac{m}{\sqrt{-\dot{x}^2}} m^2 = -\frac{1}{2} \int_I d\tau \ -2 m \sqrt{-\dot{x}^2} = \int_I d\tau \ m \sqrt{-\dot{x}^2}.
\]
(b) The criticality for $g$ is a formula for $g$ in terms of $x : g^{\tau\tau} = \frac{m^2}{\dot{x}^2}$, which says that $g^{\tau\tau} = \frac{m^2}{\dot{x}^2} x^* (-G).$ We can use it rewrite the criticality equation in $x : \frac{\partial}{\partial \tau} \left[ (g^{\tau\tau})^{\frac{1}{2}} \dot{x} \right] = 0$, as $0 = \frac{\partial}{\partial \tau} \left[ \frac{m}{\sqrt{-\dot{x}^2}} \right]$, but this is exactly the EL-equation for the length action.

3.3.4. Relativistic kinetic action for massless particles. Now we consider the case when $m = 0$. Then the action is
\[
S[x, g] \overset{\text{def}}{=} -\frac{1}{2} \int_I d\tau \ \dot{x}^2_{G, g}.
\]
The criticality equation with respect to the path $x$: $\frac{\delta S}{\delta \dot{x}} = 0$, is, as above, the requirement that the $\tau$-velocity is a multiple of some constant vector $c$
\[
\dot{x} = g^{\tau\tau} \frac{1}{2} \cdot c.
\]
The relation to the length action in the massive case was that the kinetic action reduces to the length action at the criticality in $g^{-1}$. However, in the massless case there should
be no relation since there is no length action for a massless particle. On the calculational level this is seen as the absence of solutions of the criticality equation with respect to the inverse metric – the equation is $g^{tr} = 0$.

3.4. **The philosophy of adding an auxiliary field to the action.** We are interested in

adding an auxiliary field $y$ to action $S[x]$ while keeping the physics “equivalent”.

Roughly, the new field will be eliminated either by the new criticality equation (when the action depends on it), or by the new symmetries (when it does not).

3.4.1. **Adding an auxiliary field $y$ to action $S[x]$**. By this we will mean passing to a new action $S[x, y]$ such that:

1. $S$ reduces to $S$ at the criticality for $y$. (i.e., criticality gives $y = \phi(x)$ and $S[x, \phi(x)] = S[x]$), and
2. Actions are classically equivalent, i.e., solutions of $d_x S = 0$ and $d_x S = 0 = d_y S$ are the same.
3. (???) The symmetries for $S[x, y]$ are the same as for $S[x]$.

3.4.2. **Adding a “physically meaningless” field.** This is a simpler version. The new action $S[x, y]$ simply does not depend on the new field $y$:

$$S[x, y] = S[x].$$

So, $y$ is really just a device for calculating the action $S[x]$ of $x$. 
Part II. Quantum Mechanics

0. Intro

We start with basic principles of quantum mechanics:

*The probability of an event refines to a complex number called the amplitude. Amplitudes combine (by addition) in a much simpler way than the probabilities.*

Our justification is based on (idealized) experiments with a finite number of possible evolutions of the system.

0.0.3. Lagrangian approach. When one applies these principles to systems with continuously many possible evolutions, the formula for the amplitude of a certain event transforms from a finite sum to an integral of amplitudes over all possible evolutions (=paths), i.e., a “path integral”. The key is now to know the formula for the amplitude \( Z[x] \) of a possible evolution, i.e., a path \( x \). It turns out that this is a unit complex number \( Z[x] = e^{i\hbar S[x]} \) with the phase equal to the action \( S[x] \) of the path (up to the Planck constant \( \hbar \)). So this ties with the classical Lagrangian approach.

0.0.4. Hamiltonian approach. The fundamental experimental fact here is that the quantum configuration space – the space of states of a quantum system – has a structure of a Hilbert space \( \mathcal{H} \). It usually has some functional interpretation such as \( \mathcal{H} = L^2(\mathcal{C}) \).

Moreover, the observables (in the sense of real functions on the momentum configuration space \( T^*\mathcal{C} \)), now become hermitian linear operators on \( \mathcal{H} \). The evaluation of an observable \( \hat{f} \) on a state \( v \), i.e., the measurement \( f(v) \) of \( f \) at \( v \), now becomes the eigenvalue of the operator \( \hat{f} \) on the vector \( v \). (For a more precise statement see the Dirac axioms of Q-Mechanics in 0.1 below.)

In particular, the Hamiltonian function \( H \) on \( T^*\mathcal{C} \) now becomes the Hamiltonian operator \( \hat{H} \). Hilbert space comes with a hermitian operator

*The physical law governing the system is now formulated as:

-Time evolution of the system is generated by the Hamiltonian operator \( \hat{H} \).*

The meaning is that the evolution of the system in time \( t \) is obtained by the unitary operator \( U(t) = e^{\frac{2\pi i}{\hbar} \hat{H} t} \) acting on the configuration space \( \mathcal{H} \). So, \( \hat{H} \) is the infinitesimal generator of time evolution.

0.0.5. Functional realizations of \( \mathcal{H} \). In the example of electron on a line or in a vector space \( V \), we will see that Hilbert space \( \mathcal{H} \) can be realized as \( L^2(V) \). Physicists use two traditional modes of computations (i.e., views on the Hilbert space \( \mathcal{S} \).
• Use the subspace $T$ of nice states (i.e., the test functions), and its continuous dual $D = T'$ of $T$, which is the space of idealized, i.e., distributional states.

• Use algebraic structure on a dense countable-dimensional subspace $V$ of $H$ (or of $T$). (This we will do for the harmonic oscillator, i.e., when the potential is a quadratic function of position $V = -x^2$.)

0.0.6. Historical development: Hamiltonian to Lagrangian. We will follow the historical development: we develop the operator approach to quantum mechanics, and then use it to derive the path amplitude formula.

Mathematicians often neglect the Hilbert space $H$, i.e., they do not view its appearance as the fundamental event. Rather, they are impressed by the fact that the quantization of functions on $T^*C$ to operators, happens in an organized way: the commutative algebra $O(T^*C)$ of functions ("observables") on $T^*C$, deforms to a non-commutative algebra $D_C(\hbar)$. Then they find $H$ as a natural representation of the algebra $D_C(\hbar)$, hence they recover the operator interpretation of quantized functions. (However we have most difficulty with unitarity?)

The mathematical reason that the algebra $O(T^*C)$ deforms is that $T^*C$ has a canonical Poisson structure. So, we start with the mathematical framework for the Hamiltonian approach: the deformation quantization of Poisson algebras. For instance we find that $D_C(\hbar)$ is the algebra of differential operators on $C$, so $H$ will be its natural representation $O(C)$ (in various incarnations such as $L^2(C)$).

We will calculate the amplitudes in the operator approach and find that the result is a path integral. This traces the historical discovery of the precise formula for the integrand in the path integral. (However, after making this observation, Feynman gave an independent direct heuristic justification of the path integral amplitude formula.)
0.1. Dirac axioms.

(1) Deterministic axioms
(a) To a physical system there corresponds a Hilbert space $\mathcal{H}$ (the configuration space of the system), and a Hermitian operator $H$ on $\mathcal{H}$ (the Hamiltonian).
(b) To every state $\psi$ of the system there corresponds a “wave function” $\psi \in \mathcal{H}$, defined up to $\mathbb{C}^*$.\(^8\)
(c) If at time 0 the system is in the state $\psi$, then at a time $t$ it is in the state $e^{\frac{iHt}{\hbar}}\psi$.

(2) Stochastic axioms related to measurements\(^9\)
(a) To any measurement process there corresponds a hermitian operator $Q$ (an “observable”). If $\mathcal{H}$ has a basis $v_i$ with $Q$-eigenvalues $\lambda_i$, then the only possible values of the measurement are the eigenvalues $\lambda_i$.
(b) If the system is in a state $\psi = \sum a_i v_i$ (and we normalize $v_i$’s), the probability that the measurement will give $\lambda$ is $P_{\lambda} = \frac{|a_i|^2}{\sum |a_i|^2} \frac{|a_i|^2}{|\psi|^2}$.
(c) If the measurement produces value $\lambda$ then the measurement process has moved the system from state $\psi$ to the state $p_{\lambda}^Q \psi$.

\(^8\)The terminology is based on the functional realization of $\mathcal{H}$.
\(^9\)Here, as in the original Dirac formulation, one assumes that the spectra are discrete and there are no multiplicities. A more precise version:

(a) To any measurement process there corresponds a hermitian operator $Q$ (an “observable”). So, $Q = \int_{\text{Spec}(Q)} \lambda dp^Q(\lambda)$ where $p^Q$ is the associated projector valued measure. The only possible values of the measurement are the eigenvalues $\lambda \in \text{Spec}(Q)$.
(b) If the system is in a state $\psi = \int \psi_\lambda d\lambda$ (with “$\psi_\lambda$ in the eigenspace $\lambda$”), the probability that the measurement will give a value in the interval $[\lambda_1, \lambda_2]$ is

$$\int_{\lambda_1}^{\lambda_2} d\lambda \frac{|\psi_\lambda|^2}{|\psi|^2} = \frac{\int_{\lambda_1}^{\lambda_2} |p^Q_{[\lambda_1, \lambda_2]} \psi|^2}{|\psi|^2}.$$

(c) If the measurement produces value $\lambda$ then the measurement process has moved the system from state $\psi$ to the state $p^Q_\lambda \psi$. 
1. Principles of Quantum Mechanics

Text. R.P. Feynman, *Quantum Mechanics*, volume 4 of *Feynman’s lectures on Physics*.

1.1. Amplitudes. We have no natural experience (intuition) about what happens on the small scale.\(^{10}\) By analyzing certain particle experiments one arrives at the following picture of quantum systems, i.e., systems on a small scale. This is common to the Lagrangian and Hamiltonian approach.

1.1.1. *Complex numbers and amplitudes.* The complex numbers appear in analogy with the water waves – these are known to require complex numbers in order to describe interactions. In quantum mechanics they appear as “probability amplitudes” which are refinements of the probability of an event. This gives there is a “pairing” or “inner product operation” on states, given by the

\[ \langle j | i \rangle \in \mathbb{C}, \]  

for a transition from one state \(i\) into another state \(j\).

Amplitude is sesquilinear: linear in \(i\) and anti-linear in \(j\).

1.2. Hilbert space setting (the Hamiltonian view).

1.2.1. *The linear algebra of a Hilbert space.* The notion of the “state of the system” is the analogue for quantum systems of a point of the configuration space – it is a complete description of the system at a certain time. A fundamental observation:

the set \( \mathcal{H} \) of all states of the system forms a Hilbert space,

is obtained in the following way:

1. A choice of a “filtering apparatus” \( S \) (an experimental machinery), gives the family \( B_S \) of “\( S \)-pure” states. These pure states are orthonormal for the probability amplitude pairing. Moreover, we can assume that these pure states can not be filtered into simpler states (for this, if need be, we replace a simple filtering apparatus \( S \) by a succession \( S = (S_1, ..., S_p) \) of such filtering appurata.)

2. The additive structure of the set \( \mathcal{H} \) of all states of the system comes from the further experimental fact that for any state \( \phi \), amplitudes \( \langle \phi, i \rangle \), \( i \in B_S \), behave as coefficients of a vector \( \phi \) in the basis \( B_S \).

1.2.2. *Appearance of linear operators.* Starting from the notation \( \langle x | s \rangle \) for the probability amplitude one uses symbol \( | s \rangle \) to denote state \( s \), or the corresponding vector. An operation \( A \) on states can be described in terms of the amplitudes \( \langle j | A | i \rangle \) for the probability that operation \( A \) on a particle in the state \( i \) will give a particle in the state \( j \). Again, we check experimentally that the amplitudes \( \langle j | A | i \rangle \) are sesquilinear in \( i \) and \( j \) (linear in \( i \), anti-linear in \( j \)).

\(^{10}\)It may be difficult to admit.
and anti-linear in $j$), so we can think of such operation $A$ as a linear operator $|s\rangle \mapsto A|s\rangle$ on the space of states. This interpretation of

amplitudes as matrix coefficients

is the heart of the Hamiltonian approach.

1.2.3. Time evolution by the Hamiltonian operator (Schroedinger’s equation). Next experimental act is the time evolution $\psi(t)$ of a state $\psi$ has an infinitesimal generator, a linear operator $\hat{H}$ which we call the Hamiltonian operator:

$$\frac{d}{dt}\psi(t) = \frac{2\pi i}{\hbar}\hat{H}\psi.$$  

This we call the abstract Schroedinger’s equation. To apply this formalism, one needs to say what $\hat{H}$ is in a given situation. This gives a particular Schroedinger equation, which is a differential equation that describes the evolution of the quantum system.

Historically, this breakthrough came in 1926 when Schroedinger described by such equation the motion of electrons on atomic scale. The above conceptual framework came later, and it become the foundation only with Feynman’s book.

1.3. Feynman integrals (the Lagrangian view). Feynman’s point of view on quantum mechanics is that the amplitude for an event is total of amplitudes over all possible evolutions $x$ of the system that result in this event. In the Minkowski setting (physically reasonable), this is an oscillating integral

$$\int_x d\mathbf{x} \ e^{\frac{i}{\hbar}S[\mathbf{x}]}.$$  

The amplitude for the evolution $x$ is a unit complex number $e^{\frac{i}{\hbar}S[\mathbf{x}]}$ where $S$ is the action. The relation to the classical mechanics comes from the Principle of Stationary Phase, which says that the leading contributions to an oscillating integral come from points where phase does not vary. So the classical solutions – evolutions $x$ such that $d_xS = 0$ give the main contribution to the amplitude. In other words, it is most likely that the system will evolve according to a classical solution, but other evolutions also contribute.

We will follow the historic appearance of Feynman integrals:

1.3.1. Operator approach leads to Feynman integrals. When we try to account for the amplitude of a certain transition from time $t_i$ to time $t_f$, as an accumulation of contributions over small time subintervals $dt = t_{k-1}, t_k$, we find that the basic principles on how the amplitudes interact give contributions from all possible positions $x$ during $dt$, and all possible momenta $p$ during $dt$. So, the amplitude is an integral over time $t$ and all possible choices $x(t), p(t)$ of the position and momentum at time $t$. So we really get the integral over all curves $(x, p)$ in $T^*C$. 
The integrand happens to be exactly $e^{\frac{i}{\hbar}S[x,p]}$ for some action $S$ on paths $(x, p)$ in $T^*\mathcal{C}$. This comes from the formula for the Hamiltonian operator as a quantization of the Hamiltonian function, which is of course related to the Lagrangian.
2. Amplitudes

2.1. Shooting electrons through holes leads to: probability, complex numbers, uncertainty principle and the dual nature of particles.

2.1.1. Experiment. In an idealized experiment, suppose that a source shoots something towards the wall with holes 1 and 2. A bit behind the 1st wall is the second parallel wall, with a detector that counts the hits at a position \( x \) (a coordinate on the second wall). Consider the probability distributions

- \( p(x) \) for the hit to be at \( x \), and
- \( p_i(x) \) (for \( i = 1, 2 \)), for the hit to be at \( x \) if only the hole \( i \) is open.

2.1.2. Shooting bullets. Assume we shoot some kind of bullets which can be deflected in an arbitrary direction when passing thru the hole, but they do not break. So, bullets they arrive in quanta – a bullet at a time – (no halves etc) and clearly \( p = p_1 + p_2 \) which we describe by “no interference”. (The contributions from each possibility add up, so there is no interaction between the two kinds of events.)

2.1.3. Shooting water waves. For water waves, we measure not the probability - the wave will certainly arrive – but the intensity \( I(x) \) at \( x \), which is defined as the mean square of the height (energy) of the wave. Here, the arrival is a continuous quantity. Notice that now two waves are going to interfere at \( x \).

The height of a single wave at \( x \) is of the form

\[
H(x, t) = C(x) \cos(\omega t - p) = A(x) \cos(\omega t) + B(x) \sin(\omega t)
\]

(harmonic oscillator!), which one can rewrite as

\[
H(x, t) = \text{Re}[\{A(x) - iB(x)\}e^{i\omega t}] = \text{Re}[h(x)e^{i\omega t}], \quad \text{where} \quad h(x) \triangleq A(x) - iB(x).
\]

We will call the complex number \( h(x) \) the “amplitude of intensity” at \( x \). Then the intensity at \( x \) is going to be

\[
I(x) \overset{\text{def}}{=} \frac{1}{2\pi/\omega} \int_0^{2\pi/\omega} H(x, t)^2 \, dt = \frac{1}{2\pi/\omega} \int_0^{2\pi/\omega} A^2 \cos^2(\omega t) + B^2 \sin^2(\omega t) - AB \sin(2\omega t) \, dt
\]

\[
= \frac{1}{2\pi/\omega} \left( A^2 + B^2 \right) \int_0^{2\pi/\omega} \cos^2(\omega t) \, dt = \frac{1}{2\pi/\omega} \left( A^2 + B^2 \right) \int_0^{2\pi/\omega} \frac{1}{2} 2\pi/\omega = \frac{1}{2} |h(x)|^2.
\]

The way the intensities add up when two waves (with the same frequency \( \omega \) ) interfere, is complicated, the difference \( \delta \) in the phases gives \( I = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos \delta \) (the “cosine rule in a triangle”). However, the formula for the addition of complex amplitudes is simple: \( h(x) = h_1(h) + h_2(x) \).
2.1.4. *Shooting electrons.* If one shoots electrons, they will arrive in quanta (only the whole electrons), but the probabilities \( p_i(x) \) at the position \( x \) have the same interference pattern as the intensities of water waves.

(For the comparison of the three experiments, notice that if one shoots bullets, water waves or electrons; the graphs of \( p_i \)'s will always be the same (a random distribution), while the graphs of \( p(x) \) will be the same only for water waves and electrons.

This suggests that electrons have dual nature in the sense that besides their standard behavior as particles (“bullet-like”), they also sometimes behave as waves. Moreover, we see that one will need complex numbers to describe particles. Once this “correct” point of view is taken – that the probability is originally a complex number – we return to the “no interference” situation (now for amplitudes rather then for probabilities themselves).

2.1.5. *Uncertainty principle.* This is the principle that our observation necessarily affects the experiment! We see that this has to be happening because if one can actually observe thru which of the holes the electron is passing, then clearly \( p = p_1 + p_2 \) as for the bullets.

2.2. **The amplitude principles.** These are basic principles of Quantum Mechanics:

- **(0) (Amplitude axiom)** The probability \( p \) of an event can be refined to a complex number \( \phi \) called probability amplitude, they are related by
  \[
  p = |\phi|^2 \quad \text{and} \quad \phi = \sqrt{p} e^{i\delta},
  \]
  So, the amplitude \( \phi \) consists of the probability \( p \) and the phase \( \delta \), and always \(|\phi| \leq 1\).

- **(1) (Combining amplitudes)** If an event can happen in two different ways then its amplitude is
  \[
  \phi = \phi_1 + \phi_2.
  \]
  So, probability \( p = \sqrt{(p_1)^2 + (p_2)^2 + 2p_1p_2 \cos(\delta_1 - \delta_2)} \) depends on the relative phase.

- **(2) (Concatenation)** The amplitude for the particle to arrive from \( i \) to \( k \) via \( j \) is the product of the amplitude for the particle to go from \( i \) to \( j \) and the amplitude for the particle to go from \( j \) to \( k \).

- **(3) (Reversal)** The amplitude for the particle to arrive from \( i \) to \( k \) is the complex conjugate of the amplitude for the particle to arrive from \( k \) to \( i \).

The above axioms encode a lot, so we can use them as a foundation and derive ideas (like the Uncertainty axiom), without invoking particular experiment.

2.2.1. *Complex numbers.* We see now how elementary experimental facts construct all aspects of complex numbers from nature. While (1) uses only a 2-dimensional vector space \( \mathbb{R}^2 \) with an inner product, (2) uses multiplication and (3) uses the real structure on \( \mathbb{C} \), so we have recovered the whole structure of complex numbers from physics!
2.3. **Uncertainty principle: a quantitative version.** We will now try to observe through which hole did the electron go through. For the observation we use the light, i.e., a photon, of a certain wavelength. We will see that two features of the situation

- The result of the measurement, i.e., what we it tells us has happened,
- How efficient the apparatus is;

are both dependent on the choice of wavelength. This degree of freedom gives us quantitative view on the relation between these two features. The result will be that in order to make the apparatus more efficient (so that it is likely to report a definite result), we have to affect more the result of the measurement.

2.3.1. **The experimental setting.** Behind the 1st wall, let us add a source of light (emitting photons), in the middle position, and two light detectors $D_1$ and $D_2$, near holes $H_1$ and $H_2$. So one has

For $i, j \in \{1, 2\}$, we will denote the amplitude of basic events by

- $\langle i|g \rangle = \langle H_i|g \rangle$ is the amplitude for electron to shot out of the electron gun to pass through the hole $H_i$,
- $\langle x|i \rangle = \langle x|H_i \rangle$ is the amplitude for electron that went through the hole $H_i$ to go to the position $x$,
- $a_{ij}$ is the probability amplitude for an electron which has passed through the hole $H_i$, to scatter a photon to the detector $D_j$. By symmetry,
  
$$a_{11} = a_{22} \overset{\text{def}}{=} \alpha \quad \text{and} \quad a_{21} = a_{12} \overset{\text{def}}{=} \beta.$$  

The influence of the wavelength:

1. efficient apparatus case: if the wavelength is small, then approximately
   $$\beta = 0.$$  
2. inefficient apparatus case: if the wavelength is large, then approximately
   $$|\alpha|^2 = |\beta|^2.$$  

We will also assume that the phases of $\alpha$ and $\beta$ agree.

2.3.2. **Calculation of some amplitudes.** First, the amplitude $\phi_i(x)$ and three probability $p_1(x)$ of the electron to go from the gun through the hole $H_i$ and go to the position $x$ (the second hole may as well be closed), is

$$\phi_i(x) = \langle x|H_i \rangle \cdot \langle H_i|g \rangle \quad \text{and} \quad \psi_i(x) = |\phi_1(x)|^2.$$
Next, our measuring device, will *tell us* that the electron went through the hole $H_j$ and to the position $x$ if the photon shows up at the detector $D_j$ and the electron at the position $x$. This happens in two cases ($i = 1, 2$), when

*electron goes from $g$ to $x$ via the hole $H_i$ and at $H_i$ it scatters a photon to detector $D_j$.***

The amplitude for each of these two events is

$$\langle x|i\rangle \cdot a_{ij} \cdot \langle i|g\rangle = a_{ij} \cdot \phi_i(x).$$

So, the amplitude for the detectors to tell us that the electron went to go through the hole $j = 1$ to the position $x$ is

$$\sum_i a_{ij} \cdot \langle x|i\rangle \cdot \langle i|g\rangle = \sum_i a_{i1} \cdot \phi_i(x) = a_{11} \phi_1(x) + a_{21} \phi_2(x) = \alpha \cdot \phi_1(x) + \beta \cdot \phi_2(x).$$

The corresponding probability of the measurement to tell us that the electron went thru the hole $j = 1$ to the position $x$ is therefore,

$$p = |\alpha \cdot \phi_1 + \beta \cdot \phi_2|^2.$$  

**The efficient case.** Here wavelength is small, $\beta$ is close to 0

$$p \approx |\alpha \cdot \phi_1|^2 = |\alpha|^2 \cdot p_1.$$  

**The efficient case.** Here wavelength is large, $|\alpha|^2 \approx |\beta|^2$, and if we also assume that the phases of $\alpha$ and $\beta$ agree then

$$p \approx |\alpha|^2 \cdot |\phi_1 + \phi_2|^2.$$  

2.3.3. **Conclusion.** Let us count the number of hits at $x$ in the presence of the measurement apparatus, and see whether it depends on the efficiency of the apparatus.

If we divide by the light effect $|\alpha|^2$, in the *efficient* case we get the correct result $p_1 = \phi_1(x)^2$, so the total number of hits at $x$ will be the sum

$$|\alpha|^2 \cdot (p_1(x) + p_2(x)).$$

In the *inefficient* case we get the interference formula, and the total number of hits at $x$ will be

$$|\alpha|^2 \cdot 2|\phi_1 + \phi_2|^2,$$

showing the interference pattern. So,

*The nature of our observation apparatus influences the events, a more efficient apparatus causes a more classical behavior.*
3. Operators: the Hamiltonian approach

The basic idea here is that the configuration space for a quantum system is a Hilbert space. This is deduced from experiments in 3.2. We find that the operations on states are given by linear operators, and in particular that the evolution of the system is given by one operator that we call the Hamiltonian operator (3.3).

The most relevant mathematical theory is the deformation quantization which we visit in 3.1. To mathematicians, the most reasonable aspect is the fact that the quantization of functions on $T^*\mathcal{C}$ to operators, happens in an organized way: the commutative algebra $\mathcal{O}(T^*\mathcal{C})$ of functions ("observables") on $T^*\mathcal{C}$, deforms to a non-commutative algebra $D_{\mathcal{C}}(\hbar)$. The mathematical reason that the algebra $\mathcal{O}(T^*\mathcal{C})$ deforms is that any cotangent bundle has a canonical geometric structure: the Poisson structure. Once one finds $D_{\mathcal{C}}(\hbar)$, they recover the Hilbert space $\mathcal{H}$ as its natural representation. This usually gives a functional interpretation of $\mathcal{H}$ and therefore leads to a Functional Analysis.

3.1. Deformation quantization: a mathematical notion of quantization.

3.1.1. Example: functions on the cotangent bundle deform to differential operators. For each $\hbar \in \mathbb{C}$ consider the algebra $\mathcal{A}_\hbar$ generated by $p$ and $q$ and the relation $pq - qp = \hbar \cdot 1$. In each of these, monomials $q^i p^j$ form a basis, so we can think of them as a family of associative algebras with 1 on one vector space $A$. We will think of $\mathcal{A}_1$ as the ring of differential operators $D_{A_1}$ on the line $A_1$ with coordinate $q$ and $p = \partial_q$. At the other extreme, $A_0 = \mathbb{C}[p, q]$ is the algebra of polynomial functions on $A^2$ and we will think of this $A^2$ as $T^*A^1$ with with $q$ a coordinate on the space $A^1$ and $p$ a coordinate in the vertical direction.

The same works for $A^n$ and it is invariant under a change of coordinates, so it actually works on any manifold $M$.

3.1.2. Deformation of a commutative algebra gives a Poisson structure. Consider a vector bundle $\mathcal{A}$ over a neighborhood $U$ of 0 in $\mathbb{C}$, and a family of associative algebra structures $\cdot_{\hbar}$ on vector spaces $\mathcal{A}_\hbar$, $\hbar \in U$, such that the multiplication $a \cdot_{\hbar} b$ is an analytic function on $\mathcal{A}$. We say that this is a deformation of the algebra $A = \mathcal{A}_0$ on $U$.

Since one can trivialize the vector bundle $\mathcal{A}$ near $\hbar = 0$, this is equivalent to the following simpler definition. A deformation of the algebra $A$ over $U$ is a family of associative algebra structures $\cdot_{\hbar}$ on the vector space $A$, such that (i) $(0) \cdot (A_0, \cdot_0) \cong A$ and (i) the multiplication $a \cdot_{\hbar} b$, $a, b \in A$, $\hbar \in U$, is an analytic function of $\hbar$, i.e., near 0

\[
a \cdot_{\hbar} b = \sum_{n=0}^{\infty} \hbar^n a^{\hbar} b, \quad a, b \in A.
\]
3.1.3. The ordering problem. The second point of view is more convenient for calculations. For instance now we can say that $f \in A = \mathcal{A}_0$ quantizes (deforms) to $f$ in the algebra $A = \mathcal{A}_\hbar$. However, this point of view is non-canonical since in practice the vector bundle $\mathcal{A}$ does not have a natural trivialization.

For instance in our main example, the identification of vector spaces $\mathcal{O}(T^*\mathbb{A}^n) \cong D_{\mathbb{A}^n}$ is canonical only on certain generators $q_i, p_j$ and $x_i, \partial_j$. Since there is no canonical identification for the whole vector spaces, various trivializations leads to various possible quantization $\hat{f} \in AA_1$ of a given $f \in A_0$.

We call this the ordering problem since in this example, a specification of the ordering of $q_i, p_j$’s gives a basis of all algebras $\mathcal{A}_a$, hence a trivialization of the vector bundle $\mathcal{A}$.

Our main interest is the correct quantization $\hat{H}$ of the Hamiltonian function $H$ on $T^*\mathcal{C}$. Notice that there is no problem if there are no products of $p$ and $q$: $H(q, p) = H_{\text{potential}}(q) + H_{\text{kinetic}}(p)$.

Lemma. Any deformation of a commutative algebra $A$ makes $A$ into a Poisson algebra by

$$\{ a, b \} \overset{\text{def}}{=} \lim_{\hbar \to 0} \frac{a \cdot b - b \cdot a}{\hbar}.$$

3.1.4. (Deformation) quantization of Poisson algebras. We will reverse the point of view and suppose that we start with a Poisson algebra $(A, \{, \})$ (a commutative algebra with a Poisson structure). A (deformation) quantization of the Poisson algebra $(A, \{, \})$ will be any deformation $\mathcal{A}$ of $A \cong A_0$, such that the Poisson structure on $A = \mathcal{A}_0$ induced by the deformation is $\{, \}$. One also often says imprecisely that a specific algebra $\mathcal{A}_1$ in the family $\mathcal{A}_\hbar$ is a quantization of $(A, \{, \})$.

3.1.5. Differential operators are a quantization of functions on the cotangent bundle. On any cotangent bundle, there is a canonical Poisson structure (I.2.6.2). We used it for the “canonical formalism” formulation of classical mechanics. The above deformation of functions on the cotangent bundle (3.1.1), is a quantization of the canonical Poisson structure on the cotangent bundle, since the Poisson structure on $\mathcal{O}(T^*\mathcal{A}^1)$ obtained from the deformation is the canonical Poisson structure:

$$\{ p, q \} \overset{\text{def}}{=} \lim_{\hbar \to 0} \frac{p_\hbar \cdot q - q_\hbar \cdot p}{\hbar} = \lim_{\hbar \to 0} \frac{\hbar \cdot 1}{\hbar} = 1.$$

3.1.6. Comparison with a quantization of a system in physics. In physics, the basic example of quantization is that of the harmonic oscillator. In the Hamiltonian approach it involves the deformation of the algebra of functions on the phase space $T^*\mathcal{C}$ (a Poisson space) to the algebra of differential operators $D_{\mathcal{C}}$ on the configuration space. However, more is needed:
• A Hilbert space $\mathcal{H}$ with an action of the quantization algebra $D_C$. (In this example we can use $\mathcal{H} = L^2(C).$)

• An interpretation of $\mathcal{H}$ as the space of (quantum) states – the quantum analogue of the configuration space $C$.

• A particular “quantization” of the Hamiltonian function $H$ on $T^*C$ to an “operator” $\hat{H} \in D^\mathbb{R}$ acting on $\mathcal{H}$. (It gives the time evolution of the quantum system $v \mapsto e^{it\hat{H}} v$.)

• Particular quantizations $\hat{f} \in D^\mathbb{R}$ of interesting observables $f \in \mathcal{O}(T^*C)$.

3.1.7. *Unitarity problem.* A number of questions arise, such as unitarity, i.e., we need a

1. positive definite inner product on $\mathcal{H}$, and
2. it should be compatible with the quantization of observables in the sense that for a real valued function $f$, $\hat{f}$ should be a Hermitian operator.

This is needed to ensure both that the eigenvalues of $\hat{f}$ are real and plentiful, and this is essential since the measurement of $f$ on a state $v$ of the system now makes sense only if $v \in \mathcal{H}$ is an eigenvector of $\hat{f}$. Then one says that the value of the observable $f$ on the quantum state $v$ is the corresponding eigenvalue.

The unitarity needs force the following quantization convention on generators\(^{11}\)

\[ [\hat{p}, \hat{q}] = \frac{i}{\hbar} \{ p, q \}. \]

So, since the Poisson structure on $T^*\mathbb{A}^1$ is given by $\{ p, q \} = 1$, we get $[\hat{p}, \hat{q}] = \frac{\hbar}{i}$.

3.2. *Stern-Gerlach filtering apparatus leads to Hilbert space structure on the space of states.* Consider the configuration space $\mathcal{H}$ of some quantum system. We will find that the experiments endow it with a Hilbert space structure.

3.2.1. *Dirac notation $\langle j|i \rangle$ for amplitudes.* Dirac’s notation for the probability amplitude of the particle in state $j$ to arrive into the state $i$ is $\langle j|i \rangle$.\(^{12}\) The amplitude axioms (2) and (3) are now written as

\[ \langle k|i \rangle_{\text{via } j} = \langle k|j \rangle \cdot \langle j|i \rangle \quad \text{and} \quad \langle k|i \rangle = \overline{\langle i|k \rangle}. \]

\(^{11}\)This means that $[\hat{p}, \hat{q}] = \frac{i}{\hbar} \{ p, q \}$, i.e., the natural quantization of $p$ is really $\frac{i}{\hbar} \hat{p}$. Say, the Hamiltonian function $H$ is real valued but the interesting operator is the infinitesimal generator of time, so it is the derivative of a unitary representation, and therefore it is anti-hermitian. It is denoted $\frac{i}{\hbar} \hat{H}$.

\(^{12}\)Here we do not yet worry about the time it takes, the only question is whether certain procedure will end up with the state $x$. However, it is not clear what this amplitude really is, i.e., does it depend on a particular procedure or what?
3.2.2. Pure states form a bases of the Hilbert space. We will consider the example of the spin of a particle and we will use the (idealized) Stern-Gerlach apparatus to analyze the notion of a spin. Since we are only interested in the spin, the state of a particle will mean its spin (and we design the experiment cleverly so that no other feature of the situation interferes. The final conclusion will be that the spin, i.e., any possible state of the spin of a particular particle, lies in a certain finite-dimensional vector space (an irreducible representation of $SL_2(\mathbb{C})$ or $SU(2)$).

A Stern-Gerlach apparatus consists of a (non-homogeneous) magnet thru which passes a beam of particular particles. If the apparatus splits the beam into $n$ beams, we say that this kind of particles has spin $\frac{n-1}{2}$. Say, if there is no splitting ($n = 1$), then the spin is 0. The spin is $\frac{1}{2}$ if there are two beams.

In the case of a spin one particle we get three beams which are traditionally described as

- the up beam called $+$, its angular momentum is $\hbar$,
- the down beam called $-$, with angular momentum $-\hbar$, and
- the undisturbed part of the beam called 0, with angular momentum is 0.

For a particular apparatus $S$ of such type, we can say that a particle is in the “pure state $i$ for $S$” (we denote this state $i_S$), if it “naturally belongs to the $i$th beam”. This is a property of the particle: if we apply again the apparatus $S$ to the $i$th beam, this time it will only produce the $i$th beam.

The spectacular experimental fact is that each particle will go into one of these three beams, the apparatus will find it in some pure $S$-state $i$. So, for any state $\phi$, the experiment will push it into one of the beams. We can measure the probability amplitude for a given state $\phi$ to go to the $i$th beam, i.e., the amplitude $\langle i|\phi \rangle$ for the apparatus to find $\psi$ to be in the $S$-state $i$.

It turns out that these amplitudes behave as coefficients in the decomposition of a vector $\phi$ into a linear combination of pure $S$-states. To see that this is meaningful, one compares the coefficients for various choices of the apparatus. Let us follow the apparatus $S$ with another one of the same nature but a different build (direction), call it $T$. Then the $i$th $S$-beam will split into $T$-beams and we encode this in the “transition matrix” of amplitudes, denote

$$(TS)_{ji} = \langle j_T|i_S \rangle$$

the amplitude for a particle in the $i$th $S$-beam to end up in the $j$th $T$-beam. One now checks that these transition matrices have the expected algebraic properties – they behave as the change of bases matrices for a vector space\(^{13}\).

\(^{13}\)Some algebraic properties of matrices: (0) $(SS)$ is the identity matrix 1 and in general, transition matrix satisfies (i) $(ST) = (TS)$ (reversal) and (ii) $(ST)(TS) = 1$ on the diagonal, (since $\sum_j (ST)_{ij} (TS)_{ji} = \sum_j |(TS)_{ji}|^2$ is the probability that the particle goes from $i_S$ to some $T$-state, i.e., 1).
3.2.3. The conclusion.

- The space of states is a vector space $\mathbb{H}$ over complex numbers.
- The probability amplitude is a positive definite inner product $\langle -| - \rangle$ on $\mathcal{H}$ (linear in the second and anti-linear in the first variable).
- Each Stern-Gerlach apparatus $S$ chooses an orthonormal bases $i_S$, $i \in \mathbb{B}_S$, of the space of states $\mathcal{H}$.
- Any physical operation $A$ on states is a linear operator on the space of states $\mathcal{H}$.

For the last point, starting from the notation $\langle x|s \rangle$ for the probability amplitude one uses symbol $|s \rangle$ to denote state $s$, or the corresponding vector. An operation $A$ can be described in terms of the amplitudes $\langle j|A|i \rangle$ for the probability that operation $A$ applied to a particle in the state $i$, will produce a particle in the state $j$. Once we check the sesquilinearity in $i$ and $j$, we know that these amplitudes are matrix coefficients of a linear operator.

In any situation more complicated then the above study of spins, we may have to use several apparata in succession to filter the atomic system into the pure states, that can not be decomposed further. Then the algebra is the same.

3.2.4. Notation. Starting from the notation $\langle \phi|\psi \rangle$ for amplitudes, we often denote a state $\psi$ in $\mathcal{H}$ also by $|\psi \rangle$. By $\langle \psi|$ we denote the the corresponding linear functional obtain by putting $\psi$ in the first argument of the inner product.

In this notation, the action of a linear operator $A$ on $\mathcal{H}$ and its hermitian adjoint $A^*$ are denoted

$$A|\psi \rangle \overset{\text{def}}{=} |A\psi \rangle \quad \text{and} \quad \langle \psi|A \overset{\text{def}}{=} \langle A^*\psi| \overset{\text{def}}{=} \langle \psi|A^* \rangle.$$

The meaning of the triple symbol is now clear

$$\langle \psi|A|\phi \rangle \overset{\text{def}}{=} \langle \psi|A\phi \rangle = \langle A^*\psi|\phi \rangle.$$

In particular, for a hermitian operator $H$ and a unitary operator $U$ one has

$$\langle \psi|H|\phi \rangle = \langle \psi|H\phi \rangle = \langle H\psi|\phi \rangle \quad \text{and} \quad \langle \psi|U|\phi \rangle = \langle \psi|U\phi \rangle = \langle U^{-1}\psi|\phi \rangle.$$

3.3. Time evolution of a quantum system is given by the Hamiltonian operator $\hat{H}$. Here we switch the point of view from experiments in which the time did not have any influence (the total effect of a certain apparatus we designed) to an undisturbed (by us) evolution of a system. The effect of time on states is an operation that can be thought of in terms of linear transforms $U(t_2, t_1)$ on the space of states.

One can ask does this means that the states “really” evolve under the action of $U$ if there is no disturbance (such as measurements), or that this is the “ideal” or most likely evolution of the system (while the actual evolution is not really predictable). I think that the only meaningful (measurable) quantity are the matrix coefficients $\langle \psi|U(t_2, t_1)|\phi \rangle$, i.e., the amplitudes for $\phi$ to evolve to $\phi$ from $t_1$ to $t_2$. So, $U$ is really a shorthand for a bunch
of numbers called amplitudes, that (we imagine) describe the probabilities of possible evolutions.

3.3.1. Lemma. (a) Operators $U(t_2, t_1)$ are unitary.

(b) If the physical system is time independent then $U(t) = e^{itH}$ for a (densely defined) Hermitian operator $H = \frac{d}{dt}|_{t=0} U(t, 0)$.

Proof. (a) is the claim that the time evolution preserves amplitudes $\langle U(t_2, t_1)\phi, U(t_2, t_1)\psi \rangle = \langle \phi, \psi \rangle$.

(b) Time independence: means that $U(t_2 + t, t_1 + t) = U(t_2, t_1)$. Then $U(t_2, t_1) = U(t_2 - t_1)$ for $U(t) \overset{\text{def}}{=} U(t, 0)$. The consistency condition for the time evolution $U(t_3, t_1) = U(t_3, t_2)U(t_2, t_1)$, now says that $U(t)$ is a “one-parameter unitary group”, i.e., $U(t' + t'') = U(t')U(t'')$. So $U(t)$ is determined by values on a small interval near 0, more precisely, the functional analyses says that $U(t) = e^{it\hat{H}}$ for a (densely defined) Hermitian operator $\hat{H} = \frac{d}{dt}|_{t=0} U(t)$. It is called the Hamiltonian operator since it will turn out to be the quantization of the Hamiltonian function. $\hat{H} = \frac{d}{dt}|_{t=0} U(t)$.

3.3.2. Notation: time dependence of states. $\mathcal{H}$ is the configuration space of all possible states of the system. We will relate it to time 0 by denoting

$|\psi, t\rangle \overset{\text{def}}{=} \text{the state which from time 0 to time } t \text{ evolves to } \psi\rangle$

We will call it the “state $\psi$ at time $t$”. Since $|\psi\rangle = U(t)|\psi, t\rangle$, this is

$|\psi, t\rangle = U(t)^{-1}|\psi\rangle$.

In particular, we will be interested in the “pure position state $x$ at time $t$”, i.e., $|x, t\rangle|x_{po}, t\rangle = U(-t)|x_{po}\rangle$.

3.4. Example: a single electron in a vector space. Consider an electron in a vector space $V$. We postulate that

- the Hilbert space $\mathcal{H}$ has two “continuous basis” given by the pure position states $|v_{po}\rangle, v \in V$; and the pure momentum states $|p_{po}\rangle, p \in V^*$.
- The amplitudes between position and momentum states $(u, v \in V, \alpha, \beta \in V^*)$, are

$\langle v_{po}|u_{po}\rangle = \delta(u - v), \quad \langle \beta_{mo}|\alpha_{mo}\rangle = \delta(\alpha - \beta) \quad \text{and} \quad \langle \alpha_{mo}|x_{po}\rangle = e^{-\frac{i}{\hbar}\langle \alpha, x\rangle}$.

This means that (i) $L^2(V) \cong \mathcal{H} \cong L^2(V^*)$, and that (ii) the transition between these two interpretations of $\mathcal{H}$ is given by the Fourier transform.

Let us first consider an electron on a line.

\footnote{It is called the Hamiltonian operator since it will turn out to be the quantization of the Hamiltonian function. However, here we only found a logical need for the existence of such infinitesimal generator of time evolution.}
3.4.1. Idealized position states and momentum states. One starts with the (idealized) states \( |x\rangle = |x_{\text{po}}\rangle, x \in \mathbb{R} \) in which the position is absolutely known to be \( x \) (the index \( \text{po} \) stands for position).

The states \( |x_{\text{po}}\rangle, x \in \mathbb{R} \) form a continuous analogue of the pure bases states in a quantum system with a finite dimensional space of states (3.2). We can think of these position states as a continuous basis \( |x_{\text{po}}\rangle, x \in \mathbb{R} \), of the space of states. The idealized, i.e., distributional, nature of the states \( |x_{\text{po}}\rangle, x \in \mathbb{R} \), occurs in the amplitudes

\[
\langle y_{\text{po}}|x_{\text{po}}\rangle = \delta(x - y)
\]

which make sense only as distributions. The precise meaning of this being a continuous basis is that for any state \( \psi \) we now have

\[
|\psi\rangle = \int_{\mathbb{R}} \langle x_{\text{po}}|\psi\rangle \cdot x_{\text{po}} \, dx,
\]

in the “weak” sense that this is true for amplitudes

\[
\langle \phi|\psi\rangle = \int_{\mathbb{R}} \langle x_{\text{po}}|\psi\rangle \cdot \langle \phi|x_{\text{po}}\rangle \, dx,
\]

There is another “basis” consisting of the similar (idealized) states \( |p_{\text{mo}}\rangle \) such that the momentum is known to be \( p \) (\( p \in \mathbb{R} \)).

3.4.2. Transition amplitudes. \(^{15}\) The two continuous bases are related by the amplitudes

\[
\langle p_{\text{mo}}|x_{\text{po}}\rangle = e^{-\frac{i}{\hbar}px}.
\]

3.4.3. The functional interpretation of states as wave functions. Since any state \( \psi \) is determined by the amplitudes \( \langle \psi|x_{\text{po}}\rangle, x \in \mathbb{R} \), we can think of the state \( \psi \) as a function

\[
\psi(x) \overset{\text{def}}{=} \langle x_{\text{po}}|\psi\rangle \quad \text{on} \ \mathbb{R}.
\]

For instance, \( p_{\text{mo}}(x) = \langle x_{\text{po}}|p_{\text{mo}}\rangle = \langle p_{\text{mo}}|x_{\text{po}}\rangle = e^{\frac{i}{\hbar}px} \).

Lemma. (a) The functional interpretation of states gives an identification of Hilbert spaces \( \mathcal{H} \overset{\cong}{\rightarrow} L^2(V), \psi \mapsto \psi(x) \). In terms of the continuous bases \( |x_{\text{po}}\rangle, x \in V \), of the space of states, the amplitude is the \( L^2 \)-inner product \( \langle \phi|\psi\rangle = \int_V \phi(x)\psi(x) \, dx \).

(b) The transition between writing a state in terms of pure position states \( |x_{\text{po}}\rangle \) and the pure momentum states \( |p_{\text{mo}}\rangle \) is the Fourier transform.

\(^{15}\)What is the meaning of these? For instance why is not invariant under translations in \( V \) (so there is a distinguished position)? Certainly it should not be in invariant under \( V^* \)-translations, i.e., the change of momentum.
(c) With the functional interpretation of states, any “nice” operator $A$ on states, can be written as an integral operator

$$(A\psi)(y) \int dx \, \psi(x) \cdot A(x, y) \quad \text{with the kernel} \quad A(x, y) \overset{\text{def}}{=} \langle \delta_x | A | \delta_y \rangle.$$ 

Proof. (a) $\langle \phi | \psi \rangle = \langle \int_V dx \, \langle x_{po} | \phi \rangle \cdot | x_{po} \rangle \int_V dy \, \langle y_{po} | \phi \rangle \cdot | y_{po} \rangle$

$$= \int_{V^2} dx \, dy \, \langle \phi(x) | x_{po} \rangle \cdot \langle \psi(y) | y_{po} \rangle = \int_{V^2} dx \, dy \, \overline{\phi(x)} \psi(y) \cdot \delta(x - y) = \int_V \overline{\phi(x)} \psi(x) \, dx.$$ 

(b) We interpret a state $\psi$ as two functions $\psi_{po}(q) = \langle q_{po} | \psi \rangle$ on $V$, and $\psi_{mo}(p) = \langle p_{mo} | \psi \rangle$ on $V^*$. They are related by

$$\psi_{mo}(p) = \langle p_{mo} | \psi \rangle = \langle p_{mo} | \int_V dx \, \psi_{po}(x) \cdot | x_{po} \rangle = \int_V dx \, \psi_{po}(x) \cdot \langle p_{mo} | x_{po} \rangle = \int_V dx \, \psi_{po}(x) \cdot e^{-i p \cdot x}.$$ 

(c) We need $A$ to make sense on the distributional states $\delta_x = | x_{po} \rangle$, then one should be able to pair $A \delta_x$ to $\delta_y$, and finally, operator $A$ should really be given by these infinitesimal contributions in the sense of

$$(A\psi)(y) \overset{\text{def}}{=} \langle y_{po} | A \rangle \int dx \, \psi(x) \cdot | x_{po} \rangle = \int dx \, \psi(x) \cdot \langle y_{po} | A | x_{po} \rangle = \int dx \, \psi(x) \cdot A(x, y).$$

3.4.4. Quantum view on Fourier analysis. We have found a quantum view on $L^2(V)$ and on the Fourier transform. The Hilbert space $\mathcal{H}$ – the space of states – seems to be attached directly to the phase space $T^*V$.

However, for any Lagrangian vector subspace $L$ such as $V$ or $V^*$, $\mathcal{H}$ can be identified canonically with $L^2(L)$. Once we choose $L$, we get

- (i) a “distributional basis” $|a\rangle$, $a \in L$ of $\mathcal{H}$ (we also denote $|a\rangle$ by $\delta_a$),
- (ii) an interpretation $\mathcal{H} \overset{\text{def}}{=} L^2(L)$ by $\psi \rightarrow \langle \psi | a \rangle \overset{\text{def}}{=} \langle \psi | \delta_a \rangle$, and
- (iii) the inner product on $\mathcal{H}$ given by the amplitudes can now be written in terms of this “distributional basis” as $L^2$-product written in the intriguing interaction form

$$\int_L \langle \phi | \delta_a \rangle \cdot \langle \delta_a | \phi \rangle \, da.$$ 

3.4.5. Hamiltonian operator $\hat{H}$ and the Schroedinger’s equation. The time evolution $\psi(t)$ of a quantum state $\psi = \psi(0)$ is always given by the abstract Schroedinger’s equation

$$\frac{d}{dt} \psi(t) = i \frac{\hbar}{\hbar} \hat{H} \psi.$$ 

With the functional interpretation of states, a time dependent state $\psi(t)$ is a function of two variables $\psi(x, t)$, the Hamiltonian operator $\hat{H}$ becomes an integral operator with the
kernel $H(x,y) = \langle \delta_x | \hat{H} | \delta_y \rangle$, and the Schroedinger equation becomes an integral-differential equation

$$\frac{d}{dt} \psi(x,t) = \frac{i}{\hbar} \int \psi(x,t) H(x,y) \, dx.$$  

For an electron on a line and in a potential $V(x)$, the Hamiltonian operator is

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x).$$

This is the quantization of the Hamiltonian function (the total energy)

$$H = \frac{p^2}{2m} + V(x).$$

Actually, this is the historical origin of the idea of a deformation quantization.

With this explicit Hamiltonian operator we get the original Schroedinger equation

$$i\hbar \frac{d}{dt} \psi(x,t) = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x,t) + V(x)\psi(x).$$

By itself this equation describes (in principle!) all non-relativistic phenomena, such as energy levels in an atom and chemical bindings. The relativistic form was given by Dirac a year later (for instance it covers magnetism).

3.5. **Example: Harmonic oscillator.** Physicists think of the harmonic oscillator as a mathematical formalism which covers a number of physical situations. For instance a spring-mass-gravity system, a wave, or the above electron if a potential of the form $V(x) = c \cdot x^2$ with $c > 0$. Our goal will be to understand the system in terms of the spectrum of the infinitesimal time generator $\hat{H}$.

Our phase space will be the symplectic $\mathbb{R}^2$, i.e., $T^* \mathbb{R}$. However, we will find it convenient to think of it as a complex line $\mathbb{C}$ and use the holomorphic functions.

3.5.1. **The Hamiltonian function.** On the level of the classical equation of motion, harmonic oscillator is given by $mq'' = -kq$. In terms of the frequency parameter $\omega = \sqrt{k/m}$ this is

$$q'' + \omega^2 q = 0.$$  

The Hamiltonian function is the total energy

$$H(p,q) = \frac{1}{2m} p^2 + \frac{1}{2} kq^2.$$  

If we choose the units so that we have the unit mass, this is $H(p,q) = \frac{1}{2} p^2 + \frac{1}{2} \omega^2 q^2$. 
3.5.2. **The quantization in standard coordinates.** We think of the phase space as the cotangent bundle $T^*\mathbb{R}$ to the $q$-line $\mathbb{R}_q$. Then the Hilbert space is naturally viewed as the space $\mathcal{H} = L^2(\mathbb{R}_q)$ of functions on the $q$-line. From this point of view the functions $p, q$ on $T^*\mathbb{R}$ quantize to operators on $C^\infty(T^*\mathbb{R}) \subseteq \mathcal{H}$ that satisfy the unitarity prescription $[\hat{p}, \hat{q}] = \hbar \frac{\hat{h}}{i}\{p, q\}$. So, $\{p, q\} = 1$ requires

$$[\hat{p}, \hat{q}] = \hbar \frac{\hat{h}}{i}.$$ 

This leads to the normalization

$$\hat{p} \overset{\text{def}}{=} \frac{\hbar}{i} \frac{d}{dq}, \quad \hat{q} \overset{\text{def}}{=} q.$$ 

We will introduce the Hamiltonian operator $\hat{H}$ as a naive quantization of the Hamiltonian function $H = \frac{1}{2}p^2 + \frac{1}{2}\omega^2 q^2$, so

$$\hat{H} = \frac{1}{2}\hat{p}^2 + \frac{1}{2}\omega^2 \hat{q}^2.$$ 

Its eigenvectors in $L^2(\mathbb{R}_q)$ are the Hermite polynomials.

3.5.3. **The quantization in holomorphic coordinates.** We will pass to holomorphic coordinates on $T^*\mathbb{R} = \mathbb{R}^2$ given by

$$a = \frac{1}{\sqrt{2\omega}}(\omega q + ip), \quad a^* = \frac{1}{\sqrt{2\omega}}(\omega q - ip), \quad q = \frac{1}{\sqrt{2\omega}}(a + a^*), \quad p = \frac{\sqrt{\omega/2}}{i}(a - a^*).$$

Here, $a^*$ is the complex conjugate of the function $a$. We will think of $a^*$ as the basic holomorphic coordinate $z$.

3.5.4. **Lemma.** (a) The Poisson structure $\{p, q\} = 1$, is in new coordinates

$$\{a, a^*\} = i.$$ 

(b) The quantizations of the generators are operators satisfying the commutation relation from $D_{\hbar}$:

$$[\hat{a}, \hat{a}^*] = \hbar.$$ 

(c) At $\hbar = 1$ one can realize the algebra generated by $\hat{a}, \hat{a}^*$ and the above relation, as the algebra of polynomial differential operators on holomorphic functions in $z$, by

$$a^* = z \quad \text{and} \quad a = \hbar \partial_z = \partial.$$ 

(d) The Hilbert space $\mathcal{H}$ is realized as the Hilbert space of entire functions in $z \overset{\text{def}}{=} a^*$. It is a completion of the space of polynomials $\mathbb{C}[z] = \mathbb{C}[a^*]$ with respect to the inner product

$$(f, g) \overset{\text{def}}{=} \int_\mathbb{C} f(z) g(z) e^{-z\bar{z}} \frac{dz\bar{z}}{2\pi i} = \int_\mathbb{C} f(a^*) g(a^*) e^{-a^*a} \frac{da^*da}{2\pi i}.$$
Proof. (a) The Poisson structure \( \{ p, q \} = 1 \) is now
\[
\{ \hat{a}, \hat{a}^* \} = \{ \omega q + ip, \omega q - ip \} = \frac{1}{2\omega} \{ \omega q + ip, \omega q - ip \} = \frac{1}{2\omega} 2\omega i \{ p, q \} = i.
\]
(b) The quantizations of the generators are operators satisfying the commutation relation
\[
[\hat{a}, \hat{a}^*] = \frac{i}{\hbar} \{ a, a^* \} = \frac{i}{\hbar} \hbar.
\]
(c) For the above operators
\[
[\hat{a}, \hat{a}^*] = \left[ \frac{1}{\sqrt{2\omega}} (\omega \hat{q} + i\hat{p}), \frac{1}{\sqrt{2\omega}} (\omega \hat{q} - i\hat{p}) \right] = \frac{1}{2\omega} 2\omega i [\hat{p}, \hat{q}] = i \frac{\hbar}{i} = \hbar.
\]
3.5.5. Corollary. (a) The Hamiltonian operator is
\[
\hat{H} = \omega \hat{a}^* \hat{a} = \hbar \omega z \partial.
\]
(b) The eigenvectors of \( \hat{H} \) are the powers of \( z = a^* \). We normalize them to unit vectors \( \psi_n = \frac{(a^*)^n}{\sqrt{n!}} = \frac{z^n}{\sqrt{n!}} \). Then
\[
\hat{H} \psi_n = \hbar \omega (n + \frac{1}{2}) \psi_n.
\]
The physical meaning is that the energy of the state \( \psi_n \) is \( \hbar \omega (n + \frac{1}{2}) \).

Proof. (a) We just recalculate the Hamiltonian from standard coordinates
\[
\hat{H} = \frac{1}{2} \hat{p}^2 + \frac{1}{2} \omega^2 \hat{q}^2 = \frac{1}{2} \left[ \frac{\sqrt{\omega}}{i} (\hat{a} - \hat{a}^*) \right]^2 + \frac{1}{2} \omega^2 \left[ \frac{1}{\sqrt{2\omega}} (\hat{a} + \hat{a}^*) \right]^2
\]
\[
= \frac{-\omega}{4} [\hat{a}^2 + (\hat{a}^*)^2 - \hat{a} \hat{a}^* - \hat{a}^* \hat{a}] + \frac{\omega}{4} [\hat{a}^2 + (\hat{a}^*)^2 + \hat{a} \hat{a}^* + \hat{a}^* \hat{a}] = \frac{\omega}{2} (\hat{a} \hat{a}^* + \hat{a}^* \hat{a})
\]
\[
= \omega (\hat{a}^* \hat{a} + \frac{\hbar}{2}) = \omega (\hat{a} \hat{a}^* - \frac{\hbar}{2}) = \hbar \omega (z \partial + \frac{1}{2}).
\]
(b) Now,
\[
\hat{H} z^n = \hbar \omega (z \partial + \frac{1}{2}) z^n = \hbar \omega (n + \frac{1}{2}) z^n.
\]
The change of variables \( z = x + iy = re^{i\theta} \), gives \( dz d\bar{z} = (dx + idy)(dx - idy) = 2i dy dx = 2i r dr d\theta \), hence
\[
\langle \psi_n, \psi_m \rangle = \int \frac{z^n}{\sqrt{n!}} \frac{z^m}{\sqrt{m!}} e^{-z\bar{z}} \frac{dz d\bar{z}}{2\pi i} = \int_{r=0}^{\infty} \int_{\theta=0}^{2\pi} \frac{r^{n+m}}{\sqrt{n! \sqrt{m!}}} e^{i\theta(m-n)} \cdot e^{-r^2} \frac{2i r dr d\theta}{2\pi i}
\]
\[
= \delta_{nm} \cdot \int_{r=0}^{\infty} \frac{r^{2n}}{n!} e^{-r^2} 2r dr = \delta_{nm} \cdot \int_{u=0}^{\infty} \frac{u^n}{n!} e^{-u} du
\]
\[\text{[16]} \text{better than the special functions!}\]
3.5.6. The energy spectrum of a harmonic oscillator. The energy spectrum of the theory is the set of eigenvalues of the Hamiltonian operator. We found that the energy spectrum of a harmonic oscillator with frequency $\omega = \sqrt{k/m}$ is\footnote{We have calculated with the unit mass, but this did not affect the spectrum.}

$$\hbar \omega (n + \frac{1}{2}), \quad n = 0, 1, \ldots \quad \text{i.e.,} \quad \frac{\omega \hbar}{2}, \frac{3\omega \hbar}{2}, \frac{5\omega \hbar}{2}, \ldots$$

3.5.7. The representation theoretic organization of the Hilbert space (language of creation and annihilation). Instead of considering a particular realization of the Hilbert space of states $\mathcal{H}$, one often considers the way it is organized by the action of the quantized algebra of observables, i.e., an algebraic representation of the algebra of observables on a dense vector subspace $V$ of $\mathcal{H}$.

The action of the algebra $D_{\hbar}(\hbar) = \bigoplus \mathbb{C}\tilde{p}^{i}\tilde{q}^{j} = \bigoplus \mathbb{C}\tilde{a}^{i}\tilde{a}^{*j}$ on $\mathcal{H}$ is given by the action of two operators $a$ and $a^*$. The space is graded by the action of the Hamiltonian operator $\tilde{H}$ which corresponds to energy. We have observed that

\textbf{Lemma.} (a) Eigenvalues of $\tilde{H}$ are bounded from below and there is a unique vector (line) on which it has the smallest eigenvalue. We call it the vacum vector $|\text{vac}\rangle$.

(b) The vacum is killed by $a$ and applying $a^*$ to the vacum creates a bases $(a^*)^n|\text{vac}\rangle$, $n \geq 0$ of $\mathcal{H}$.

So we call $a^*$ a creation operator, and $a$ the annihilation operator, since $a$ kills the vacum and each vector is killed by some power of $a$. From this algebraic point of view, the interesting space is the subspace $V = \bigoplus_{n\geq 0} \cdot (a^*)^n|\text{vac}\rangle \subseteq \mathcal{H}$.

3.5.8. Planck constant. To simplify formulas one often uses a choice of units for which $\hbar = 1$. However, one will sometimes think of $\hbar$ as a variable. \textit{Physically} it corresponds to the idea that when we change the scale that we are interested in, our view of the size of $\hbar$ changes. \textit{Formally}, it is a useful trick to have an additional variable.

3.5.9. Harmonic oscillator in a metric vector space. We consider a real vector space $M$ with a metric $G$ which we diagonalize into

$$G = \sum (x^\mu)^2 - \sum (y^\nu)^2.$$
4. Path integrals: the Lagrangian approach

Feynman’s idea of a path integral formula for the amplitude of an event is that one integrates over all possible evolutions (histories) of the system that result in this event. This is said to be the most physically fundamental and useful point of view. However, we will follow the historical development and deduce the path integral formula as a consequence of the operator approach.

4.0.10. Theorem. One can write this amplitude as a Feynman integral over paths in the tangent bundle $\langle x_f, t_f | x_i, t_i \rangle = \int_{(x, y): [t_i, t_f] \rightarrow T^C, x(t_i) = x_i, x(t_f) = x_f} d(x, y) e^{\frac{i}{\hbar} S[x, y]}$.

Here, one uses the action on such paths $S[x, y] \stackrel{\text{def}}{=} \int_{t_i}^{t_f} dt L(x(t), \dot{x}(t), y(t)) = \int_{t_i}^{t_f} dt \frac{1}{2} \dot{x} \cdot p - V(x)$, which for velocity curves $(x, \dot{x})$ reduces to the standard action $S[x]$.

In the remainder we check this claim. What we get is rather a formula involving a limit over some $N$-step piecewise curves $\lim_{N \to \infty} \int_{\text{"N-curves"}} d(x, p)_{(t_*, x_*, p_*)} e^{\frac{i}{\hbar} S[(x, p)_{(t_*, x_*, p_*})]}$.

The path integral formula in the theorem is an interpretation (or shorthand) for this.

4.0.11. The setting for the formula. We consider a quantum mechanical system in a Hamiltonian formalism, and we calculate the amplitude that the system evolves from the initial pure position state $x_i$ at the initial time $t_i$, to the final pure position state $x_f$ at the final time $t_f$. This is a matrix coefficient of the evolution operator $U(\Delta t) = e^{\frac{i}{\hbar} \Delta t \cdot \hat{H}}$,

$$\langle x_f, t_f | x_i, t_i \rangle = \langle x_f | U(t_f - t_i) | x_i \rangle.$$ 

We will write it in the form of a Feynman integral over all paths into $TC$ (or $T^*C$).

4.0.12. The appearance of the Feynman integral. It is obtained by trying to account for the amplitude that $x$ at time $t_i$ will evolve to $x_f$ at time $t_f$, as an integral over the time interval $[t_i, t_f]$, i.e., as an accumulation of contributions over small time subintervals $[t_{k-1}, t_k]$.

(1) Inserting points $t_1, \ldots, t_{N-1}$ between $t_i$ and $t_f$ leads to an integral $\int_{R^{N-1}} dx_1 \cdots dx_{N-1}$ over all possible positions $x_k$ at moments $t_k$. 

---

18Hopefully (?) this is an example of the philosophy of adding an auxiliary field to the action from I.3.4.1. Here, one is adding $y$ to $x$. 
(2) For short time intervals \([t_{k-1}, t_k]\), the linear approximation (in \(t_k - t_{k-1}\)) of the amplitude is given by an integral \(\int_{\mathbb{R}^N} dp_1 \cdots dp_N\) of contributions from all possible momenta \(p_k\).

(3) So, an approximation of the amplitude based on the partition \(t_i = t_0, \ldots, t_N = t_f\), ends up being an integral over \(N\) possible positions \((x_k, p_k)\) in \(T^*\mathcal{C}\).

(4) Interestingly, the integrand is exactly \(e^{i\hat{S}}\) for an obvious action \(S\) of the piecewise linear path in \(T^*\mathcal{C}\) given by these points. So we get the Feynman integral over all “\(N\)-step curves” in \(T^*\mathcal{C}\).

(5) Finally, as \(N\) gets large we get the precise formula for the amplitude – and we interpret the limit of \(N\)-step curves to be the moduli of all curves in \(T^*\mathcal{C}\).

4.1. **Matrix coefficients of \(\hat{f}\) for the normal ordering quantization.** Recall that for a state \(\psi\) we denote by 
\[
|\psi, t\rangle = U(t^{-1})|\psi\rangle,
\]
the “state \(\psi\) at time \(t\)”, i.e., the state which in the period from time 0 to time \(t\) evolves to \(\psi\).

4.1.1. **Problem.** Find the amplitude that the system evolves from the initial pure position state \(x_i\) at the initial time \(t_i\), to the final pure position state \(x_f\) at the final time \(t_f\). This is the matrix coefficient
\[
\langle x_f, t_f|x_i, t_i\rangle = \langle U(-t_f)x_f|U(-t_i)x_i\rangle = \langle x_f|U(t_f - t_i)|x_i\rangle.
\]

4.1.2. **Assumptions on the physical system.** We will do all calculations for a system with the Hamiltonian function and the Hamiltonian operator
\[
H = \frac{\hat{p}^2}{2m} + V(x) \quad \text{and} \quad \hat{H} = \frac{\hat{p}^2}{2m} + V(x) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x).
\]
Say, for an electron on a line and in a potential \(V\).

The Hilbert space \(\mathcal{H}\) has distributional basis given by pure position states \(|a_{p_0}\rangle, a \in \mathbb{R}\), and the pure momentum states \(|b_{m_0}\rangle, b \in \mathbb{R}\). The algebra of operators on \(\mathcal{H}\) is generated by the position and momentum operators \(\hat{x} = x\) and \(\hat{p} = \frac{i\hbar}{\partial_x}\). They act by
\[
\hat{x}|a_{p_0}\rangle = a|a_{p_0}\rangle, \quad \hat{p}|b_{m_0}\rangle = b|b_{m_0}\rangle.
\]
i.e., in terms of the functional interpretation:
\[
x \cdot \delta_a = a \cdot \delta_a \quad \text{and} \quad \frac{\hbar}{i} \partial_x e^{ixb} = b e^{ixb}.
\]
4.1.3. Normal ordering and Weyl ordering. In general the Hamiltonian operator is obtained from the Hamiltonian function by a (purely defined) quantization procedure which requires some choices. Here we notice that the above Hamiltonian function is simple enough, so that its quantization is independent of the choices.

Recall that the identification of functions on $T^\ast \mathbf{C}$ and $D_{\mathbf{C}}$ is not canonical - it is only defined canonically on the generators $x$ and $p$. The way one assigns operators to functions, $f \mapsto \hat{f}$ is therefore not canonical, so the choice of $\hat{H}$ is not canonical in general.

The simplest identification is obtained by $p^i q^j \leftrightarrow \hat{p}^i \hat{q}^j$, hence $\hat{(p^i q^j)} = \hat{p}^i \hat{q}^j$. This prescription can be described by “$p$ before $q$”, or “annihilation before creation”. It is called normal ordering.

Another useful choice is the Weyl ordering which chooses symmetric expressions, say, 

$$\hat{p} \hat{q} = \frac{1}{2} (\hat{p} \hat{q} + \hat{q} \hat{p}).$$

All such prescriptions work the same on the Hamiltonian function if does not mix the variables, i.e $H = H_{\text{kinetic}}(p) + H_{\text{potential}}(x)$, as we have assumed above.

4.1.4. Lemma. For $f \in \mathbb{C}[x,p] = \mathbb{C}[q,p]$

$$(a) \quad \langle p_\text{mo} | \hat{f} | x_\text{po} \rangle = e^{-\hat{p}x} f(x,p),$$

$$(b) \quad \langle y_\text{po} | \hat{f} | x_\text{po} \rangle = \int_{\mathbb{R}} dp \ e^{\frac{i}{\hbar} p(y-x)} f(x,p).$$

$$(c) \quad \langle y_\text{po} | e^{\hat{f}} | x_\text{po} \rangle = \int_{\mathbb{R}} dp \ e^{\frac{i}{\hbar} p(y-x)+sf(x,p)} \pmod{s^2}.$$ 

Proof. (a) It suffices to check (a) for $f = p^ix^j$. For the normal ordering quantization the LHS is $\langle p_\text{mo} | \hat{p}^i \hat{x}^j | x_\text{po} \rangle = \langle p_\text{mo} | x_\text{po} \rangle \cdot \hat{p}^i \hat{x}^j = e^{-\hat{p}x} f(x,p)$. Now (b) follows 

$$\langle y_\text{po} | \hat{f} | x_\text{po} \rangle = \int_{\mathbb{R}} dp \ \langle y_\text{po} | p_\text{mo} \rangle \cdot \langle p_\text{mo} | \hat{f} | x_\text{po} \rangle = \int_{\mathbb{R}} dp \ e^{\frac{i}{\hbar} p(y-x)} e^{-\frac{i}{\hbar} p^ix^j} f(x,p).$$

In (c) the claim is that the RHS is the linearization of the LHS. At $s = 0$ both sides equal $\delta(x-y)$, applying $\frac{d}{ds}|_{s=0}$ to (c) yields known claim (b).

4.1.5. Remark. (1) Our formal calculations are meaningful in the appropriate spaces of distributions. For instance, the integral in (b) does not converge absolutely. The only meaning of this integral is that it is a distribution on the $xy$-plane, hence absolute convergence holds only for integrals

$$\int_{\mathbb{R}^2} dx \ dy \ \phi(x,y) \cdot \langle y_\text{po} | \hat{H} | x_\text{po} \rangle = \int_{\mathbb{R}^3} dx \ dy \ dp \ \phi(x,y) \cdot e^{\frac{i}{\hbar} p(y-x)} H(x,p)$$

where function $\phi$ is in the Schwartz space $\mathcal{S}(\mathbb{R}^2)$. 
(2) Claim (c) says that the short time evolution of pure position states equals, modulo $(t'' - t')^2$,
\[
\langle x''_{po}, t'' | x'_{po}, t' \rangle = \langle x''_{po} | U(t'' - t') | x'_{po} \rangle = \langle x''_{po} | e^{-\frac{i}{\hbar} (t'' - t') \hat{H}} | x'_{po} \rangle
\]
\[
= \int_{\mathbb{R}} dp \ e^{\frac{i}{\hbar} \left[p(x'' - x') - (t'' - t')H(x, p)\right]}.
\]

4.2. Localization of the amplitude on the time interval.

4.2.1. Insertion of a point in the time interval. For times $a < b < c$ and positions $x, z$,
the formula
\[
\langle z, c | x, a \rangle = \int_{\mathbb{R}} dy \ \langle z, c | y, b \rangle \cdot \langle y, b | x, a \rangle,
\]
can be justified on the level of physics (any evolution from $x$ at time $a$ to $z$ at time $c$ has
to go through some position $y$ at time $b$), or mathematics (the states $|y, b\rangle$, $y \in \mathbb{R}$, form
a distributional basis of $\mathcal{H}$).

4.2.2. Amplitude as an integral over possible intermediate positions. Now we insert $N - 1$
points into the time interval: $t_i = t_0 < t_1 < \cdots < t_N = t_f$ and we denote $x_0 = x_i$, $x_N = x_f$. Then
\[
\langle x_f, t_f | x_i, t_i \rangle = \int_{\mathbb{R}^{N-1}} dx_1 \cdots dx_{N-1} \ \prod_{1 \leq k \leq N} \langle x_k, t_k | x_{k-1} t_{k-1} \rangle.
\]
(There are more factors in the product then variables of integration.)

4.2.3. Amplitude as an integral over possible intermediate positions and momenta. Suppose that
the intervals are of the same length: $t_k - t_{k-1} = \varepsilon \overset{\text{def}}{=} \frac{t_f - t_i}{N}$. According to the
linearization formula of the time evolution (4.1.5.2), modulo $\varepsilon^2$ we have
\[
\langle x_k, t_k | x_{k-1} t_{k-1} \rangle = \int_{\mathbb{R}} dp_k \ e^{\frac{i}{\hbar} \left[p_k(x_k - x_{k-1}) - \varepsilon \cdot H(x_{k-1}, p_k)\right]}.
\]
Therefore, modulo $\varepsilon$\n\[
\langle x_f, t_f | x_i, t_i \rangle = \int_{\mathbb{R}} dx_1 \cdots dx_{N-1} \ \int_{\mathbb{R}^N} dp_1 \cdots dp_N \ e^{\frac{i}{\hbar} \left[\sum_{1}^{N} p_k(x_k - x_{k-1}) - \varepsilon \cdot H(x_{k-1}, p_k)\right]}
\]
(since $\prod_{1}^{N} (c_i + \varepsilon^2) = \prod_{1}^{N} c_i + \varepsilon^2 \sum_{1}^{N} \prod_{j \neq i} c_j + \cdots$, and $\varepsilon^2 \sum_{1}^{N}$ behaves like 1).

4.2.4. Interpretations. The integral over all choices of $k$ positions and $k$ momenta will be
interpreted as an integral over all curves $(x, p)$ in $T^*C$. Then the phase of the integrand
will be viewed as an action for curves in $T^*C$. 
4.2.5. An action for paths in the (co)tangent bundle. In a traditional manner we start
with paths \( (x, p) : [t_i, t_f] \rightarrow TC \) into the tangent bundle. For these we choose a version of
the usual Lagrangian function on \( TC \) as
\[
\mathcal{L}(x, \dot{x}, y) = T(\dot{x}, y) - V(x) = \frac{m}{2} \dot{x} \cdot y - V(x).
\]
When we will translate this into a path \((x, p)\) in the phase space \( T^*C \), (using \( p = ym \)), we
get a version of the usual Hamiltonian function
\[
\mathcal{H}(x, \dot{x}, p) = \dot{x} \cdot p - L(x, \dot{x}, y) = \frac{1}{2} \dot{x} \cdot p + V(x).
\]
Observe that for velocity and momentum curves (i.e., \( y = \dot{x} \)), these are the usual La-
grangian and Hamiltonian of a path \( x \) in \( C \).

The corresponding action for paths in \( TC \) (or \( T^*C \)) is
\[
S(x, y) = \left\{ \int_{t_i}^{t_f} dt \mathcal{L}(x(t), \dot{x}(t), y(t)) \right\} = \left\{ \int_{t_i}^{t_f} dt \frac{1}{2} \dot{x} \cdot p - V(x) \right\} = \left\{ \int_{t_i}^{t_f} dt \dot{x} \cdot p - \mathcal{H}(x, \dot{x}, p) \right\} = S(x, p)
\]

4.2.6. The integrand and the piecewise linear curves in \( T^*C \). We can use the data
\((t_s, x_s, p_s) = (t_0, ..., t_N, x_0, ..., x_N, p_1, ..., p_N)\), to cook up a piecewise linear (discontinuous)
curve \((x, p)_{(t_s, x_s, p_s)} : [t_i, t_f] \rightarrow T^*C \), such that for \( t_{k-1} \leq t < t_k \),
\[
x(t) \text{ is a line from } x_{k-1} \text{ to } x_k, \text{ and}
\]
\[
p(t) \text{ is a constant } p_k.
\]
We will call these curves the \( N \)-curves, and think of them as the level \( N \) approximation
of the moduli of curves in \( T^*C \). Then the phase of the integrand is the action of this
\( N \)-curve
\[
\text{phase} = \sum_{k=1}^{N} p_k (x_k - x_{k-1}) = \frac{1}{\varepsilon} \mathcal{H}(x_{k-1}, p_k) = \sum_{k=1}^{N} (t_k - t_{k-1}) \cdot |p_k \frac{x_k - x_{k-1}}{t_k - t_{k-1}} - \mathcal{H}(x_{k-1}, p_k)|
\]
\[
= S[(x, p)_{(t_s, x_s, p_s)}]
\]
If \( p \) would coincide with \( m \cdot \dot{x} \) this would be the formula for the action, en

4.2.7. The limit of the integrand. For a curve \((x, p)\) in \( T^*C \), we define \( x_k, p_k \) as values at
\( t_k \), and we get a piecewise linear discontinuous approximation \((x, p)_{(t_s, x_s, p_s)}\) of the curve
\((x, p)\). We notice that the action is continuous in these approximations, i.e.,
\[
\lim_{N \rightarrow \infty} S[(x, p)_{(t_s, x_s, p_s)}] = \lim_{N \rightarrow \infty} \sum_{k=1}^{N} (t_k - t_{k-1}) \cdot |p_k \frac{x_k - x_{k-1}}{t_k - t_{k-1}} - H(x_{k-1}, p_k)|
\]
\[
= \int_{t_i}^{t_f} dt \ p(t) \cdot \dot{x}(t) - H(x(t), p(t)) = S[x, p].
\]

\[^{19}\text{However, our } \mathcal{L} \text{ and } \mathcal{H} \text{ are functions on } TC \times C TC \text{ and } T^*C \times C T^*C, \text{ rather then } TC \text{ and } T^*C.\]
4.2.8. Now we have interpreted the approximate (modulo $\varepsilon = (t_f - t_i)/N$), calculation of the amplitude $\langle x_f, t_f | x_i, t_i \rangle$ as the integral over all "$N$-curves" in $T^*C$

$$\langle x_f, t_f | x_i, t_i \rangle \approx \lim_{N \to \infty} \int_{\mathbb{R}^{N-1}} dx_1 \cdots dx_{N-1} \int_{\mathbb{R}^N} dp_1 \cdots dp_N e^{\frac{i}{\hbar} \left[ \sum_{k=1}^{N} p_k (x_k - x_k - 1) - \varepsilon \cdot H(x_k-1, p_k) \right]}$$

$$= \lim_{N \to \infty} \int \text{"$N$-curves" } (x, p)_{(t_i, t_f, p_i)} \text{ in } T^*C \, d(x, p)_{(t_i, t_f, p_i)} e^{\frac{i}{\hbar} S[(x, p)_{(t_i, t_f, p_i)}]} .$$

Finally, when we pass to the limit $N \to \infty$, we get an exact formula for the amplitude

$$\langle x_f, t_f | x_i, t_i \rangle = \lim_{N \to \infty} \int \text{"$N$-curves" } (x, p)_{(t_i, t_f, p_i)} \text{ in } T^*C \, d(x, p)_{(t_i, t_f, p_i)} e^{\frac{i}{\hbar} S[(x, p)_{(t_i, t_f, p_i)}]} .$$

To interpret it we imagine that the moduli of $N$-curves should converge to the moduli of all curves in $T^*C$. This gives

$$\langle x_f, t_f | x_i, t_i \rangle = \int_{(x, p): [t_i, t_f] \to T^*C} d(x, p) e^{\frac{i}{\hbar} S[(x, p)]}$$

$$= \int_{(x, y): [t_i, t_f] \to T^*C} d(x, y) e^{\frac{i}{\hbar} S[(x, y)]} .$$

4.2.9. Remark. The action here involves curves $(x, y)$ into $TC$ (the same as $(x, p)$ into $T^*C$), but this is different then just the velocity curves of curves into $C$. 

5. Example: Relativistic free particle

Here we repeat the calculation of amplitude for a free particle, except that in the relativistic setting we work in a larger space with auxiliary fields: an arbitrary parameterization $x$ (rather than the parameterization by time), and a worldsheet metric $g$. In terms of the analogy with strings we are now working with the (relativistic!) Polyakov-type action for particles.

5.1. Amplitude calculation. For two events $x_i, x_f$ in the spacetime, the amplitude of $x_i$ evolving to $x_f$ is

$$\langle x_f | x_i \rangle = N \cdot \int_{x_i}^{x_f} Dx \int_{\text{metrics } g} Dg \ e^{\frac{i}{\hbar} S[x,g]}.$$ 

Here $N$ is an unknown constant. The notation in the first integral means that we fix an interval $I = [\tau_i, \tau_f]$ and we integrate over all paths $x : I \rightarrow M$ with the boundary conditions $x(\tau_i) = x_i$ and $x(\tau_f) = x_f$.

5.2. Gauge fixing (= Equivariant integration). Consider an integral $\int_X f$ with a symmetry group $\Gamma$, i.e., $\Gamma$ acts on $X$ and preserves the function $f$ and the measure on $X$. Then one can try to define the equivariant integral as

$$\int^\Gamma_X f \overset{\text{def}}{=} \int_{\Gamma \backslash X} \bar{f},$$

where $\bar{f}$ is the descent of $f$ to the quotient $X/\Gamma$ (and one uses the quotient measure on $X/\Gamma$). Actually, if the symmetry is infinite, this is the only meaning one can attach to the original integral $\int_X f$.

Next, if one knows a section $s : \Gamma \backslash X \rightarrow X$ of the quotient map (in physics section is called gauge), then one can rewrite the integral on the quotient as an integral $\int_{s(\Gamma \backslash X)} f$ on the section $s(\Gamma \backslash X)$ which is a subspace of $X$, hence less abstract.

5.2.1. Problems. One needs

(1) a “good” quotient $\Gamma \backslash X$,
(2) the correct measure (the Faddeev-Popov measure), on the section $s(\Gamma \backslash X)$.

---

\(^{20}\)FIND SOURCE OR SIMPLIFY THE STRING VERSION

\(^{21}\)In spacetime we will use terminology event rather then position, to emphasize that the difference from the space position.
5.2.2. The standard example. Let $\Gamma = \mathbb{Z}$ act on $X = \mathbb{R}$ by translations and consider an invariant (i.e., periodic) function $f(x + n) = f(x)$, $x \in \mathbb{R}, n \in \mathbb{Z}$. Now the $\mathbb{Z}$-equivariant integration of a periodic function on $\mathcal{R}$ is just the integration on the circle $\Gamma \backslash \mathbb{R} = S^1$:

$$
\int_{\mathbb{R}} dx \ f(x) \overset{\text{def}}{=} \int_{S^1} dx \ f(e^{2\pi i x}) = \int_{0}^{1} dx \ f(x),
$$

and in the last equality we use the standard (discontinuous) section $S^1 \to [0, 1) \subseteq \mathbb{R}$.

5.3. Gauge fixing of the metric $g$ and length as a Teichmüller parameter (left-over parameter).

5.4. Amplitudes.
Part III. (Quantum) Field Theory

0. Intro

Text. Borcherds, Quantum Field Theory.
Quantum Field Theory is an extension of the formalism of Quantum Mechanics from particles to fields.

0.0.1. Field theories. Mechanics deals with a particle like objects that are supported at a point of the spacetime. At this point a particle information (mass, velocity, charge,...) is a vector.
Field theories deal with objects ("fields") that are distributed over some space. So, a field is something like a function on a manifold. More generally fields are distributions with values in a vector bundle (or something like that).
The most obvious are the fields on the spacetime, say, the electric fields, etc. However, there are also fields on the worldsheet, say, a metric on the worldsheet.

0.0.2. $\sigma$-models. These are the worldsheet field theories where the main field is the evolution map $x$ (or its component functions), viewed as a function on the worldsheet $\Sigma$ with values in the spacetime $(M,G)$, i.e., a manifold $M$ with a metric $G$ (which need not be positive definite).
So mathematically this is the science of maps between manifolds. The ones of physical interest are the ones that minimize certain action such as (i) volume or (ii) energy. In mathematics such maps are called (i) minimal submanifolds, and (ii) harmonic maps.
The terminology $\sigma$-model comes simply from the habit of denoting the source of the map (the worldsheet) by $\Sigma$.
So, classical mechanics can be viewed as a $\sigma$-model worldline field theory

0.0.3. Lagrangian view: Feynman’s path integrals. The key notion is the action $S[x]$. A classical theory formulates physical laws by:

\[
evolution x \text{ is a solution of the criticality equation, i.e., } d_x S = 0.
\]
A quantum theory in ...

0.0.4. Calculus of variations. This is the mathematical framework for the classical Lagrangian approach, i.e., in classical physics the evolution is a solution of an EL-equation, we say that it is a classical solution.
The classical worldsheet field theories (i.e., classical $\sigma$-models), for strings and branes will require calculus of variations for any worldsheet manifold $\Sigma$. (See I.1.3 for the case $\text{dim}(\Sigma) = 1$.)
1. Classical Field Theory: Calculus of Variations

1.0.5. Goals. Classically, physical laws are expressed in the Lagrangian approach by requiring that the certain quantity $S[x]$ attached to a possible evolution $x$ of the system is minimal. One calls $S[x]$ the action of $x$, the criticality equation $d_x S = 0$ is called the Euler-Lagrangian equation, and from the quantum point of view the solutions of the EL-equation are called the classical solutions.

Below we calculate the EL-equation in a standard special setting and in local coordinates.

1.0.6. The setting. Here, the field $x$ is a map $x: \Sigma \rightarrow M$ from a manifold (with boundary) $\Sigma$ to a manifold $M$. The action $S[x]$ is an integral

$$S[x] \overset{\text{def}}{=} \int_\Sigma d\tau \, L(x(\tau), \dot{x}(\tau), \tau),$$

of a function $L(x,v,\tau)$ which is a (possibly $\Sigma$-dependent) function of the position $x$ in $M$, and velocity $v \in T_xM$. So the Lagrangian $L(x,v,\tau)$ is a function on $TM \times \Sigma$.

1.0.7. Reduction to the standard local setting. The criticality equation will be local in $\Sigma$, and we calculate it in local coordinates $\tau^i$ on an open piece $\Sigma \subseteq \Sigma$, which we identify via the coordinates $\tau^i$ with a coordinate box $\Sigma \subseteq \mathbb{R}^p$. In the process, the measure $d\tau$ on $\Sigma$ becomes a multiple $\mu(\tau) \cdot d\tau^0 \cdot \cdots \cdot d\tau^p$ of the standard measure on $\Sigma \subseteq \mathbb{R}^p$, and we incorporate it into a new Lagrangian $L(x,v,\tau) = \mu(\tau) \cdot L(x,v,\tau)$. We do all our calculations in the setting of a box $\Sigma$ and the corresponding Lagrangian $L = \mu \cdot L$.

1.0.8. EL-equations. In the above local coordinates the criticality equation becomes

(1) on $\Sigma$, there are ingredients $L_x$ and $L_{\frac{\partial}{\partial \tau}}$ that come from the non-dynamical dependence on the field $x$ (dependence of $L$ on $x$, no $\Sigma$-derivatives), and the dynamical dependence on $x$ (dependence of $L$ on the $\Sigma$-derivatives of $x$):

$$L_x = \frac{d}{d\tau}(L_{\frac{\partial}{\partial \tau}}),$$

(2) on $\partial \Sigma$ we have the Von Neumann Boundary condition

$$L_{\frac{\partial}{\partial n}} = 0.$$

1.0.9. Mathematical analysis. Here we really do calculus in topological vector spaces, but I will omit the relevant technology.

1.1. Variation $\frac{\delta S}{\delta x}$ and pointwise variation $\frac{\delta S}{\delta x(t)}$ of the action.
1.1.1. **Action.** We are interested in the critical points of a functional \( S[x] \) on the space of all paths \( x: \Sigma \to M \) in a manifold \( M \). We will call \( S[x] \) the *action* of the path \( x \), it will be an integral

\[
S[x] \overset{\text{def}}{=} \int_{\Sigma} d\tau \ L(x(\tau), \dot{x}(\tau), \tau) ,
\]

of a function \( L(x, v, \tau) \) which is a (possibly time dependent) function of the position \( x \) in \( M \), and velocity \( v \in T_xM \). So the *Lagrangian* \( L(x, v, \tau) \) is a function on \( TM \times \Sigma \).

1.1.2. **Variation \( \frac{\delta S}{\delta x} \) and pointwise variation \( \frac{\delta S}{\delta x(t)} \).** The variation means just the differential \( \frac{\delta S}{\delta x} = d_x S \in T^*_x[C^\infty(\Sigma, M)] \).

Its distributional version is the variation with respect to the value of \( x \) at a fixed point \( t \in \Sigma \). For a tangent vector \( v \in T_{x(t)}M \) we consider the variation of \( S[x] \) when one moves \( x \) by a distributional path \( u = \delta_t(\tau) \cdot v \) concentrated at the point \( t \in \Sigma \), so

\[
\frac{\delta S}{\delta x(t)} v \overset{\text{def}}{=} \frac{\delta S}{\delta x} \delta_t(\tau) \cdot v = (d_x S) \delta_t(\tau) \cdot v.
\]

1.1.3. **Local situation.** For simplicity we will calculate in the local situation, so \( \Sigma \) will be a rectangle \( \prod I_k \subseteq \mathbb{R}^p \) for some intervals \( I_k = \{a_i \leq \tau^i \leq b_i\} \), and the target \( M \) will be a vector space. Now

\[
S[x] = \int_{\Sigma} d\tau \ L(x, \dot{x}, \tau) = \int_{I_1} \cdots \int_{I_p} d\tau_1 \cdots d\tau_p \ L(x(\tau), \dot{x}(\tau), \tau).
\]

The dot denotes the derivative with respect to \( \tau \). So, \( \dot{x}(\tau): T_x \Sigma \to T_{x(\tau)}M \), i.e., \( \dot{x} \) is a section of \( x^*TM \) and in local coordinates it is a matrix \( \frac{\delta x^\mu}{\delta \tau^i} \).

Bellow we calculate the criticality equation (Euler-Lagrange equation) of this situation.

1.1.4. **Theorem.** The variation (differential) of the action \( S[x] = \int_{\Sigma} d\tau \ L(x(\tau), \dot{x}(\tau), \tau) \), in the direction of \( u \in T_x(C^\infty(\Sigma, M)) = C^\infty(\Sigma, M) \), is

\[
\frac{\delta S}{\delta x} u = \sum_j \left[ \prod_{k \neq j} \int_{I_k} d\tau^k \ L_{\frac{\partial}{\partial \tau^j}} u(\tau) \right]_{\tau^j = a_j}^{\tau^j = b_j} + \int_{\Sigma} d\tau \ [L_x(x, \dot{x}, \tau) - \frac{d}{d\tau} L_{\dot{x}}(x, \dot{x}, \tau)] \cdot u
\]

\[
= \int_{\partial \Sigma} d\sigma \ L_{\frac{\partial}{\partial n}} (x(\sigma), \dot{x}(\sigma)) \cdot u(\sigma) + \int_{\Sigma} d\tau \ [L_x - \frac{d}{d\tau} L_{\dot{x}}](x(\tau), \dot{x}(\tau), \tau) \cdot u(\tau)
\]

for the normal vector \( n \) to \( \partial \Sigma \).

(b) The EL-equations (criticality equations) are

\[
L_x = \frac{d}{d\tau}(L_{\frac{d}{d\tau}}) \text{ on } \Sigma, \text{ and } L_{\frac{\partial}{\partial n}} = 0 \text{ on } \partial \Sigma.
\]
(c) The pointwise variation at \( t \in \Sigma \) is the integral kernel for the variation operator \( \frac{\delta S}{\delta x} \)

\[
\frac{\delta S}{\delta x(t)} = [L_x - \frac{d}{d\tau} L_{\dot{x}}][x(t), \dot{x}(t), t] + \delta_{t\in\partial\Sigma} \cdot L_{\partial\Sigma}(x(t), \dot{x}(t), t), \quad t \in \Sigma.
\]

**Proof.** (a) The variation \( \frac{\delta S}{\delta x} \) of \( S \) at \( x \) and with respect to \( u \), is the differential

\[
\frac{\delta S}{\delta x} u = (d_x S) u \overset{\text{def}}{=} \frac{d}{d\varepsilon}|_{\varepsilon=0} S(x + \varepsilon u) = \int T d\tau \frac{d}{d\varepsilon}|_{\varepsilon=0} L(x + u, \dot{x} + \dot{u}, \tau).
\]

By differentiating under the integral and using the chain rule this becomes

\[
= \int \int d\tau L_x(x, \dot{x}, \tau) u + L_{\dot{x}}(x, \dot{x}, \tau) \dot{u}.
\]

Since \( \dot{x} \) is a vector \( \frac{\partial x}{\partial \tau} \), this is

\[
= \int \int d\tau^1 \cdots d\tau^p L_x \cdot u + L_{\dot{x}} \cdot \frac{\partial u}{\partial \tau^j}.
\]

It remains to use integration by parts for the “dynamic” ingredients \( \frac{\partial u}{\partial \tau^j} \) (i.e., the Stokes theorem)

\[
= \int_{\Sigma} d\tau \ [L_x - \frac{\partial}{\partial \tau^j} L_{\dot{x}}] \cdot u + \sum_j \left( \prod_{k \neq j} \int_{I_k} d\tau^k \ [L_{\dot{x}}(x, \dot{x}, \tau) u(\tau)]_{\tau^j=a_j}^{\tau^j=b_j} \right) \cdot \delta_v (\tau^j).
\]

(b) The equation on \( \Sigma \) is obtained by using all \( u \)'s supported in the interior of \( \Sigma \) and the continuity at the boundary. The equation on \( \partial\Sigma \) is then obtained by varying the restriction of \( u \) to the boundary. In (c),

\[
(d_x S) \delta_i(\tau) v = \int_{\partial\Sigma} d\sigma L_{\partial\Sigma}(x(\sigma), \dot{x}(\sigma), \sigma) \cdot \delta_i(\sigma) v + \int_{\Sigma} d\tau [L_x - \frac{d}{d\tau} L_{\dot{x}}](x(\tau), \dot{x}(\tau), \tau) \cdot \delta_i(\tau) v
\]

1.1.5. **Remarks.** (a) The variation with respect to \( x \) consists of two parts, due to the appearance of \( x \) as a non-dynamical variable (i.e., the appearance of the field \( x \) itself without derivatives), and as a dynamical variable (i.e., the appearance of the derivative \( \dot{x} \) of the field). The non-dynamical term of the variation has no time derivative, while the dynamical term has time derivative with a minus (since the dynamical variable causes integration by parts).

(b) In general, \( \Sigma \) will be a manifold with boundary and a Riemannian metric that will give the measure used to define the action \( S \). The boundary term in the variation formula will disappear in some cases:

- (i) If \( \Sigma \) is a closed manifold and \( L = L(x, \dot{x}) \), i.e., \( L(x, \dot{x}, \tau) \) does not depend on \( \tau \).
- (ii) If we consider maps \( x \) with the fixed values on \( \partial\Sigma \).
Here, (i) is clear, either by invoking the Stokes theorem or its proof (integral \( S[x] \) over \( \Sigma \) can be calculated as a sum of contributions from pieces \( \Sigma_k \) on which one has local coordinates. If we calculate the contribution to \( (d_x S)u \) from all pieces, the boundary contributions will cancel. The second case is also obvious since \( u = 0 \) on \( \partial \Sigma \).

(c) We can think of \( \dot{x} = \frac{dx}{d\tau} \) as a vector \( \frac{dx}{d\tau} \) or a matrix \( \frac{\partial x^\mu}{\partial \tau} \). Then the criticality equations can be written as

\[
L_x = \frac{\partial}{\partial \tau^j} L \frac{\partial x^j}{\partial \tau}, \quad L_x^\mu = \frac{\partial}{\partial \tau^j} L \frac{\partial x^\mu}{\partial \tau^j}, \quad \text{i.e.,,}
\]

the variation with respect to the value of \( x \) at \( \tau \) is

\[
\frac{\delta S}{\delta x(\tau)} \overset{\text{def}}{=} (d_x S) \delta \tau = [L_x - \frac{\partial}{\partial \tau^j} L \frac{\partial x^j}{\partial \tau^j}](x(\tau), \dot{x}(\tau), \tau),
\]

and if we vary only the \( \mu \) coordinate of \( x \) at \( \tau \)

\[
\frac{\delta S}{\delta x^\mu(\tau)} \overset{\text{def}}{=} (d_x S) \delta \tau = [L_{x^\mu} - \frac{\partial}{\partial \tau^j} L \frac{\partial x^\mu}{\partial \tau^j}](x(\tau), \dot{x}(\tau), \tau).
\]

1.1.6. The case \( S(x) = \int_{t_1}^{t_2} L(x(t)) \, dt \).


The difference between Lagrangians and EL-equations is that a Lagrangian is to be integrated, so it is rather a measure then a function. (?)

2. Perturbative expansion of Feynman integrals by Feynman graphs

This is one basic idea in quantum field theory.

**Appendix A. Geometry**

A.1. Anomaly. Let us say that a classical symmetry has an anomaly if it does not survive the quantization. From the point of view of path integrals, a symmetry of the action still need not preserve the measure in the path integral (we have been vague on this measure), hence need not preserve the path integral.

This may be OK – a classical symmetry may be replaced by its own quantization. For the Polyakov action the diffeomorphism symmetry quantizes to the Virasoro symmetry, and the Weyl symmetry (i.e., the conformal change of metric), has an anomaly which cancels for \( D = 26 \).
A.2. **Spinors.** These are the sections of any spinor bundle $S$ on $\Sigma$. Recall that we have $(\Sigma, g, or)$, hence a complex structure $C$ on $\Sigma$. A spinor bundle is any holomorphic line bundle which is a square root of the canonical bundle $K$. Now, spinors are the sections of a spinor bundle $S$ on $\Sigma$.

Spinor bundles form a torsor for $H^1(\Sigma, \pm) \cong \mathbb{Z}_2^{2g}$.

Dept. of Mathematics and Statistics, University of Massachusetts at Amherst, Amherst MA 01003-4515, USA

*E-mail address:* mirkovic@math.umass.edu