MATHEMATICAL METHODS IN QUANTUM MECHANICS

Lecture notes by Răzvan Gelca

Contents

This course was taught for the first time in the Department of Mathematics and Statistics at Texas Tech University in Spring 2015 and the second time in Fall 2017. Most of the materials were taken from the lecture notes of the Quantum Mechanics course that I took at University of Michigan with Alejandro Uribe, and from the books *Quantum Theory for Mathematicians* by Brian Hall, *Lectures on Quantum Mechanics for Mathematics Students* by L.D. Faddeev and O.A. Yakubovskii, and *Theta Functions and Knots* by myself.

Chapter 1

Classical Mechanics

This chapter is inspired by V.I. Arnol'd's *Mathematical Methods of Classical Mechanics*, and the students can consult this book for further details and examples. Because classical mechanics we can "see", it is phrased in the language of geometry. Originally geometry was an intuitive science of points, lines, and circles, but the coordinates introduced by Descartes were a powerful tool, and they led to differential geometry. It is in the language of differential geometry that we will tell the story of classical mechanics, and we do it from two points of view: the lagrangian and the hamiltonian formalism.

1.1 The Lagrangian formalism of classical mechanics

1. Lagrangian mechanics describes the motion of a system using the configuration space. A Lagrangian mechanical system is defined by a *finite dimensional manifold* (which parametrizes all possible configurations of the mechanical system) and a *function* on the tangent bundle to this manifold (the Lagrangian).

A smooth n-dimensional manifold M looks locally like $\mathbb{R}^n$. In formulas we will use a local coordinate $\mathbf{q} = (q_1, q_2, \ldots, q_n)$. By Whitney's embedding theorem, there is a smooth embedding $f : M \to \mathbb{R}^{2n}$. Thus we arrive at Poincarè's description of a manifold: a set in $\mathbb{R}^N$ that is locally the graph of a map $\phi : \mathbb{R}^r \to \mathbb{R}^s$.

Definition. The abstract definition of a smooth manifold (due to O. Veblen) is that it is that it is a topological space M together with an open covering $(U_\alpha)_{\alpha \in A}$ and a collection of bijections $\phi_\alpha : U_\alpha \to \mathbb{R}^n$ such that $\phi_\beta \circ \phi_\alpha^{-1}$ is smooth (whenever this map is defined).

The maps $\phi_\beta \circ \phi_\alpha^{-1}$ are the changes of coordinates from one local chart to another, if the two charts overlap.

The tangent space at a point $\mathbf{q}_0$ to the manifold M, $T_{\mathbf{q}_0}M$, can be defined abstractly as the set of equivalence classes of curves that have the same derivative at $\mathbf{q}_0$ in some system of local coordinates. Concretely, it is the vector space of tangent vectors at $f(\mathbf{q}_0)$ to $f(M)$ (here f is the embedding of M into some $\mathbb{R}^n$. A vector is of the form $\frac{d}{dt}(f \circ \mathbf{q})(t)|_{t=0}$, were $\mathbf{q}(t)$ is a curve, $t \in (-\epsilon, \epsilon)$, with $\mathbf{q}(0) = \mathbf{q}_0$. It is customary to denote the tangent vector to the curve $\mathbf{q}$ by $\dot{\mathbf{q}}$, or more precisely, the tangent vector to the curve $\mathbf{q}$ at point $\mathbf{q}(t)$ by $\dot{\mathbf{q}}(t)$.

The coordinates of a vector are computed in a local chart as $\frac{d}{dt}(\phi_\alpha \circ f \circ \mathbf{q})dt$. Then, if in the local chart U_α we use coordinates $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, we identify $T_{\mathbf{q}_0}M$ with

the vector space with basis $\partial/\partial x_1, \partial/\partial x_2, \ldots, \partial/\partial x_n$. Also, if in the local chart U_β we use coordinates $\mathbf{y} = (y_1, y_2, ..., y_n)$ and we identify $T_{\mathbf{q}_0}M$ with the vector space with basis $\partial/\partial y_1, \partial/\partial y_2, \ldots, \partial/\partial y_n$, then for $\mathbf{v} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ in the coordinate chart U_α then in the coordinate chart U_β, $\mathbf{v}$ has coordinates

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \cdots \\ \beta_n \end{pmatrix} = \left(\frac{\partial(\phi_\beta \circ \phi_\alpha^{-1})_j}{\partial x_k} \right)_{jk} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \cdots \\ \alpha_n \end{pmatrix}.$$

Formally we can think that the coordinates in the chart ϕ_α are $\mathbf{x}$ and in the chart ϕ_β are $\mathbf{y}$, and then the matrix for the change of coordinates of vectors is $\frac{\partial \mathbf{y}}{\partial \mathbf{x}}$.

The tangent bundle, TM, is the union of all tangent spaces, which is given a structure of a smooth manifold using the charts defined by ϕ_α and the coordinates of the tangent vector in the chart ϕ_α. The transition functions (from one coordinates system to another) are

$$\left(\phi_\beta \circ \phi_\alpha^{-1}, \left(\frac{\partial(\phi_\beta \circ \phi_\alpha^{-1})_j}{\partial x_k} \right)_{jk} \right)$$

Locally it is diffeomorphic to $\mathbb{R}^n \times \mathbb{R}^n$. Heuristically we think of it as being the "system of positions and velocities of a system of several particles with constraints".

2. Let $L : TM \times \mathbb{R} \to \mathbb{R}$ be a smooth function, which plays the role of the Lagrangian of the system; it is particular to the given mechanical system and it is used to describe the evolution of the mechanical system as follows. Fix points $\mathbf{q}_0$ and $\mathbf{q}_1$ on M and consider the infinite dimensional space of smooth curves γ given by functions of the form $\mathbf{q} : [t_0, t_1] \to M$ such that $\mathbf{q}(t_0) = \mathbf{q}_0$ and $\mathbf{q}(t_1) = \mathbf{q}_1$. On this space define the functional

$$\Phi(\gamma) = \int_{t_0}^{t_1} L(\mathbf{q}, \dot{\mathbf{q}}, t) dt.$$

Here as the variable of L you plug in the point $\mathbf{q}(t) \in M$, the tangent vector $\dot{\mathbf{q}}(t)$ to the curve at this point, and the time t.

Hamilton's Minimal Action Principle: The motions of a mechanical system described by L coincide with the extrema of the functional, $\Phi(\gamma)$, called *action*.

A curve is an extremum of Φ if the "derivative" of Φ zero at this curve. This is made precise within the framework of *variational calculus*.

Here we view M as a subset of some $\mathbb{R}^n$, and then consider a paths γ in M and variations of these paths, written as $\gamma + \delta$, also in M. The differential of Φ is a linear functional F such that

$$\Phi(\gamma + \delta) - \Phi(\gamma) = F(\delta) + R$$

where $R(\gamma, \delta) = O(\delta^2)$ (if δ and its derivative are less than ϵ in absolute value, then $|R| < \epsilon^2$). We consider only variations in the space of paths in M that connect $\mathbf{q}_0$ and $\mathbf{q}_1$. So δ itself,

which is a path in $\mathbb{R}^n$, is zero at t_0 and t_1. We have

$$\Phi(\gamma+\delta)-\Phi(\gamma)=\int_{t_0}^{t_1}[L(\mathbf{q}+\mathbf{h},\dot{\mathbf{q}}+\dot{\mathbf{h}},t)-L(\mathbf{q},\dot{\mathbf{q}},t)]dt$$
$$=\int_{t_0}^{t_1}\left[\frac{\partial L}{\partial \mathbf{q}}\mathbf{h}+\frac{\partial L}{\partial \dot{\mathbf{q}}}\dot{\mathbf{h}}\right]dt+O(\mathbf{h}^2)=\int_{t_0}^{t_1}\left[\frac{\partial L}{\partial \mathbf{q}}-\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}}\right)\right]\mathbf{h}dt+O(\mathbf{h}^2).$$

where for the last step we used integration by parts (and the fact that $\mathbf{h}$ is zero at the endpoints).

So the differential of Φ is this last integral, and the extrema of the action are those trajectories $t\mapsto\mathbf{q}(t)$ that satisfy the *Euler-Lagrange equations:*

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}}\right)-\frac{\partial L}{\partial \mathbf{q}}=0.$$

3. Here is a reformulation of Newtonian mechanics (in $\mathbb{R}^n$). Let $L=T-V$ where $T=\sum_j m_j\dot{q_j}^2/2$ is the kinetic energy and $V=V(\mathbf{q})$ is the potential energy. Then

$$\frac{\partial L}{\partial \dot{q}_j}=\frac{\partial T}{\partial \dot{q}_j}=m_j\dot{q}_j,\quad \frac{\partial L}{\partial q_j}=-\frac{\partial V}{\partial q_j}.$$

We obtain Newton's equations

$$m_j\ddot{q}_j+\frac{\partial V}{\partial q_j}=0,$$

which can be written in the more familiar form $ma_j=F_j$, were a_j is the jth component of the acceleration.

Remark 1.1.1. Newton's equations show that the *total energy* $E=T+V$ is conserved (i.e. does not change it time).

Example 1.1.1. A free particle moving in $\mathbb{R}$ has $L=T=\frac{m\dot{q}^2}{2}$. The Euler-Lagrange equation is

$$\frac{d}{dt}(m\dot{q})=0$$

which is equivalent to $\ddot{q}=0$. This means we are in the presence of uniform motion.

Example 1.1.2. The harmonic oscillator with no damping has a single force acting, and this is given by Hooke's law: $F=-kq$. We (forcefully) introduce the potential $V=\frac{kq^2}{2}$. Then

$$L=\frac{m\dot{q}^2}{2}-\frac{kq^2}{2}.$$

Then

$$\frac{\partial L}{\partial q}=-kq,\quad \frac{\partial L}{\partial \dot{q}}=m\dot{q},$$

so we obtain the equation of the harmonic oscillator

$$m\ddot{q}+kq=0,$$

whose solutions are of the form $q(t)=c_1\cos\omega t+c_2\sin\omega t$ with $\omega=\sqrt{k/m}$.

1.2 The Hamiltonian formalism of classical mechanics

4. To pass from the Lagrangian to the Hamiltonian formulation of classical mechanics we use the Legendre transform. This transform is still part of what is traditionally called Lagrangian mechanics, but it establishes the transition from one formulation to the other and turns Lagrangian mechanics into a subtheory of Hamiltonian mechanics.

First we assume, as it is usually the case in real life applications, that the potential energy depends on $\mathbf{q}$ only and the kinetic energy is a positive definite quadratic function in $\dot{\mathbf{q}}$. Then $L(\mathbf{q}, \dot{\mathbf{q}}, t)$ is convex in $\dot{\mathbf{q}}$.

In general for a convex function $f(x)$ the Legendre transform is a function g defined as follows. For a number p, the function $F(x, p) = px - f(x)$ has a maximum at some point $x(p)$. Then $g(p) = F(p, x(p))$.

So in our case $F(\dot{\mathbf{q}}, \mathbf{p}) = \mathbf{p}\dot{\mathbf{q}} - L(\mathbf{q}, \dot{\mathbf{q}}, t)$. To maximize we set its derivative with respect to $\dot{\mathbf{q}}$) equal to 0, and get $\mathbf{p} - \frac{\partial L}{\partial \dot{\mathbf{q}}} = 0$, that is $\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}}$. So the Legendre transform of L (with respect to the variable $\dot{\mathbf{q}}$) is the Hamiltonian function

$$H(\mathbf{p}, \mathbf{q}, t) = \mathbf{p}\dot{\mathbf{q}} - L(\mathbf{q}, \dot{\mathbf{q}}, t),$$

with variables $\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}}$ the (conjugate) *momenta*.

The Euler-Lagrange equations are equivalent to Hamilton's equations

$$\begin{aligned} \dot{\mathbf{q}} &= \frac{\partial H}{\partial \mathbf{p}} \\ \dot{\mathbf{p}} &= -\frac{\partial H}{\partial \mathbf{q}}. \end{aligned}$$

Indeed, $\dot{\mathbf{q}} = \partial H/\partial \mathbf{p}$ is just from definition of H, while $\dot{\mathbf{p}} = -\partial H/\partial \mathbf{q}$ are the Euler-Lagrange equations. Aha, so by introducing the momenta we turn a system of $n = \dim M$ second order differential equations into a system of $2n$ first order differential equations!

In the setting of Newtonian mechanics, the Hamiltonian is $H = T + V$, the total energy of the system.

Example 1.2.1. For the harmonic oscillator,

$$p = \frac{\partial L}{\partial \dot{q}} = \frac{\partial}{\partial \dot{q}} \left(\frac{m\dot{q}^2}{2} - \frac{kq^2}{2} \right) = m\dot{q}.$$

Hence

$$H = \frac{p^2}{2m} + \frac{kq^2}{2}.$$

Hamilton's equations are

$$\begin{aligned} \dot{q} &= \frac{p}{m} \\ \dot{p} &= -kq. \end{aligned}$$

So the second order Euler-Lagrange equation now becomes Hamilton's system of first order differential equations. To solve this system you set

$$A = \begin{pmatrix} 0 & \frac{1}{m} \\ -q & 0 \end{pmatrix}.$$

and then write the system as $\frac{d\mathbf{x}}{dt} = A\mathbf{x}$. The solution is of the form $\mathbf{x}(t) = e^{tA}\mathbf{x}(0)$, and of course to find the exponential of the matrix you need to diagonalize it, and in the process of diagonalization you get the same characteristic equation as for the second order equation, etc...

5. Now we discuss a subtle point, namely the fact that one of the main differences between the Lagrangian and the Hamiltonian point of view is that one of them happens on the *tangent* bundle while the other happens on the *cotangent* bundle of the configuration space.

In general for a function $f(\mathbf{x})$, the gradient

$$\frac{\partial f}{\partial \mathbf{x}} = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right)$$

is a covector. You can think of it as being the differential form $\frac{\partial f}{\partial \mathbf{x}} d\mathbf{x}$.

Because of the chain rule

$$\frac{\partial f}{\partial \mathbf{y}} = \frac{\partial f}{\partial \mathbf{x}} \frac{\partial \mathbf{x}}{\partial \mathbf{y}},$$

the matrix that changes the "vector" $\frac{\partial f}{\partial \mathbf{x}}$ to the "vector" $\frac{\partial f}{\partial \mathbf{y}}$ is

$$\frac{\partial \mathbf{x}}{\partial \mathbf{y}}^T = \left(\left(\frac{\partial \mathbf{y}}{\partial \mathbf{x}}\right)^{-1}\right)^T = \left(\left(\frac{\partial (\phi_\beta \circ \phi_\alpha^{-1})_j}{\partial x_k}\right)^{-1}_{jk}\right)^T.$$

We can take the union of the cotangent space at every point of the configuration space M to produce the cotangent bundle T^*M. This can be endowed with the structure of a smooth manifold using the charts and the coordinates of forms. Locally T^*M looks like some $\mathbb{R}^{2n}$, where n is the number of degrees of freedom (parameters) of the classical system. The transition maps from one system of coordinates to another are

$$\left(\phi_\beta \circ \phi_\alpha^{-1}, \left(\left(\frac{\partial (\phi_\beta \circ \phi_\alpha^{-1})_j}{\partial x_k}\right)^{-1}_{jk}\right)^T\right).$$

T^*M a manifold, of *even dimension.*

So to conclude, the Legendre transform maps functions that depend of the position and velocity, i.e. point and tangent vector, to functions that depend on the position and the momentum, i.e. point and cotangent vector.

You can think of this in terms of directional derivatives as follows. If V is a vector space and $f : V \to \mathbb{R}$ is a function, then for $v \in V$, $df_v \in V^*$. Indeed,

$$df_v(w) = \frac{d}{dt}(f(v + tw))|_{t=0}, \quad w \in V$$

which is a linear functional on V. We thus get a linear map $V \mapsto V*$. For $f = L(\dot{\mathbf{q}})$, and $V = T_{\mathbf{q_0}}$ we obtain a map from $T_{\mathbf{q}_0}M \mapsto T^*_{\mathbf{q}_0}M$, which defines the Legendre transform $TM \to T^*M$.

So Hamiltonian mechanics happens on the *cotangent bundle* of the manifold of configurations. If M is the manifold of configurations, we denote by T^*M its cotangent bundle. Locally T^*M looks like some $\mathbb{R}^{2n}$, where n is the number of degrees of freedom (parameters) of the classical system. Thus T^*M is itself a manifold, of *even dimension.* In fact, as we will see below, it is a manifold of special type.

We use the term *phase space* of the classical mechanical system to denote the cotangent bundle of the manifold of configurations. The phase space parametrizes the positions and momenta of the mechanical system.

Definition. A symplectic manifold is a manifold endowed with a globally defined symplectic form. A symplectic form is a nondegenerate closed 2-form.

We could think intuitively of a nondegenerate 2-form ω as an oriented area element (it tells us how to integrate functions on surfaces in M). A 2-form is an object that, when restricted to a point on the manifold, it yields a bilinear antisymmetric map $T_{\mathbf{q}}M \times T_{\mathbf{q}}M \to \mathbb{R}$, and these bilinear maps vary smoothly with the point. A 2-form is called nondegenerate if for every vector $v \neq 0$ there is a vector w such that $\omega(v, w) \neq 0$. The form is called closed if its differential is zero.

Theorem 1.2.1. The cotangent bundle T^*M is a symplectic manifold with the symplectic form

$$\omega = d\mathbf{p} \wedge d\mathbf{q} = \sum_{j=1}^{n} dp_j \wedge dq_j.$$

Proof. In local coordinates it is not hard to see that this form is nondegenerate and closed. It is harder to see that it is well defined. To prove this, we define a 1-form. Let v be a vector tangent to T^*M at a point $(\mathbf{q}, \mathbf{p})$. The differential of the projection $\pi : T^*M \to M$ maps v to $\pi_* v$. Let $\theta(v) = \mathbf{p}(\pi_* v)$. In local coordinates $\theta = \mathbf{p}d\mathbf{q}$, so $d\theta = \omega$. Note that $d\omega = d^2\theta = 0$. The form θ is called the canonical 1-form. □

6. Hamiltonian mechanics can be defined on a general symplectic manifold, not necessarily of the form T^*M, so we allow constraints at the level of momenta as well (not just at the level of positions).

Therefore in Hamiltonian mechanics the phase space is a pair (M, ω), where M is an even dimensional real manifold and ω is a symplectic form on M.

Theorem 1.2.2. (Darboux) Given a symplectic manifold (M, ω), every point has a neighborhood and a system of coordinates of that neighborhood in which the symplectic form is $d\mathbf{p} \wedge d\mathbf{q}$.

So for local computations we can assume that we are in $\mathbb{R}^{2n}$ with the standard symplectic form. For local computations we use the formulas

$$\omega\left(\frac{\partial}{\partial p_j}, \frac{\partial}{\partial q_k}\right) = \delta_{jk}, \omega\left(\frac{\partial}{\partial q_j}, \frac{\partial}{\partial p_k}\right) = -\delta_{jk}, \omega\left(\frac{\partial}{\partial p_j}, \frac{\partial}{\partial p_k}\right) = \omega\left(\frac{\partial}{\partial q_j}, \frac{\partial}{\partial q_k}\right) = 0,$$

The symplectic form defines an isomorphism

$$T_{\mathbf{q}_0}M \mapsto T^*_{\mathbf{q}_0}M$$

by $v \mapsto \omega(\cdot, v)$ where the latter maps a vector w to $\omega(w, v)$ (there is a fancy notation: $i_v\omega = \omega(v, \cdot)$).

Now we develop a formalism for **observable quantities**. This will bring us closer to the points of view of quantum mechanics. In classical mechanics we can see the particle, so we can determine its position and velocity precisely, or, in the present formalism, its position and momentum. Then we associate various quantities that are physically relevant and sometimes arise from conservation laws, such as angular momentum, total energy, kinetic energy, etc. But in quantum mechanics we cannot see particles. All we can do is set up some experiment which detects the existence of the particle and its "quantum" properties. Usually the outcome of the experiment is some data. We can reason the same way in the classical world, and think that what we measure are some numerical quantities that depend on the state of the particle. We thus are looking at functions of $(\mathbf{p}, \mathbf{q})$. We now develop the formalism for these functions.

The symplectic form associates to each function f a Hamiltonian vector field $\mathbf{X}_f$ defined as the inverse through this isomorphism of df, that is

$$df = \omega(\cdot, \mathbf{X}_f).$$

In local coordinates

$$\mathbf{X}_f = \left(\frac{\partial f}{\partial \mathbf{p}}\right)^T \frac{\partial}{\partial \mathbf{q}} - \left(\frac{\partial f}{\partial \mathbf{q}}\right)^T \frac{\partial}{\partial \mathbf{p}}.$$

There is a fancy way of writing this using the notation $i_{\mathbf{X}}\omega = \omega(\mathbf{X}, \cdot)$. Then $\mathbf{X}_f$ is defined by

$$df = -i_{X_f}\omega.$$

Example 1.2.2. For the coordinate functions q_j, p_j, we have

$$\mathbf{X}_{q_j} = -\frac{\partial}{\partial p_j}, \quad \mathbf{X}_{p_j} = \frac{\partial}{\partial q_j}.$$

7. There is a *Poisson bracket* for smooth functions on the symplectic manifold M which is defined by

$$\{f, g\} = -\omega(\mathbf{X}_f, \mathbf{X}_g).$$

In local coordinates

$$\begin{aligned}
\{f, g\} &= -\omega(\mathbf{X}_f, \mathbf{X}_g) = -\omega\Big(\sum_j \frac{\partial f}{\partial p_j}\frac{\partial}{\partial q_j} - \frac{\partial f}{\partial q_j}\frac{\partial}{\partial p_j}, \sum_k \frac{\partial g}{\partial p_k}\frac{\partial}{\partial q_k} - \frac{\partial g}{\partial q_k}\frac{\partial}{\partial p_k}\Big) \\
&= -\sum_{jk} \frac{\partial f}{\partial p_j}\frac{\partial g}{\partial p_k}\omega\Big(\frac{\partial}{\partial q_j}, \frac{\partial}{\partial q_k}\Big) + \sum_{jk} \frac{\partial f}{\partial p_j}\frac{\partial g}{\partial q_k}\omega\Big(\frac{\partial}{\partial q_j}, \frac{\partial}{\partial p_k}\Big) \\
&\quad + \sum_{jk} \frac{\partial f}{\partial q_j}\frac{\partial g}{\partial p_k}\omega\Big(\frac{\partial}{\partial p_j}, \frac{\partial}{\partial q_k}\Big) - \sum_{jk} \frac{\partial f}{\partial q_j}\frac{\partial g}{\partial q_k}\omega\Big(\frac{\partial}{\partial p_j}, \frac{\partial}{\partial p_k}\Big) \\
&= \left(\frac{\partial f}{\partial \mathbf{q}}\right)^T \frac{\partial g}{\partial \mathbf{p}} - \left(\frac{\partial f}{\partial \mathbf{p}}\right)^T \frac{\partial g}{\partial \mathbf{q}}.
\end{aligned}$$

Theorem 1.2.3. The Poisson bracket satisfies

1. $\{af + bg, h\} = a\{f, h\} + b\{g, h\}$ for a, b constants and f, g functions;
2. $\{f, g\} = -\{g, f\}$;
3. $\{fg, h\} = \{f, h\}g + f\{g, h\}$.
4. **(The Jacobi identity)** $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$.

The conditions can be checked by working in local coordinates. The four conditions define in general what is called a Poisson bracket structure on $C^\infty(T^*M)$. The first, second, and fourth conditions define a *Lie bracket* structure.

Definition. A manifold whose algebra of functions is endowed with a Poisson bracket is called a Poisson manifold.

Example 1.2.3. Not all Poisson manifolds are symplectic. For example $\mathbb{R}^3$ (which is not symplectic because it has odd dimension), has the a Poisson bracket defined by

$$\{f, g\} = \langle x, \nabla f_x \times \nabla g_x \rangle .$$

Recall that a Lie algebra is a vector space endowed with a *Lie bracket* (an antisymmetrical bilinear 2-form satisfying the Jacobi identity).

Example 1.2.4. The *Heisenberg Lie algebra* $\mathfrak{H}(\mathbb{R}^n)$ is the algebra generated by the coordinate and momentum functions $q_1, q_2, \ldots, q_n, p_1, p_2, \ldots, p_n$, satisfying

$$\{q_j, q_k\} = \{p_j, p_k\} = 0, \quad \{p_j, q_k\} = -\delta_{jk}.$$

These equations are the canonical commutation relations for the classical positions and momenta. This can be modeled using matrices as

$$q_j = \begin{pmatrix} 0 & e_j^T & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad p_k = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & e_k \\ 0 & 0 & 0 \end{pmatrix}, \quad 1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

We can also think of the Poisson bracket infinitesimally, as a Lie bracket of Hamiltonian vector fields:

$$[\mathbf{X}_f, \mathbf{X}_g] = \mathbf{X}_{\{f,g\}}.$$

Using the Poisson bracket we can write Hamilton's equations as

$$\dot{q_j} = \{q_j, H\}, \quad \dot{p_j} = \{p_j, H\}.$$

In general, if $f(\mathbf{p}, \mathbf{q})$ is a function of position and momentum, which we interpret as a observable (i.e. measurable) quantity then its value, as the particle moves along the trajectory $(\mathbf{q}, \mathbf{p})$, satisfies the system of first order partial differential equations

$$\frac{df}{dt} = \{f, H\}.$$

This is the general form of *Hamilton's equations*. This formula can be proved as follows

$$\frac{df}{dt} = \frac{df}{d\mathbf{p}}\dot{\mathbf{p}} + \frac{df}{d\mathbf{q}}\dot{\mathbf{q}} = -\frac{df}{d\mathbf{p}}\frac{\partial H}{\partial q} + \frac{df}{d\mathbf{q}}\frac{\partial H}{\partial p} = \{f, H\}.$$

where we have used Hamilton's equations.

8. Diffeomorphisms of the symplectic manifold that preserve the symplectic form are called *symplectomorphisms*. So $\phi : M \to M$ is a symplectomorphism if $\phi^*\omega = \omega$. (Here $phi^*\omega$ is the pull back of ω defined as $phi^*\omega_x(X, Y) = \omega_x(d\phi_x(X), d\phi_x(Y))$, where $d\phi_x(X)$ is the push forward of the vector X defined by $\phi_x(\gamma'(0)) = (\phi \circ \gamma)'(0)$.) Symplectomorphisms also preserve the Poisson bracket, so they preserve the Hamiltonian formalism. Symplectomorphisms are therefore the "isomorphisms" of classical mechanics.

Example 1.2.5. For a free 1-dimensional particle, the phase space is $T^*\mathbb{R} = \mathbb{R}^2$ with coordinates q the position and p the momentum. The symplectic form is $dp \wedge dq$. This is the same as the area element in the plane, taken with the sign that specifies orientation.

The linear transformations that preserve area and orientation form the group

$$SL(2, \mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid ad - bc = 1 \right\},$$

which are therefore the linear syplectomorphisms.

Indeed, if ϕ is the symplectomorphism defined by the matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, then the pullback of ω through ϕ is the symplectic form $\phi^*\omega$ defined by the condition

$$\begin{aligned} \phi^*\omega\left(\frac{\partial}{\partial p}, \frac{\partial}{\partial q}\right) &= \omega\left(a\frac{\partial}{\partial p} + b\frac{\partial}{\partial q}, c\frac{\partial}{\partial p} + d\frac{\partial}{\partial q}\right) = ad\omega\left(\frac{\partial}{\partial p}, \frac{\partial}{\partial q}\right) + bc\omega\left(\frac{\partial}{\partial q}, \frac{\partial}{\partial p}\right) \\ &= ad - bc = 1 = \omega\left(\frac{\partial}{\partial p}, \frac{\partial}{\partial q}\right). \end{aligned}$$

Example 1.2.6. For n free 1-dimensional particles, the phase space is $T^*\mathbb{R}^n = \mathbb{R}^{2n}$, with coordinates $\mathbf{q} = (q_1, q_2, \ldots, q_n)$ (positions) and $\mathbf{p} = (p_1, p_2, \ldots, p_n)$ (momenta). The symplectic form is $d\mathbf{p} \wedge d\mathbf{q} = \sum_j dp_j \wedge dq_j$.

The linear transformations that preserve the symplectic form are the elements of the symplectic group

$$Sp(2n, \mathbb{R}) = \left\{ h \in GL(2n, \mathbb{R}) \mid h^T \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} h = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} \right\}.$$

A particular example of symplectomorphisms, which play an important role in mechanics, are *Hamiltonian flows*. Let $H(\mathbf{p}, \mathbf{q})$ be a *time-independent Hamiltonian function* and let $\mathbf{X}_H$ be its associated vector field.

The Hamiltonian flow defined by H is a family of diffeomorphisms $\phi_t : M \to M$ indexed by some interval I containing 0 such that for every $x \in M$, $\phi_t(x) : I \to M$ is defined by the differential equation

$$\frac{d}{dt}\phi_t(x) = (\mathbf{X}_H)_{\phi_t(x)},$$

where the right term denotes the vector (of the vector field) at the point $\phi_t(x)$.

Note that the uniqueness of the solution to the differential equation implies $\phi_t(\phi_s(x)) = \phi_{t+s}(x)$. Thus the Hamiltonian flow is a one parameter group of diffeomorphisms.

Theorem 1.2.4. The Hamiltonian flow defines a one parameter group of symplectomorphisms which conserve H.

Proof. As a consequence of the definition, each observable varies under the flow by the Hamilton equation $df/dt = \{f, H\}$. So the conservation of the energy follows from

$$\begin{aligned}\frac{dH}{dt} &= \left(\frac{\partial H}{\partial \mathbf{q}}\right)^T \frac{d\mathbf{q}}{dt} + \left(\frac{\partial H}{\partial \mathbf{p}}\right)^T \frac{d\mathbf{p}}{dt} = \left(\frac{\partial H}{\partial \mathbf{q}}\right)^T \dot{\mathbf{q}} + \left(\frac{\partial H}{\partial \mathbf{p}}\right)^T \dot{\mathbf{p}} \\ &= \left(\frac{\partial H}{\partial \mathbf{q}}\right)^T \frac{\partial H}{\partial \mathbf{p}} - \left(\frac{\partial H}{\partial \mathbf{p}}\right)^T \frac{\partial H}{\partial \mathbf{q}} = 0.\end{aligned}$$

This is the same as

$$\frac{dH}{dt} = \{H, H\} = 0.$$

To show that Hamiltonian flows are symplectomorphisms, let us prove that

$$\frac{d}{dt}\phi^*\omega = 0.$$

We have

$$\frac{d}{dt}\phi^*\omega = \phi_t^* \lim_{h\to 0} h^{-1}[\phi_h^*\omega - \omega] = \phi_t^* \mathcal{L}_{\mathbf{X}_H}\omega.$$

So let us check that

$$\mathcal{L}_{\mathbf{X}_H}\omega = \lim_{h\to 0} h^{-1}[\phi_h^*\omega - \omega] = 0.$$

We use the Cartan formula

$$\mathcal{L}_{\mathbf{X}} = i_{\mathbf{X}} \circ d + d \circ i_{\mathbf{X}},$$

to get

$$\mathcal{L}_{\mathbf{X}_H}\omega = i_{\mathbf{X}_H} \circ d\omega + d\omega(\mathbf{X}_H, \cdot) = 0 - d^2 H = 0.$$

Cartan's formula can be proved as follows. Both sides satisfy the "product rule" so they are derivations. So it suffices to check on functions, where it is trivial, and on 1-forms:

$$[i_X d + d i_{\mathbf{X}}]df = d i_{\mathbf{X}}(df) = d[i_{\mathbf{X}}(df)] = d(\mathbf{X}(f)) = d\mathcal{L}_{\mathbf{X}}(f) = \mathcal{L}_{\mathbf{X}} df.$$

Here is a different way to check that Hamiltonian flows are symplectomorphisms. Consider a surface Σ. It suffices to show that

$$\int_\Sigma \omega = \int_{\phi_t(\Sigma)} \omega$$

Using Stokes' Theorem we can write

$$\int_{\phi_t(\Sigma)} \omega - \int_{\Sigma} \omega = \int_{F\Sigma} d\omega = 0,$$

where $F\Sigma$ is the "flow" of Σ between time 0 and time t. So the "area element" is preserved. □

In the 2-dimensional case this shows that the Hamiltonian flow conserves the area (Liouville's Theorem).

Example 1.2.7. The torus $\mathbb{T}^2 = S^1 \times S^1 = \mathbb{R}^2/\mathbb{Z}^2$ is a symplectic manifold with symplectic form $dp \wedge dq$, where (p, q) mod 1 are the coordinates of the point on the torus.

Let us consider the Hamiltonian function $H(p, q) = \sin 2\pi p$. Then

$$X_H = 2\pi \cos 2\pi p \frac{\partial}{\partial q}.$$

The Hamiltonian flow is

$$\phi_t(p, q) = 2\pi(tq \bmod 1) \cos 2\pi p.$$

On the other hand the family of translations

$$\phi_t(p, q) = (p + t, q)$$

which preserve the symplectic structure, are not a Hamiltonian flow. Indeed, we should have

$$\frac{\partial H}{\partial p} = 0 \text{ and } \frac{\partial H}{\partial q} = -1,$$

which would imply $H = -q + C$, $C = \text{const}$. But this function is not well defined on the torus!

The map $(p, q) \mapsto (-q, p)$ is also a symplectomorphism, but it is not part of a one-parameter group, so it does not arise from a Hamiltonian flow.

Chapter 2

"Matrix Mechanics" - A First Encounter with Quantum Mechanics

"I think I can safely say that nobody understands quantum mechanics." (Richard Feynman)

This chapter and the rest of the course are heavily inspired by Brian Hall's book *Quantum Theory for Mathematicians*, L.D. Faddeev and O.A. Yakubovskii Lectures on Quantum Mechanics for Mathematics Students, as well as on the lecture notes I took in the class of Alejandro Uribe at University of Michigan.

2.1 Problems that have led to the discovery of quantum mechanics

1. One should start with a quote by Lord Kelvin: "There is nothing new to be discovered in physics now, all that remains is more and more precise measurement". This was the universally accepted truth at the end of the 19th century. But then, as more and more precise measurements were made, problems started to appear.

There was the Michaelson-Morley experiment detecting the speed of light. It showed that the speed of light was always the same regardless of the orientation and the motion of the reference frame. To solve this inconsistency of physics, Einstein formulated a new set of postulates that fundamentally changed our understanding of physics at macroscopic level.

But then there were several experiments that raised questions about the microscopic world.

- The *black body radiation* is a first example where classical theory of physics produces nonsense. A black body is an object that emits radiation only when heated. An example is a stove element that turns red when heated. Experiments and classical physics predict that the black body emits electromagnetic radiation in all wavelengths, and the power of the radiation emitted by the black body is proportional to λ^{-4}, where λ is the wavelength. Thus as the frequency of the radiation increases, so does the power of the radiation, tending towards infinity (this is known as the utraviolet catastrophe).

  Experiments verify this law at low frequencies, but as the frequencies grow, the law fails badly. The problem was solved by Max Planck. In 1894, he has been comissioned

by the industry to increase the efficiency of light bulbs. This led him to the study of black body radiation. To predict correctly the radiation emitted by a black body, he had to postulate, in year 1900, what has since been known as Planck's law:

$$E = hn\nu,$$

namely that in a certain frequency, the energy emitted can only be an integer multiple of a certain quantity. Moreover, that quantity is proportional to the frequency. He called this assumption "an act of despair", but... it works.

- The *photoelectric effect* is a phenomenon in which a material emits electrons when light is shone upon it. A curious discovery was made at the end of 19th century, namely that the intensity of the electric current generated by the photoelectric effect does not depend on the intensity of light, but on the color of the light. Albert Einstein explained this in 1905 by making use of Planck's law: light can only be emitted in quanta, called photons, and only one photon with enough energy could make the material eject one electron, thus producing the observed effect.

- The distribution of *spectral lines of the hydrogen atom* is governed by a formula that was partially found by Johann Balmer in 1885, and then in its full extent by Johannes Rydberg in 1888. This formula gives the reciprocals of the wavelengths of the light emitted by the hydrogen atom as

$$\lambda^{-1} = R_H \left(\frac{1}{m^2} - \frac{1}{n^2} \right),$$

where R_H is a constant, and $m < n$ vary among the positive integers. This distribution of spectral lines was explained in 1913 by the model of the atom given by Niels Bohr.

2. Starting with the Rydberg formula, let us see how "matrix mechanics". I will present you the explanation given by Werner Heisenberg in his book *The Physical Principles of the Quantum Theory*, as it is always advised to learn the great ideas directly from the great masters who have discovered them.

First, let us place ourselves in the classical setting from the end of 19th century. The view was that the hydrogen atom has a nucleus surrounded by electrons, and the motion of the electrical charges which are the electrons produces, according to Maxwell's equations, electromagentic waves. These electromagnetic waves is what we see in the spectrum. Now, the classical theory predicts that the motion of an electron should produce electromagnetic waves in one fundamental frequency together with harmonics in frequencies that are integer multiples of this fundamental frequency. Basically you have some periodic phenomenon which creates periodic waves, and these waves can be expanded into a Fourier series. The first sinusoidal wave of the expansion gives the fundamental frequency, and the others are the harmonics.

But the Balmer series contradicts this. We do not see a Fourier series expansion. Let us examine this closely.

In classical mechanics we observe functions on the phase space of the system. But what do we observe in the case of quantum theory? Well, for the hydrogen atom we observe its spectral lines, or more precisely, the frequencies given by the formula

$$\nu(m,n) = \frac{R_H}{m^2} - \frac{R_H}{n^2} = T_m - T_n.$$

These satisfy the Rydberg-Ritz combination principle:

$$\nu(m,k) + \nu(k,n) = \nu(mn).$$

Planck's law tells us that energy is emitted in quanta, and that the energy of one quantum of light is proportional to the frequency: $E = h\nu$.

Thus what we observe are energies of quanta of light, and we have a sequence of numbers $W_1, W_2, \ldots$ such that

$$\nu(m,n) = \frac{1}{h}(W_n - W_m).$$

We arrange the observed values $\nu(mn)$ in an infinite matrix.

Now we observe several waves with certain frequencies and certain amplitudes. These waves can be represented by exponential functions, and the amplitudes by real numbers. Using the above intuition we arrange these waves corresponding to the spectral lines as a matrix these in an array:

$$\left(q(m,n)e^{2\pi i\nu(m,n)t}\right), \quad m,n.$$

This departs from the classical situation where a wave is decomposed into elementary oscillations by the Fourier transform, but then in the model the elementary oscillations are summed (with the inverse Fourier transform).

So, instead of the sum of elementary oscillations, here we have a matrix of elementary oscillations. Now, $\nu(m,m) = 0$, and moreover, we can allow $\nu(m,n) = -\nu(n,m)$ because of the differences, so that the entries of the matrix are defined for all pairs of integers.

...

The axioms of quantum mechanics are phrased based on these speculations. They define an abstract mathematical formalism, which tells us how to predict the results of particular experiments, and it turns out that these predictions match the results of the experiments. We should point out that the mathematical formalism of quantum mechanics is more amorphous than that of classical mechanics; once the general axioms are established there are many fixes and guesses that have to be applied to particular situations. Moreover, one should realize that the mathematics of quantum mechanics is not a rephrasing in rigorous terms of common language and intuition. While it comes from our intuition of the world, it introduces mathematical tools that we cannot identify with our senses.

2.2 The axioms of quantum mechanics

3. The states of a quantum system are in one-to-one correspondence with complex lines in a separable complex Hilbert space.

A Hilbert space is first a $\mathbb{C}$-vector space H endowed with a map $\langle , \rangle : H \times H \to \mathbb{C}$, called inner product, that is linear in the first variable, satisfies $\langle \phi, \phi \rangle \geq 0$ with equality if and only if $\psi = 0$, and $\langle \phi, \psi \rangle = \overline{\langle \psi, \phi \rangle}$. The inner product defines a norm $\|\phi\| = \sqrt{\langle \phi, \phi \rangle}$ and the norm induces the distance between ϕ and ψ as $\|\phi - \psi\|$. With this distance we can define a metric space structure on H, and for H to be a Hilbert space we require it to be complete as a metric space, namely that every Cauchy sequence has a convergent subsequence. A good example is the Hilbert space $L^2(\mathbb{R})$ with the inner product $\langle f, g \rangle = \int f(t)\overline{g(t)}dt$.

The quantum phase space is

$$\mathbb{P}H = \{[\psi] \,|\, \psi \in H \backslash \{0\}\} \text{ where } [\psi] = \{\lambda\psi \,|\, \lambda \in \mathbb{C}^*\}.$$

4. *The observables are self-adjoint operators on the Hilbert space.*

To define the concept of a self-adjoint operator, the key ingredient is the Riesz lemma. This lemma states that every continuous linear functional $\phi : H \to \mathbb{C}$ is of the form $\phi \mapsto \langle \phi, \psi \rangle$.

Now we look at densely defined linear operators, namely at linear operators A that are defined on some dense subspace $\mathcal{D}(A) \subset H$ (here $\mathcal{D}(A)$ is called the domain of A). Whenever the linear functional $\phi \mapsto \langle A\phi, \psi \rangle$ is continuous on the dense subset $\mathcal{D}(A)$ it can be extended, by continuity to the whole H, and then by the Riesz representation theorem it is of the form $\phi \mapsto \langle \phi, \xi \rangle$. Then we can define $A^*\psi = \xi$. This is the adjoint of A and its domain is the set of all ψ for which this construction works.

$A : H \to H$, is self-adjoint if

- A is symmetric, that is $\langle A\psi, \phi \rangle = \langle \psi, A\phi \rangle$,
- A is densely defined ($\overline{\mathcal{D}(A)} = H$).
- $\mathcal{D}(A^*) = \mathcal{D}(A)$.

Example 2.2.1. If A is continuous and symmetric than A is selfadjoint.

Example 2.2.2. If $H = L^2(\mathbb{R})$ and $A = i\frac{d}{dx}$ then A is self-adjoint without being continuous.

5. Recall that the spectrum of an operator is the complement of the set of complex numbers λ for which the operator $\lambda I - A$ is continuous and has a continuous inverse (where I is the identity operator). If A is continuous, it is known that the spectrum is compact (i.e. closed and bounded in $\mathbb{C}$) and it is also known that for any operator the spectrum is nonempty. Unlike the case of matrices, it might be that the space of a continuous operator is not discrete, or that the spectrum of an operator that is not continuous is unbounded.

We focus now on observables with discrete spectrum. *If A is an observable with discrete spectrum*, then A is diagonalizable. So

$$A\psi_j = \lambda_j \psi_j,$$

where ψ_j, $j \in J$, is an orthonormal basis.

In this situation,

- *An observation of A always results in some λ_j for some j.*

- *If $\psi = \sum_j c_j \psi_j$ and the system is in state $[\psi]$ and you observe A, the probability of obtaining the result λ_j is $|c_j^2|/||\psi||^2$.*

Consequently, if you measure A repeatedly with the system in state ψ, the average of the measurements is $\langle A\psi, \psi\rangle / \|\psi\|^2$ (the expected value of the observable):

$$Exp(A)_{[\psi]} = \frac{1}{\|\psi\|^2} \sum_j \lambda_j |c_j|^2.$$

6. *If A and B are quantum observables, then AB is the quantum observable: "observe B then observe A then multiply the resulting numbers".*

Then

$$\frac{1}{i\hbar}(AB - BA) = \frac{1}{i\hbar}[A, B]$$

is the self-adjoint measure of the lack of commutativity of observations.

7. Here is a new look at states. We determine that the system is in a certain state by performing experiments. In those experiments we measure certain observables. In classical mechanics, a given state yields always the same measurement for the observable. In quantum mechanics this is not the case. So each state ω of the system associates to a given quantum observable a probability distribution (and we have explained above what it is for observables that have discrete spectrum). So for an observable A the state ω defines a map $E \mapsto \omega_A(E)$ from the Borel sets in $\mathbb{R}$ to $[0, 1]$ such that

$$\omega_A(\sqcup E_j) = \sum_j \omega_A(E_j), \quad \omega_A(\emptyset) = 0, \omega_A(\mathbb{R}) = 1.$$

We should also have a relationship of the form: if observable B is equal to $f(A)$ for some function f, then $\omega_B(E) = \omega_A(f^{-1}(E))$.

Note that a convex combination of probability measures is a probability measure, so if $\omega = t\omega_{1,A} + (1-t)\omega_{2,A}$ then ω should also correspond to a state, and we let this state be $t\omega_1 + (1-t)\omega_2$. A state that can be decomposed this way is called a *mixed state*, a state that cannot be decomposed this way is called a *pure state*.

The expected value of an observable A is

$$\omega(A) = \int_{-\infty}^{\infty} \lambda d\omega_A(\lambda).$$

Physical intuition suggests we should have the following properties

$$\begin{aligned}
&\omega(cI) = c \\
&\omega(\alpha A + \beta B) = \alpha\omega(A) + \beta\omega(B) \\
&\omega(A^2) = \omega(A^*A) \geq 0 \\
&\omega(A) \in \mathbb{R}.
\end{aligned}$$

Now we would like to have enough observables to be able to separate states, so let us assume that all self-adjoint operators are quantum observables. Extend by linearity ω to

all operators, and in particular to the C^* algebra of bounded linear operators. Then we have a *positive linear functional* on the C^* algebra of bounded linear operators on H (whose operator norm is equal to 1). In fact every positive linear functional of norm 1 should be thought of as a (mixed) state. The *pure states* are precisely the linear functionals of the form

$$A \mapsto \langle A\psi, \psi \rangle ,$$

for some $\psi \in H$ of norm 1.

Well, we are only concerned with pure states, we call these simply states, and hence the first axiom of quantum mechanics. Note that $\psi \mapsto e^{i\theta}\psi$ does not change the state, this is why we work with $\mathbb{P}H$.

In general, for a not necessarily diagonalizable operator A, the expected value of A in the state ψ ($\|\psi\| = 1$) is $\langle A\psi, \psi \rangle$.

8. A consequence of de Broglie's formula $p = \frac{h}{\lambda}$ is the Heisenberg uncertainty principle. Let us formulate the *Heisenberg uncertainty principle* in dimension 1. If q is the position and p is the momentum then

$$\Delta q \Delta p \geq h,$$

where Δq and Δp are In other words, the standard deviations σ_q and σ_p satisfy

$$\sigma_q \sigma_p \geq \frac{\hbar}{2}.$$

Theorem 2.2.1. Two operators that satisfy the *canonical commutation relation*

$$PQ - QP = -i\hbar I$$

satisfy the Heisenberg uncertainty principle

$$\sigma_Q \sigma_P \geq \frac{\hbar}{2}.$$

Proof. The standard deviation of the observable A in state ψ (here ψ is a unit vector) is

$$\sigma_A^2 = \left\langle (A - A_{avg})^2 \psi, \psi \right\rangle = \left\langle A^2\psi, \psi \right\rangle - \langle A\psi, \psi \rangle^2 ,$$

where $A_{avg} = \langle A\psi, \psi \rangle$. Now we mimic the proof of the Cauchy-Schwarz inequality. We first assume $P_{avg} = Q_{avg} = 0$, by translating P and Q by some multiples of the identity operator. Then we start with

$$\langle (Q + itP)\psi, (Q + itP)\psi \rangle \geq 0, \quad \|\psi\| = 1, \alpha \in \mathbb{R},$$

and rewrite it as

$$\left\langle Q^2\psi, \psi \right\rangle + t^2 \left\langle P^2\psi, \psi \right\rangle + it \left\langle (QP - PQ)\psi, \psi \right\rangle \geq 0.$$

This is the same as

$$t^2 \left\langle P^2\psi, \psi \right\rangle - t\hbar \left\langle \psi, \psi \right\rangle + \left\langle Q^2\psi, \psi \right\rangle \geq 0.$$

This is a quadratic in t which is nonnegative, so its discriminant is nonpositive. We get

$$\left\langle P^2\psi, \psi \right\rangle \left\langle Q^2\psi, \psi \right\rangle \geq \frac{\hbar^2}{4}.$$

For the last step we used $\langle \psi, \psi \rangle = 1$. The conclusion follows by taking the square root. □

Remark 2.2.1. We have more generally, for two observables A and B measured in state ψ the inequality

$$\sigma_A \sigma_B \geq \frac{\hbar}{2} |\langle \{A, B\}_\hbar \psi, \psi \rangle|$$

where

$$\{A, B\}_\hbar = \frac{1}{i\hbar}[A, B],$$

is the *quantum Poisson bracket.*

9. We need to explain the name quantum Poisson bracket. If there is an association of operators P, Q to the classical observables (coordinate functions on the phase space) p and q, then the linear map from the Heisenberg Lie algebra to the space of linear operators on H, $L(H)$, defined by

$$p \mapsto P, \quad q \mapsto Q, \quad 1 \mapsto I$$

is a Lie algebra isomorphism onto the image, if we endow $L(H)$ with the Lie bracket $\{\cdot, \cdot\}_\hbar$.

2.3 In search of a nice quantization scheme

10. Quantum mechanics should be obtained by using the intuition of classical mechanics and classical mechanics should be the large scale behavior of quantum mechanics. Thus there should exists an algorithm for passing from classical mechanics to quantum mechanics (aiding our understanding of quantum mechanics) and an algorithm for passing from quantum mechanics to classical mechanics (which should allow us to recover the mechanics we already know).

The quantization "algorithm" should associate

- to the classical phase space a Hilbert space;
- to classical observables self-adjoint operators on the Hilbert space.

This association should satisfy Dirac's quantization conditions, as outlined by Paul Dirac in 1926 in *Proc. Royal Soc. London A*:

(1) $\mathrm{op}(1) = I$, where I is the identity operator;

(2) if $f = c_1 f_1 + c_2 f_2$ then $\mathrm{op}(f) = c_1 \mathrm{op}(f_1) + c_2 \mathrm{op}(f_2)$;

(3) (the correspondence principle) $\mathrm{op}(\{f, g\}) = \frac{1}{i\hbar}[\mathrm{op}(f), \mathrm{op}(g)]$ (!);

(4) the representation of quantum observables on the Hilbert space is irreducible.

There is one model for which this "algorithm" is specified as an axiom, the Heisenberg quantization conditions:

- $C^\infty(\mathbb{R}^n \times \mathbb{R}^n)$ becomes $L^2(\mathbb{R}^n)$;
- the coordinate function q_j becomes the operator $Q_j = M_{q_j}$ of multiplication by q_j;
- the momentum function p_j becomes the operator $P_j = -i\hbar\frac{\partial}{\partial q_j}$.

The choice of these operators is quite intuitive, the position operator is the multiplication by the position function, and the momentum is the instantaneous translation (the exponential of the differentiation operator is translation). Unfortunately both operators are unbounded. This means that they are discontinuous maps on $L^2(\mathbb{R})$ and in particular they are not defined everywhere. One should point out that the "eigenvalues" of M_{q_j} are Dirac's delta functions and they are not in the Hilbert space. Can we do any better?

Clearly finite dimensional operators don't work, because if $[Q_j, P_j] = i\hbar I$ then

$$0 = \text{trace}([Q_j, P_j]) = \text{trace}(i\hbar I) = i\hbar \dim H.$$

Before the proof, let us recall that a continuous linear operator $A : \mathcal{H} \to \mathcal{H}$ is also called a bounded operator, and for a good reason: The $\epsilon - \delta$ definition of continuity on metric spaces implies that a linear operator $A : \mathcal{H} \to \mathcal{H}$ is continuous if and only if there is $C > 0$ such that $\|Ax\| \leq C\|x\|$ for all x. The infimum of all C with this property is called the norm of A and is denoted by $\|A\|$. The norm satisfies $\|A + B\| \leq \|A\| + \|B\|$ and $\|AB\| \leq \|A\|\|B\|$. Bounded operators are nice, for one thing because they are defined everywhere not just on a dense subset, but also because there is a rich theory of bounded linear operators. But we have the following result.

Theorem 2.3.1. (Wintner-Wielandt) The relation $PQ - QP = -i\hbar I$ cannot be satisfied by bounded linear operators P and Q.

Proof. Assume that this is possible and set $A = Q$, $B = i/\hbar P$. Then $AB - BA = I$. We prove by induction that

$$A^n B - BA^n = nA^{n-1}.$$

This is clearly true for $n = 1$. If we assume that this is true for n, then for $n + 1$ we have

$$A^{n+1}B - BA^{n+1} = A^n(AB - BA) + (A^n B - BA^n)A = A^n + nA^{n-1}A = (n + 1)A^n.$$

Then

$$n\|A^{n-1}\| = \|A^n B - BA^n\| \leq 2\|A^n\|\|B\| \leq 2\|A^{n-1}\|\|A\|\|B\|.$$

So $n \leq 2\|A\|\|B\|$ for all positive integers n. This is clearly impossible. □

So this definition of the position and momentum seems to be the best choice we have. In fact there is a "theorem" due to M. Stone saying that this is the "only" choice. We will make this theorem precise in the guise of the Stone-von Neumann theorem.

11. We therefore have the Lie algebra isomorphism onto the image

$$\mathfrak{H}(\mathbb{R}^n) \to L(L^2(\mathbb{R}^n)), \quad f \mapsto \text{op}(f).$$

Can this be extended to a Lie algebra isomorphism onto the image from the Lie algebra $C^\infty(\mathbb{R}^n \times \mathbb{R}^n)$ or in general from the Lie algebra of smooth functions on the phase space of a classical system to the algebra of linear operators on the Hilbert space of the associated quantum system?

The question is how to extend the definition to functions of p_j, q_j.

Example 2.3.1. The operator associated to the kinetic energy of one particle is

$$\mathrm{op}\left(\sum_{j=1}^{3}\frac{1}{2m}p_j^2\right) = \sum_{j=1}^{3}\frac{1}{2m}\mathrm{op}(p_j^2) = \text{guess} = \sum_{j=1}^{3}\frac{1}{2m}\mathrm{op}(p_j)^2 = -\frac{\hbar^2}{2m}\Delta,$$

where Δ is the Laplacian, $\Delta = \sum_{j=1}^{3}\frac{\partial^2}{\partial x_j^2}$. This is the *quantum kinetic energy.*

Using intuition, we can define the quantum Hamiltonian of a Newtonian system as

$$\mathrm{op}(H) = \mathrm{op}(T) + \mathrm{op}(V) = -\frac{\hbar^2}{2m}\Delta + \mathrm{op}(V) = -\frac{\hbar^2}{2m}\Delta + M_{V(x)}.$$

But we run into a big problem when trying to define $\mathrm{op}(p_jq_k)$. The problem is that the naive definition $\mathrm{op}(q_jp_k) = \mathrm{op}(q_j)\mathrm{op}(p_k)$ fails because the latter product is not a self-adjoint operator. So we can try

$$\mathrm{op}(q_jp_k) = \frac{1}{2}\left(\mathrm{op}(q_j)\mathrm{op}(p_k) + \mathrm{op}(p_k)\mathrm{op}(q_j)\right).$$

But then how should we quantize $q_j^m p_k^n$. Should we average over all possible ways to write the product? And is Dirac's quantization condition (3) still satisfied? As the following result shows, this works for quadratic polynomials.

Proposition 2.3.1. The quantization scheme from the Lie algebra of quadratic polynomials on $\mathbb{R}^{2n}$ in the variables $q_1, q_2, \ldots, q_n, p_1, p_2, \ldots, p_n$, defined by

$$\mathrm{op}(1) = I, \quad \mathrm{op}(q_j) = M_{q_j}, \quad \mathrm{op}(p_j) = -i\hbar\tfrac{\partial}{\partial q_j},$$
$$\mathrm{op}(q_jq_k) = M_{q_jq_k}, \quad \mathrm{op}(p_jp_k) = -\hbar^2\tfrac{\partial^2}{\partial q_j\partial q_k}, \quad \mathrm{op}(q_jp_k) = -i\hbar\left(M_{q_j}\tfrac{\partial}{\partial q_k} + \tfrac{1}{2}\delta_{jk}\right)$$

acting on $L^2(\mathbb{R}^{2n})$, which are then extended linearly to the Lie algebra of quadratic polynomials, satisfies Dirac's quantization conditions (1), (2), (3).

Proof. Conditions (1) and (2) are straightforward by the linearity requirement. Condition (3) is straightforward for Poisson brackets of functions depending only on q_j's or only on p_j's. For the others the computations are tedious but involve just differentiation and multiplication. Here is an example:

$$\{q_jp_k, p_\ell\} = \delta_{j\ell}p_k.$$

and

$$\begin{aligned}\frac{1}{i\hbar}[\mathrm{op}(q_jp_k), \mathrm{op}(p_\ell)] &= -\left[-i\hbar\left(M_{q_j}\frac{\partial}{\partial q_k} + \frac{\delta_{jk}}{2}\right)\frac{\partial}{\partial q_\ell} + \frac{\partial}{\partial q_\ell}i\hbar\left(M_{q_j}\frac{\partial}{\partial q_k} + \frac{\delta_{jk}}{2}\right)\right]\\ &= -i\hbar\delta_{j\ell}\frac{\partial}{\partial q_k} = \mathrm{op}(p_k).\end{aligned}$$

□

12. But we should not get too excited!

Theorem 2.3.2. (The Weak No-Go Groenewold Theorem) The above quantization scheme cannot be extended to polynomials of degree less than or equal to 4.

Proof. (following Mark J. Gotay) For all quadratic polynomials $R_1(x)$ and all cubic polynomials $R_2(x)$ we must have

$$\begin{aligned}
&\mathrm{op}(R_2(q_j)) = R_2(\mathrm{op}(q_j)), \quad \mathrm{op}(R_2(p_j)) = R_2(\mathrm{op}(p_j)) \\
&\mathrm{op}(R_1(q_j)p_k) = \frac{1}{2}[R_1(\mathrm{op}(q_j))\mathrm{op}(p_k) + \mathrm{op}(p_k)R_1(\mathrm{op}(q_j))] \\
&\mathrm{op}(q_j R_1(p_k)) = \frac{1}{2}[\mathrm{op}(q_j)R_1(\mathrm{op}(p_k)) + R_1(\mathrm{op}(p_k))\mathrm{op}(q_j)].
\end{aligned}$$

Again all situations can be checked in a similar fashion and we only check the case $R(q_j) = q_j^3$. Write $\mathrm{op}(q_j^3) = \mathrm{op}(q_j)^3 + T$. Then $\{q_j^3, q_k\} = 0$ and $\{q_j^3, p_k\} = -3q_j^2\delta_{jk}$ implies that T commutes with all $\mathrm{op}(q_k)$, $\mathrm{op}(p_k)$, $k = 1, 2, \ldots, n$. But then T also commutes with $\mathrm{op}(q_j)\mathrm{op}(p_k) + \mathrm{op}(p_k)\mathrm{op}(q_j)$. Then

$$\begin{aligned}
-\mathrm{op}(q_j^3) &= \frac{1}{3}\mathrm{op}(\{p_j q_j, q_j^3\}) = \frac{1}{3i\hbar}[\mathrm{op}(p_j q_j), \mathrm{op}(q_j^3)] \\
&= \frac{1}{3i\hbar}\left[\frac{1}{2}(\mathrm{op}(q_j)\mathrm{op}(p_j) + \mathrm{op}(p_j)\mathrm{op}(q_j), \mathrm{op}(q_j)^3 + T\right] \\
&= \frac{1}{6i\hbar}[\mathrm{op}(q_j)\mathrm{op}(p_j) + \mathrm{op}(p_j)\mathrm{op}(q_j), \mathrm{op}(q_j)^3] = -\mathrm{op}(q_j)^3.
\end{aligned}$$

With this fact at hand, consider the equality

$$\frac{1}{9}\{q_j^3, p_j^3\} = \frac{1}{3}\{q_j^2 p_j, p_j^2 q_j\} \quad \left(= q_j^2 p_j^2\right).$$

Quantizing the equality we obtain

$$\frac{1}{9}[\mathrm{op}(q_j)^3, \mathrm{op}(p_j)^3] = \frac{1}{3}[\mathrm{op}(q_j^2 p_j), \mathrm{op}(p_j^2 q_j)].$$

The left-hand side is

$$-i\hbar^3\left(M_{q_j}^2\frac{\partial}{\partial q_j^2} + 2M_{q_j}\frac{\partial}{\partial q_j} + \frac{2}{3}I\right)$$

while the right-hand side is

$$\frac{1}{3}\left(\mathrm{op}(q_j^2 p_j)\mathrm{op}(p_j^2 q_j) - \mathrm{op}(p_j^2 q_j)\mathrm{op}(q_j^2 p_j)\right).$$

This is a bit messy, so we compute separately

$$\begin{aligned}
\mathrm{op}(q_j^2 p_j) &= \frac{1}{2}\left(\mathrm{op}(q_j)^2\mathrm{op}(p_j) + \mathrm{op}(p_j)\mathrm{op}(q_j)^2\right) = -\frac{i\hbar}{2}\left(M_{q_j^2}\frac{\partial}{\partial q_j} + \frac{\partial}{\partial q_j}M_{q_j^2}\right) \\
&= -i\hbar\left(M_{q_j}^2\frac{\partial}{\partial q_j} + M_{q_j}\right),
\end{aligned}$$

and

$$\mathrm{op}(p_j^2 q_j) = \frac{1}{2}\left(\mathrm{op}(p_j)^2\mathrm{op}(q_j) + \mathrm{op}(q_j)\mathrm{op}(p_j)^2\right) = -\hbar^2\left(\frac{\partial}{\partial q_j} + M_{q_j}\frac{\partial^2}{\partial q_j^2}\right).$$

So the right-hand side is

$$-i\hbar^3\left(M_{q_j}^2\frac{\partial}{\partial q_j^2} + 2M_{q_j}\frac{\partial}{\partial q_j} + \frac{1}{3}I\right).$$

And they don't coincide. □

We would like to use this symmetric quantization scheme. To this end, let us note that the error is of the order of $\hbar^2$ (this after dividing by $\frac{1}{i\hbar}$). And $\hbar$ is already very small, and $\hbar^2$ is even smaller. We are led to allow an error in Dirac's third condition. It is nowadays standard to let the correspondence principle be

- (3) $\mathrm{op}(\{f,g\}) = \frac{1}{i\hbar}[\mathrm{op}(f), \mathrm{op}(g)] + O(\hbar)$.

This modification will allow the construction of a quantization scheme, as we will see below, but it will also permit the coexistence of an entire menagerie of non-equivalent quantization schemes.

2.4 Weyl quantization

13. This quantization method was introduced by Hermann Weyl in 1931. Let us consider first the 1-dimensional case.

Definition. For monomials in the position q and momentum p the Weyl quantization is defined by

$$\mathrm{op}(q^m p^n) = \frac{1}{(m+n)!}\sum_\sigma \sigma(\mathrm{op}(q), \ldots, \mathrm{op}(q), \mathrm{op}(p), \ldots, \mathrm{op}(p))$$

where the sum is taken over all possible permutations of $m+n$ objects and σ of m copies $\mathrm{op}(q)$ and n copies of $\mathrm{op}(p)$ means the product of those $\mathrm{op}(q)$ and $\mathrm{op}(p)$ multiplied in the order specified by σ. Weyl quantization is then extended linearly to all polynomials.

Proposition 2.4.1. Weyl quantization satisfies

$$\mathrm{op}(aq+bp)^m) = (a\,\mathrm{op}(q) + b\,\mathrm{op}(p))^m.$$

Proof. This is just an easy consequence of the definition. Note for example that

$$\begin{aligned}(a\,\mathrm{op}(q) + b\,\mathrm{op}(p))^3 &= a^3\mathrm{op}(q)^3 + a^2b\,\mathrm{op}(q)^2\mathrm{op}(p) + a^2b\,\mathrm{op}(q)\mathrm{op}(p)\mathrm{op}(q) + a^2b\,\mathrm{op}(p)\mathrm{op}(q)^2\\ &\quad + ab^2\mathrm{op}(q)\mathrm{op}(p)^2 + ab^2\mathrm{op}(p)\mathrm{op}(q)\mathrm{op}(p) + ab^2\mathrm{op}(p)^2\mathrm{op}(2) + b^3\mathrm{op}(p)^3\\ &= \mathrm{op}(a^3q^3) + 3\mathrm{op}(a^2bq^2p) + 3\mathrm{op}(ab^2qp^2) + \mathrm{op}(p^3) = \mathrm{op}((aq+bp)^3).\end{aligned}$$

□

This means that we can define the quantization of a power series in $aq+bp$ as well through a limiting process, if the partial sums of operators converges. In particular we can define

$$\operatorname{op}\left(e^{2\pi i(aq+bp)}\right) = \sum_{m=0}^{\infty} \frac{1}{m!}[2\pi i(a\operatorname{op}(q) + b\operatorname{op}(p)]^m = e^{2\pi i(a\operatorname{op}(q)+b\operatorname{op}(p))}.$$

This operator is defined by

$$\operatorname{op}\left(e^{2\pi i(aq+bp)}\right)\psi = \lim_{N\to\infty} \sum_{m=0}^{N} \frac{1}{m!}[2\pi i(a\operatorname{op}(q) + b\operatorname{op}(p))]^m\psi$$

for all smooth functions $\psi \in L^2(\mathbb{R})$, whenever the series converges in $L^2(\mathbb{R})$.

But now a miracle happens! The operator $A = a\,op(q)+b\operatorname{op}(p)$ is *self-adjoint*, so $U = e^{iA}$ can be extended to a *unitary* operator on the entire Hilbert space. Recall that a unitary operator is one that is continuous, one-to-one, onto, and preserves the Hilbert space norm (in fact it preserves the inner product). Said differently, U is unitary if it is bounded and $UU^* = U^*U = I$. That U preserves the inner product follows from

$$\left\langle e^{iA}\psi_1, e^{iA}\psi_2\right\rangle = \left\langle \psi_1, e^{-iA}e^{iA}\psi_2\right\rangle = \left\langle \psi_1, \psi_2\right\rangle.$$

Here we used the fact that $(iA)^* = -iA$ and then applied it term by term to the series expansion of the exponential. Now extend U to the entire Hilbert space. Its inverse is the continuous extension e^{-iA}, so we have $U : H \to H$ unitary.

We conclude that although $a\operatorname{op}(q)+b\operatorname{op}(p)$ is not a bounded operator and is only defined on a dense subset of the Hilbert space, $\operatorname{op}\left(e^{2\pi i(aq+bp)}\right)$ is a unitary operator that is defined everywhere.

14. Now we can define Weyl quantization in $\mathbb{R}^{2n}$. We first do it for exponential functions:

$$\operatorname{op}\left(e^{2\pi i(\mathbf{a}^T\mathbf{q}+\mathbf{p}^T\mathbf{p})}\right) = e^{2\pi i(\mathbf{a}^T\operatorname{op}(\mathbf{q})+\mathbf{b}^T\operatorname{op}(\mathbf{p}))},$$

where $\operatorname{op}(\mathbf{q}) = (\operatorname{op}(q_1), \operatorname{op}(q_2), \ldots, \operatorname{op}(q_n))$ and $\operatorname{op}(\mathbf{p}) = (\operatorname{op}(p_1), \operatorname{op}(p_2), \ldots, \operatorname{op}(p_n))$.

Then for $f \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{C})$ expand f into elementary oscillations using the Fourier transform and replace each oscillation (exponential function) by its Weyl quantization. This means that we set

$$\hat{f}(\mathbf{a}, \mathbf{b}) = \int_{\mathbb{R}^{2n}} f(\mathbf{q}, \mathbf{p})e^{-2\pi i(\mathbf{a}^T\mathbf{q}+\mathbf{b}^T\mathbf{p})}d\mathbf{q}d\mathbf{p},$$

so that

$$f(\mathbf{q}, \mathbf{p}) = \int_{\mathbb{R}^{2n}} \hat{f}(\mathbf{a}, \mathbf{b})e^{2\pi i(\mathbf{a}^T\mathbf{q}+\mathbf{b}^T\mathbf{p})}d\mathbf{a}d\mathbf{b},$$

and then set

$$\operatorname{op}(f) = \int_{\mathbb{R}^{2n}} \hat{f}(\mathbf{a}, \mathbf{b})e^{2\pi i(\mathbf{a}^T\operatorname{op}(\mathbf{q})+\mathbf{b}^T\operatorname{op}(\mathbf{p}))}d\mathbf{a}d\mathbf{b}.$$

The fact that the exponential operators are unitary makes the convergence issues parallel those in the real-valued case.

15. As we explained before we have the following result.

Proposition 2.4.2. If R_1 and R_2 are polynomials of total degree at most 2 in $q_1, q_2, \ldots, q_n$, $p_1, p_2, \ldots, p_n$, then

$$\mathrm{op}(\{R_1, R_2\}) = \frac{1}{i\hbar}[\mathrm{op}(R_1), \mathrm{op}(R_2)].$$

We also know that this condition fails for certain polynomials of degree 3, by the weak form of the Groenewold No-Go Theorem.

Theorem 2.4.1. (Groenewold's No-Go Theorem) There exists no linear map from the Lie algebra of polynomials of degree at most 4 to the Lie algebra of differential operators (with bracket $\frac{1}{i\hbar}[\cdot,\cdot]$ that extends the quantization scheme for the Heisenberg Lie algebra.

Proof. Here is a sketch of the proof. Let

$$D = \sum_k f_k(\mathbf{q}) \left(\frac{\partial}{\partial \mathbf{q}}\right)^{\mathbf{k}}.$$

The first observation is that if D commutes with all operators of the form M_{q_j} and $\frac{\partial}{\partial q_j}$, then it is a multiple of the identity operator.

One can show that every polynomial of degree at most 2 is a linear combination of Poisson brackets of polynomials of degree at most 2, and every polynomial of degree at most 3 is a linear combination of Poisson brackets of polynomial of degree at most 3. Then like in the first step in the proof of Theorem 2.2.2, we deduce that the quantization scheme coincides with Weyl quantization for polynomials of total degree less than or equal to 3.

And then we redo the last step on the proof of Theorem 2.2.2. □

2.5 A menagerie of quantizations

First let us fix the notation: $\mathbf{a}^{\mathbf{j}} = a_1^{j_1} a_2^{j_2} \cdots a_n^{j_n}$. Here is a list of important quantization schemes:

1. Pseudodifferential operator quantization:
$$\mathrm{op}(\mathbf{q}^{\mathbf{j}}\mathbf{p}^{\mathbf{k}}) = \mathrm{op}(\mathbf{q})^{\mathbf{j}}\mathrm{op}(\mathbf{p})^{\mathbf{k}}.$$
2. Symmetrized pseudodifferential operator quantization:
$$\mathrm{op}(\mathbf{q}^{\mathbf{j}}\mathbf{p}^{\mathbf{k}}) = \frac{1}{2}(\mathrm{op}(\mathbf{q})^{\mathbf{j}}\mathrm{op}(\mathbf{p})^{\mathbf{k}} + \mathrm{op}(\mathbf{p})^{\mathbf{k}}\mathrm{op}(\mathbf{q})^{\mathbf{j}}).$$
3. Weyl quantization:
$$\mathrm{op}((\mathbf{a}^T\mathbf{q} + \mathbf{b}^T\mathbf{p})^m) = (\mathbf{a}^T\mathrm{op}(\mathbf{q}) + \mathbf{b}^T\mathrm{op}(\mathbf{p}))^m.$$
4. Wick quantization:
$$\mathrm{op}((\mathbf{q} + i\mathbf{p})^{\mathbf{j}}(\mathrm{op}(\mathbf{q} - i\mathbf{p})^{\mathbf{k}}) = (\mathrm{op}(\mathbf{q}) - i\mathrm{op}(\mathbf{p}))^k(\mathrm{op}(\mathbf{q}) + i\mathrm{op}(\mathbf{p}))^{\mathbf{j}}.$$
5. Anti-Wick quantization, also known as Toeplitz quantization:
$$\mathrm{op}((\mathbf{q} + i\mathbf{p})^{\mathbf{j}}(\mathrm{op}(\mathbf{q} - i\mathbf{p})^{\mathbf{k}}) = (\mathrm{op}(\mathbf{q}) + i\mathrm{op}(\mathbf{p}))^{\mathbf{j}}(\mathrm{op}(\mathbf{q}) - i\mathrm{op}(\mathbf{p}))^{\mathbf{k}}.$$

Chapter 3

The Heisenberg Group

3.1 The Heisenberg group

16. It is now time to have our first encounter with Lie groups (named after Sophus Lie) and their representations. Lie groups lie at the intersection of algebra and geometry. They are both

· groups: they have a multiplication that is associative, has an identity element, and every element has an inverse

· manifolds: locally they look like $\mathbb{R}^n$ (well, the manifold can also be infinite dimensional but in this class we do not care about that case)

Moreover, the algebraic and geometric aspects are related by the fact that multiplication: $(x, y) \mapsto xy$ and the operation of taking the inverse: $x \mapsto x^{-1}$ are continuous.

Since Lie groups are manifolds, they have tangent spaces at each point, and in particular at the origin. The multiplication of the Lie group induces an operation on the tangent space at the origin, called Lie bracket, which is the infinitesimal form of the multiplication of the Lie group. The Lie bracket $(X, Y) \mapsto [X, Y]$ satisfies the same properties that the Poisson bracket satisfies (see Theorem 1.2.3).

Returning to Lie groups, Cayley has noticed that every group is a group of transformations. But we do not like transformations (i.e. bijections of sets) at large, we prefer transformations of some geometric space, such as $\mathbb{R}^n$. So then the bigger question arises whether groups can be described as groups of linear transformations of a vector space. Well, sometimes they can, sometimes they cannot. We make a compromise: study the homomorphisms from our group to a group of linear transformations. This is called representation theory; we say that our group is represented on a vector space, the homomorphism is called a representation.

While Lie groups arose in classical physics, quantum theory has a surprisingly deep connection to Lie groups and their representation theory.

17. We can also start with a Lie algebra of linear transformations of a vector space. Its elements are linear transformations and the Lie bracket is the commutator $[A, B] = AB - BA$. Then the exponentials

$$e^A = I + \frac{1}{1!}A + \frac{1}{2!}A^2 + \cdots + \frac{1}{n!}A^n + \cdots$$

form a group because by the Baker-Campbell-Hausdorff formula

$$e^A e^B = e^{A+B+c(A,B)}$$

where $c(A, B)$ is an expression obtained from A and B by applying only addition and commutators. So the exponent on the right also lies in the Lie algebra, and thus we are in the presence of a group. The Lie algebra is a vector space, so it has a manifold structure, and the exponential is a local diffeomorphism, inducing a manifold structure on the group of exponentials. Thus we obtain a Lie group.

If we start with the Heisenberg Lie algebra generated by the position and momentum operators, then the associated Lie group is the Heisenberg Lie group. More precisely, we start with the Lie algebra spanned as a real vector space by

$$Q_j = 2\pi i M_{q_j}, \quad P_j = 2\pi\hbar \frac{\partial}{\partial q_j}, \quad I,$$

acting on $L^2(\mathbb{R}^n)$. We exponentiate these as dictated by Weyl quantization. What we obtain is a Lie group. This is the Heisenberg group.

18. So let us study the Heisenberg group, which is determined by the Weyl quantizations of the exponential functions.

Let us denote

$$\exp(\mathbf{a}^T\mathbf{Q} + \mathbf{b}^T\mathbf{P} + tI) = e^{2\pi i t}\mathrm{op}\left(e^{2\pi i(\mathbf{a}^T\mathbf{q}+\mathbf{b}^T\mathbf{p})}\right).$$

Theorem 3.1.1. We have the following action on $L^2(\mathbb{R}^n)$:

$$\exp(\mathbf{a}^T\mathbf{Q} + \mathbf{b}^T\mathbf{P} + tI)\psi(\mathbf{q}) = e^{2\pi i\mathbf{a}^T\mathbf{q}+\pi i h\mathbf{a}^T\mathbf{b}+2\pi i t}\psi(\mathbf{q} + h\mathbf{b})$$

One should stress out that in the statement of the theorem we have Planck's constant not the reduced Planck's constant.

Proof. The exponential of a differentiation operator is a translation and the exponential of the multiplication by a function is the multiplication by the exponential of that function. The second of these statements is easier to understand. Here is an explanation of the first. The expression

$$e^{t\frac{d}{dx}}\psi(x) = \sum_{n=0}^{\infty} \frac{t^n}{n!}\frac{d^n\psi}{dx^n}(x)$$

is just the MacLaurin expansion of ψ about x, so for all analytical functions it equals $f(x + t)$. But analytical functions are dense in $L^2(\mathbb{R})$ (because for examples functions of the form $e^{-x^2}P(x)$ where P is a polynomial are). So the exponential of differentiation equals translation for a dense set of functions. But both operators are continuous so they coincide everywhere.

So we have

$$\exp(b_j P_j)\psi(q_1,\dots,q_n) = \exp(2\pi i(-i\hbar b_j \frac{\partial}{\partial q_j}))\psi(q_1,\dots,q_n) = \psi(q_1,\dots,q_j + hb_j,\dots,q_n)$$
$$\exp(a_j Q_j)\psi(q_1,\dots,q_n) = e^{2\pi i a_j q_j}\psi(q_1,\dots,q_n)$$
$$\exp(2\pi i t I)\psi(q_1,\dots,q_n) = e^{2\pi i t}\psi(q_1,\dots,q_n).$$

Whenever two operators commute, their exponentials commute, so it is of no difficulty to prove the result for $\mathbf{a} = 0$ or for $\mathbf{b} = 0$. The general case follows from the Baker-Campbell-Hausdorff formula for operators whose second commutators are zero ($[A,[A,B]] = 0$):

$$e^A e^B = e^{A+B+\frac{1}{2}[A,B]}. \quad \square$$

Corollary 3.1.1. We have the following multiplication rule for quantized exponentials

$$\begin{aligned}&\exp(\mathbf{a}^T\mathbf{Q} + \mathbf{b}^T\mathbf{P} + tI)\exp(\mathbf{a}'^T\mathbf{Q} + \mathbf{b}'^T\mathbf{P} + t'I)\\ &\quad = \exp\left((\mathbf{a}+\mathbf{a}')^T\mathbf{Q} + (\mathbf{b}+\mathbf{b}')^T\mathbf{P} + t + t' - \frac{h}{2}(\mathbf{a}^T\mathbf{b}' - \mathbf{b}^T\mathbf{a}')\right).\end{aligned}$$

19. This prompts us to make the following definition:

Definition. The *Heisenberg group* with real entries $\mathbf{H}(\mathbb{R}^n)$ is $\mathbb{R}^n \times \mathbb{R}^n \times U(1)$ with multiplication

$$(\mathbf{a},\mathbf{b},e^{2\pi i t})\cdot(\mathbf{a}',\mathbf{b}',e^{2\pi i t'}) = \left(\mathbf{a},\mathbf{a}',\mathbf{b}+\mathbf{b}',e^{2\pi i(t+t'-\frac{h}{2}(\mathbf{a}^T\mathbf{b}'-\mathbf{b}^T\mathbf{a}'))}\right).$$

For further use, we let

$$\omega((\mathbf{a},\mathbf{b}),(\mathbf{a}',\mathbf{b}')) = -h(\mathbf{a}^T\mathbf{b}' - \mathbf{b}^T\mathbf{a}')),$$

where h is Planck's constant.

The Heisenberg group is precisely the group of Weyl quantizations of exponential functions in the plane. One should emphasize that the Heisenberg group is a Lie group whose Lie algebra is the Heisenberg Lie algebra. Theorem 3.1.1 provides a faithful infinite dimensional representation

$$\rho_0 : \mathbf{H}(\mathbb{R}^n) \to U(L^2(\mathbb{R}^n))$$

of the Heisenberg group as a group of unitary operators on $L^2(\mathbb{R}^n)$. Here faithful means that it is injective. This is sometimes called the Schrödinger representation.

We have the following exponential form for the canonical commutation relations

$$\begin{aligned}&\exp P_j \exp Q_k = e^{2\pi i h\delta_{jk}}\exp Q_k \exp P_j\\ &\exp P_j \exp P_k = \exp P_k \exp P_j\\ &\exp Q_j \exp Q_k = \exp Q_k \exp Q_j.\end{aligned}$$

3.2 The Stone-von Neumann theorem

Theorem 3.2.1. (a) The representation ρ_0 of the Heisenberg group from Theorem 3.1.1 is irreducible.
(b) Any unitary irreducible representation ρ of $\mathbf{H}(\mathbb{R}^n)$ such that $\rho(\exp(tI)) = e^{2\pi it}I$ is unitary equivalent to ρ_0.

The proof (which we borrowed from Gerard Lion's book) is quite long and we divide it in steps. Let therefore ρ be a unitary irreducible representation of $\mathbf{H}(\mathbb{R}^n)$ onto some Hilbert space H.

20. First, for two compactly supported functions f, g on

$$\mathbf{H}(\mathbb{R}^n) = \mathbb{R}^{2n} \times U(1) = B \times U(1),$$

we define their convolution

$$f * g(u) = \int_{\mathbf{H}(\mathbb{R}^n)} f(v)g(v^{-1}u)dv,$$

where dv is the translation-invariant measure on $\mathbb{R}^n \times U(1)$.

21. Now every element of $\mathbf{H}(\mathbb{R}^n)$ is of the form $b\exp(tI)$, with $b \in B = \mathbb{R}^{2n}$. Identify B with $\mathbf{H}(\mathbb{R}^{2n})/U(1)$. Using this splitting we can turn ρ into a "representation" of the *additive group* $B = \mathbb{R}^{2n}$.

Restrict the convolution to functions that satisfy the equivariance condition

$$f(ue^{2\pi it}) = e^{-2\pi it}f(u).$$

We can identify these with functions on B, but we should always keep in mind the equivariance condition. The convolution defined above becomes a convolution for functions on B:

$$(f * g)(b) = \int_B f(b')g(b - b')e^{\pi i\omega(b',b)}db.$$

Define also the actions of the Heisenberg group on functions

$$(u * f)(b) = f(u^{-1}b), \quad (f * u)(b) = f(bu^{-1}).$$

For a function f with compact support in $B = \mathbb{R}^{2n}$, which can be thought of as a compactly supported equivariant function on the Heisenberg group, we define the bounded operator $W_\rho(f) : H \to H$ by

$$\langle W_\rho(f)x, y\rangle = \int_B f(b) \langle \rho(b)x, y\rangle \, db.$$

This is called the Weyl transform of f. We can write it as

$$W_\rho(f) = \int_B f(b)\rho(b)db.$$

It has the following properties, which are not hard to check:

(i) $W_\rho(f * g) = W_\rho(f)W_\rho(g)$,

(ii) $W_\rho(f^*) = W_\rho(f)^*$, where $f^*(v) = \overline{f(-v)}$,

(iii) $W_\rho(u * f) = \rho(u)W_\rho(f)$,

(iv) $W_\rho(f * u) = W_\rho(f)\rho(u)$.

22. A *Hilbert-Schmidt operator* on H is an operator A satisfying

$$\|A\|_{HS}^2 = \sum_j \|Ae_j\|^2 < \infty$$

for some orthonormal basis $(e_j)_j$. This sum does not depend on the orthonormal basis, and defines the Hilbert-Schmidt norm $\|A\|_{HS}$. The Hilbert-Schmidt operators form a Hilbert space themselves (wow!), with the inner product

$$\langle A, B\rangle_{HS} = \sum_j \langle Ae_j, Be_j\rangle,$$

and with orthonormal basis the rank one operators

$$E_{j,k}(x) = \langle x, e_j\rangle e_k.$$

This means that every Hilbert-Schmidt operator is of the form

$$\sum_{jk} a_{jk} e_k \otimes e_j^* = \sum_j a_{jk} < \cdot, e_k > e_j,$$

with

$$\|A\|_{HS} = \left(\sum_{jk} |a_{jk}|^2\right)^{1/2}.$$

We denote by $\mathcal{HS}(H)$ the Hilbert space of Hilbert-Schmidt operators on H. Note for example that $E_{j,j}$ are the projections onto the axes of coordinates. We should point out that every finite rank operator is Hilbert-Schmidt, and that finite rank operators form a dense set in the Hilbert space of Hilbert-Schmidt operators. So all Hilbert-Schmidt operators are compact, meaning that they map bounded sets to sets whose closure is compact (but not vice-versa).

We can define in general $E_{x,y} = \langle \cdot, y\rangle x$. Then

$$\langle E_{x,y}, E_{x',y'}\rangle_{HS} = \langle x, x'\rangle \overline{\langle y, y'\rangle}.$$

The map

$$(x, y) \mapsto E_{x,y}$$

is an isometry from $H \otimes H^*$ to $\mathcal{HS}(H)$.

23. For proving (a) we start with the observation that $W_{\rho_0}(f)$ is a "sum" of operators of the form $\rho(u)$, $u \in \mathbf{H}(\mathbb{R}^n)$, so if ρ_0 has in invariant subspace, then so do all $W_{\rho_0}(f)$. Thus it suffices to show that there is no subspace that is invariant under all operators $W_{\rho_0}(f)$, or under all operators in the closure of this set of operators (see Proposition 3.2.2 below). For that we need the following result:

Proposition 3.2.1. W_{ρ_0} extends to an isomorphism $L^2(\mathbb{R}^{2n}) \to \mathcal{HS}(H)$, where $H = L^2(\mathbb{R}^n)$.

Proof. Step 1. We show that the Weyl transform of a function is an integral operator.

Set $\ell = \mathbb{R}^n \oplus 0$, $\ell' = 0 \oplus \mathbb{R}^n$. We identify $L^2(\mathbb{R}^n)$ with $L^2(\ell')$. Then for $\psi \in L^2(\mathbb{R}^n)$, we have

$$(W_{\rho_0}(f)\psi)(\mathbf{q}) = \iint f(\mathbf{y}+\mathbf{x})e^{\pi i\mathbf{x}^T(2\mathbf{q}+h\mathbf{y})}\psi(\mathbf{q}+h\mathbf{y})d\mathbf{x}d\mathbf{y},$$

where we used Theorem 3.1.1. Set $\xi = \mathbf{q} + h\mathbf{y}$ to turn this into

$$\iint e^{i\pi\mathbf{x}^T(\mathbf{q}+\xi)} f\left(\frac{\mathbf{q}-\xi}{h}+\mathbf{x}\right) e^{2\pi i\mathbf{x}^T\xi}\psi(\xi)d\mathbf{x}d\xi = \iint e^{i\pi\mathbf{x}^T(\mathbf{q}+\xi)} f\left(\frac{\mathbf{q}-\xi}{h}+\mathbf{x}\right) d\mathbf{x}\psi(\xi)d\xi.$$

Define

$$K_f(\mathbf{q},\xi) = \int e^{i\pi\mathbf{x}^T(\xi+\mathbf{q})} f\left(\frac{\mathbf{q}-\xi}{h}+\mathbf{x}\right) d\mathbf{x}.$$

This is what is called a *kernel*, which defines an *integral operator*

$$\psi(\mathbf{q}) \mapsto \int K(\mathbf{q},\mathbf{y})\psi(\mathbf{y})d\mathbf{y}.$$

So we recognize the Weyl transform of f to be an integral operator.

Step 2. We show that the Hilbert space of Hilbert-Schmidt operators on $L^2(\mathbb{R}^n)$ is the same as the space of integral operators with L^2 kernels, and moreover, the L^2 norm of the kernel is the Hilbert-Schmidt norm of the operator.

Indeed, if A is a Hilbert-Schmidt operator on $L^2(\mathbb{R}^n)$ then

$$\begin{aligned}(A\psi)(\mathbf{x}) &= \sum_{jk} a_{jk} < \psi, e_k > e_j = \sum_{jk} a_{jk}\left(\int_{\mathbb{R}^n} \psi(\mathbf{y})\overline{e_k(\mathbf{y})}d\mathbf{y}\right) e_j(x) \\ &= \int_{\mathbb{R}^n} \psi(\mathbf{y})\left(\sum_{jk} a_{jk}e_j(\mathbf{x})\overline{e_k(\mathbf{y})}\right) d\mathbf{y}.\end{aligned}$$

If we denote

$$K_A(\mathbf{x},\mathbf{y}) = \int_{\mathbb{R}^n} \sum_{jk} a_{jk}e_j(\mathbf{x})\overline{e_k(\mathbf{y})}d\mathbf{y},$$

then it is not hard to see that

$$\|K_A(\mathbf{x},\mathbf{y})\|_2 = \|A\|_{HS}.$$

Therefore we wrote the operator A as an integral operator

$$(A\psi)(\mathbf{x}) = \int_{\mathbb{R}^n} \psi(\mathbf{y}) K_A(\mathbf{x}, \mathbf{y}) d\mathbf{y}.$$

Conversely, if $K(\mathbf{x}, \mathbf{y})$ is any L^2 kernel then we can define

$$(A\psi)(\mathbf{y}) = \int \psi(\mathbf{y}) K(\mathbf{x}, \mathbf{y}) d\mathbf{y},$$

and standard inequalities about integrals show that $A\psi$ is L^2, and that for an orthonormal basis $\sum \|Ae_n\|^2$ is finite. Moreover $\|A\|_{HS} = \|K\|_2$.

Step 3. We show that the L^2 norm of a function is the same as $2^{n/2}$ times the L^2 norm of the kernel K_f of its Weyl transform.

The partial Fourier transform

$$(\mathcal{F}_{\mathbf{x}} f)(\mathbf{y}, \mathbf{q}) = \int e^{-\pi i \mathbf{x}^T \mathbf{q}} f\left(\frac{\mathbf{y}}{h} + \mathbf{x}\right) d\mathbf{x}$$

is, up to multiplication by a constant, a unitary isomorphism of $L^2(\mathbb{R}^{2n})$. And

$$K_f(\mathbf{q}, \mathbf{y}) = (\mathcal{F}_x f)\left(\frac{\mathbf{q} - \mathbf{y}}{2}, -\frac{\mathbf{y} + \mathbf{q}}{2}\right).$$

So a multiple of W_ρ is isometry which therefore can be extended to the entire $L^2(\mathbb{R}^n)$. And because the Fourier transform is an isomorphism, this extension is onto. Thus the proposition is proved. □

24. Now (a) is a direct corollary of the following result.

Proposition 3.2.2. If T is a bounded operator that commutes with all operators of the form $\rho(u)$, $u \in \mathbf{H}(\mathbb{R}^n)$, then T is a multiple of the identity operator.

Proof. If T commutes with every operator of the form $\rho(u)$, the T also commutes with operators of the form $W_\rho(f)$ with f compactly supported. But then by continuity it also commutes with all operators of the form $W_\rho(f)$, $f \in L^2(\mathbb{R}^n)$. In particular it commutes with the operators of the form $E_{jk} = e_j \otimes e_k^*$.

If T is not a multiple of the identity operator, then there are noncollinear vectors ψ_1, ψ_2 such that $T\psi_1 = \psi_2$. Using Gram-Schmidt we can construct a basis such that $< Te_1, e_2 >= \lambda \neq 0$. But then

$$0 = Te_1 \otimes e_2^*(e_1) = e_1 \otimes e_2^* T(e_1) = \lambda \neq 0.$$

This is a contradiction and the proposition is proved. □

Now if the Schrödinger representation were reducible, then we could write $L^2(\mathbb{R}^n) = H_1 \oplus H_2$, with H_1 and H_2 invariant subspaces (here it is important that the representation is unitary). But then the orthogonal projection onto H_1 would commute with $\rho(u)$ for all u, and this would contradict the above proposition.

25. We now start the proof of (b). Let ρ be a representation of $\mathbf{H}(\mathbb{R}^n)$ on a Hilbert space H such that $\rho(\exp tI) = e^{2\pi it}I$. Then the Weyl transform defines a homomorphism W_ρ from the algebra of rapidly decreasing functions $\mathcal{S}(\mathbb{R}^{2n})$ ($f \in \mathcal{S}(\mathbb{R}^{2n})$ if $\sup_{x\in R}^n |x^\alpha D^\beta f(x)| < \infty$ for all $\alpha, \beta \in \mathbb{Z}_+^n$) endowed with the convolution $*$ into a subalgebra of operators $\mathcal{A}$ on H such that $W_\rho(f^*) = (W_\rho(f))^*$. Moreover, $\mathcal{S}(\mathbb{R}^n)$ is isomorphic to the algebra $\mathcal{A}_0$ of integral operators with rapidly decreasing kernel via the Weyl transform W_{ρ_0} defined by the Schrödinger representation. Then $\phi = W_\rho \circ W_{\rho_0}^{-1}$ is an isomorphism of operator algebras satsfying $\phi(A^*) = (\phi(A))^*$, for all A, and $\phi(I) = I$.

Let $f_1 \in \mathcal{S}(\ell') = \mathcal{S}(0 \oplus \mathbb{R}^n)$ such that $\|f_1\|_2 = 1$. Consider the orthogonal projection operator $P_1 : L^2(\mathbb{R}^n) \to \mathbb{C}f_1$. Then P_1, being finite rank, is Hilbert-Schmidt, so it is defined by a kernel. In fact this kernel is $f_1(x)\overline{f_1(y)}$, so P_1 is of the form $W_{\rho_0}(g_1)$, for $g_1 \in \mathcal{S}(\mathbb{R}^{2n})$. Using the properties of projectors we have

$$g_1 * g_1 = g_1, \quad g_1^* = g_1, \quad g_1 * u * g_1 = \langle \rho_0(u)f_1, f_1\rangle\, g_1.$$

These relations and the multiplicative properties of the Weyl transform imply that $W_\rho(g_1)$ is a projector onto the space $H_1 = W_\rho(g_1)H$.

Lemma 3.2.1. The Hilbert space H is spanned by elements of the form $\rho(u)W_\rho(g_1)x$ for $u \in \mathbf{H}(\mathbb{R}^n)$ and $x \in H$.

Proof. Since ρ is a unitary representation, it suffices to show that the Hilbert space H is spanned by elements of the form $\rho(u)W_\rho(g_1)\rho(u^{-1})x$. Let $y \in H$ be orthogonal to all those elements. We have

$$\begin{aligned}
&\left\langle y, \rho(u)W_\rho(g_1)\rho(u^{-1})x\right\rangle = 0\\
&= \int_{\mathbb{R}^{2n}} \left\langle y, \rho(\exp(\mathbf{a}^T\mathbf{Q} + \mathbf{b}^T\mathbf{P})\rho(\exp(\mathbf{a}'^T\mathbf{Q} + \mathbf{b}'^T\mathbf{P}))\rho(\exp(-\mathbf{a}^T\mathbf{Q} - \mathbf{b}^T\mathbf{P}))x\right\rangle g_1(\mathbf{a}', \mathbf{b}')d\mathbf{a}'d\mathbf{b}'\\
&= \int_{\mathbb{R}^{2n}} \left\langle y, \rho(\exp(\mathbf{a}'^T\mathbf{Q} + \mathbf{b}'^T\mathbf{P})x\right\rangle g_1(\mathbf{a}', \mathbf{b}')e^{2\pi i\omega((\mathbf{a},\mathbf{b}),(\mathbf{a}',\mathbf{b}'))}d\mathbf{a}'d\mathbf{b}'.
\end{aligned}$$

This is a Fourier transform of $F(\mathbf{a}', \mathbf{b}') = \left\langle \left\langle y, \rho(\exp(\exp(\mathbf{a}'^T\mathbf{Q} + \mathbf{b}'^T\mathbf{P})\right\rangle g_1(\mathbf{a}', \mathbf{b}')\right\rangle$, which is identically equal to zero if and only if F is identically equal to zero. Since g_1 is not identically equal to zero, it follows that there is an element of the Heisenberg group, $u = \exp(\mathbf{a}'^T\mathbf{Q} + \mathbf{b}'^T\mathbf{P})$, such that $\langle y, \rho(u)x\rangle = 0$ for all $x \in H$. But $\rho(u)$ is unitary, so $\langle y, x\rangle = 0$ for all $x \in H$. Hence $y = 0$ and the lemma is proved. □

26. Before we proceed with the last step, we recall the notion of the tensor product of two vector spaces, and then of two Hilbert spaces.

For vector spaces V and W, construct the set $V \times W$. Now formally build a vector space whose basis consists of the elements of $V \times W$. Then factor this space by the relations
$(v_1 + v_2, w) = (v_1, w) + (v_2, w)$,
$(v, w_1 + w_2) = (v, w_1) + v(w_2)$,
$c(v, w) = (cv, w) = (v, cw)$,
for all $v, v_1, v_2 \in V$, $w, w_1, w_2 \in W$, $c \in C$. The resulting vector space is the tensor product $V \otimes W$. For example $\mathbb{C}^3 \otimes \mathbb{C}^2 = \mathbb{C}^6$.

If H_1, H_2 are Hilbert spaces, endow $H_1 \otimes H_2$ with the tensor product

$$\langle v_1 \otimes w_1, v_2 \otimes w_2 \rangle = \langle v_1, v_2 \rangle \langle w_1, w_2 \rangle .$$

The completion of this space is the Hilbert space $H_1 \otimes H_2$. For example $L^2(\mathbb{R}^n) \otimes \mathbb{C}^2 = L^2(\mathbb{R}^n) \oplus L^2(\mathbb{R}^n)$ and $L^2(\mathbb{R}^n) \otimes L^2(\mathbb{R}^n) = L^2(\mathbb{R}^{2n})$.

27. Let us return to the space $H_1 = W_\rho(g_1)H$. We will prove that there exists a unitary isomorphism U between $H_0 \otimes H_1$ and H, where $H_0 = L^2(\mathbb{R}^n)$ is the Hilbert space onto which the Heisenberg group is represented via ρ_0 and H is the Hilbert space onto which the Heisenberg group is represented via ρ. We will define U by

$$U(\rho_0(u)f_1 \otimes w) = \rho(u)w, \quad u \in \mathbf{H}(\mathbb{R}^n), w \in H_1.$$

then extend this linearly. Because the representation ρ_0 is irreducible, the set of linear combinations $\sum c_j \rho_0(u_j) f_1$ is dense in $L^2(\mathbb{R}^n)$. So if we check that U is an isometry, then we can extend it to the whole $H_0 \times H_1$.

We check that for $w_1 = W_\rho(g_1)x_1$, $w_2 = W_\rho(g_1)x_2$, $u_1, u_2 \in \mathbf{H}(\mathbb{R}^n)$,

$$\langle \rho(u_1)w_1, \rho(u_2)w_2 \rangle_H = \langle \rho_0(u_1)f_1, \rho_0(u_2)f_1 \rangle_{H_0} \langle w_1, w_2 \rangle_{H_1} .$$

Indeed

$$\begin{aligned}
&\langle \rho(u_1)W_\rho(g_1)x_1, \rho(u_2)W_\rho(g_1)x_2 \rangle_H = \left\langle W_\rho(g_1)\rho(u_2)^{-1}\rho(u_1)W_\rho(g_1)x_1, x_2 \right\rangle_H \\
&\quad = \left\langle W_\rho(g_1 * (u_2^{-1}u_1) * g_1)x_1, x_2 \right\rangle_H = \langle \rho(u_2)^*\rho(u_1)f_1, f_1 \rangle_{H_0} \langle W_\rho(f_1)x_1, x_2 \rangle_H \\
&\quad = \left\langle \rho_0(u_2^{-1}u_1)f_1, f_1 \right\rangle_{H_0} \langle W_\rho(f_1)W_\rho(f_1)x_1, x_2 \rangle_H \\
&= \langle \rho_0(u_1)f_1, \rho_0(u_2)f_1 \rangle_{H_0} \langle W_\rho(f_1)x_1, W_\rho(f_1)x_2 \rangle_{H_1} ,
\end{aligned}$$

as claimed.

Now using Lemma 3.2.1, we deduce that U is surjective, so it is a unitary isomorphism between $H_0 \otimes H_1$ and H. The way U was defined shows that $U(\rho_0(u) \otimes I_{H_1})U^{-1} = \rho(u)$. So ρ is a multiple of the representation ρ_0.

3.3 The projective representation of the symplectic group

28. There is actually another irreducible representation of the Heisenberg group on $L^2(\mathbb{R}^n)$, given by

$$\exp(\mathbf{a}^T\mathbf{Q} + \mathbf{b}^T\mathbf{P} + tI)\psi(\mathbf{p}) = \exp(2\pi i(i\hbar\mathbf{a}^T\frac{\partial}{\partial \mathbf{p}} + \mathbf{b}^T\mathbf{p} + tI))\psi(\mathbf{p}),$$

where the position and momentum operators are

$$\mathrm{op}(q_j)\psi(\mathbf{p}) = i\hbar\frac{\partial \psi}{\partial p_j}, \quad \mathrm{op}(p_j)\psi(\mathbf{p}) = p_j\psi(\mathbf{p}).$$

This is the quantization in the *momentum representation*.

In this case

$$\exp(\mathbf{a}^T\mathbf{Q} + \mathbf{b}^T\mathbf{P} + tI)\psi(\mathbf{p})\psi(\mathbf{p}) = e^{2\pi i\mathbf{b}^T\mathbf{p} - \pi ih\mathbf{a}^T\mathbf{b} + 2\pi it}\psi(\mathbf{p} - h\mathbf{a}),$$

which is the Schrödinger representation of the Heisenberg group in the momentum representation.

By the Stone-von Neumann Theorem, these two representations are unitary equivalent, in fact the unitary equivalence is the Fourier transform

$$(\mathcal{F}_h\psi)(\mathbf{p}) = h^{n/2}\int_{\mathbb{R}^n}\psi(\mathbf{q})e^{-2\pi ih\mathbf{p}^T\mathbf{q}}d\mathbf{q}.$$

29. Note that the above representation of the Heisenberg group, which arises from the quantization of the position and the momentum in the momentum representation can be obtained from the standard representation ρ_0 as follows: First apply to $(\mathbf{a}, \mathbf{b})$ the symplectomorphism with matrix

$$\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

to change the element of the Heisenberg group, then apply ρ_0. We can do this for every element f of the symplectic group (the group of linear symplectomorphisms). By the Stone-von Neumann theorem, the resulting representation, which is irreducible, is unitary equivalent to ρ_0. We thus obtain a map $f \mapsto U_f$, from symplectomorphisms to unitary operators on $L^2(\mathbb{R}^n)$, which because of the Schur Lemma from representation theory satisfies

$$U_{f_1\circ f_2} = cU_{f_1}U_{f_2},$$

where c is a complex number of absolute value 1. This is the projective representation of the symplectic group. There is a way to get rid of the "projectivity" and turn this into a true group representation if we replace the symplectic group by its double cover, the *metaplectic group*. We obtain the Segal-Shale-Weil representation.

We should point out that this means that the symplectomorphisms which are elements of the symplectic group can be *quantized*. Note that the axioms of quantum mechanics make no provisions about the existence of a quantum version of the symplectomorphisms. In the next chapter we will encounter other symplectomorphisms that can be quantized.

Chapter 4

The Schrödinger Equation

4.1 The Schrödinger equation as the time evolution of a state

30. So let us assume that we know how to quantize every Hamiltonian, for example by using the method of Weyl quantization. We will use the short-hand notation $\hat{H}$ for $\mathrm{op}(H)$, the operator associated to H.

In classical Hamiltonian mechanics, the state of a system is given by the position and momentum coordinates of the particle. The time evolution is described by Hamilton's equations.

In quantum mechanics, the quantum state of a system is given by a ray $[\psi]$ in a Hilbert space H (or by abuse of language by ψ itself). The *Schrödinger equation* describes the time evolution of the quantum state of a particle.

It has the form

$$i\hbar\frac{\partial\psi}{\partial t} = \hat{H}\psi.$$

"Where did we get that from? Nowhere. It is not possible to derive it from anything you know. It came out of the mind of Schrödinger." (R. Feynman)

Example 4.1.1. If we take the quantization of the total energy of a particle in one-dimensional space, as it appears in Newtonian mechanics

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

then we have

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial q^2} + M_{V(q)}.$$

So the Schrödinger equation is

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial q^2} + V(q)\psi(q).$$

This is a second order partial differential equation.

31. We want to connect Schrödinger's equation with the Hamiltonian formalism of classical mechanics. Recall that the result of measuring an observable A at time t in state $\psi(t)$ should be

$$\langle A\psi(t), \psi(t)\rangle .$$

In the same vein as for Hamilton's equations, we want to understand the instantaneous rate of change of the result of the measurement. We compute

$$\langle A\psi(t), \psi(t)\rangle = \left\langle Ae^{-(i/\hbar)t\hat{H}}\psi(0), e^{-(i/\hbar)t\hat{H}}\psi(0)\right\rangle = \left\langle e^{(i/\hbar)t\hat{H}}Ae^{-(i/\hbar)t\hat{H}}\psi(0), \psi(0)\right\rangle .$$

Using the product rule for differentiation, we obtain

$$\frac{d}{dt}\langle A\psi(t), \psi(t)\rangle \mid_{t=0} = \left\langle [(i/\hbar)\hat{H}A - (i/\hbar)A\hat{H}]\psi(0), \psi(0)\right\rangle = \left\langle \frac{1}{i\hbar}[A, \hat{H}]\psi(0), \psi(0)\right\rangle .$$

We recognize the quantum Poisson bracket.

Now we can do this at every moment in time, not just $t = 0$, and write the Schrödinger equation for observables

$$\frac{d}{dt}\langle A\psi(t), \psi(t)\rangle = \left\langle \{A, \hat{H}\}_q\psi(t), \psi(t)\right\rangle .$$

This should remind us of Hamilton's equation

$$\frac{df}{dt} = \{f, H\},$$

where f is a classical observable. We arrive at the Heisenberg picture. We consider a situation where the state (wave function) does not change, but the observable (operator) changes as

$$A(t) = e^{(i/\hbar)t\hat{H}}Ae^{-(i/\hbar)t\hat{H}}.$$

Then we get the quantum version of Hamilton's equation, which is

$$\frac{dA}{dt} = \{A, \hat{H}\}_q.$$

32. Let us return to the problem of solving the Schrödinger equation, and let us assume that H is time independent. Then so is $\hat{H}$. We try to separate variables by writing $\psi(t, q) = \phi(q)\tau(t)$. Then the equation becomes

$$i\hbar\phi\frac{d\tau}{dt} = (\hat{H}\phi)\tau.$$

So

$$\frac{i\hbar\frac{d\tau}{dt}}{\tau} = \frac{\hat{H}\phi}{\phi}.$$

Each of the two sides of the equality depends on a different set of variables (time on the left, space on the right), so both should be equal to a constant called E. Hence

$$\hat{H}\phi = E\phi.$$

This is the time-independent Schrödinger equation. It's solutions are also referred to as *stationary-state* solutions. We see that E is an eigenvalue of $\hat{H}$, so E is an energy level that can be measured by an experiment (according to the third axiom of quantum mechanics).

So Schrödinger's equation leads us naturally to considering the energy levels of the quantum Hamiltonian.

There is another way to look at this. Schrödinger's equation should remind us of the well known equation

$$\frac{dx}{dt} = ax.$$

This equation has the solution $x(t) = e^{at}x(0)$. For a system of first order equations

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x},$$

we still have the solution $\mathbf{x}(t) = e^{At}\mathbf{x}(0)$.

So, unlike Hamilton's equations, Schrödinger's equation can be solved easily:

$$\psi(\mathbf{q}, t) = e^{(-i/\hbar)t\hat{H}}\psi(\mathbf{q}, 0).$$

But what does this mean? If $\hat{H}$ is diagonalizable, then we can compute $e^{(-i/\hbar)t\hat{H}}$ the way we exponentiate a diagonalizable matrix. But for that we need to diagonalize $\hat{H}$, so we need to find the stationary state solutions.

Once we have diagonalized $\hat{H}$ in a basis ψ_j with eigenvalues E_j, we can write $\psi(\mathbf{q}, 0) = \sum_j c_j\psi_j$ and then the solution to Schrödinger's equation is

$$\psi(\mathbf{q}, t) = \sum_j c_j\psi_j e^{-\frac{i}{\hbar}E_j t}.$$

Example 4.1.2. The free particle in $\mathbb{R}^3$ has the Schrödinger equation

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\Delta\psi.$$

We try to apply the above considerations to this case. Now we are in trouble already! The Laplacian does not have discrete spectrum, in fact the spectrum of the Laplacian is $[0, \infty)$. So no eigenfunctions in $L^2(\mathbb{R}^3)$. We look for "eigenfunctions" outside of this space.

We solve

$$-\Delta\phi = \lambda\phi, \quad \lambda \geq 0$$

, where $\lambda = \frac{2mE}{\hbar^2}$. This is not so hard, the solutions are

$$\phi(\mathbf{q}) = e^{i\mathbf{k}^T\mathbf{q}}, \quad \mathbf{k} \in \mathbb{R}^3, \|\mathbf{k}\|^2 = \lambda.$$

These are referred to as plane waves. Here $\mathbf{k}$ is the momentum of the plane wave, in agreement with

$$-i\hbar\frac{\partial}{\partial q_j}e^{i\mathbf{k}^T\mathbf{q}} = \hbar k_j e^{i\mathbf{k}^T\mathbf{q}}.$$

We get the solution to Schrödinger's equation given by

$$\psi(\mathbf{q},t) = \int_{\mathbb{R}^3} e^{-\frac{i\hbar}{m}t\|\mathbf{k}\|^2}e^{i\mathbf{k}^T\mathbf{q}}f(\mathbf{k})d\mathbf{k}.$$

Example 4.1.3. Let us consider a free particle confined to a box. This means that we assume that there is a potential which is 0 inside the box and is infinite outside. Let the box be also 1-dimensional, say the interval $[0, L]$. Then the time-independent Schrödinger equation takes the form of the initial value problem

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dq^2} = E\psi(q), \quad \psi(0) = \psi(L) = 0.$$

This only has solutions if

$$E = E_j = \frac{j^2\pi^2\hbar^2}{2mL^2}$$

for some positive integer j. We obtain the corresponding eigenfunctions

$$\psi_j(q) = \sqrt{\frac{2}{L}}\sin\left(\frac{j\pi q}{L}\right), \quad j = 1, 2, 3, \ldots.$$

It is important to observe that ψ_j, $j = 1, 2, 3, ...$ form an orthonormal basis of the Hilbert space $L^2([0, L])$.

4.2 The harmonic oscillator

33. For the classical harmonic oscillator without damping, Hooke's law yields the following Hamiltonian

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

Recall that the solutions are sinusoidal with frequency ω. The quantum Hamiltonian is

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial q^2} + \frac{m\omega^2}{2}M_{q^2} = \frac{1}{2m}\left(-\hbar^2\frac{\partial^2}{\partial q^2} + (m\omega)^2 M_{q^2}\right)$$

We have the Schrödinger equation

$$i\hbar\frac{\partial\psi}{\partial t} = \frac{1}{2m}\left(-\hbar^2\frac{\partial^2\psi}{\partial q^2} + (m\omega)^2 q^2\psi\right)$$

34. Here is an approach to solve the Schrödinger equation without actually solving the differential equation, using Dirac's "ladder operator" method. Based on the formula

$$(a+ib)(a-ib) = a^2 + b^2$$

we define the *annihilation operator* (a.k.a. *raising operator*)

$$a = \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{q} + \frac{i}{m\omega}\hat{p}\right) = \sqrt{\frac{m\omega}{2\hbar}}\left(M_q + \frac{\hbar}{m\omega}\frac{\partial}{\partial q}\right)$$

and the *creation operator* (a.k.a. *raising operator*)

$$a^* = \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{q} - \frac{i}{m\omega}\hat{p}\right) = \sqrt{\frac{m\omega}{2\hbar}}\left(M_q - \frac{\hbar}{m\omega}\frac{\partial}{\partial q}\right).$$

The $*$ is not accidental, one operator is the adjoint of the other. Moreover, using the canonical commutation relation we compute

$$a^*a = \frac{1}{\hbar\omega}\frac{1}{2m}\left(\hat{p} + (m\omega)^2\hat{q}\right) - \frac{1}{2}I.$$

Therefore

$$\hat{H} = \hbar\omega\left(a^*a + \frac{1}{2}I\right).$$

Let us ignore the $\frac{1}{2}I$, known as the quantum correction, and proceed with understanding the spectral properties of a^*a. Right now the elimination of the term $\frac{1}{2}I$ seems ad hoc, but we point out that there is a rigorous procedure called the metaplectic correction for removing this term and adjusting the spectrum of the Hamiltonian to its correct values. Note that $a^*a \geq 0$, so its spectrum lies in $[0,\infty)$. We work in the assumption that a^*a has at least one eigenvalue (which is not necessarily true, but we assume it to be true for our considerations). First, we compute

$$[a^*, a] = I,$$

and then

$$[a, a^*a] = a, \quad [a^*, a^*a] = -a^*.$$

Lemma 4.2.1. Suppose ψ is an eigenvector of a^*a with eigenvalue λ. Then

$$a^*a(a\psi) = (\lambda - 1)(a\psi) \text{ and } a^*a(a^*\psi) = (\lambda + 1)(a^*\psi).$$

Thus $a\psi$ is either equal to zero, or it is an eigenvector of a^*a with eigenvalue $\lambda - 1$, and $a^*\psi$ is either equal to zero, or it is an eigenvector of a^*a with eigenvalue $\lambda + 1$.

Proof. We have

$$a^*a(a\psi) = (a(a^*a) - a)\psi = (\lambda - 1)a\psi$$

and the same for $a^*\psi$. □

As a corollary, since a^*a has only non-negative eigenvalues, the "lowering" of eigenvalues must end somewhere. So for an eigenvalue ψ there is $N > 0$ such that $a^N\psi \neq 0$ but $a^{N+1}\psi = 0$. Let

$$\psi_0 = a^N\psi/\|a^N\psi\|, \quad \psi_n = (a^*)^n\psi_0, \quad n \geq 0.$$

Theorem 4.2.1. The vectors ψ_n, $n \geq 0$ satisfy the following relations

$$\begin{aligned} a^*\psi_n &= \psi_{n+1} \\ a^*a\psi_n &= n\psi_n \\ \langle\psi_n, \psi_m\rangle &= n!\delta_{m,n} \\ a\psi_{n+1} &= n\psi_n. \end{aligned}$$

Proof. The only things to check are the third and fourth properties. For different indices the eigenvectors correspond to different eigenvalues, so they are orthogonal (a property of Hermitian operators). For equal indices have (inductively)

$$\begin{aligned} \langle\psi_n, \psi_n\rangle = \langle a^*\psi_{n-1}, a^*\psi_{n-1}\rangle = \langle\psi_{n-1}, aa^*\psi_{n-1}\rangle = \langle\psi_{n-1}, (a^*a + I)\psi_{n-1}\rangle \\ = ((n-1)+1)\langle\psi_{n-1}, \psi_{n-1}\rangle = n \cdot (n-1)! = n!. \end{aligned}$$

Also $a\psi_{n+1} = aa^*\psi_n = (a^*a + I)\psi_n = (n+1)\psi_n$. □

So if we have exactly one one solution for the equation $a\psi_0 = 0$, then the vectors $\frac{1}{\sqrt{n!}}\psi_n$ form an orthonormal basis of the Hilbert space and they diagonalize $\hat{H}$ with diagonal entries $n + \frac{1}{2}$, $n \geq 0$. In this case we solved the time-independent Schrödinger equation

$$\hat{H}\psi = E\psi,$$

and moreover, we can make sense of the solution to the Schrödinger equation

$$\psi(q, t) = e^{-\frac{i}{\hbar}t\hat{H}}\psi(q, 0)$$

since we know how to exponentiate diagonal operators.

35. Now we turn to analytical methods. The equation $a\psi_0 = 0$ reads

$$\frac{\hbar}{m\omega}\frac{\partial\psi_0}{\partial q} = -q\psi(q),$$

with solution

$$\psi_0(q) = Ce^{-\frac{m\omega}{2\hbar}q^2}.$$

Normalize this to have norm 1 by setting $C = \sqrt{\frac{\pi m\omega}{\hbar}}$. Thus

$$\psi_0(q) = \sqrt{\frac{\pi m\omega}{\hbar}}e^{-\frac{m\omega}{2\hbar}q^2}.$$

This shows that both assumptions in the ladder method work: that the Hamiltonian operator has eigenvalues, and that 0 is an eigenvalue with multiplicity one. The state ψ_0 is called the *ground state*, and the states ψ_n, $n > 0$, are called the *excited states*. For the *Hermite polynomials* defined by

$$H_n(q) = (-1)^n e^{q^2} \frac{d^n}{dq^n} e^{-q^2}, \quad n \geq 0$$

we have

$$\psi_n(q) = \sqrt[4]{\frac{m\omega}{\pi\hbar}} \frac{1}{2^n} e^{-\frac{m\omega}{2\hbar}q^2} H_n\left(\sqrt{\frac{m\omega}{\hbar}} q\right).$$

Note that since the Hermite polynomials form an orthonormal basis of $L^2(\mathbb{R}^2, e^{-q^2/2}dq)$, we deduce that $\frac{1}{\sqrt{n!}}\psi_n$ is an orthonormal basis of $L^2(\mathbb{R}, dq)$, so indeed $\hat{H}$ is diagonalizable and the diagonal entries have one-dimensional eigenspaces.

36. We now describe a *holomorphic model* that produces the creation and annihilation operators. We consider the space of holomorphic functions for which

$$\frac{i}{2\pi} \int_{\mathbb{C}} |f(z)|^2 e^{-|z|^2} dz \wedge d\bar{z} < \infty.$$

Note that the quantity on the left is positive. Endow this space with the inner product

$$\langle f, g \rangle = \frac{i}{2\pi} \int_{\mathbb{C}} f(z)\overline{g(z)} e^{-|z|^2} dz \wedge d\bar{z} < \infty.$$

so that that the previous expression is the square of the norm. Then this is a Hilbert space called the *Segal-Bargmann space*. An orthonormal basis is

$$f_n(z) = \frac{z^n}{\sqrt{n!}}, \quad n = 0, 1, \ldots.$$

We set

$$a^* = M_z, \quad a = \frac{d}{dz}.$$

These satisfy the same condition as the creation and annihilation operators:

$$[a, a^*] = I.$$

The inner product is chosen so that a and a^* are one the adjoint of the other! Note that we can reverse engineer $\hat{q}$ and $\hat{p}$ as

$$\hat{q} = \sqrt{\hbar}\frac{a + a^*}{\sqrt{2m\omega}} \quad \hat{p} = i\sqrt{2\hbar m\omega}(a^* - a),$$

obtaining another representation of the position and momentum operators, which of course is unitary equivalent to the Schrödinger representation because of the Stone-von Neumann Theorem.

4.3 The WKB method

37. This method is named after Wentzel, Krammers, and Brillouin. The idea is to construct approximate solutions to the Schrödinger equation by trying to find solutions of the form

$$\psi(t, \mathbf{q}) = e^{(i/\hbar)S(t,\mathbf{q})}.$$

Substituting in the Schrödinger equation

$$i\hbar\frac{\partial\psi}{\partial t} = \left(-\frac{\hbar^2}{2m}\Delta + V(\mathbf{q})\right)\psi$$

we obtain

$$-\frac{\partial S}{\partial t} = -\frac{\hbar^2}{2m}\left(\frac{i}{\hbar}\Delta_{\mathbf{q}}S - \hbar^{-2}\|\nabla_{\mathbf{q}}S\|^2\right) + V.$$

We can view this as

$$-\frac{\partial S}{\partial t} = \frac{1}{2}\|\nabla_{\mathbf{q}}S\|^2 + V + O(\hbar).$$

Ignoring $O(\hbar)$ we obtain

$$-\frac{\partial S}{\partial t}(\mathbf{q}, t) = H(\mathbf{q}, \nabla_{\mathbf{q}}S).$$

This is the Hamilton-Jacobi equation from classical mechanics. Note that when S is the action ($S = \int L dt$), then this yields an equivalent formulation of classical mechanics.

38. Let us now turn to the time-independent Schödinger equation, and in order to simplify computations we work in the 1-dimensional case:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dq^2} + V(q)\psi = E\psi.$$

The substitution $\psi = e^{(i/\hbar)S(q)}$ yields the differential equation

$$\frac{1}{2m}\left(\frac{dS}{dq}\right)^2 + V(q) = E + \frac{i\hbar}{2m}\frac{d^2S}{dq^2}.$$

As above, we ignore the $\hbar$ term and turn this into the (time-independent version of the) Hamilton-Jacobi equation

$$H(q, S'(q)) = \frac{(S'(q))^2}{2m} + V(x) = E.$$

This gives

$$S'(x) = \pm\sqrt{2m(E - V(x))}.$$

Let us turn to the classical phase space $\mathbb{R}^2 = T^*\mathbb{R}$. Then the differential $dS = S'dq$ can be viewed as a map $dS : \mathbb{R} \to T^*\mathbb{R}$, $q \mapsto S'(q)dq$. Then S satisfies the Hamilton-Jacobi equation if and only if the image of dS lies in the level manifold $H^{-1}(E)$. This establishes a fundamental relationship between classical and quantum mechanics: "When the image of dS lies in a level manifold of the classical Hamiltonian, then $e^{(i/\hbar)S}$ can be taken as a first-order approximate solution to the Hamilton equation."

Here is another way to look at this. Using the expansion

$$S(q) = S(q, \hbar) = \sum_{n=0}^{\infty} (-i\hbar)^n S_n(q),$$

we obtain

$$\frac{1}{2m}(S_0')^2 = E - V(q), \quad S_0'S_1' = -\frac{1}{2}S_0''.$$

Let $p(q) = \sqrt{2m(E - V(x)}$ be the classical momentum of a particle with potential V and energy E. The solution to the first differential equation is

$$S_0 = \pm \int p(q)dq,$$

and from the second we get $S_1 = \frac{1}{2}\ln p$. So

$$\psi(q) = \frac{1}{\sqrt{|p(x)|}} \left(c_1 e^{(i/\hbar)\int p(q)dq} + c_2 e^{-(i/\hbar)\int p(q)dq} \right) (1 + O(\hbar))$$

39. We can go to a more general *Ansatz*, namely that

$$\psi(q) = A(q)e^{\pm(i/\hbar)S(q)},$$

where $S'(q) = p(q) = \sqrt{2m(E - V(q))}$. The amplitude function $A(q)$ is chosen to be independent of $\hbar$.

Proposition 4.3.1. For any two numbers E_1 and E_2 with $E_1 > \inf V(q)$ there exists a constant C and a nonzero compactly supported smooth function $A(q)$ with the following property: For every $E \in [E_1, E_2]$ the support of A is contained in the classically allowed region at energy E and the function ψ given by

$$\psi(q) = A(q) \exp\left(\pm \frac{i}{\hbar} \int p(q)dq \right)$$

satisfies

$$\|\hat{H}\psi - E\psi\| \leq C\hbar\|\psi\|.$$

Here by classically allowed region we mean the region where a particle with energy E can lie in the configuration space.

Proof. We will actually choose A to by any function with support in the classically allowed region for E_1 (which is included in the classically allowed region for E). We compute

$$\hat{H}\psi - E\psi = -\frac{\hbar^2}{2m}\left(A''(q) \pm 2\frac{i}{\hbar}A'(q)p(q) \pm \frac{i}{\hbar}p'(q)A(q)\right)e^{\pm(i/\hbar)\int p(q)dq}.$$

Thus

$$\|\hat{H}\psi - E\psi\| \leq \frac{\hbar^2}{2m}\|A''\| + \frac{\hbar}{2m}\|2A'p + Ap'\|.$$

Notice that the right-hand side is of order $\hbar\|\psi\|$ and $\|2A'p + Ap'\|$. □

We can even improve on the error by forcing $2A'p = -p'A$ so as to obtain

$$\hat{H}\psi - E\psi = -\frac{\hbar^2}{2m}A''\psi.$$

Chapter 5

The Hydrogen Atom. The Spin

This chapter follows closely the book of Faddeev and Yakubovskii.

5.1 The classical Kepler problem

40. We start with a short detour through classical mechanics. The classical Kepler problem asks to find the trajectories of planets orbiting the sun. The sun exterts the force

$$\mathbf{F} = -k\frac{\mathbf{q}}{\|\mathbf{q}\|^3},$$

which is given by the potential

$$V(\mathbf{q}) = -\frac{k}{\|\mathbf{q}\|}.$$

If m is the mass of the planet and M is the mass of the sun, then actually $k = GMm$.

Because the potential is invariant under rotations, the angular momentum $\mathbf{J} = \mathbf{q} \times \mathbf{p}$ is conserved (and we require it to be nonzero to avoid collision with the sun). Note that the planet moves in a plane orthogonal to $\mathbf{J}$. We can compute explicitly

$$\mathbf{J} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ q_1 & q_2 & q_3 \\ p_1 & p_2 & p_3 \end{vmatrix} = (q_2p_3 - q_3p_2)\mathbf{i} + (q_3p_1 - q_1p_3)\mathbf{j} + (q_1p_2 - q_2p_1)\mathbf{k},$$

and the conservation of the momentum follows by an easy computation in which we use Hamilton's equations.

Definition. The *Runge-Lenz vector* is the vector-valued function on $\mathbb{R}^3\backslash\{0\} \times \mathbb{R}^3$ given by

$$\mathbf{A}(\mathbf{q}, \mathbf{p}) = \frac{1}{mk}\mathbf{p} \times \mathbf{J} - \frac{\mathbf{q}}{\|\mathbf{q}\|}.$$

Proposition 5.1.1. The Runge-Lenz vector is conserved.

Proof. Using the fact that $\dot{\mathbf{p}} = \mathbf{F}$, we compute

$$\begin{aligned}\dot{A}(t) &= \frac{1}{mk}\mathbf{F}\times\mathbf{J} - \frac{1}{\|\mathbf{q}\|}\frac{\mathbf{p}}{m} + \frac{\mathbf{q}}{\|\mathbf{q}\|^2}\sum_{j=1}^{3}\frac{\partial\|\mathbf{q}\|}{\partial q_j}\frac{dq_j}{dt}\\ &= -\frac{1}{m}\frac{1}{\|\mathbf{q}\|^3}\mathbf{q}\times(\mathbf{q}\times\mathbf{p}) - \frac{1}{\|\mathbf{q}\|}\frac{\mathbf{p}}{m} + \frac{\mathbf{q}}{\|\mathbf{q}\|^2}\sum_{j=1}^{3}\frac{\partial\|\mathbf{q}\|}{\partial q_j}\frac{dq_j}{dt}\\ &= \frac{1}{m}\left(-\frac{1}{\|\mathbf{q}\|^3}\mathbf{q}(\mathbf{q}\cdot\mathbf{p}) + \frac{1}{\|\mathbf{q}\|^3}\mathbf{p}(\mathbf{q}\cdot\mathbf{q}) - \frac{\mathbf{p}}{\|\mathbf{q}\|} + \frac{\mathbf{q}(\mathbf{q}\cdot\mathbf{p})}{\|\mathbf{q}\|^3}\right) = 0.\end{aligned}$$

Here we used $\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = \mathbf{b}(\mathbf{a}\cdot\mathbf{c}) - \mathbf{c}(\mathbf{a}\cdot\mathbf{b})$. □

41. Using the Runge-Lenz vector as an auxiliary tool we will now prove that the trajectory is either an *ellipse*, a *parabola*, or a *hyperbola*. Note that $\mathbf{A}$ lies in the plane of motion. Choose cartesian coordinates in this plane of motion such that $\mathbf{A}$ defines the x-axis, then switch to *polar coordinates*.

Theorem 5.1.1. In polar coordinates, the trajectory of the planet is given by

$$r(t) = \frac{\|\mathbf{J}\|^2}{mk}\frac{1}{1+\|\mathbf{A}\|\cos\theta(t)}.$$

Proof. Let $\mathbf{r}(t)$ be the trajectory. We compute

$$\mathbf{A}\cdot\mathbf{r} = Ar\cos\theta = \frac{1}{mk}\mathbf{r}\cdot(\mathbf{p}\times\mathbf{J}) - r.$$

But

$$\mathbf{r}\cdot(\mathbf{p}\times\mathbf{J}) = (\mathbf{r}\times\mathbf{p})\cdot\mathbf{J} = \mathbf{J}\cdot\mathbf{J} = \|\mathbf{J}\|^2.$$

Now solve for r. □

Note that the equation from the statement is the equation of a conic

$$r = \frac{ep}{1+e\cos\theta}$$

where e is the excentricity and $|p|$ is the distance between the focus and the directrix.

5.2 Angular momentum in quantum mechanics

42. The angular momentum $\mathbf{J}$ is a vector valued function, so it is not an observable in the way we define observables. But its coordinates themselves are observables. Its coordinates in classical mechanics are $q_2p_3 - q_3p_2$, $q_3p_1 - q_1p_3$, and $q_1p_2 - q_2p_1$. We quantize these as

$$\begin{aligned}L_1 &= \hat{q}_2\hat{p}_3 - \hat{q}_3\hat{p}_2 = i\hbar\left(M_{q_3}\frac{\partial}{\partial q_2} - M_{q_2}\frac{\partial}{\partial q_3}\right)\\ L_2 &= \hat{q}_3\hat{p}_1 - \hat{q}_1\hat{p}_3 = i\hbar\left(M_{q_1}\frac{\partial}{\partial q_3} - M_{q_3}\frac{\partial}{\partial q_1}\right)\\ L_3 &= \hat{q}_1\hat{p}_2 - \hat{q}_2\hat{p}_1 = i\hbar\left(M_{q_2}\frac{\partial}{\partial q_1} - M_{q_1}\frac{\partial}{\partial q_2}\right).\end{aligned}$$

Here it is important to notice that for $j \neq k$, $\hat{q_j}$ commutes with $\hat{p_k}$, so we do get self-adjoint operators. It is also customary to work with

$$L^2 = L_1^2 + L_2^2 + L_3^2.$$

We have the following commutation relations

$$[L_1, L_2] = i\hbar L_3, \quad [L_2, L_3] = i\hbar L_1, \quad [L_3, L_1] = i\hbar L_2, \quad [L_1, L^2] = [L_2, L^2] = [L_3, L^2] = 0.$$

The operator L^2 is positive, so $L = \sqrt{L^2}$ can be defined, in case it is needed.

Lemma 5.2.1. If in spherical coordinates ψ depends on $r = \sqrt{q_1^2 + q_2^2 + q_3^2}$ only, then $L_j\psi = 0$, $j = 1, 2, 3$.

Proof. We only check L_3, the others are exactly the same.

$$L_3\psi(r) = L_3\psi(q_1^2 + q_2^2 + q_3^2) = i\hbar\left(q_2\frac{\partial\psi}{\partial q_1} - q_1\frac{\partial\psi}{\partial q_2}\right) = i\hbar\psi'(q_1^2 + q_2^2 + q_3^2)(2q_2q_1 - 2q_1q_2) = 0.$$

□

43. Since we will need the Laplacian, which defines the quantum kinetic energy, in the Schrödinger equation, we have the following result.

Lemma 5.2.2. In spherical coordinates the Laplacian is given by the formula

$$\Delta = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{1}{\hbar^2 r^2}L^2.$$

Proof. We compute

$$-\frac{1}{\hbar^2}L^2 = \left(M_{q_3}\frac{\partial}{\partial q_2} - M_{q_2}\frac{\partial}{\partial q_3}\right)^2 + \left(M_{q_1}\frac{\partial}{\partial q_3} - M_{q_3}\frac{\partial}{\partial q_1}\right)^2 + \left(M_{q_2}\frac{\partial}{\partial q_1} - M_{q_1}\frac{\partial}{\partial q_2}\right)^2$$
$$= \sum_j q_j^2\Delta - \sum_j q_j^2\frac{\partial^2}{\partial q_j^2} - 2\sum_{j<k} q_jq_k\frac{\partial^2}{\partial q_j\partial q_k} - 2\sum_j q_j\frac{\partial}{\partial q_j}.$$

The chain rule gives

$$\frac{\partial}{\partial q_j} = \frac{\partial}{\partial r}\frac{\partial r}{\partial q_j} = \frac{\partial}{\partial r}\frac{q_j}{r}.$$

Multiply this by q_j and add over $j = 1, 2, 3$ to obtain

$$r\frac{\partial}{\partial r} = \sum_j q_j\frac{\partial}{\partial q_j}.$$

Square to obtain

$$\left(r\frac{\partial}{\partial r}\right)^2 = \sum_j q_j^2\frac{\partial^2}{\partial q_j^2} + 2\sum_{j<k} q_jq_k\frac{\partial^2}{\partial q_j\partial q_k} + \sum_j q_j\frac{\partial}{\partial q_j}.$$

Combine this with the previous computation to obtain the conclusion. □

5.3 The rotation group $SO(3)$ and its Lie algebra

44. Let us recall a few facts about the rotation group $SO(3)$ and its Lie algebra $so(3)$. The Lie group consists of orthogonal matrices of determinant 1. The Lie algebra $so(3)$ consists of the skew symmetric 3×3 matrices. It is generated by the matrices

$$A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

These are the infinitesimal rotations about the three coordinate axes. They satisfy

$$[A_j, A_{j+1}] = A_{j+2},$$

where indices are taken modulo 3. Every other skew symmetric matrix is a linear combination of these. The rotations by angle α about the coordinate axes are

$$e^{\alpha A_1}, \quad e^{\alpha A_2}, \quad e^{\alpha A_3}.$$

In general, if $\mathbf{n} = (n_1, n_2, n_3)$ is a unit vector, then the rotation by angle α about the axis of this vector is

$$g(\alpha \mathbf{n}) = e^{\alpha(n_1 A_1 + n_2 A_2 + n_3 A_3)}$$

Proposition 5.3.1. The Lie algebra $so(3)$ is isomorphic with the Lie subalgebra of quantum observables generated by L_1, L_2, L_3.

Proof. Recall that the Lie bracket of operators is $\frac{1}{i\hbar}[\cdot, \cdot]$. With this in mind, the isomorphism is $A_j \mapsto L_j$, $j = 1, 2, 3$. □

Theorem 5.3.1. The representation of $SO(3)$ on $L^2(\mathbb{R}^3)$ defined by

$$W(g)\psi(\mathbf{q}) = \psi(g^{-1}\mathbf{q})$$

satisfies

$$W(g(\alpha \mathbf{n})) = e^{-(i/\hbar)\alpha(n_1 L_1 + n_2 L_2 + n_3 L_3)}.$$

Proof. Let us check this for the case where $\mathbf{n} = \mathbf{k}$ is unit vector defining the z-axis. In this case we have to prove that

$$W(g(\alpha \mathbf{k}))\psi(\mathbf{q}) = \psi(q_1 \cos\alpha + q_2 \sin\alpha, -q_1 \sin\alpha + q_2 \cos\alpha, q_3) = e^{-i\hbar\alpha L_3}\psi(\mathbf{q}).$$

Consider the function

$$\Psi_{\mathbf{q}}(\alpha) = e^{-(i/\hbar)\alpha L_3}\psi(\mathbf{q}).$$

We have

$$\frac{d\Psi_{\mathbf{q}}(\alpha)}{d\alpha} = -(i/\hbar)L_3\phi_{\mathbf{q}}(\alpha) = \left(q_2 \frac{\partial}{\partial q_1} - q_1 \frac{\partial}{\partial q_2}\right)\Psi_{\mathbf{q}}(\alpha),$$

and $\Psi_{\mathbf{q}}(0) = \psi(\mathbf{q})$. It is not hard to check that $\psi(g(\alpha \mathbf{k})\mathbf{q})$ satisfies the same initial value problem, so they are equal and the theorem is proved. □

45. Let us turn to spherical coordinates:

$$x = r\sin\phi\cos\theta, \quad y = r\sin\phi\sin\theta, \quad z = r\cos\phi.$$

The convention follows the Texas Tech calculus text book. Strangely there are two notational conventions, exchanging θ and ϕ. So be very careful!

We introduce a normalized version of the angular momentum operators, the so called *dimensionless momentum operators*, by

$$\tilde{L}_j = \frac{1}{\hbar} L_j.$$

Let $\tilde{L}^2 = \tilde{L}_1^2 + \tilde{L}_2^2 + \tilde{L}_3^2$.

As Lemma 5.2.1 showed, the angular momentum operators depend only on the angle variables and not on r. Thus it is wise to consider the decomposition

$$L^2(\mathbb{R}^3) = L^2((0,\infty)) \otimes L^2(S^2)$$

where $L^2(0,\infty)$ is endowed with the integration measure $r^2 dr$ and $L^2(S^2)$ is endowed with the integration measure $\sin\phi d\phi d\theta$.

Switching to spherical coordinates, we have

$$\begin{aligned}
\tilde{L}_1 &= i\left(\sin\theta\frac{\partial}{\partial\phi} + \cot\phi\cos\theta\frac{\partial}{\partial\theta}\right)\\
\tilde{L}_2 &= -i\left(\cos\theta\frac{\partial}{\partial\phi} - \cot\phi\sin\theta\frac{\partial}{\partial\theta}\right)\\
\tilde{L}_3 &= -i\frac{\partial}{\partial\theta}.
\end{aligned}$$

Set

$$L_\pm = \tilde{L}_1 \pm i\tilde{L}_2 = e^{\pm i\theta}\left(\pm\frac{\partial}{\partial\phi} + i\cot\phi\frac{\partial}{\partial\theta}\right).$$

Note that a representation of the Lie algebra generated by L_1, L_2, L_3 is the same a representation for the Lie algebra generated by $L_-, L_+, \tilde{L}_3$.

Let us understand what we want. We have a representation of $SO(3)$ on the space of states by unitary operators, that arises naturally from the rotations of the space. We want to understand the spectral theory of each operator. It suffices to choose a direction in space, and declare it the z-axis, so that the operator is now rotation about the z-axis.

Now, as it is customary in the representation theory of Lie groups, instead of addressing the question of representing the group, we address the question of representing the Lie algebra. And indeed, we can relate the spectral theory of L_3 and e^{iL_3} to each other by the spectral mapping theorem. Thus we focus on L_3, or rather on $\tilde{L}_3$.

First, note that if $\tilde{L}_3$ and $\tilde{L}^2$ are diagonalizable, then because they commute, they are simultaneously diagonalizable. We proceed as before, assuming that they are indeed diagonalizable, and then by understanding the properties of eigenvectors we compute them

explicitly and thus prove that the operators are indeed diagonalizable. Let therefore $Y(\theta, \phi)$ be a common eigenvalue of $\tilde{L}_3$ and $\tilde{L}^2$:

$$\tilde{L}^2 Y = \lambda Y, \quad \tilde{L}_3 Y = mY.$$

Note that m is necessarily an integer, since Y is periodic in θ with period 2π. Also, note that $\tilde{L}^2 = \tilde{L}_1^2 + \tilde{L}_2^2 + \tilde{L}_3^2$ implies $\lambda \geq m^2$. Denote by $E_{\lambda,m}$ the common eigenspace of $\tilde{L}^2$ and $\tilde{L}_3$ with eigenvalues λ, m. We have

$$[\tilde{L}_3, L_+] = L_+, \quad [\tilde{L}_3, L_-] = L_-, \quad [L_+, L_-] = 2\tilde{L}_3.$$

Consequently

$$L_+ : E_{\lambda,m} \to E_{\lambda,m+1}, \quad L_- : E_{\lambda,m} \to E_{\lambda,m-1},$$

so these are raising and lowering operators in the sense of representation theory. Because $\lambda \geq m^2$, there must be a moment when $L_+ = 0$, that is $L_+^{k+1} E_{\lambda,m} = 0$ where k is chosen to be the smallest with this property. So $L_+^{k+1} Y = 0$ and $L_+^k E_{\lambda,m} \subset E_{\lambda,m+k}$. Then since $L_- L_+ = L^2 - \tilde{L}_3^2 - \tilde{L}_3$,

$$0 = L_- L_+ L_+^k Y = (\lambda - (m+k)^2 - (m+k))Y = (\lambda - (m+k)(m+k+1))Y.$$

So $\lambda = \ell(\ell+1)$, where $\ell = m + k$. For $\lambda = m(m+1) = \ell(\ell+1)$, let $Y_{\ell,\ell} \in E_{m(m+1),m}$. We should have

$$\begin{aligned} -i\frac{\partial Y_{\ell,\ell}}{\partial \theta} &= \ell Y_{\ell,\ell} \\ \frac{\partial Y_{\ell,\ell}}{\partial \phi} + i\cot\phi \frac{\partial Y_{\ell,\ell}}{\partial \theta} &= 0. \end{aligned}$$

From the first equation we see that

$$Y_{\ell,\ell}(\theta, \phi) = e^{i\ell\theta} F_{\ell,\ell}(\phi).$$

Now substitute in the second equation to obtain

$$\frac{\partial F_{\ell,\ell}(\phi)}{\partial \phi} = l\cot\phi F_{\ell,\ell}(\phi).$$

So $F_{\ell,\ell}(\phi) = C\sin^\ell \phi$. Thus

$$Y_{\ell,\ell}(\theta, \phi) = C\sin^\ell \phi e^{i\ell\theta}.$$

We deduce that the spaces $E_{\lambda,m}$ are 1-dimensional. The lowering process gives, after a normalization,

$$\begin{aligned} Y_{\ell,m-1} &= -\frac{1}{\sqrt{(\ell+m)(\ell-m+1)}} L_- Y_{\ell,m} \\ &= -\frac{1}{\sqrt{(\ell+m)(\ell-m+1)}} e^{-i\theta}\left(-\frac{\partial}{\partial\phi} + i\cot\phi\frac{\partial}{\partial\theta}\right) Y_{\ell,m}. \end{aligned}$$

Some computations yield

$$Y_{\ell,m} = -\frac{1}{\sqrt{2\pi}} e^{im\theta} P_\ell^m(\cos\phi),$$

where

$$P_\ell^m(t) = \sqrt{\frac{(\ell+m)!}{(\ell-m)!}} \sqrt{\frac{2\ell+1}{2}} \frac{1}{2^\ell \ell!} (1-t^2)^{-\frac{m}{2}} \frac{d^{\ell-m}(t^2-1)^\ell}{dt^{\ell-m}}$$

are the (normalized) Legendre polynomials. The functions $Y_{\ell,m}$ are called *spherical functions*.

We conclude that a basis for an irreducible representation of the rotation group consists of

$$Y_{\ell,m}(\theta,\phi), \quad m = -\ell, -\ell+1, \ldots, \ell-1, \ell,$$

for a fixed ℓ. Note that spherical functions are dense in $L^2(S^2)$, so the representation of $SO(3)$ on $L^2(S^2)$ can be decomposed into irreducible representations on spaces of spherical functions. Note that since we were able to solve the above differential equations, and since spherical functions span a dense subspace of $L^2(S^2)$, we deduce that L^2 and $\tilde{L}_3$ are indeed diagonalizable. What is the most amazing part of this investigation of the issue of diagonalizing $\tilde{L}_3$ is that it gives rise all finite dimensional irreducible representations of $SO(3)$.

5.4 The Schrödinger equation for the hydrogen atom

46. Here we have a proton (of mass m_1) and an electron (of mass m_2). So the quantum Hamiltonian of this system is

$$H = -\frac{\hbar}{2m_1}\Delta_1 - \frac{\hbar}{2m_2}\Delta_2 + V(\mathbf{q_1} - \mathbf{q_2}),$$

where $\mathbf{q_1}$ and $\mathbf{q}_2$ are the coordinates of the proton and the electron. We use the new variables

$$\mathbf{Q} = \frac{m_1\mathbf{q}_1 + m_2\mathbf{q}_2}{m_1+m_2}, \quad \mathbf{q} = \mathbf{q}_1 - \mathbf{q}_2,$$

where $\mathbf{Q}$ is the coordinate of the center of inertia, and $\mathbf{q}$ is the relative coordinate. In these coordinates

$$H = -\frac{\hbar^2}{2M}\Delta_{\mathbf{Q}} - \frac{\hbar^2}{2\mu}\Delta_{\mathbf{q}} + V(\mathbf{q}),$$

with $M = m_1 + m_2$, $\mu = m_1 m_2/(m_1+m_2)$. The time-independent Schrödinger equation

$$H\psi = E\psi$$

can be separated into the $\mathbf{Q}$ and $\mathbf{q}$ variables, and only the second has something interesting, as the first is just the Schrödinger equation for free motion.

47. So we focus on the Schrödinger equation

$$-\frac{\hbar}{2\mu}\Delta\psi + V(\mathbf{q})\psi = E\psi.$$

The hydrogen atom has the potential with spherical symmetry

$$V(\mathbf{q}) = V(r) = -\frac{e^2}{r}.$$

Using Lemma 5.2.2 we can write the time-independent Schrödinger equation as

$$-\frac{\hbar^2}{2\mu r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)\psi + \frac{L^2}{2\mu r^2}\psi - \frac{e^2}{r}\psi = E\psi$$

We solve the equation in *atomic units* with $\hbar = 1$, $\mu = 1$, $e^2 = 1$. We look for solutions of the form

$$\psi(r, \theta, \phi) = R_\ell(r)Y_{\ell,m}(\theta, \psi).$$

They are eigenvectors of $L^2 = \tilde{L}^2$ and $L_3 = \tilde{L}_3$, so they describe states of the particle with definite values of the square of the angular momentum and its z-projection. We obtain the following equation for R_ℓ:

$$-\frac{1}{2r^2}\frac{d}{dr}\left(r^2\frac{dR_\ell}{dr}\right) + \frac{\ell(\ell+1)}{2r^2}R_\ell - \frac{1}{r}R_\ell = ER_\ell.$$

48. We introduce the function $f_\ell = rR_\ell$. This satisfies the radial Schrödinger equation

$$-\frac{1}{2}f_\ell''(r) + \frac{l(l+1)}{2r^2}f_\ell(r) - \frac{1}{r}f_\ell = Ef_\ell.$$

The radial equation coincides with the time-independent Schrödinger equation

$$-\frac{1}{2}\frac{d^2\psi}{dr^2} + V(r)\psi = E\psi$$

if we introduce the effective potential

$$V_{eff}(r) = \begin{cases} -\frac{1}{r} + \frac{\ell(\ell+1)}{2r^2} & \text{if } r > 0 \\ \infty & \text{if } r \leq 0. \end{cases}$$

Here we take into account that $f_\ell(r)$ is only defined for $r > 0$.

Let us examine the behavior of the solution when $r \to \infty$ and $r \to 0$. The first limit produces the equation

$$\frac{1}{2}f_\ell'' + Ef_\ell = 0.$$

For $E > 0$ this has two linearly independent solutions e^{-ikr} and e^{-ikr}, where $k^2 = 2E$. If $E < 0$, then this has two linearly independent solutions $e^{-\kappa r}$, $e^{\kappa r}$, $\kappa^2 = -2E$.

When $r \to 0$, note that the term with denominator r^2 dominates the term with denominator r, so we should have something like

$$f_l'' - \frac{\ell(\ell+1)}{r^2} f_\ell = 0.$$

This has solutions $r^{-\ell}$ and $r^{\ell+1}$.

We are looking for solutions of the form $\psi(\mathbf{q}) = \frac{f_\ell(r)}{r} Y_{\ell m}(\theta, \phi)$ which are continuous (so that they satisfy the Schrödinger equation), and are either square integrable, in which case they yield eigenfunctions, or are bounded, in which case they they give a point in the spectrum that belongs to the continuous spectrum.

The continuity of ψ implies $f_\ell(0) = 0$, thus near zero the solution should be like $Cr^{\ell+1}$. Next, for $E > 0$, the solution f_ℓ is always bounded, all we need is to impose the right behavior at 0, so this gives the continuous spectrum. For $E < 0$, we should have a solution that behaves like $Ce^{-\kappa r}$ when $r \to \infty$. We will see that this situation gives rise to the eigenvalues.

49. It is convenient to look for a solution of the form

$$f_\ell(r) = r^{l+1} e^{-\kappa r} \Lambda_\ell(r).$$

Substituting in the radial Schrödinger equation we obtain the second order equation for Λ_ℓ:

$$\Lambda_\ell'' + \left(\frac{2(\ell+1)}{r} - 2\kappa \right) \Lambda_\ell' + \left(\frac{2}{r} - \frac{2\kappa(\ell+1)}{r} \right) \Lambda_\ell = 0.$$

Now we set

$$\Lambda_\ell(r) = \sum_{j=0}^{\infty} a_j r^j.$$

The second order differential equation yields a first order recursive relation in the coefficients (because of the way the powers of r appear). This recursive relation is

$$a_{j+1} = 2 \frac{\kappa(j+\ell+1) - 1}{(j+1)(j+2\ell+2)} a_j.$$

The ratio test shows that the series converges for all r. Let us examine the large r behavior of the resulting function. Note that

$$a_{j+1} \approx \frac{2\kappa}{j+1} a_j \text{ so } a_j \approx c \frac{(2\kappa)^j}{j!}.$$

Thus $\Lambda_\ell \approx ce^{2\kappa r}$. But then f_ℓ grows too fast. So the only hope is that the recursive relation hits a zero! This happens exactly when

$$\kappa = \frac{1}{j+\ell+1} \text{ for some } j.$$

Thus we obtain the eigevalues of the Hamiltonian as the values E such that $-2E = \frac{1}{n^2}$. These are

$$E_n = -\frac{1}{2n^2},$$

with eigenfunctions of the form

$$\psi = r^\ell e^{-\frac{1}{k+\ell+1}r}\lambda_{k\ell}(r)Y_{lm}(\theta,\phi).$$

And we obtain the formula for the frequencies of the spectral lines

$$h\nu_{mn} = E_n - E_m, \quad E_n > E_m,$$

where, when converting back from standard units units, we have

$$E_n = -\frac{\mu e^4}{2n^2h^2}.$$

5.5 The Lie group $SU(2)$ and the spin

50. Recall that

$$SU(2) = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \mid |a|^2 + |b|^2 = 1 \right\}.$$

We immediately notice that as a manifold $SU(2)$ is homeomorphic to the 3-dimensional sphere. Writing $a = x_1 + iy_1$, $b = x_2 + iy_2$, we have

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} = x_1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + y_1 \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} + x_2 \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + y_2 \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

This establishes an isomorphism between $SU(2)$ and the group U of unit quaternions, with the four matrices corresponding to $1, i, j, k$.

Now we construct a $2-1$ surjective group homomorphism

$$\pi : SU(2) \to SO(3).$$

For this, let $\mathbb{H}$ be the set of quaternions and let us consider the Lie group homomorphism

$$U \to Aut(\mathbb{H}), \quad g \mapsto (x \mapsto gxg^{-1}).$$

Note that image of U consists of isometries, and because it fixes the reals, it also fixes the orthogonal space to the reals, that is the imaginaries:

$$\text{Im } \mathbb{H} = \{bi + cj + dk \mid a, b, c \in \mathbb{R}\}.$$

We identify Im $\mathbb{H}$ with $\mathbb{R}^3$, thus we see that $SU(2)$ is represented by isometries of $\mathbb{R}^3$. Because $SU(2)$ is simply connected, these isometries must lie in $SO(3)$. Moreover, the kernel is $\{\pm 1\}$,

and since $SU(2)$ and $SO(3)$ are of the same dimension the map π that we constructed is onto.

51. As a corollary, the Lie algebras of $SU(2)$ and $SO(3)$ are isomorphic. But we can construct the Lie algebra $su(2)$ directly; it consists of the matrices that are trace zero and antihermitian:

$$su(2) = \left\{ \begin{pmatrix} ia & z \\ -\bar{z} & -ia \end{pmatrix} \mid a \in \mathbb{R}, z \in \mathbb{C} \right\}.$$

It is generated by the matrices

$$u_1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \quad u_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad u_3 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

Note that $u_1 = i\sigma_1$, $u_2 = -i\sigma_2$, $u_3 = i\sigma_3$, where $\sigma_1, \sigma_2, \sigma_3$ are the Pauli matrices. We have the isomorphism

$$su(2) \to so(3), \quad u_j \mapsto 2A_j, \quad j = 1, 2, 3.$$

52. We now describe all irreducible representations of $SU(2)$. Let V^n be the space of homogeneous polynomials of degree n in two variables z_1 and z_2. The dimension of V^n is $n+1$, and we endow it with the inner product that makes $P_k(z_1, z_2) = z_1^k z_2^{n-k}$ an orthonormal basis. We let $SU(2)$ act by

$$(gP)(z_1, z_2) = P\left((z_1, z_2)g\right), \quad g \in SU(2).$$

Theorem 5.5.1. The representations V^n, $n \geq 0$ are irreducible and unitary and there are no other finite dimensional unitary irreducible representations of $SU(2)$.

Proof. It is not hard to check that the representations are unitary, because they preserve lengths. To show that they are irreducible, we follow the book of Bröcker and tom Dieck. It suffices to show that every endomorphism A of V^n that is $SU(2)$ invariant is a multiple of identity. First let

$$g = \begin{pmatrix} e^i & 0 \\ 0 & e^{-i} \end{pmatrix} \in SU(2).$$

Then $gP_k = e^{(2k-n)i} P_k$, so $gAP_k = e^{(2k-n)i} AP_k$. Since the $e^{(2k-n)i}$-eigenspace of g acting on V^n is $\mathbb{C}P_k$, we must have $AP_k = c_k P_k$ for some $c_k \in \mathbb{C}$.

We now consider the real rotations

$$r_t = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}.$$

We have

$$\begin{aligned} Ar_t P_n = A(z_1 \cos t + z_2 \sin t)^n &= \sum_k \binom{n}{k} \cos^k t \cdot \sin^{n-k} t \cdot AP_k \\ &= \sum_k \binom{n}{k} \cos^k t \cdot \sin^{n-k} t \cdot c_k P_k. \end{aligned}$$

Also

$$r_t A P_n = \sum_k \binom{n}{k} \cos^k t \cdot \sin^{n-k} t \cdot c_n P_k.$$

This means that $c_k = c_n$, so $A = c_n I$.

For the converse, we use the fact that every representation of a Lie group into $GL(V)$ is smooth. With the following lemma we transfer the question about representations of the Lie group to that of representations of the Lie algebra.

Lemma 5.5.1. Let $SU(2) \to GL(V)$ be a smooth representation, and let $su(2) \to gl(V)$ be the induced representation of Lie algebras. Then V is an irreducible representation of $SU(2)$ if and only if it is an irreducible representation of $su(2)$.

Proof. Note that $SU(2)$ is obtained by exponentiating elements of $su(2)$. Let $v \in V \backslash \{0\}$. Then

$$\lim_{t\to 0} \frac{e^{Dt}v - v}{t} = Dv \text{ and } e^D v = \sum_{n=0}^{\infty} \frac{1}{n!} D^n v.$$

So every point in a minimal (closed) invariant subspace containing v of $SU(2)$ is in a minimal (closed) invariant subspace containing v of $su(2)$ and vice-versa. □

Let us study the invariant subspaces of $su(2)$. We complexify

$$sl(2,\mathbb{C}) = su(2) \otimes_{\mathbb{R}} \mathbb{C}$$

and work on the easier task of finding the irreducible representations of $sl(2,\mathbb{C})$.

Next, let $L_3 = -iu_3$, $L_+ = \frac{1}{\sqrt{2}}(-u_2 - iu_1)$, $L_- = \frac{1}{\sqrt{2}}(u_2 - iu_1)$, so that $[L_3, L_\pm] = \pm L_\pm$, $[L_+, L_-] = L_3$. Next, let v be an eigenvector of L_3 in some irreducible representation of $sl(2,\mathbb{C})$, with eigenvalue λ. Then

$$L_3 L_+ v = (\lambda + 1) L_+ v \text{ and } L_3 L_- v = (\lambda - 1) L_- v.$$

So L_+ and $L+-$ act as raising/lowering operators. Then there is k such that $L_+^k v = 0$. We may actually assume $L_+ v = 0$ (by changing notation). Then the irreducible representation is the span of $v, L_- v, L_-^2 v, \ldots, L_-^n v$ for some n. But then the trace of L_3 is the sum of its eigenvalues, which are λ, $\lambda - 1$, ...,$\lambda - n$. L_3 being a commutator has trace zero, which implies $\lambda = \frac{n}{2}$. This means that there is only one irreducible representation of $sl(2,\mathbb{C})$ in each dimension, and we already know that there is one such representation that comes from $SU(2)$. So we have recovered the representations V^{n+1} of $SU(2)$, $n \geq 0$. □

53. We index the irreducible representations by $\ell = \frac{\dim\ V-1}{2} = \frac{n}{2}$, so we change the notation to $V_\ell = V^{2\ell+1}$. The number ℓ is called *spin*.

The eigenvector v is called the highest weight vector. It's eigenvalue is $\ell = \frac{n}{2}$. Note that $L_+(L_-^k v) = k(\ell - \frac{1}{2}(k-1))(L_-^{k-1} v)$. We define the inner product on V_ℓ by asking that the following vectors form an orthonormal basis:

$$|\ell, k\rangle = \sqrt{\frac{2^{\ell-k}(\ell-k)!}{(2\ell)!(\ell-k)!}} L_-^{\ell-k} v, \quad k = -\ell, -\ell+1, \ldots, \ell.$$

Example 5.5.1. If $\ell = 1/2$ we have the orthonormal basis

$$\left|\frac{1}{2}, \frac{1}{2}\right\rangle, \quad \left|\frac{1}{2}, -\frac{1}{2}\right\rangle.$$

with

$$L_+ \left|\frac{1}{2}, \frac{1}{2}\right\rangle = 0, \quad L_+ \left|\frac{1}{2}, -\frac{1}{2}\right\rangle = \frac{1}{\sqrt{2}} \left|\frac{1}{2}, \frac{1}{2}\right\rangle$$
$$L_- \left|\frac{1}{2}, -\frac{1}{2}\right\rangle = 0, \quad L_- \left|\frac{1}{2}, \frac{1}{2}\right\rangle = \frac{1}{\sqrt{2}} \left|\frac{1}{2}, -\frac{1}{2}\right\rangle.$$

The irreducible representations of $SO(3)$ are in bijective correspondence with the irreducible representations V^n of $SU(2)$ in which $-I$ acts as the identity. These are precisely the representations with integer spin. The other representations of $SU(2)$ in which $-I$ does not act as the identity only define projective representations of $SO(3)$. The representation (projective or exact) is defined by

$$gv = \pi^{-1}(g)v, \quad g \in SO(3).$$

By definition, a particle with integer spin is called a *boson* and a particle with half-integer spin is called a *fermion*.

Based on experiments, physicists have proposed a modification of the Hilbert space $L^2(\mathbb{R}^3)$ so as to include internal degrees of freedom. The idea is that for each particle with "spin" ℓ, the Hilbert space should be

$$L^2(\mathbb{R}^3)\hat{\otimes}V_\ell,$$

where V_ℓ is an irreducible projective representation of $SO(3)$ of dimension $2\ell + 1$ (which we dissambiguate as an irreducible representation of $SU(2)$). Here the hat denotes the Hilbert space tensor product.

54. Let us open a parenthesis and discuss the situation of *composite systems*. Based on the fact that for two 3-dimensional particles the Hilbert space is $L^2(\mathbb{R}^6) = L^2(\mathbb{R}^3)\hat{\otimes}L^2(\mathbb{R}^3)$, we introduce the following axiom of quantum mechanics.

The Hilbert space of a composite system made up of two subsystems is the Hilbert tensor product of the Hilbert spaces describing the two systems.

We recall that the Hilbert tensor product of $\mathcal{H}_1$ and $\mathcal{H}_2$, $\mathcal{H}_1\hat{\otimes}\mathcal{H}_2$ is obtained by endowing $\mathcal{H}_1 \otimes \mathcal{H}_2 = \text{Span}\{x_1 \otimes x_2 \,|\, x_1 \in \mathcal{H}_1, x_2 \in \mathcal{H}_2\}$ with the inner product

$$\langle x_1 \otimes x_2, y_1 \otimes y_2 \rangle = \langle x_1, y_1 \rangle \langle x_2, y_2 \rangle$$

and then completing it in the induced norm topology.

The quantum Hamiltonian for a non-interacting composite system is

$$\hat{H} = \hat{H}_1 \otimes I + I \otimes \hat{H}_2$$

where $\hat{H}_1, \hat{H}_2$ are the quantum Hamiltonian of the two systems. This is because of the product rule in the differentiation of the tensor product

$$\begin{aligned} i\hbar\frac{d}{dt}(\psi_1 \otimes \psi_2) &= i\hbar\frac{d\psi_1}{dt} \otimes \psi_2 + i\hbar\psi_1 \otimes \frac{d\psi_2}{dt} \\ &= \hat{H}_1\psi_1 \otimes \psi_2 + \psi_1 \otimes \hat{H}_2\psi_2 = \hat{H}(\psi_1 \otimes \psi_2). \end{aligned}$$

Let us return to the case of a particle with spin ℓ. Its state is $\psi(\mathbf{x}) \otimes v \in L^2(\mathbb{R}^3)\hat{\otimes}V_\ell$. The system consisting of two such particles has the Hilbert space equal to

$$L^2(\mathbb{R}^6)\hat{\otimes}(V_\ell \otimes V_\ell).$$

Now one of the postulates of quantum mechanics is that *identical particles are indistinguishable*. This means that the state associated to the system formed by the first and the second particle should coincide, up to multiplication by a constant, with the system formed by the second and the first particle, that is

$$\psi(\mathbf{x}, \mathbf{y}) = c\psi(\mathbf{y}, \mathbf{x}).$$

Applying this reasoning twice, we see that $c^2 = 1$, so $c = \pm 1$.

Experimental considerations suggest that $c = (-1)^{2\ell}$. So $c = 1$ for bosons and $c = -1$ for fermions. So we have the following axiom of quantum physics:

Axiom: Consider a collection of N identical particles moving in $\mathbb{R}^3$ and having integer spin ℓ. Then the Hilbert space of such a collection is the subspace of $L^2(\mathbb{R}^{3N})\hat{\otimes}(V_\ell)^{\otimes N}$ consisting of those functions ψ that satisfy

$$\psi(\mathbf{x}_{\sigma(1)}, \mathbf{x}_{\sigma(2)}, \ldots, \mathbf{x}_{\sigma(N)}) = \psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$$

for every permutation σ. Consider a collection of N identical particles moving in $\mathbb{R}^3$ and having half-integer spin ℓ. Then the Hilbert space of such a collection is the subspace of $L^2(\mathbb{R}^{3N})\hat{\otimes}(V_\ell)^{\otimes N}$ consisting of those functions ψ that satisfy

$$\psi(\mathbf{x}_{\sigma(1)}, \mathbf{x}_{\sigma(2)}, \ldots, \mathbf{x}_{\sigma(N)}) = \mathrm{sign}(\sigma)\psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$$

for every permutation σ.

55. This now leads to the natural question of decomposing $V_\ell \otimes V_m$ into sums of irreducible representations.

Theorem 5.5.2. (Clebsch-Gordan Theorem) For any ℓ and m,

$$V_\ell \otimes V_m = \bigoplus_{j=|\ell-m|}^{\ell+m} V_j = V_{|\ell-m|} \oplus V_{|\ell-m|+1} \oplus \cdots \oplus V_{\ell+m-1} \oplus V_{\ell+m}.$$

Proof. Here are two sketches of proofs:
1. We argue on the particular example $V_1 \otimes V_{5/2}$, in which case we represent $e_j \otimes e_k = |1, j\rangle \otimes |5/2, k\rangle$ as the nodes of the diagram in the figure. Examining the eigenvectors of L_3 in $V_1 \otimes V_{5/2}$ we see that they are the subspaces that are spanned by the nodes of the diagonal

lines drawn in the figure. It is now not hard to decompose $V_1 \otimes V_{5/2}$ as the direct sum of irreducibles: there is one that runs from the lower-left corner to the upper-right corner which is isomorphic to $V_{7/2}$ (note that it has dimension 8, since there are 7 steps to take you from the lower-left to the upper right). You can also compute its highest weight by hand and see that it is 7/2. The next "diagonal" has dimension 2, so there will be a vector that is left out and that one will determine a 6-dimensional irreducible representation, with highest weight 5/2, so we get a copy of $V_{5/2}$. Then we look at the next "diagonal", and we get the $V_{3/2}$. So

$$V_1 \otimes V_{5/2} = V_{3/2} \otimes V_{5/2} \otimes V_{7/2},$$

as claimed.

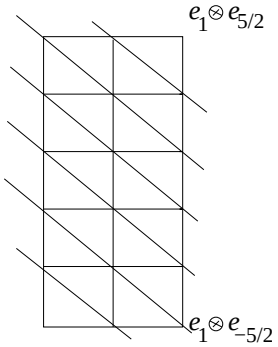

Figure 5.1:

2. Note that every element in $SU(2)$ is conjugate to an element of the form

$$e(t) = \begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix}.$$

Now we use the notion of a character of a representation, which is the trace of the representation. This is a function on the Lie group, or rather on the conjugacy classes of elements of the Lie group. In our case, each representation V will induce the character $\chi_V(t) = \text{trace}e(t)$. In the case of the representation V_ℓ, the value of this character is

$$\chi_\ell = \sum_{k=0}^{2\ell} e^{2i(\ell-k)t}.$$

Now characters are multiplicative with respect to the tensor product and additive with respect to sum. And they determine representations. Thus to check the Clebsch-Gordan theorem, it suffices to check the equality of characters, and this follows from

$$\left(\sum_{k=0}^{2\ell} e^{2i(\ell-k)}\right)\left(\sum_{n=0}^{2m} e^{2i(m-n)}\right) = \sum_{j=|\ell-m|}^{\ell+m}\left(\sum_{t=0}^{2j} e^{2i(j-t)}\right).$$

□

Chapter 6

Quantum Mechanics on Manifolds

Now let us assume that we have a classical mechanical system with constraints, and that we want to quantize it. As seen in the first chapter, the phase space is the cotangent bundle of the manifold which is the configuration space.

6.1 The prequantization line bundle and geometric quantization

56. So we want to apply the quantization scheme to the cotangent line bundle, or more generally to a symplectic manifold (M, ω).

Thus we want

$$\begin{array}{ccc} (M,\omega) & \longrightarrow & \mathcal{H} \\ C^\infty(M) & \longrightarrow & L(\mathcal{H}). \end{array}$$

so that

$$\mathrm{op}(\{f,g\}) = \frac{1}{i\hbar}[\mathrm{op}(f), \mathrm{op}(g)].$$

To make this condition work we can set $\mathrm{op}(f) = -i\hbar X_f$, because $[X_f, X_g] = -X_{\{f,g\}}$. But then $\mathrm{op}(1) = 0$, which conflicts with the other quantization conditions. Then we could set $\mathrm{op}(f) = i\hbar X_f + f$, but then the above quantization condition fails.

Now comes the geometric trick: If ω, the symplectic form, has what is called a symplectic potential θ, meaning that

$$d\theta = \omega,$$

then if we set

$$\mathrm{op}(f) = -i\hbar\left(X_f - \frac{i}{\hbar}\theta(X_f)\right) + f,$$

then this works! This formula is referred to as *geometric quantization* of observables.

Example 6.1.1. In the case of a free 1-dimensional particle,

$$X_q = -\frac{\partial}{\partial p}, \quad X_p = \frac{\partial}{\partial q}.$$

Then $\theta = -qdp$, so

$$\mathrm{op}(q)\phi = -i\hbar\left(-\frac{\partial\phi}{\partial p} + \frac{i}{\hbar}q\phi\right) + q\psi = i\hbar\frac{\partial\phi}{\partial p}$$
$$\mathrm{op}(p)\phi = -i\hbar\frac{\partial\phi}{\partial q} + p\phi.$$

Wow, what is this? In fact you get the standard Schrödinger representation if you set $\phi(q,p) = e^{-(i/\hbar)qp}\psi(q)$. Then

$$\mathrm{op}(q)\psi = q\psi, \quad \mathrm{op}(p)\psi = -i\hbar\frac{\partial\psi}{\partial q}.$$

57. Two problems arise in this example. The first is that the state ϕ depends on both p and q. In fact, for a general manifold we cannot separate the variables, and so we are forced to work with functions of both p and q. Well, not quite, as we will see below there is an elegant way to separate variables, and thus to restrict ourselves to only some functions, in this case to $\phi = \psi(q)e^{-(i/\hbar)pq}$, with $\psi \in L^2(\mathbb{R})$.

But there is a second problem, which is more subtle. This comes from the fact that while every symplectic form is locally of the form $\sum_j dp_j \wedge dq_j$, and so locally we can find the potential $\sum_j p_j dq_j$, or $-\sum_j q_j dp_j$, this does not work globally. We only have a family of local potentials, and we have to patch those together. We do this by means of a *line bundle*.

Definition. A line bundle over a manifold M is a manifold L and a map $\pi : L \to M$ such that there exists a cover of M by open sets $(U_j)_j$, and homomorphisms $\phi_j : \pi^{-1}(U_j) \to U_j \times \mathbb{C}$ such that $\pi_j(\phi_j(x)) = \pi(x)$ for every $x \in L$ where $\pi_j : U_j \times \mathbb{C} \to U_j$ is $\pi_j(y,z) = y$. Moreover, the transition functions $\phi_j \circ \phi_k^{-1}$ should be of the form $(y,z) \mapsto (y, t_{jk}(x)z)$ with $t_{jk}(x)$ a linear map depending smoothly on x.

Here is another way to say this definition:

Definition. A (complex) *line bundle* $\mathcal{L}$ on a manifold M is defined by a cover $(U_j)_{j\in J}$ of M by contractible open sets such that $U_j \cap U_k$ is either empty or contractible for all j and k, and for every pair (j,k) with $U_j \cap U_k \neq \emptyset$ a smooth map

$$c_{jk} : U_j \cap U_k \to \mathbb{C}\backslash\{0\}.$$

The maps c_{jk} should satisfy the conditions

$$c_{kj} = (c_{jk})^{-1} \text{ and } c_{jk}c_{kl}c_{lj} = 1, \text{ for all } j,k,l \in J. \tag{6.1.1}$$

The line bundle itself is the quotient of the disjoint union of the sets $U_j \times \mathbb{C}$ by the equivalence relation which identifies $(x,z) \in U_j \times \mathbb{C}$ with $(x, c_{jk}z)$ for all $x \in U_j \cap U_k$, $z \in \mathbb{C}$.

The conditions (6.1.1) mean that c_{jk}, $j,k \in J$, is a *Čech cocycle*, more precisely a Čech 2-cocycle. If we consider a family of functions

$$d_j : U_j \to \mathbb{C}\backslash\{0\},$$

then the cocycle $d_j c_{jk}$ defines an equivalent line bundle.

Cocycles form an abelian group under the multiplication,

$$((c_{jk}),(c'_{jk})) \mapsto (c_{jk}c'_{jk})$$

which we denote by $\check{Z}^2((U_j),\mathbb{C})$. The tensor product of two line bundles over the same manifold is obtained by considering a collection of charts common to both and the associated cocycles. Then the cocycle of the tensor product line bundle is the product of the cocycles of the two line bundles.

Let $\check{B}^2((U_j),\mathbb{C})$ be the subgroup consisting of cocycles of the form $d_j d_k^{-1}$, $j,k \in J$, where the functions d_j are as above. If the quotient of two cocycles is in $\check{B}^2((U_j),\mathbb{C})$ then they define the same line bundle. In fact, it can be proved that this is a necessary and sufficient condition.

The quotient group

$$\check{H}^2((U_j),\mathbb{C}) = \check{Z}^2((U_j),\mathbb{C})/\check{B}^2((U_j),\mathbb{C})$$

is called the second *Čech cohomology group*. Each element of $\check{H}^2((U_j),\mathbb{C})$ defines a line bundle up to equivalence.

In this case we want to construct a line bundle with curvature $(1/\hbar)\omega$. We should point out that there is an obstruction to the existence of such a line bundle, known as the *Weil integrality condition*. This requires that

$$\frac{1}{2\pi}\omega \in H^2(M,\mathbb{Z}).$$

This means that when you integrate ω over every closed oriented surface the result is an integral multiple of 2π.

Here is the standard construction: The form ω being symplectic is closed, and so on each open set U_j it is exact, by Poincaré's Lemma. Hence there are real 1-forms θ_j on U_j called potentials, such that

$$d\theta_j = \omega.$$

We have $d(\theta_j - \theta_k) = 0$ on $U_j \cap U_k$, so by using again Poincaré's Lemma on the contractible domains $U_j \cap U_k$, we deduce that there are smooth functions

$$f_{jk} : U_j \cap U_k \to \mathbb{R}$$

such that $df_{jk} = \theta_j - \theta_k$.

Assume that

$$c_{jk} = e^{if_{jk}} \tag{6.1.2}$$

satisfies the cocycle condition (6.1.1). Then it defines a line bundle $\mathcal{L}$. The line bundle $\mathcal{L}$ is said to have curvature ω.

The 1-forms $-i\theta_j$ are the local expressions of a connection form on M. Indeed, by differentiating the relation (6.1.2) we obtain

$$dc_{jk} = ie^{if_{jk}}df_{jk} = ic_{jk}(\theta_j - \theta_k).$$

This can be rewritten as

$$-i\theta_k = c_{jk}^{-1}dc_{jk} + c_{jk}^{-1}(-i\theta_j)c_{jk},$$

and we recognize the formula expressing the change of the connection form under changes of coordinates.

The line bundle constructed above is called the *prequantization line bundle*. The Hilbert space will consist of sections of this line bundle.

Definition. A *section* of the line bundle L is a map $s : M \to L$ such that $\pi \circ s = 1_M$.

The operators defined by geometric quantization now act on smooth L^2 sections of L. There is however a problem with this, in that the Hilbert space is too big. For example for the case of a free 1-dimensional particle we get $L^2(\mathbb{R}^2)$ and not $L^2(\mathbb{R})$. We will learn in the next section how to pick the right sections.

6.2 Polarizations

Three examples should clarify the general situation. One is the standard example, where we work on $L^2(\mathbb{R}^n)$, with the variable $\mathbf{q}$, and have

$$\mathrm{op}(q_j) = M_{q_j}, \quad \mathrm{op}(p_j) = -i\hbar\frac{\partial}{\partial q_j}.$$

The second example corresponds to working in the momentum representation. This time $L^2(\mathbb{R}^n)$, with variable $\mathbf{p}$, and

$$\mathrm{op}(q_j) = i\hbar\frac{\partial}{\partial p_j}, \quad \mathrm{op}(p_j) = M_{p_j}.$$

The third situation was encountered when we studied the harmonic oscillator, where the Hilbert space was a space of harmonic functions. What these situations had in common was that the functions that were states could be characterized by the conditions that they are annihilated by some differential operators: in the first case $\frac{\partial}{\partial p_j}\psi = 0$, in the second case $\frac{\partial}{\partial q_j}\psi = 0$, while in the third case $\frac{\partial}{\partial \bar{z}}\psi = 0$.

We now bring these conditions in a general unified framework, which works on manifolds. Let (M, ω) be a symplectic manifold of dimension $2n$. The tangent space $T_xM = \mathbb{R}^{2n}$ is itself endowed with the symplectic form ω. For a subspace W of T_xM, we define the orthogonal

$$W^\perp = \{v \in T_x \,|\, \omega(v, w) = 0 \text{ for all } w \in W\}.$$

The subspace W is called isotropic if $W \subset W^\perp$ and Lagrangian if it is maximal isotropic. Alternatively, Lagrangian subspaces are isotropic subspaces of dimension n. $\mathbf{L}$ is Lagrangian if and only if $\mathbf{L}^\perp = \mathbf{L}$.

We complexify the tangent space to $T_xM \otimes \mathbb{C} = \mathbb{C}^{2n}$, by placing complex coefficients in front of the basis vectors. The form ω extends to a symplectic form on $T_xM \otimes \mathbb{C}$. The notion of Lagrangian subspace extends to the complex situation as well, by considering complex subspaces.

Definition. A *polarization of a symplectic real vector space* (V, ω) is a Lagrangian subspace $\mathbf{L}$ of $V \otimes \mathbb{C}$.

A *complex distribution* $\mathbf{F}$ on M is a subbundle of the complexification of the tangent bundle of M. In that sense, at each point $x \in M$, the distribution associates a subspace of $T_xM \otimes \mathbb{C}$, and these subspaces vary smoothly with p and have all the same dimension. A polarization is called involutive if the Poisson bracket of two vector fields in $\mathbf{F}$ is also in $\mathbf{F}$.

A complex distribution $\mathbf{F}$ is called *Lagrangian* if

$$\dim_{\mathbb{C}} \mathbf{F} = \frac{1}{2} \dim_{\mathbb{R}} M \text{ and } \omega|\mathbf{F} \times \mathbf{F} = 0.$$

Definition. A *polarization of the symplectic manifold* (M, ω) is a complex Lagrangian involutive distribution $\mathbf{F}$ such that the function $p \mapsto \dim(\mathbf{F}_p \cap \overline{\mathbf{F}}_p)$ is constant on M, $\overline{\mathbf{F}}$ being the complex conjugate of $\mathbf{F}$.

Example 6.2.1. $M = \mathbb{R}^{2n}$ with the standard symplectic form. Then the distribution on M,

$$F = \text{Span}\left(\frac{\partial}{\partial p_1}, \frac{\partial}{\partial p_2}, \ldots, \frac{\partial}{\partial p_n}\right)$$

is a polarization.

Example 6.2.2. $M = \mathbb{R}^{2n}$ with the standard symplectic form. Then

$$F = \text{Span}\left(\frac{\partial}{\partial q_1}, \frac{\partial}{\partial q_2}, \ldots, \frac{\partial}{\partial q_n}\right)$$

is a polarization.

Example 6.2.3. $M = \mathbb{R}^2$ with the standard symplectic form. Then

$$F = \text{Span}\left(\frac{\partial}{\partial \bar{z}}\right)$$

is a polarization.

Now let us return to the prequantization line bundle L over the phase space M. This line bundle has a connection

$$\nabla = d - i\theta,$$

which is a differential operator that acts on sections. Locally sections are functions, and d is just the operator that turns a function into its differential. The operator d is not independent of the coordinates, and thus it is not globally defined. By adding the "correction" term $-i\theta$ we obtain a globally defined operator, meaning that the new differential operator behaves well under changes of coordinates.

Definition. The Hilbert space of the quantization consists of those sections of $\mathcal{L}$ that are *covariantly constant* in the direction of $\mathbf{F}$, meaning that

$$\nabla_{\mathbf{v}} s = 0, \text{ for all } \mathbf{v} \in \mathbf{F}.$$

Here $\nabla_{\mathbf{v}}$ is the directional derivative in the direction of $\mathbf{v}$.

Example 6.2.4. Let $M = \mathbb{R}^{2n}$ with the standard symplectic form and with polarization

$$F = \text{Span}\left(\frac{\partial}{\partial p_1}, \frac{\partial}{\partial p_2}, \ldots, \frac{\partial}{\partial p_n}\right).$$

Let

$$\theta = -(1/\hbar)\sum_{j=1}^{n} q_j dp_j$$

so that $d\theta = (1/\hbar)\omega$. Then the covariantly constant sections $s \in L^2(\mathbb{R}^{2n})$ are those that satisfy

$$\frac{\partial}{\partial p_j} s + (i/\hbar) q_j s = 0, \quad j = 1, 2, \ldots, n.$$

Solving the differential equation we obtain

$$s(\mathbf{q}, \mathbf{p}) = \psi(\mathbf{q}) e^{-(i/\hbar)\mathbf{p}^T \mathbf{q}}.$$

This should be compared with Example 6.1.1.

Example 6.2.5. Set $z_j = q_j - ip_j$. We consider the polarization

$$F = \text{Span}\left(\frac{\partial}{\partial \bar{z}_1}, \frac{\partial}{\partial \bar{z}_2}, \frac{\partial}{\partial \bar{z}_n}\right),$$

where

$$\frac{\partial}{\partial \bar{z}_j} = \frac{1}{2}\left(\frac{\partial}{\partial q_j} + i\frac{\partial}{\partial p_j}\right).$$

Take also

$$\theta = (1/\hbar)\sum_j p_j dq_j.$$

Then

$$\theta\left(\frac{\partial}{\partial \bar{z}_j}\right) = \frac{p_j}{2}.$$

So

$$\nabla_{\partial/\partial \bar{z}_j} = \frac{\partial}{\partial \bar{z}_j} - \frac{i}{2\hbar} p_j.$$

Thus the sections of the quantization line bundle (which are just functions over $\mathbb{R}^{2n} = \mathbb{C}^n$) should satisfy the system of differential equations

$$\frac{\partial \psi}{\partial \bar{z}_j} - \frac{i}{2\hbar} p_j \psi = 0, \quad j = 1, 2, \ldots, n.$$

If we set

$$\psi(\mathbf{z}, \bar{\mathbf{z}}) = \phi(\mathbf{z}, \bar{\mathbf{z}}) e^{-\|Im\, \mathbf{z}\|^2/(2\hbar)},$$

then an easy computation shows that

$$\nabla_{\partial/\partial \bar{z}_j} \phi = e^{-\|Im\, \mathbf{z}\|^2/(2\hbar)} \frac{\partial \phi}{\partial \bar{z}_j}.$$

Thus $\nabla_{\partial/\partial \bar{z}_j} = 0$ for all j if and only if ϕ is a holomorphic function in the variables $z_j = q_j - ip_j$. If we require that the Hilbert space consists of functions that are square integrable over $\mathbb{R}^{2n}$ then ϕ should be in the Segal-Bargmann space of holomorphic square integrable functions on $\mathbb{C}^n$ with respect to the measure

$$e^{-\|Im\, \mathbf{z}\|^2/(2\hbar)} d\mathbf{p} d\mathbf{q}.$$

The argument for placing a Hilbert space structure on the space of sections that we used in the second example cannot be applied to the first example. So we need to be more "clever" when we choose the Hilbert space structure. Moreover, experiments show however that the line bundle constructed above is not quite right.

To solve both problems, there is a procedure called metalinear (or metaplectic) correction, which when applied to the prequantization line bundle yields a new line bundle, and that is the correct line bundle to work with. In short, the idea is the following.

First, for a point $x \in M$, we let $(v_1, v_2, \ldots, v_n)$ be a basis of F_x, called a frame at x. When changing from one coordinate system to another, specifically from U_j to U_k, the frame changes by multiplication by an element $C_{jk}(x) \in GL(n, \mathbb{C})$. The elements C_{jk} satisfy the cocycle condition that we have seen before, but they are matrices. Now consider the cocycle $c_{jk} = \det C_{jk}$. This cocycle defines the line bundle $\Lambda^n F$. Assume that we can find a line bundle L' over M such that $L' \otimes L' = \Lambda^n F$. Then we replace L by $L \otimes L'$ and consider covariantly constant sections of this new line bundle.

Now what happens is that there is an inner product at each point such that

$$d \langle s, s' \rangle_x = \langle \nabla s, s' \rangle_x + \langle s, \nabla s' \rangle_x .$$

This inner product is constant on D where the complexification of D is $F \cap \overline{F}$. It is natural to define $\langle s, s' \rangle = \int M \langle s, s' \rangle_x \omega^n$, but this diverges. So instead we write $\langle s, s' \rangle_x \omega_x^n$ as $\langle s, s' \rangle_x \alpha \wedge \beta$ where α is the "volume form" of $\Lambda^n F$ and only compute $\int_{M/D} \langle s, s' \rangle x\beta$. This decomposition of the integrand is clearly possible locally, and it is precisely because of the alteration of the line bundle that we can decompose it globally.

Example 6.2.6. Let us consider the case of a free one-dimensional particle, whose phase space is $\mathbb{R}^2$. If we take $\theta = pdq$, and the polarization $\partial/\partial p$, then $\theta(\partial/\partial p) = 0$, so

$$\nabla_{\partial/\partial p} \psi = \frac{\partial \psi}{\partial p} = 0,$$

so ψ is a function of q only. Using the formula for geometric quantization we compute

$$\mathrm{op}(q)\psi = q\psi, \quad \mathrm{op}(p)\psi = -i\hbar\frac{\partial\psi}{\partial p}.$$

This gives exactly the formulas predicted by the Schrödinger representation of the Heisenberg Lie algebra. So we should be very happy. But...

$$\mathrm{op}(p^2) = -2i\hbar p\frac{\partial}{\partial q} - p^2.$$

And this introduces the variable p, thus it does not act on $L^2(\mathbb{R})$ with the variable q. This shows the limitations of geometric quantization. Even for the simplest example of a one-dimesional free particle you cannot quantize quadratic polynomials (which we did so easily in Chapter 2). Thus we cannot even quantize kinetic energy.

We conclude this chapter with the observation that we have on the one hand the Weyl quantization, which can only be performed when the phase space is $\mathbb{R}^{2n}$ with the standard symplectic form, but which can be applied to a large variety of observables. On the other hand we have a method of quantizing geometrically every phase space, but this method is very restrictive when it comes to decide which observables (Hamiltonians) can be quantized.