# From Classical Mechanics to Quantum Mechanics

Richard Froese

Department of Mathematics
University of British Columbia

May 4, 2018

These are the notes for five one hour lectures delivered at the 2016 CRM summer school in Spectral Theory at Laval University. They were intended to provide some general background on classical and quantum mechanics to advanced undergraduate and beginning graduate students in mathematics. The first four lectures contain an outline of these two theories from a mathematical point view and a comparison of the classical and quantum descriptions of some simple systems. The last lecture is about hidden variables and is meant to illustrate the essential strangeness of the quantum description of nature. I have tried to make these notes more accessible by including an outline of some basic ideas in measure theory and operator theory. A similar outline of the basics of manifolds, tangent and cotangent spaces and differential forms would have been useful for understanding the sections on Lagrangian submanifolds and Hamilton-Jacobi equations. Unfortunately I did not have time to include this. I am grateful to Cyrus Bhiladvala and the referee for making suggestions for improvement.

## Table of Contents

**1. Classical Mechanics**

Newton's equations and Hamilton's equations ..... 2
Example 1: Free motion ..... 4
Example 2: Harmonic Oscillator ..... 5
Example 3: Two bump potential $n = 1$ ..... 6
Symplectic form, Poisson bracket and time evolution of observables ..... 7
Symplectic flow and Poincaré recurrence ..... 10
Summary of Classical Mechanics (so far) ..... 12
The Hamilton Jacobi Equation ..... 12
Lagrangian submanifolds ..... 13
Legendre transforms ..... 15
Flow outs ..... 15

Extended phase space and the geometry of the Hamilton Jacobi equations ..... 16
Remark on solving for the flow using the Hamilton Jacobi equation ..... 18
An example ..... 19

**2. Review of Probability and Operator theory**

Probability ..... 21
Two theorems in measure theory ..... 23
Mixed states in Classical Mechanics ..... 24
Background and self-adjoint operators ..... 24
Spectral theorem for self-adjoint operators ..... 25
The Fourier transform ..... 26
Uncertainty principle ..... 27
One parameter strongly continuous unitary groups and Stone's theorem ..... 27

**3. Quantum Mechanics**

Abstract Quantum description of a physical system ..... 29
Quantum particle in an external potential ..... 30
Example 1: Free motion ..... 32
Using the Hamilton Jacobi equation ..... 34
Example 2: The Harmonic Oscillator ..... 37
Example 3: The two bump potential ..... 41

**4. Hidden variables and non-locality**

Introduction ..... 43
Spin observables $A$ and $B$ ..... 44
Tensor Product ..... 46
Hardy's example ..... 46
Bell's second theorem: non-locality ..... 48

**References** ..... 51

## 1. Classical Mechanics

### Newton's equations and Hamilton's equations

We begin by showing how Classical Mechanics describes the motion of a single particle of mass $m$ moving in configuration space $\mathbb{R}^n$ under the influence of an external conservative force

field. The force felt by the particle at the point $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ is $F(x) = -\nabla V(x)$, where the potential $V$ is a real valued function. The motion is described by the trajectory $x(t)$, where $t \in \mathbb{R}$ is time, and $x(t)$ is the position at time $t$. According to Newton's law $F = m\mathbf{a}$, the trajectory satisfies the equation

$$m\ddot{x} = -\nabla V(x). \tag{1.1}$$

This is a system of $n$ second order ODEs so under suitable conditions on $V$, there is a unique solution $x(t)$ if we impose initial conditions

$$\begin{aligned} x(0) &= x_0 \\ \dot{x}(0) &= v_0. \end{aligned} \tag{1.2}$$

There is a standard trick for turning a system of $n$ second order equations into an equivalent system of $2n$ first order equations. If $x(t)$ solves Newton's equation (1.1) with initial conditions (1.2), then, defining the momentum as $p(t) = m\dot{x}(t)$, we find that $(x(t), p(t)) \in \mathbb{R}^n \times \mathbb{R}^n$ satisfies Hamilton's equations

$$\begin{aligned} \dot{x} &= \frac{1}{m}p \\ \dot{p} &= m\ddot{x} = -\nabla V(x), \end{aligned} \tag{1.3}$$

with initial conditions

$$\begin{aligned} x(0) &= x_0 \\ p(0) &= mv_0 = p_0. \end{aligned} \tag{1.4}$$

On the other hand, if $(x(t), p(t))$ solves Hamilton's equations (1.3) with initial condition (1.4), then $x(t)$ solves Newton's equation (1.1) with initial conditions (1.2).

Hamilton's equations are a first order system in phase space $\mathbb{R}^n \times \mathbb{R}^n$. The right side of (1.3) defines a vector field on phase space, namely

$$(x, p) \mapsto \begin{bmatrix} \frac{1}{m}p \\ -\nabla V(x) \end{bmatrix}.$$

The trajectories $(x(t), p(t))$ are the corresponding flow.

Although it is not obvious, the vector field on the right of (1.3) is very special. Let $H(x, p)$ be the real valued function on phase space defined by

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

Here $p^2 = \langle p, p \rangle$ is the square of the standard Euclidean norm of $p$ in $\mathbb{R}^n$. The function $H(x, p)$, called the Hamiltonian, represents the total energy (kinetic plus potential) of the particle. Written in terms of $H$, Hamilton's equations have the form

$$\begin{aligned} \dot{x} &= \frac{\partial H}{\partial p}(x, p) \\ \dot{p} &= -\frac{\partial H}{\partial x}(x, p), \end{aligned} \tag{1.5}$$

where

$$\frac{\partial H}{\partial x} = \begin{bmatrix} \frac{\partial H}{\partial x_1} \\ \vdots \\ \frac{\partial H}{\partial x_n} \end{bmatrix}, \quad \frac{\partial H}{\partial p} = \begin{bmatrix} \frac{\partial H}{\partial p_1} \\ \vdots \\ \frac{\partial H}{\partial x_p} \end{bmatrix}$$

denote the gradients with respect to the position variables $x_1, \ldots, x_n$ and momentum variables $p_1, \ldots, p_n$ respectively.

To get a feeling for Hamilton's equations let's look at three simple examples. For a Hamiltonian of the form $H(x,p) = \frac{p^2}{2m} + V(x)$ we can get a mental image of the motion by imagining a particle sliding without friction on a hill whose height at the point $x$ is $V(x)$.

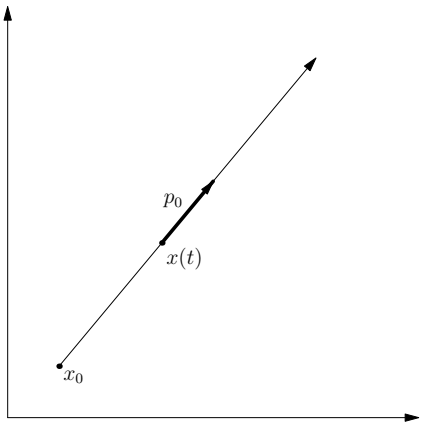
In these examples we will pay attention to the presence of bound states and scattering states. Bound states are trajectories for which $x(t)$ remains bounded for all time, while scattering states are those where $|x(t)|$ tends to infinity when $t \to \pm\infty$.

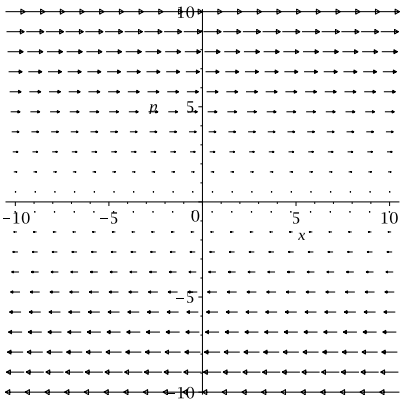
### Example 1: Free motion

If the force $F = 0$ then the motion is free. This happens if $V(x) = V_0$ is constant. In this case Hamilton's equations are

$$\dot{x} = \frac{1}{m}p$$
$$\dot{p} = 0.$$

The second equation says that the momentum $p(t) = p_0$ is constant. Then $x(t) = x_0 + \frac{1}{m}p_0 t$. So the trajectories in configuration space are straight lines in the direction of the momentum.




Trajectory in position space $n = 2$ Vector field in phase space $n = 1$

In this example, states with $p_0 \neq 0$ move with uniform velocity. Clearly $|x(t)| \to \infty$ for these states, so they are scattering states. States with $p_0 = 0$ are bound states that stay fixed at $x_0$ for all time.

**Example 2: Harmonic Oscillator**

The harmonic oscillator potential in $n$ dimensions is given by

$$V(x) = \frac{m}{2}\big(\omega_1^2 x_1^2 + \cdots + \omega_n^2 x_n^2\big).$$

where $\omega_1, \ldots, \omega_n$ are positive constants. Hamilton's equations for the harmonic oscillator are

$$\begin{aligned} \dot{x}_i &= \frac{1}{m} p_i \\ \dot{p}_i &= -m\omega_i^2 x_i. \end{aligned}$$

for $i = 1, \ldots, n$. For simplicity let's take $m = \omega_i = 1$. Then we have $n$ independent $2 \times 2$ linear systems

$$\begin{bmatrix} \dot{x}_i \\ \dot{p}_i \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_i \\ p_i \end{bmatrix}$$

whose solutions are given by clockwise rotations in phase space

$$\begin{bmatrix} x_i \\ p_i \end{bmatrix} = \exp\left( t \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right) \begin{bmatrix} x_{i,0} \\ p_{i,0} \end{bmatrix} = \begin{bmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{bmatrix} \begin{bmatrix} x_{i,0} \\ p_{i,0} \end{bmatrix}.$$

Another way to solve this system is to consider the complex functions

$$a_k = \frac{1}{\sqrt{2}} \left( x_k + i p_k \right)$$

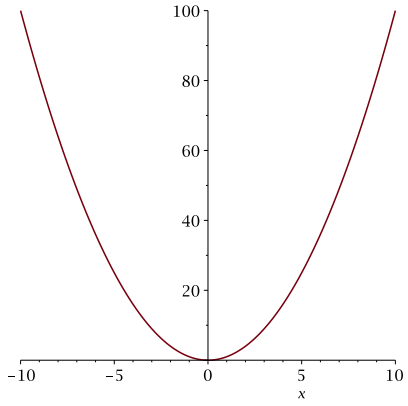
for $k = 1, \ldots, n$. Then

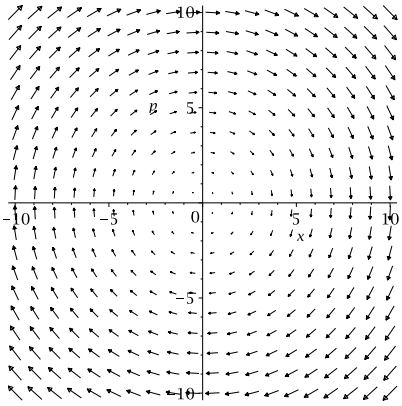
$$\dot{a}_k = \frac{1}{\sqrt{2}} \left( \dot{x}_k + i \dot{p}_k \right) = \frac{1}{\sqrt{2}} \left( p_k - i x_k \right) = -i a_k.$$

Thus $a_k(t) = e^{-it} a_{k,0}$ for some initial conditions $a_{k,0}$, from which we can recover $x_k(t)$ and $p_k(t)$ by taking the real and imaginary parts. So if $a_{k,0} = A_k e^{i\alpha_k}$ then

$$\begin{aligned} x_k(t) &= \quad A_k \cos(t - \alpha_k) \\ p_k(t) &= -A_k \sin(t - \alpha_k). \end{aligned}$$

Solutions are circles in phase space traversed counter-clockwise. All trajectories are bound states.



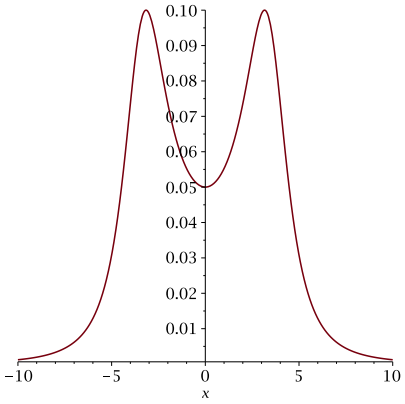




Harmonic oscillator potential $n = 1$ Harmonic oscillator vector field in phase space

*Problem 1.1:* What are the solutions for general $m$ and $\omega_i$?

**Example 3: Two bump potential** $n = 1$

Now consider a potential $V(x)$ that looks like this:

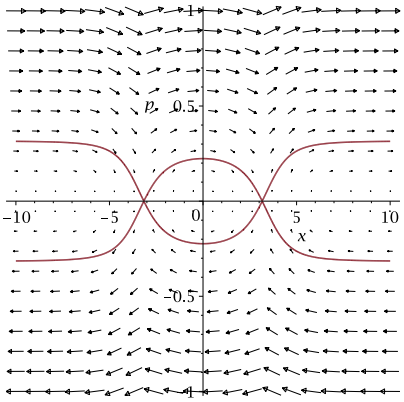



Two bump potential

Hamilton's equations are

$$\dot{x} = \frac{1}{m}p$$
$$\dot{p} = -V'(x).$$

Although we cannot write down the exact solution, we can get a qualitative idea of the motion by sketching the Hamiltonian vector field. Here is a plot, with some exceptional solution curves




Hamiltonian vector field for two bump potential

*Problem 1.2:* Which states are bound states, which are scattering states, and which are neither?

### Symplectic form, Poisson bracket and time evolution of observables

The vector field $X_H$ on the right side of Hamilton's equations (1.5) has the form

$$X_H = \begin{bmatrix} \dfrac{\partial H}{\partial p} \\ -\dfrac{\partial H}{\partial x} \end{bmatrix} = J\nabla H$$

where

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}, \quad \nabla H = \begin{bmatrix} \dfrac{\partial H}{\partial x} \\ \dfrac{\partial H}{\partial p} \end{bmatrix}.$$

Replacing $H$ in this construction with any smooth function $f$ on phase space, we can produce vector fields $X_f$. These will define a corresponding flow. But the Hamiltonian function $H$ is special because the flow it generates is the time evolution.

In classical mechanics, (sufficiently nice) functions $f(x, p)$ on phase space are called *observables*. They correspond to measurable quantities. Here are some examples.

| Observable | $f(x,p)$ |
|---|---|
| position | $x_i$ |
| momentum | $p_i$ |
| angular momentum in 2-d | $x_1p_2 - x_2p_1$ |
| energy | $H(x,p)$ |

To determine how the value of an observable changes in time as the particle moves along a trajectory we use the chain rule to compute

$$\frac{d}{dt}f(x(t),p(t)) = \left\langle \nabla f, \begin{bmatrix} \dot{x}(t) \\ \dot{p}(t) \end{bmatrix} \right\rangle = \langle \nabla f, J\nabla H \rangle \tag{1.6}$$

where the gradients are evaluated at $(x(t),p(t))$. We now introduce some notation to rewrite this equation. Define $\omega$ to be the antisymmetric bilinear form whose value on vectors $X$ and $Y$ (thought of as tangent vectors to phase space) is given by

$$\omega[X,Y] = \langle X, JY \rangle . \tag{1.7}$$

This $\omega$ is called the *symplectic form*. Then the *Poisson bracket* of two observables $f$ and $g$ is defined as

$$\{f,g\} = \omega[\nabla f, \nabla g] = \sum_i \frac{\partial f}{\partial x_i}\frac{\partial g}{\partial p_i} - \frac{f\partial}{\partial p_i}\frac{\partial g}{\partial x_i}.$$

The Poisson bracket is an antisymmetric bilinear form on (differentiable) observables. It is a derivation, which means that it satisfies the product rule

$$\{f,gh\} = g\{f,h\} + \{f,g\}h.$$

It also satisfies the Jacobi identity

$$\{\{f,g\},h\} + \{\{h,f\},g\} + \{\{g,h\},f\} = 0.$$

Using this notation, equation (1.6) can be written

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

with all functions evaluated at $(x(t),p(t))$. Since $\{f,f\} = 0$ for any $f$ we see that energy is conserved:

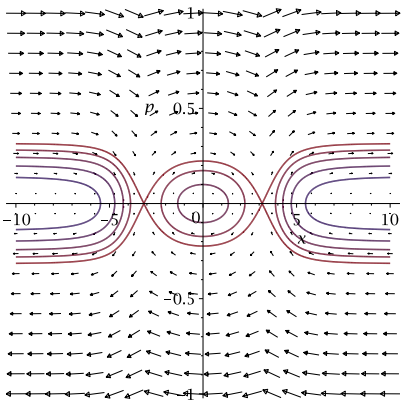
$$\frac{d}{dt}H = \{H,H\} = 0.$$

Other observables $f$ will be constants of motion provided $\{f,H\} = 0$. Here is an example. For the harmonic oscillator in $n$ dimensions, we can define $n$ observables $I_1,\ldots,I_n$, called action variables, as

$$I_k(x,p) = \frac{1}{2m}p_k^2 + \frac{m}{2}\omega_k^2 x_k^2.$$

Clearly $H = I_1 + \cdots + I_n$ and for $j \neq k$, $\{I_k, I_j\} = 0$ since they depend on disjoint sets of variables. Thus $\{I_k, H\} = \{I_k, I_1 + \cdots + I_n\} = \{I_k, I_k\} = 0$. So the action variables are constants of the motion.

When $n = 1$ the constant energy surfaces are curves in $\mathbb{R}^2$. In this case we can see the shape of the orbits simply by plotting them. For the harmonic oscillator with $n = 1$ the curves are circles $\{x^2 + p^2 = 2E\}$ when $E > 0$ or the point $(0, 0)$ when $E = 0$. For the two-bump potential a plot of the constant energy curves looks like this.




Some constant energy curves for two bump potential

So we can get a pretty good picture of the flow without much work.

When there are $n$ constants of the motion $f_1, \ldots, f_n$, as for the harmonic oscillator, the orbits lie on the intersection of all the level surfaces. When $\{f_i, f_j\} = 0$, such systems (under some regularity conditions) are called completely integrable. The intersections of the level surfaces turn out to be $n$ dimensional tori and the motion is straight line motion.

The other extreme is if the orbit is spread out evenly on the $n-1$ dimensional constant energy surface. In this case the system is called ergodic.

*Problem 1.3:* Verify that action functions for the Harmonic oscillator verify $\{I_j, I_k\} = 0$ for every $j$ and $k$.

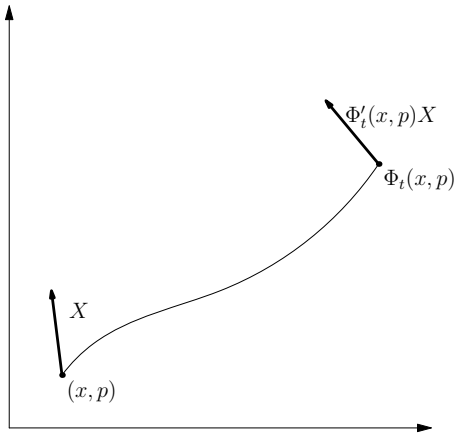
*Problem 1.4:* What happens to the position $x_j$ under the flow generated by momentum $p_k$. What about the other way around?

### Symplectic flow and Poincaré recurrence

The Hamiltonian flow $\Phi_t$ associated with $X_H$ is the map $\Phi_t : (x,p) \mapsto (x(t),p(t))$, where $(x(t),p(t))$ solves Hamilton's equations with intial condition $(x,p)$. In other words

$$\frac{d}{dt}\Phi_t(x,p) = X_H(\Phi_t((x,p))), \qquad \Phi_0((x,p)) = (x,p)$$

We want to show that the Hamiltonian flow preserves the symplectic form. To formulate this result recall that for fixed $t$ the differentiable map $\Phi_t : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$ induces a map on tangent vectors, sending $X$ in the tangent space at $(x,p)$ to $\Phi_t'(x,p)X$ in the tangent space at $\Phi_t((x,p))$, where $\Phi_t'(x,p)$ is the Jacobian matrix (also denoted $d\Phi_t(x,p)$).




The flow $\Phi_t$ and the map $\Phi_t'(x,p)$

If we write

$$\Phi_t((x,p)) = (f(t,x,p), g(t,x,p)).$$

then

$$\Phi_t'((x,p)) = \begin{bmatrix} \dfrac{\partial f}{\partial x}(t,x,p) & \dfrac{\partial f}{\partial p}(t,x,p) \\ \dfrac{\partial g}{\partial x}(t,x,p) & \dfrac{\partial g}{\partial p}(t,x,p) \end{bmatrix}$$

Saying that the flow preserves the symplectic form means

$$\omega[X,Y] = \omega[\Phi_t'((x,p))X, \Phi_t'((x,p))Y] \tag{1.8}$$

for every $t$.

**Proposition 1.1** *The Hamiltonian flow $\Phi_t$ preserves the symplectic form*

*Proof:* We will show that $\Phi_t'$ is a symplectic matrix, which means that

$$\Phi_t'^T J \Phi_t' = J \tag{1.9}$$

If (1.9) holds then $\omega[\Phi_t' X, \Phi_t' Y] = \langle \Phi_t' X, J\Phi_t' Y\rangle = \langle X, \Phi_t'^T J\Phi_t' Y\rangle = \langle X, JY\rangle = \omega[X, Y]$. Thus (1.9) implies (1.8).

Writing out the equation for the flow, we have

$$\frac{d}{dt}\Phi_t(x,p) = J\nabla H(\Phi_t(x,p)).$$

Computing the Jacobian of both sides, and exchanging the order of the partial differentiation with respect to $x$ and $p$ with the time derivative gives

$$\frac{d}{dt}\Phi_t'(x,p) = JH''(\Phi_t(x,p))\Phi_t'(x,p)$$

where $H''$ is the Hessian of $H$. Now

$$\begin{aligned}\frac{d}{dt}(\Phi_t'^T J\Phi_t') &= (JH''\Phi_t')^T J\Phi_t' + \Phi'^T J(JH''\Phi_t') \\ &= \Phi_t'^T H''^T J^T J\Phi_t' + \Phi_t'^T JJH''\Phi_t' \\ &= \Phi_t'^T H''\Phi_t' - \Phi_t'^T H''\Phi_t' \\ &= 0\end{aligned}$$

Here we used $J^T J = I$, $J^2 = -I$ and $H''^T = H''$. This implies that $\Phi_t'^T J\Phi_t'$ is constant.When $t = 0$ we have $\Phi_0' = I$ so the constant value is $J$. □

**Corollary 1.2** *(Liouville's theorem) The Hamiltonian flow preserves the phase space volume.*

*Proof:* Taking the determinant of both sides of (1.9) yields $\det(\Phi_t')^2 = 1$ so $|\det(\Phi_t')| = 1$. Now if $A$ is a (measurable) set in phase space then

$$\text{Vol}(\Phi_t(A)) = \int_A |\det(\Phi_t')(x,p)|\, dx^n dp^n = \int_A dx^n dp^n = \text{Vol}(A)$$

□

**Corollary 1.3** *(Poincaré's recurrence theorem) Suppose that $\Phi_t(V) \subset V$ for all times $t$ and for some set $V$ in phase space with Vol$(V) < \infty$. Let $v \in V$ and $T > 0$. Then any neighbourhood $U \subseteq V$ of $v$ contains a point that returns to $U$ after time $T$.*

*Proof:* Let $F = \Phi_T$ and denote the $k$ fold iterate $F \circ F \circ \cdots \circ F$ by $F^k$. Then $U, F(U), F^2(U), \ldots$ are contained in $V$ and have the same non-zero volume. If they were all disjoint then $V$ would have infinite volume. Thus $F^k(U) \cap F^j(U) \neq \emptyset$ for some $k > j$ which implies $F^{k-j}(U) \cap U \neq \emptyset$. □

*Problem 1.5:* Show that if $M$ is a symplectic matrix $M^T JM = J$ then $\det(M) = 1$.

### Summary of Classical Mechanics (so far)

A *pure state* of the system is a point $(x, p)$ in phase space. We call these pure states to distinguish them from mixed states, introduced later. A pure state is meant to describe the system completely with no uncertainty.

*Observables* are smooth real valued functions $f$ on phase space. There is a pairing between states and observables given by

$$\langle f|(x,p)\rangle = f(x,p).$$

This is a real number representing the result of making a measurement of the observable $f$ when the system is in the state $(x, p)$.

There is a distinguished observable $H(x, p)$ representing the total energy of the system. The time evolution of states is given by Hamiltonian flow $\Phi_t$. This flow preserves the symplectic form. The time evolution on observables is defined via the pairing. Explicitly, $f(x, p, t) = \Phi_t^* f$, where $\langle \Phi_t^* f|(x,p)\rangle = \langle f|\Phi_t(x,p)\rangle$. The time evolution on observables obeys

$$f' = \{f, H\},$$

where $\{\cdot, \cdot\}$ is the Poisson bracket.

### The Hamilton Jacobi Equation

One of the points of contact between Classical and Quantum mechanics is the Hamilton Jacobi equation.

$$H(x, \nabla S(x,t)) + \frac{\partial S}{\partial t}(x,t) = 0 \tag{1.10}$$

and its variant

$$H(x, \nabla s(x,E)) - E = 0 \tag{1.11}$$

We want to explain how these equation can be solved using the Hamiltonian flow for $H$. Since the solution to these equations can be used to construct approximate solutions to the Schrödinger equation in various situations, the equations provide a link between classical and quantum mechanics. We will consider an example of this later in the course.

Although it is not really important for us, it is remarkable that sometimes solutions to (1.11) can be used to determine the Hamiltonian flow for $H$.

It is maybe not surprising that first order equations can be solved using flows. But there is some pretty geometry connected with these equations that I'll try to describe. A reference for this material (in fact for everything in this section) is Arnold's classic book on Classical Mechanics [A]. I have tried to give an informal account of the main ideas. This section uses the language of

differential forms. In the context of classical mechanics, these are described in Arnold [A] and in Marsden and Ratiu [MR].

Our phase space is $\mathcal{P} = \mathbb{R}^n \times \mathbb{R}^n$ with co-ordinates $(x_1, \ldots, x_n, p_1, \ldots, p_n)$. In the language of differential forms the symplectic form we introduced above is

$$\omega = \sum_{i=1}^{n} dp_1 \wedge dx_i \tag{1.12}$$

A more general setting for classical mechanics is a symplectic manifold. This is an even dimensional manifold, typically the co-tangent bundle of configuration space, equipped with a closed non-degenerate 2-form $\omega$. (Closed means that $d\omega = 0$ and nondegenerate means that if $\omega[X, Y] = 0$ for every $Y$, then $X = 0$.) Darboux's theorem states that near every point one can find a co-ordinate system so that $\omega$ is given by (1.12). Actually there are always many such co-ordinate systems. They are called symplectic co-ordinates, or canonical co-ordinates.

Then the connection between (1.12) and (1.7) is this. Let $(x_1, \ldots, x_n, p_1, \ldots, p_n)$ be a symplectic co-ordinate system. Any vector $\xi$ in the tangent space of $\mathcal{P}$ at $q$ can be associated with the derivative at $t = 0$ of a curve $q(t)$ in $\mathcal{P}$ with $q(0) = q$. If $(x_1(t), \ldots, x_n(t), p_1(t), \ldots, p_n(t))$ are the co-ordinates of $q(t)$ then $\xi$ has the co-ordinate representation

$$X = \begin{bmatrix} \dot{x}_1(0) \\ \vdots \\ \dot{x}_n(0) \\ \dot{p}_1(0) \\ \vdots \\ \dot{p}_n(0) \end{bmatrix}.$$

Given two such vectors tangent vectors $\xi$ and $\eta$ with co-ordinate representations $X$ and $Y$ we find that $\sum_{i=1}^{n} dp_1 \wedge dx_i[\xi, \eta] = \langle X, JY \rangle$.

### Lagrangian submanifolds

We now introduce a special class of submanifolds of phase space. They have the important property that often they can be written as the graph of a gradient. Let's give the definition and explain this.

A submanifold $\mathcal{L}$ of $2n$ dimensional phase space $\mathcal{P}$ with symplectic form $\omega$ is *Lagrangian* if

(i) $\dim(\mathcal{L}) = n$

(ii) $\omega$ vanishes on $\mathcal{L}$.

Condition (ii) means that for any 2 vectors $X, Y$ tangent to $\mathcal{L}$ at some point, $\omega[X, Y] = 0$. We can check this as follows. Take any two curves $(x(t), p(t))$ and $(\tilde{x}(t), \tilde{p}(t))$ that stay in $\mathcal{L}$ and both pass

through $q \in \mathcal{L}$ when $t = 0$. Let $X = \begin{bmatrix} \dot{x}(0) \\ \dot{p}(0) \end{bmatrix}$ and $Y = \begin{bmatrix} \dot{\tilde{x}}(0) \\ \dot{\tilde{p}}(0) \end{bmatrix}$. Then we must have $\langle X, JY \rangle = 0$ . If $\mathcal{L}$ satisfies (ii) but not (i) it is called *isotropic*.

Here are some examples. When $n = 1$ then a $n$-dimensional submanifold of phase space $\mathbb{R} \times \mathbb{R}$ is a curve. There is only one tangent direction, so all tangent vectors are a multiple of a single vector $X$. By the antisymmetry of $\omega[X, X] = 0$. Another example is the submanifold

$$\mathcal{L}_1 = \{(x_0, p) : p \in \mathbb{R}^n\}.$$

containing all points in phase space lying above a fixed $x_0$ in configuration space. Tangent vectors to $\mathcal{L}_1$ have the form $\begin{bmatrix} 0 \\ A \end{bmatrix}$ and $\left\langle \begin{bmatrix} 0 \\ A \end{bmatrix}, J \begin{bmatrix} 0 \\ B \end{bmatrix} \right\rangle = 0$, for any $A, B \in \mathbb{R}^n$. A final simple example is

$$\mathcal{L}_2 = \{(x, p) \in \mathbb{R}^n \times \mathbb{R}^n : x = p\}.$$

In this case tangent vectors have the form $\begin{bmatrix} A \\ A \end{bmatrix}$ for $A \in \mathbb{R}^n$ and $\left\langle \begin{bmatrix} A \\ A \end{bmatrix}, J \begin{bmatrix} B \\ B \end{bmatrix} \right\rangle = 0$, for any $A, B \in \mathbb{R}^n$.

An $n$-dimensional submanifold $\mathcal{L}$ of phase space is a *graph over* $x = (x_1, \ldots, x_n)$ if there are functions $p(x) = (p_1(x), \ldots, p_n(x))$ so that $(x, p) \in \mathcal{L} \Leftrightarrow p = p(x)$. In this case $x = (x_1, \ldots, x_n)$ are co-ordinates for $\mathcal{L}$. In the examples above $\mathcal{L}_1$ is a graph over $p_1, \ldots, p_n$ but not over $x_1, \ldots, x_n$. On the other hand $\mathcal{L}_2$ is a graph over both $x_1, \ldots, x_n$ and $p_1, \ldots, p_n$.

Here is the main point of this section.

**Proposition 1.4** *Let $(x_1, \ldots, x_n, p_1, \ldots, p_n)$ be symplectic co-ordinates. Suppose that $\mathcal{L}$ is a graph over $(x_1, \ldots, x_n)$. Then $\mathcal{L}$ is Lagrangian $\Leftrightarrow$ there is a function $S(x)$ such that $p(x) = \nabla S(x)$.*

*Remark:* This proposition holds more generally for co-ordinates $(\tilde{x}_1, \ldots, \tilde{x}_n, \tilde{p}_1, \ldots, \tilde{p}_n)$ where $\sum d\tilde{p}_i \wedge d\tilde{x}_i = f\omega$ is a multiple of $\omega$. So $(\tilde{x}_1, \ldots, \tilde{x}_n, \tilde{p}_1, \ldots, \tilde{p}_n)$ could be $(\tilde{x}_1, \ldots, \tilde{x}_n, \tilde{p}_1, \ldots, \tilde{p}_n) = (p_1, \ldots, p_n, x_1, \ldots, x_n)$ in which case $f = -1$. Sometimes this extra flexibility is useful.

*Proof:* If $p(x) = \nabla S(x)$, then tangent vectors to $\mathcal{L}$ at the point $(x, p(x))$ have the form $\begin{bmatrix} A \\ S''(x)A \end{bmatrix}$ where $A \in \mathbb{R}^n$ and $S''(x) = \left[ \dfrac{\partial^2 S(x)}{\partial x_i \partial x_j} \right]$ is the Hessian of $S$. Then

$$\left\langle \begin{bmatrix} A \\ S''(x)A \end{bmatrix}, J \begin{bmatrix} B \\ S''(x)B \end{bmatrix} \right\rangle = \langle A, S''(x)B \rangle - \langle S''(x)A, B \rangle = 0$$

follows from the symmetry $S'' = S''^T$ of $S''$. Thus $\mathcal{L}$ is Lagrangian.

Conversely, if $\mathcal{L}$ is Lagrangian then the one form $\alpha = \sum p_i dx_i$ satisfies $d\alpha = \sum dp_i \wedge dx_i = \omega$. Thus $d\alpha = 0$ when restricted to $\mathcal{L}$. Since we are in a situation where every closed loop is spanned by a two dimensional surface (topologically $\mathcal{L} = \mathbb{R}^n$), this implies that $\alpha = dS$ for some function $S$ on $\mathcal{L}$. We can think of $S$ as function of the co-ordinates $(x_1, \ldots, x_n)$. Then

$$\alpha = \sum p_i(x) dx_i = dS = \sum \frac{\partial S}{\partial x_i}(x) dx_i$$

which implies $p_i(x) = \dfrac{\partial S}{\partial x_i}(x)$. $\square$

The function $S(x)$ is called a generating function.

If we know the functions $p_i(x)$ that determine $\mathcal{L}$, we can compute $S(x)$ (which is only determined up to a constant $S(x_0)$) by choosing a base point $x_0$ and curve $\gamma = ((x(t), p(x(t)))$ in $\mathcal{L}$ from $(x_0, p(x_0))$ to $(x, p(x))$. Then

$$\begin{aligned} S(x) &= S(x_0) + \int_\gamma dS \\ &= S(x_0) + \int_\gamma \sum p_i dx_i \\ &= S(x_0) + \int_0^1 \sum p_i(x(t))\dot{x}(t)dt \end{aligned}$$

### Legendre transforms

Suppose that the Lagrangian manifold $\mathcal{L}$ is simultaneously a graph over $(x_1, \ldots, x_n)$ and $(p_1, \ldots, p_n)$. Then there are functions $S(x)$ and $s(p)$ such that

$$(x, p) \in \mathcal{L} \Leftrightarrow p = \nabla S(x) \Leftrightarrow x = \nabla s(p).$$

In this situation $s$ is called a Legendre transform of $S$. Notice that if $\gamma$ is a path connecting $(x_0, p_0)$ to $(x, p)$ in $\mathcal{L}$ then

$$\begin{aligned} S(x) - S(x_0) &= \int_\gamma \sum p_i dx_i \\ &= \int_\gamma \sum d(p_i x_i) - x_i dp_i \\ &= \langle p, x\rangle - \langle p_0, x_0\rangle - s(p) + s(p_0) \end{aligned}$$

If we choose constants and co-ordinates such that $S(x_0) + s(p_0) - \langle p_0, x_0\rangle = 0$ then

$$S(x) = \langle p, x\rangle - s(p)$$

with $p$ determined by $x = \nabla s(p)$. This is the classical formula for the Legendre transform (at least for differentiable functions).

We can also consider partial Legendre transforms. If $\mathcal{L}$ is a graph over say $p_1, x_2, \cdots, x_n$ then we can apply our proposition to the co-ordinate system $p_1, x_2, \ldots, x_n, -x_1, p_2, \ldots, p_n$. Then there is a generating function $s(p_1, x_2, \ldots, x_n)$ such that $ds = -x_1 dp_1 + \sum_{i=2}^n p_i dx_i = -d(x_1 p_1) + \sum_{i=1}^n p_i dx_i$ so that (up to constants)

$$s(p_1, x_2, \ldots, x_n) = -x_1 p_1 + S(x_1, \ldots, x_n).$$

In this situation $s$ is a partial Legendre transform of $S$.

### Flow outs

In this section we show how to use the Hamiltonian flow to enlarge a constant energy isotropic manifold to a constant energy Lagrangian manifold.

**Proposition 1.5** *Let $H(x,p)$ be a Hamiltonian and suppose $\mathcal{L}_0$ is an $(n-1)$ dimensional submanifold of phase space such that*

*(i) $\mathcal{L}_0$ is isotropic, i.e., $\omega[X,Y]=0$ for $X,Y$ tangent to $\mathcal{L}_0$,*

*(ii) $\mathcal{L}_0$ is contained in a constant energy surface $\{(x,p) : H(x,p)=E\}$*

*(iii) The Hamiltonian vector field $X_H$ is nowhere tangent to $\mathcal{L}_0$.*

*Let $\mathcal{L}_0^+$ be the $n$-dimensional submanifold in phase space swept out by the Hamiltonian flow for $H$ with initial conditions on $\mathcal{L}_0$. Then $\mathcal{L}_0^+$ is is Lagrangian and also contained in the constant energy surface $\{(x,p) : H(x,p)=E\}$.*

*Proof:* Since Hamiltonian flow preserves $\omega$ we need only check $\omega[X,Y]=0$ for vectors $X,Y$ tangent to $\mathcal{L}_0^+$ at the initial points in $\mathcal{L}_0$. Such tangent vectors have the form $X=X_1+aX_H, Y=Y_1+bX_H$ where $X_1, Y_1$ are tangent to $\mathcal{L}_0$. Since $\mathcal{L}_0$ is isotropic, $\omega[X_1,Y_1]=0$ and $\omega[X_H,X_H]=0$ by antisymmetry. Finally, $\omega[X_1,X_H]=\langle X_1, \nabla H\rangle = 0$ because $X_1$ points in a direction in the constant energy surface. Similarly $\omega[Y_1,X_H]=0$. This implies $\omega[X,Y]=0$.

Since the Hamiltonian flow preserves energy, and all the initial points are contained in $\{(x,p) : H(x,p)=E\}$, the flow out $\mathcal{L}_0^+$ is contained in $\{(x,p) : H(x,p)=E\}$ too. □

### Extended phase space and the geometry of the Hamilton Jacobi equations

Consider a two dimensional phase space $\mathcal{P}_1 = \mathbb{R}\times\mathbb{R}$ with variables $(\tau, E)$ and symplectic form $\omega_1 = dE\wedge d\tau$ and Hamiltonian $H_1(\tau,E)=E$. Then we can form extended phase space $\mathcal{P}'$ as the product

$$\mathcal{P}' = \mathcal{P}\times\mathcal{P}_1 = \mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}$$

with variables $(x,p,\tau,E)$ and symplectic form

$$\omega' = \omega - \omega_1 = \sum_{i=1}^n dp_i\wedge dx_i - dE\wedge d\tau.$$

We define a new Hamiltonian on $\mathcal{P}'$ as

$$H'(x,p,\tau,E) = H(x,p) - H_1(\tau,E) = H(x,p) - E$$

Here is the geometric meaning of solving the Hamilton Jacobi equations (1.10) and (1.11):

> *Finding solutions amounts to finding a Lagrangian submanifold of $\mathcal{L}'\subset\mathcal{P}'$ contained in the constant energy surface $H'(x,p,\tau,E)=0$.*

To see this, suppose we have found such an $\mathcal{L}'$ that is a graph over $(x,\tau)$. Then $\mathcal{L}' = \{(x,p,\tau,E : p=\nabla_x S, E=\partial S/\partial\tau\}$ for some generating function $S(x,\tau)$. Since $\mathcal{L}'$ lies in the

surface $H' = 0$, the generating function $S(x, \tau)$ will satisfy $H'(x, \nabla_x S, \tau, \partial S/\partial \tau) = 0$. But this is the Hamilton-Jacobi equation (1.10) (with $t = \tau$). Similarly, if $\mathcal{L}'$ is a graph over $(x, E)$ then the generating function $s(x, E)$ will satisfy (1.11). Once we know $\mathcal{L}'$, the generating functions $S(x, \tau)$ and $s(x, E)$ can be found as described above. They will be partial Legendre transforms.

How can we find a Lagrangian submanifold of extended phase space that does the job? One way of constructing $\mathcal{L}'$ is to start with a Lagrangian submanifold $\mathcal{L}$ of $\mathcal{P}$ and imbed it as an isotropic submanifold in $\mathcal{P}'$ contained in the constant energy surface $H' = 0$ via $(x, p) \mapsto (x, p, 0, H(x, p))$. We can then flow out these points under the Hamiltonian flow for $H'$, to obtain a Lagrangian submanifold in $\mathcal{P}'$ in the constant enery surface $H' = 0$.

*Problem 1.6:* Show that the Hamiltonian vector field $X_{H'}$ is never tangent to this imbedded isotropic submanifold.

To perform the flow out, we can use the following proposition, which we state for more general products.

**Proposition 1.6** *Consider the extended phase space*

$$\mathcal{P}' = \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^k \times \mathbb{R}^k$$

*with variables* $(x_1, \ldots, x_n, p_1, \ldots, p_n, X_1, \ldots, X_k, P_1, \ldots, P_k)$ *and symplectic form*

$$\omega' = \sum_{i=1}^{n} dp_i \wedge dx_i - \sum_{j=1}^{k} dP_j \wedge dX_j.$$

*Suppose we have a Hamiltonian of the form* $H'(x, p, X, P) = H(x, p) - H_1(X, P)$. *Let* $\mathcal{L}'$ *be a Lagrangian submanifold in* $\mathcal{P}'$ *that is contained in the constant energy surface* $H' = 0$. *Let* $(x(t), p(t), X(t), P(t))$ *be a path in* $\mathcal{L}'$. *Consider the following statements:*

*(i)* $(x(t), p(t), X(t), P(t))$ *is a Hamiltonian trajectory for* $H'$

*(ii)* $(x(t), p(t))$ *is a Hamiltonian trajectory for* $H$

*(iii)* $(X(t), P(t))$ *is a Hamiltonian trajectory for* $H$

*Then (i)* $\Leftrightarrow$ *(ii) and (iii). Let* $\pi : \mathcal{L}' \to \mathbb{R}^k \times \mathbb{R}^k$ *be the projection* $(x, p, X, P) \mapsto (X, P)$ *and* $d\pi$ *the induced map on tangent spaces. If* $d\pi$ *is onto (this requires* $k \leq n$*) then (ii)* $\Rightarrow$ *(iii) and (i).*

*Proof:* Condition (i) can be written

$$\begin{bmatrix} \dot{x} \\ \dot{p} \\ \dot{X} \\ \dot{P} \end{bmatrix} = \begin{bmatrix} J & 0 \\ 0 & -J \end{bmatrix} \nabla H' = \begin{bmatrix} J & 0 \\ 0 & -J \end{bmatrix} \begin{bmatrix} \dfrac{\partial H}{\partial x} \\ \dfrac{\partial H}{\partial p} \\ -\dfrac{\partial H_1}{\partial X} \\ -\dfrac{\partial H_1}{\partial P} \end{bmatrix}$$

which is equivalent to $((x(t), p(t))$ and $(X(t), P(t))$ satisfying Hamiltions equations for $H$ and $H_1$ respectively. This shows (i) $\Leftrightarrow$ (ii) and (iii).

To show that (ii) $\Rightarrow$ (iii) and (i) assume that (ii) holds. Let $X' = \begin{bmatrix} a \\ b \\ A \\ B \end{bmatrix}$ be in the tangent space to $\mathcal{L}'$. The constant energy condition yields $\langle \nabla H', X' \rangle = 0$ which implies $\left\langle \nabla H, \begin{bmatrix} a \\ b \end{bmatrix} \right\rangle = \left\langle \nabla H_1, \begin{bmatrix} A \\ B \end{bmatrix} \right\rangle$. Since $\mathcal{L}'$ is Lagrangian and $\begin{bmatrix} \dot{x} \\ \dot{p} \\ \dot{X} \\ \dot{P} \end{bmatrix}$ is also a tangent vector we also know that $\left\langle \begin{bmatrix} \dot{x} \\ \dot{p} \\ \dot{X} \\ \dot{P} \end{bmatrix}, \begin{bmatrix} J & 0 \\ 0 & -J \end{bmatrix} \begin{bmatrix} a \\ b \\ A \\ B \end{bmatrix} \right\rangle = 0$ so that $\left\langle \begin{bmatrix} \dot{x} \\ \dot{p} \end{bmatrix}, J \begin{bmatrix} a \\ b \end{bmatrix} \right\rangle = \left\langle \begin{bmatrix} \dot{X} \\ \dot{P} \end{bmatrix}, J \begin{bmatrix} A \\ B \end{bmatrix} \right\rangle$. Condition (ii) says that $(x(t), y(t))$ is Hamiltonian flow for $H$. Thus

$$\begin{aligned} \left\langle \begin{bmatrix} \dot{X} \\ \dot{P} \end{bmatrix}, J \begin{bmatrix} A \\ B \end{bmatrix} \right\rangle &= \left\langle \begin{bmatrix} \dot{x} \\ \dot{p} \end{bmatrix}, J \begin{bmatrix} a \\ b \end{bmatrix} \right\rangle \\ &= \left\langle J\nabla H, J \begin{bmatrix} a \\ b \end{bmatrix} \right\rangle \\ &= \left\langle \nabla H, \begin{bmatrix} a \\ b \end{bmatrix} \right\rangle \\ &= \left\langle \nabla H_1, \begin{bmatrix} A \\ B \end{bmatrix} \right\rangle \\ &= \left\langle J\nabla H_1, J \begin{bmatrix} A \\ B \end{bmatrix} \right\rangle \end{aligned}$$

The condition on $d\pi$ means that we hit every vector $\begin{bmatrix} A \\ B \end{bmatrix}$ in $\mathbb{R}^k \times \mathbb{R}^k$ as $X'$ ranges over the tangent space of $\mathcal{L}'$. Thus $J \begin{bmatrix} A \\ B \end{bmatrix}$ also ranges over all vectors in $\mathbb{R}^k \times \mathbb{R}^k$. So we find that $\left\langle \begin{bmatrix} \dot{X} \\ \dot{P} \end{bmatrix} \right\rangle = J\nabla H_1$. Thus $(iii)$ holds. We already saw that (ii) and (iii) $\Rightarrow$ (i).

□

#### Remark on solving for the flow using the Hamilton Jacobi equation

As an aside, we can now see how the Hamilton Jacobi equation can be used to solve for the flow. Suppose we can solve a family of Hamilton Jacobi equations ,

$$H(x, \nabla S(x, P)) - H_1(P) = 0.$$

indexed by $P \in \mathbb{R}^k$ with $k = n$. Then $\mathcal{L}' = \{(x, \nabla_x S(x, P), \nabla_P S(x, P), P), x \in \mathbb{R}^n, P \in \mathbb{R}^k\}$ is a Lagrangian submanifold in the surface $H(x, p) - H_1(P) = 0$. In favourable situations one can solve for $(x, p)$ as as function of $(X, P)$ when $(x, p, X, P) \in \mathcal{L}'$. Call this function $(x(X, P), p(X, P))$. The proposition above tells us that if $(X(t), P(t))$ is a Hamiltonian trajectory for $H_1$ then the image

$(x(X(t),P(t)),p(X(t),P(t))$ is a Hamiltonian trajectory for $H$. But the equations for $(X(t),P(t))$ are easy to solve. We have $\dot{X} = \dfrac{\partial H_1}{\partial P}$ and $\dot{P} = 0$. Thus $P = P_0$ is constant in time and $X(t) = \dfrac{\partial H_1}{\partial P}(P_0)t$.

### An example

We now construct the solution to (1.10) corresponding to the set of free trajectories starting at a given point $x_0$ in configuration space. So the Hamiltonian is $H(x,p) = \dfrac{p^2}{2m}$.

Start with the Lagrangian submanifold of $\mathbb{R}^n \times \mathbb{R}^n$ given by

$$\mathcal{L} = \{(x_0, p) : p \in \mathbb{R}^n)\}$$

Then we imbed this as an isotropic zero energy (for $H' = H(p) - E$) surface in extended phase space $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}$ via

$$(x_0, p) \mapsto (x_0, p, 0, \frac{p^2}{2m}).$$

Flowing out using the flow for $H'$ from the initial point $(x_0, p_0, 0, \dfrac{p_0^2}{2m})$ yields the trajectory $(x_0 + \dfrac{1}{m}p_0 t, p_0, t, \dfrac{p_0^2}{2m})$. Thus a typical point in $\mathcal{L}'$ can be written

$$(x, p, \tau, E) = (x_0 + \frac{1}{m}p_0\tau, p_0, \tau, \frac{1}{2m}p_0^2)$$

for some choice of $p_0$ and $\tau$. Here we have written it as a graph over $(p_0, \tau)$. Solving $x = x_0 + \dfrac{1}{m}p_0\tau$ for $p_0 = \dfrac{m(x - x_0)}{\tau}$ We can also write points in $\mathcal{L}'$ as

$$(x, p, \tau, E) = (x, \frac{m(x - x_0)}{\tau}, \tau, \frac{m(x - x_0)^2}{2\tau^2})$$

A path $\gamma$ from $(x_0, p_0, 0, \dfrac{1}{2m}p_0^2)$ to $(x, p, \tau, E)$ is provided by

$$(x(t), p(t), \tau(t), E(t)) = (x_0 + \frac{1}{m}p_0 t, p_0, t, \frac{1}{2m}p_0^2), \quad t \in [0, \tau]$$

Since $dS = pdx - Ed\tau$ we find

$$\begin{aligned} S(x,\tau) &= \int_\gamma pdx - Ed\tau \\ &= \int_0^\tau (\langle p_0, \dot{x}(t)\rangle - \frac{1}{2m}p_0^2)dt \\ &= \frac{1}{2m}p_0^2\tau \end{aligned}$$

Since $p_0 = \dfrac{x - x_0}{m\tau}$,

$$S(x,\tau) = m\frac{|x - x_0|^2}{2\tau}.$$

The singularity at $\tau = 0$ is connected with the fact that $\mathcal{L}'$ fails to be a graph over $(x, \tau)$ at $\tau = 0$.

There is a connection with the classical action. If we carry out this construction with a more general Hamiltonian we will find that if $x(t)$ is the path in configuration space corresponding to a Hamiltonian trajectory we have

$$S(x(\tau), \tau) = \int_{t=0}^{\tau} \langle p(t), \dot{x}(t) \rangle - H(x(t), p(t)) dt$$

where $\dot{x}(t) = \dfrac{\partial H}{\partial p}(x(t), p(t))$. This implies that $\langle p(t), \dot{x}(t) \rangle - H(x(t), p(t))$ is the Legendre transform of $H(x, p)$ with respect to $p$, namely the Lagrangian $L(x(t), \dot{x}(t))$. Thus $S(x(\tau), \tau)$ is the action integral for this path.

## 2. Review of Probability and Operator theory

### Probability

We now want to consider the situation in Classical Mechanics where we have incomplete knowledge of our system. This incomplete knowledge is described by a probability measure on phase space. Probability measures also are basic to Quantum Mechanics, where they appear even when we have a complete knowledge of the system. So in this section we will review some basic definitions.

A *probability space* is a triple

$$(\Omega, \mathcal{F}, \mathbb{P})$$

where $\Omega$ is a set (e.g., phase space), $\mathcal{F}$ is a $\sigma$–algebra of subsets of $\Omega$ and $\mathbb{P}$ is a probability measure. A *$\sigma$-algebra on* $\Omega$ is a collection $\mathcal{F}$ of subsets of $\Omega$ satisfying

(i) $\Omega \in \mathcal{F}$, $\emptyset \in \mathcal{F}$,

(ii) $A \in \mathcal{F} \Rightarrow \Omega \backslash A \in \mathcal{F}$ (closed under complements),

(iii) $\{A_i\}_{i\in\mathbb{N}} \subset \mathcal{F} \Rightarrow \bigcup_{i\in\mathbb{N}} A_i \in \mathcal{F}$ (closed under countable unions),

(iv) $\{A_i\}_{i\in\mathbb{N}} \subset \mathcal{F} \Rightarrow \bigcap_{i\in\mathbb{N}} A_i \in \mathcal{F}$ (closed under countable intersections).

In fact, for a non-empty collection of subsets, (ii) and (iii) $\Rightarrow$ (i) and (iv). Subsets in $\mathcal{F}$ are called measurable sets and the pair $(\Omega, \mathcal{F})$ is called a *measurable space*. Sets in $\mathcal{F}$ are also called *events*.

For any collection $\mathcal{F}_0$ of subsets of $\Omega$ there is a smallest $\sigma$–algebra on $\Omega$ containing $\mathcal{F}_0$, namely the intersection of all such $\sigma$–algebras. This is the $\sigma$–algebra generated by $\mathcal{F}_0$. If $\Omega$ has a topology then there is a smallest $\sigma$–algebra $\mathcal{B}$ containing the open sets. This is called the *Borel* $\sigma$–algebra and its elements are called Borel sets.

A *positive measure* on a measurable space $(\Omega, \mathcal{F})$ is a countably additive function $\mu : \mathcal{F} \to [0, \infty]$. This means

(i) $\mu(\emptyset) = 0$

(ii) $\mu\left(\bigcup_{i\in\mathbb{N}} A_i\right) = \sum_{i\in\mathbb{N}} \mu(A_i)$ for any countable collection sets in $\{A_i\}_{i\in\mathbb{N}} \subset \mathcal{F}$ that is pairwise disjoint, i.e., $A_i \cap A_j = \emptyset$ for $i \neq j$.

If $\mathcal{F} = \mathbb{B}$, the Borel sets, then $\mu$ is called a Borel measure.

*Example:* a point mass (or delta function)

$$\delta_{x_0}(A) = \begin{cases} 1 & \text{if } x_0 \in A \\ 0 & \text{if } x_0 \notin A \end{cases}$$

If $\Omega = \mathbb{R}^n$ this is also denoted $\delta(x - x_0)$. A *pure point* (or discrete) measure is one of the form

$$\mu = \sum_i \alpha_i \delta_{x_i}$$

for some countable collection of points $\{x_i\}$ and positive numbers $\alpha_i$ with $\sum_i \alpha_i < \infty$.

The standard measure on $\mathbb{R}^n$ defined on Borel sets that reproduces the volume of rectangular boxes and is invariant under translations and rotations is called *Lebesgue measure*.

A *probability measure* on a measurable space $(\Omega, \mathcal{F})$ is a positive measure $\mathbb{P}$ with $\mathbb{P}(\Omega) = 1$. The number $\mathbb{P}(A)$ indicates the likelihood that an outcome in $A$ occurs. In this case $(\Omega, \mathcal{F}, \mathbb{P})$ is called a *probability space*.

A function $f : \Omega_1 \to \Omega_2$ where $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ are measurable spaces is called *measurable* if the pre-image of any set in $\mathcal{F}_2$ is contained in $\mathcal{F}_1$, that is

(i) $f^{-1}(F_2) \in \mathcal{F}_1$ for any $F_2 \in \mathcal{F}_2$.

Measurable functions on probability spaces are also called *random variables*.

If $f : \Omega_1 \to \Omega_2$ is a measurable function with respect to $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ and $\mu$ is a measure on $\mathcal{F}_1$ then we can define the image measure $f_*[\mu]$ as

$$f_*[\mu](F_2) = \mu(f^{-1}(F_2))$$

If $f : \Omega_1 \to \Omega_2$ is a measurable function with respect to $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ then the $\sigma$–algebra generated by $f$ is $\{f^{-1}(A) : A \in \mathcal{F}_2\}$. It is the smallest $\sigma$-algebra for which $f$ is measurable.

An important case is when $\Omega_2 = \mathbb{R}$ and $\mathbb{P}$ is a probability measure. Then the image measure $f_*[\mathbb{P}]$ is a measure on $\mathbb{R}$ and called the distribution measure of $f$. The number $f_*[\mathbb{P}](I)$ gives the probability that the value of $f$ lies in $I$.

We now discuss *integration*. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Define the indicator function

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}.$$

Then function of the form $s = \sum_{i=1}^n \alpha_i \chi_{A_i}$, for $\alpha_i \in \mathbb{R}$ and $A_i \in \mathcal{F}$ a finite collection of disjoint sets, is called a simple function. The integral of $s$ is defined as

$$\int_\Omega s d\mu = \sum_{i=1}^n \alpha_i \mu(A_i)$$

The integral of a positive measurable function $f$ is defined using monotone limits of simple functions. For a measurable function $f$ that is not necessary positive we can write $f = f_+ - f_-$ and define the integral of $f$ is the difference of the integrals of $f_-$ and $f_+$.

If $f$ is a random variable on a probability space the *expected value* of $f$, denoted $\mathbb{E}[f]$ is the average or mean

$$\mathbb{E}[f] = \int_\Omega f d\mathbb{P}.$$

We have

$$\mathbb{E}[\chi_A] = \mathbb{P}(A).$$

The *variance* of $f$ is the expected value of $(f - \mathbb{E}[f])^2$

$$\text{Var}(f) = \mathbb{E}[(f - \mathbb{E}[f])^2] = \mathbb{E}[f^2] - (\mathbb{E}[f])^2.$$

Two events $A$ and $B$ are *independent* if

$$\mathbb{P}[A \cap B] = \mathbb{P}[A]\,\mathbb{P}[B].$$

If $A$ is an event with $\mathbb{P}[A] \neq 0$ then the *conditional probability* of an event $B$ given $A$ is

$$\mathbb{P}[B|A] = \frac{\mathbb{P}[A \cap B]}{P[A]}$$

In this case $A$ and $B$ are independent if and only if $\mathbb{P}[B|A] = \mathbb{P}[B]$. Conditional probabilities appear again in the last lecture where the idea of conditioning with respect to a $\sigma$–algebra is also discussed.

### Two theorems in measure theory

We will refer to the following results:

**Theorem 2.1** *[The Riesz-Markov theorem]. Let X be a locally compact Hausdorff space. Every positive linear functional $\lambda$ on $C_c(X)$, is represented by a unique Borel measure $\mu$ on X. This means*

$$\lambda(f) = \int_X f(x) d\mu(x)$$

*for all $f \in C_c(X)$.*

A Borel measure $\mu$ on $\mathbb{R}^n$ is *absolutely continuous* with respect to Lebesgue measure $m$ if

$$m(A) = 0 \Rightarrow \mu(A) = 0.$$

In this case we write $\mu << m$. If $\mu << m$ then there exists an integrable function $f(x)$ such that for every Borel set $A$

$$\mu(A) = \int_A f(x) dm(x).$$

Two measures $\mu_1$ and $\mu_2$ on $(\Omega, \mathcal{F})$ are *singular*, denoted $\mu_1 \perp \mu_2$ if there are disjoint sets $A, B \in \mathcal{F}$ with $A \cup B = \Omega$ such that $\mu_1(A) = \mu_2(B) = 0$. A measure $\mu$ is *continuous* if $\mu(\{x\}) = 0$ for every $x \in \Omega$.

**Theorem 2.2** *[Lebesgue decomposition theorem]. Let $\mu$ be a regular Borel measure on $\mathbb{R}^n$ and let $m$ denote Lebesgue measure. Then*

$$\mu = \mu_{ac} + \mu_{sc} + \mu_{pp}$$

*where $\mu_{ac} << m$, $\mu_{sc}$ is continuous with $\mu_{sc} \perp m$ and $\mu_{pp}$ is pure point.*

### Mixed states in Classical Mechanics

We now define a *mixed state* in Classical Mechanics to be a probability measure $\mu$ on phase space. This represents a situation when we have incomplete knowledge of the system. The pairing of an observable $f$ and a state $\mu$ now does not yield a real number. Instead the value is a measure on $\mathbb{R}$, namely the distribution measure of $f$:

$$\langle f|\mu\rangle = f_*[\mu]$$

A pure state $(x, p)$ now identified with the mixed state $\delta_{(x,p)}$.

The time evolution of a mixed state is defined in the obvious way. If $\Phi_t$ is the Hamiltonian flow on phase space, then the flow on mixed states is $\Phi_{t*}$. In other words

$$\mu_t(A) = \mu(\Phi_t^{-1}(A))$$

### Background and self-adjoint operators

I'll assume you know the definitions of *Hilbert space* (with complex scalars), *bounded operators*, *adjoints*, *projections*, *spectrum*, *linear functionals* and *unitary operators*.

A bounded operator $A$ is self-adjoint if $A^* = A$. I won't give the definition of unbounded self-adjoint operators, and will try to get by with examples. An unbounded self-adjoint operator $A$ acting in a Hilbert space $\mathcal{H}$ is only defined on a dense set $\mathcal{D}(A) \subset \mathcal{H}$ called the domain. The spectrum $\sigma(A)$ of a self-adjoint operator $A$ is a closed subset of $\mathbb{R}$. The set $\sigma(A)$ is bounded iff the operator $A$ is bounded. All eigenvalues of $A$ lie in $\sigma(A)$, but there may be more points. Here are some examples.

*Examples:*

(i) $\mathcal{H} = \mathbb{C}^n$, $A$ is Hermitian $n \times n$, $\sigma(A) = \{\text{eigenvalues}\}$.

(ii) $\mathcal{H} = L^2(\mathbb{R}^n, d^n x)$, $A\psi = V(x)\psi(x)$ is multiplication by a bounded measureable function $V(x)$, $\sigma(A) = \text{ess range}(V)$ (or $\text{range}(V)$ if $V$ is continuous).

(iii) $\mathcal{H} = L^2(\mathbb{R}^n, d^n x)$, $A\psi = x_i\psi(x)$ is multiplication by $x_i$, $\mathcal{D}(A) = \{\psi \in \mathcal{H} : x_i\psi \in \mathcal{H}\}$, $\sigma(A) = \mathbb{R}$.

(iv) $\mathcal{H} = L^2(\mathbb{R}^n, d^n k)$, $A\phi = -i\partial\phi/\partial k_i$ , $\mathcal{D}(A) = \{\phi : x_i(\mathcal{F}^{-1}\phi)(x) \in L^2(\mathbb{R}^n, d^n x\}$, $\sigma(A) = \mathbb{R}$.

In example (iv), $\mathcal{F}$ is the Fourier transform (see below). If $\phi$ is not differentialble $\partial\phi/\partial k_i$ can be taken to mean $\mathcal{F}x_i\mathcal{F}^{-1}$, where $x_i$ is the multiplication operator.

**Spectral theorem for self-adjoint operators**

Let $A$ be self-adjoint and (for the moment) bounded. If $p(x)$ is a polynomial then we may define $p(A)$ in the obvious way by substituting $A$ for $x$. This is possible for any bounded $A$, but if $A$ is self-adjoint then one can show that

$$\|p(A)\| \leq \sup_{x\in\sigma(A)} |p(x)|.$$

This alllows us to define $f(A)$ for any $f \in C(\sigma(A))$ by taking limits. Now pick $\psi \in \mathcal{H}$. Then the map

$$f \mapsto \langle \psi, f(A)\psi\rangle$$

defines a bounded positive linear functional on $C(\sigma(A))$. Thus, by the Riesz-Markov theorem, there exists a regular positive Borel measure $\mu_\psi^A$ such that

$$\langle \psi, f(A)\psi\rangle = \int_{\sigma(A)} f(\lambda)\ d\mu_\psi^A(\lambda).$$

By definition, $\mu_\psi^A$ is the spectral measure for $\psi$ and $A$. Now notice that the right side of this equation still makes sense if $f$ is bounded Borel function. This allows us to use the equation (in reverse) to define and operator $f(A)$ for any bounded Borel functions. The map $f \mapsto f(A)$ is a $*$–algebra homomorphism. This means (among other things) that $\bar{f}(A) = f(A)^*$ and $(fg)(A) = f(A)g(A)$. Now let $f(x) = \chi_I(x)$ be the indicator function for a Borel set $I$. Then the operator $\chi_I(A)$ is a projection, called the spectral projection for $I$ and $A$.

The proof of these facts for unbounded self-adjoint $A$ is different, but the results are the same.

*Examples:*

(i) $\mathcal{H} = \mathbb{C}^n$, $A$ is Hermitian $n \times n$: Let $\{\lambda_k\}$ be the eigenvalues and $\{e_k\}$ an orthonormal basis of eigenvectors. Then $f(A) = \sum_k f(\lambda_k)\ e_k \otimes e_k^*$ so

$$\begin{aligned}
\langle \psi, f(A)\psi\rangle &= \sum_k f(\lambda_k)\ |\ \langle e_k, \psi\rangle\ |^2 \\
\mu_\psi^A &= \sum_k \delta(\lambda - \lambda_k)\ |\ \langle e_k, \psi\rangle\ |^2 \\
\chi_I(A) &= \sum_{\{k:\lambda_k\in I\}} e_k \otimes e_k^*
\end{aligned}$$

(ii) $\mathcal{H} = L^2(\mathbb{R}, dx)$, $A$ is multiplication by $x$ then

$$\begin{aligned}
\langle \psi, f(A)\psi\rangle &= \int_{\mathbb{R}} f(x)\ |\psi(x)|^2 dx \\
\mu_\psi^A &= |\psi(x)|^2 dx \\
\chi_I(A) &= \text{multiplication by } \chi_I(x)
\end{aligned}$$

### The Fourier transform

Let $\mathcal{S}$ be the Schwartz space of rapidly decreasing functions. The Fourier transform $\mathcal{F}\psi$ of $\psi \in \mathcal{S}$ is the function

$$(\mathcal{F}\psi)(k) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-i\langle k,x\rangle} \psi(x) d^n x.$$

The Fourier transform maps $\mathcal{S} \to \mathcal{S}$ and has an inverse given by

$$(\mathcal{F}^{-1}\phi)(x) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{i\langle x,k\rangle} \phi(k) d^n k.$$

We will use the notation $\widehat{\psi}$ for $\mathcal{F}\psi$ amd $\widetilde{\phi}$ for $\mathcal{F}^{-1}\phi$. Here are some properties of $\mathcal{F}$.

(i) $i\dfrac{d}{dk}\widehat{\psi}(k) = \widehat{x\psi(x)}$,

(ii) $k\widehat{\psi}(k) = -i\widehat{\dfrac{d}{dx}\psi(x)}$,

(iii) $(\widetilde{\widehat{\psi}})(x) = \psi(x)$, $(\widehat{\widetilde{\psi}})(k) = \psi(k)$.

(iv) $\int_{\mathbb{R}^n} |\psi(x)|^2 d^n x = \int_{\mathbb{R}^n} |\widehat{\psi}(k)|^2 d^n k$

(v) Let $\varphi_0(x)$ be the Gaussian $\varphi_0(x) = \pi^{-1/4} e^{-x^2/2}$. Then $\mathcal{F}\varphi_0 = \varphi_0$.

Properties (i) and (ii) say that the Fourier transform exchanges multiplication and differentiation by the co-ordinate functions. Property (iii) says that $\mathcal{F}$ and $\mathcal{F}^{-1}$ really are inverses. Property (iv) lets us extend the definition of $\mathcal{F}$ from $\mathcal{S}$ to $L^2(\mathbb{R}^n, d^n x)$ by taking limits. The extension is a unitary operator $\mathcal{F} : L^2(\mathbb{R}^n, d^n x) \to L^2(\mathbb{R}^n, d^n k)$.

The unitary operator $\mathcal{F}$ allows us to define differentiation operators. Let $p_i$ denote the operator acting in $\mathcal{H} = L^2(\mathbb{R}^n, d^n x)$ as

$$(p_i\psi)(x) = \mathcal{F}^{-1} k_i \mathcal{F}\psi$$

with domain $\mathcal{D}(p_i) = \{\psi \in L^2(\mathbb{R}^n, d^n x) : k_i\widehat{\psi} \in L^2(\mathbb{R}^n, d^n k)$. Then $p_i$ is unitarily equivalent to a multiplication operator by a real valued function, and hence self-adjoint. If $\psi \in \mathcal{S}$, then $p_i\psi = -i\partial\psi/\partial x_i$.

We now define three families of unitary operators $T_x$, $M_k$ and $D_a$ as

$$(T_{x_0}\psi)(x) = \psi(x - x_0) \quad \text{(translation by } x_0 \in \mathbb{R}^n\text{)}$$
$$(M_{k_0}\psi)(x) = e^{i\langle k_0,x\rangle}\psi(x) \quad \text{(multiplication by } e^{i\langle k_0,x\rangle}\text{, } p_0 \in \mathbb{R}^n\text{)}$$
$$(D_a\psi)(x) = a^{-n/2}\psi(x/a) \quad \text{(dilation by } a > 0\text{)}$$

and note that

(i) $T_{x_0}M_{k_0} = e^{-i\langle k_0,x_0\rangle} M_{k_0}T_{x_0}$

(ii) $\mathcal{F}T_{x_0} = M_{-x_0}\mathcal{F}$

(iii) $\mathcal{F} M_{k_0} = T_{k_0} \mathcal{F}$

(iv) $\mathcal{F} D_a = D_{a^{-1}} \mathcal{F}$.

We also have (at least when acting on suitable $\psi$),

(v) $T_{x_0} x T_{-x_0} = x - x_0, \quad T_{x_0} p T_{-x_0} = p$

(vi) $M_{k_0} p M_{-k_0} = p - k_0, \quad M_{k_0} x M_{-k_0} = x$

(vii) $D_a x D_{a^{-1}} = a^{-n/2} x, \quad D_a p D_{a^{-1}} = a^{n/2} p$

### Uncertainty principle

For simplicity let $n = 1$ and $\psi \in \mathcal{S}$ with $\|\psi\| = 1$. Define

$$\mathrm{Var}_\psi(x) = \langle \psi, x^2 \psi \rangle - \big( \langle \psi, x\psi \rangle \big)^2$$
$$\mathrm{Var}_\psi(p) = \langle \psi, p^2 \psi \rangle - \big( \langle \psi, p\psi \rangle \big)^2 = \Big\langle \widehat{\psi}, k^2 \widehat{\psi} \Big\rangle - \Big( \Big\langle \widehat{\psi}, k\widehat{\psi} \Big\rangle \Big)^2$$

Then

$$\mathrm{Var}_\psi(x) \mathrm{Var}_\psi(p) \geq 1/4$$

and equality holds iff $\psi = M_{p_0} T_{x_0} D_a \varphi_0$ (up to a possible phase factor $e^{i\eta}$) where $\varphi_0$ is the normalized Gaussian.

*Proof:* Using (iv), (v) and (vi) above we can find $x_0$, $k_0$ and $a$ such that $\phi = D_{a^{-1}} T_{-x_0} M_{-k_0} \psi$ verifies $\mathrm{Var}_\psi(x) = \langle \phi, x^2 \phi \rangle = 1/2$ and $\mathrm{Var}_\psi(p) = \langle \phi, p^2 \phi \rangle$. By the unitarity of the operators, $\|\phi\| = 1$.

The key ingredient is the commutator formula $i[p, x] = 1$. This implies

$$1 = \langle \phi, \phi \rangle = \langle \phi, i[p, x]\phi \rangle = -\operatorname{Re} \langle \phi', x\phi \rangle \leq \|\phi'\| \|x\phi\| = \sqrt{\mathrm{Var}_\psi(x) \mathrm{Var}_\psi(p)}$$

To get equality in the Cauchy-Schwarz inequality we must have $\langle \phi', x\phi \rangle$ real and the two factors proportional. This implies $\phi' = -\lambda x \phi$ for $\lambda > 0$. Integrating, we get $\phi(x) = C e^{-\lambda x^2/2}$ and then $\|\phi\| = 1$ and $\|x\phi\|^2 = 1/2$ imply $\lambda = 1$ and $C = \pi^{-1/4}$. So $\phi = \varphi_0$ and $\psi = M_{p_0} T_{x_0} D_a \varphi_0$. $\square$

### One parameter strongly continuous unitary groups and Stone's theorem

A map $\mathbb{R} \to \{\text{unitary operators on } \mathcal{H}\}$ is called a one parameter strongly continuous unitary group if $U(0) = I$, $U(t_1 + t_2) = U(t_1)U(t_2)$ and $t \mapsto U(t)\psi$ is continuous in $\mathcal{H}$ for any fixed $\psi \in \mathcal{H}$.

**Theorem 2.3** *[Stone's theorem] Let $U(t)$ be a one parameter strongly continuous unitary group. Then there is a unique self-adjoint operator $A$ such that $U(t) = e^{itA}$. Conversely, every self-adjoint operator $A$ generates a one parameter strongly continuous unitary group $U(t) = e^{itA}$. The limit $\lim_{t\to 0} \dfrac{(U(t) - U(0))\psi}{t}$ exists and equals $iA\psi$ iff $\psi \in \mathcal{D}(A)$. So we see that in general $U(t)\psi$ is continuous, but if $\psi \in \mathcal{D}(A)$ then $U(t)\psi$ is differentiable with $\dfrac{d}{dt} U(t)\psi = iAU(t)\psi$.*

*Problem 2.1:* The family of operators $T_{ta}$ for $a \in \mathbb{R}^n$ and $M_{tb}$ for $b \in \mathbb{R}^n$ are both one parameter strongly continuous unitary groups. Their generators are $\langle a, p \rangle$ and $\langle b, x \rangle$. What can you say about $D(a)$ for $a > 0$?

## 3. Quantum Mechanics

### Abstract Quantum description of a physical system

Start with a Hilbert space $\mathcal{H}$. *Observables* are self-adjoint operators acting in $\mathcal{H}$. For an observable $A$, the spectrum $\sigma(A)$ represents possible outcomes of a measurement. *Pure states* are one dimensional subspaces of $\mathcal{H}$. We can identify these with normalized vectors $\psi \in \mathcal{H}$ with $\|\psi\| = 1$, but then $\psi$ and $e^{i\eta}\psi$ (for real $\eta$) represent the same state. Given an observable $A$ and a state $\psi$, the pairing is given by the spectral measure.

$$\langle A|\psi\rangle = \mu_\psi^A$$

This is interpreted as the distribution measure for the value of $A$. So the expected value of $A$ in the state $\psi$ is

$$\int \lambda d\mu_\psi^A(\lambda) = \langle \psi, A\psi\rangle$$

If we repeatedly measure $A$ when the system is in the state $\psi$ this will be the average value. Similarly the probability of finding the value of $A$ in a Borel set $I$ is

$$\mathbf{P}_\psi[A \in I] = \langle \psi, \chi_I(A)\psi\rangle\,.$$

We are using the symbol $\mathbf{P}_\psi$ rather than $\mathbb{P}$ is to emphasize that there is no underlying probability measure on the set of states. There is only a single state $\psi$ involved. This number represents the proportion of times a repeated experiment to measure $A$ lands in $I$ when the system is in the state $\psi$.

It is important to realize that a pure state $\psi \in \mathcal{H}$, represents complete knowledge of the system, despite the fact that we cannot predict with certainty the outcome of all measurements of observables. In situations where we have incomplete knowledge, one can define quantum mixed states , as we did in Classical Mechanics . Briefly, a mixed state is given by a trace class positive definite operator $M$ with trace equal to 1, called a density matrix. Such an operator has a spectral representation $M = \sum_{i=1}^\infty \mu_i \psi_i \otimes \psi_i^*$, where $\psi_i \otimes \psi_i^*$ denotes the projection given by $\psi_i \otimes \psi_i^* \phi = \langle \psi_i, \phi\rangle \psi_i$, and the $\mu_i$ are postitive numbers summing to 1. The number $\mu_i$ is interpreted as the probabilty that the system is in state $\psi_i$. In this setup, a pure state is represented by a rank one density matrix.

We will not consider mixed quantum states in this course, except to point out that the superposition of two states $\psi_1$ and $\psi_2$, given by $\dfrac{\psi_1 + \psi_2}{\|\psi_1 + \psi_2\|}$ is a pure state, represented by a single unit vector and not a mixed state.

Every observable $A$ generates a flow on states given by $e^{-itA}$. There is a distinguished observable, the Hamiltonian $H$ that generates the time evolution $e^{-itH}$. The trajectory $\psi_t = e^{-itH}\psi$ is defined for all states, but if $\psi \in \mathcal{D}(H)$ then $\psi_t$ satisfies the Schrödinger equation

$$i\frac{\partial}{\partial t}\psi_t = H\psi_t.$$

We can transfer the time evolution to the observables in such a way that

$$\langle \psi_t, A\psi_t \rangle = \langle \psi, A_t\psi \rangle$$

by choosing

$$A_t = e^{itH}Ae^{-itA}$$

Notice that

$$\dot{A}_t = i[H, A_t]$$

where $[A, B] = AB - BA$.

Now we consider the case of commuting observables. Two bounded operators are said to commute if $[A, B] = 0$. Two unbounded self-adjoint operators $A$ and $B$ commute if $f(A)$ and $g(B)$ commute for bounded Borel functions $f$ and $g$. For commuting self-adjoint operators $A$ and $B$, $\chi_I(A)\chi_J(B)$ is a projection and

$$\mathbf{P}[A \in I \text{ and } B \in J] = \langle \psi, \chi_I(A)\chi_J(B)\psi \rangle$$

gives the joint distribution for $A$ and $B$ if we measure them simultaneously. If $A$ and $B$ don't commute, then $\chi_I(A)\chi_J(B)$ need not be a projection. We interpret this as saying that we cannot measure $A$ and $B$ simultaneously.

#### Quantum particle in an external potential

To describe the motion of a particle we need to choose a Hilbert space, and operators to represent the important observables. For a particle moving in configuration space $\mathbb{R}^n$ we may choose

$$\mathcal{H} = L^2(\mathbb{R}^n, d^n x).$$

Then it is natural to take the position operator $x_i$ to be multiplication by $x_i$. Classical observables that are functions of $x$ are then represented by multiplication by the same function. The probability of finding the particle in the set $I$ when the system is in the state $\psi$ is

$$\langle \psi, \chi_I(x)\psi \rangle = \int_I |\psi(x)|^2 d^n x.$$

In other words, the distribution measure for position is $|\psi(x)|^2 d^n x$. To determine what operator should represent momentum $p_i$ we recall that in Classical Mechanics the flow generated by $p$ is a translation in $x$. Therfore it is natural to choose $p_i$ to be the generator of the unitary group of translations. Thus we set

$$p_i = -i\frac{\partial}{\partial x_i} = \mathcal{F}^{-1} k_i \mathcal{F}.$$

The distribution measure for the momentum will then be $|\widehat{\psi}(k)|^2 d^n k$.

We can already see peculiar feature about Quantum Mechanics: the distributions for position and momentum are not independent. The uncertainty principle for the Fourier transform implies that if the distribution for the position is sharply peaked in some state $\psi$, so that $\text{Var}_\psi(x)$ is small, then $\text{Var}_\psi(p)$ will have to be large.

Finally we define the Hamiltonian operator as

$$H = \frac{p^2}{2m} + V = -\frac{1}{2m}\Delta + V$$

where $p^2 = -\Delta$ is (minus) the Laplacian operator and $V$ is multiplication by $V(x)$.

The first task is to find a domain that makes $H$ self-adjoint. This can be complicated if $V$ is unbounded. But if $V$ is bounded $H$ is self-adjoint on the domain of $\Delta$ given by

$$\mathcal{D}(H) = \{\psi \in L^2(\mathbb{R}^n, d^n x) : k^2 \widehat{\psi}(k) \in L^2(\mathbb{R}^n, d^n k)\}.$$

Once we have defined $H$ as a self-adjoint operator, we can ask about the properties of the time evolution $e^{-itH}$. These are closely connected to the spectral properties of $H$.

For example, if $H$ has an eigenvalue $E$ so that $H\psi = E\psi$, then the eigenfunction $\psi$ evolves in time as

$$e^{itH}\psi = e^{itE}\psi.$$

This expression represents the same state for all $t$, since $e^{itE}$ is a number with modulus 1. So eigenvalues correspond to bound states.

What can we say about the time evolution of $\psi$ if $\mu_\psi^H$ is absolutely continuous (with respect to Lebesgue measure).

To answer this question, consider the observable given by the projection $P\psi = \langle \phi, \psi \rangle \phi$ onto the state $\phi$. This observable has spectrum $\{0, 1\}$ corresponding to whether or not a state in our system is in the state $\phi$. So the expected value of $P$, which is the probability that $\psi$ is in the state $\phi$ is

$$\langle \psi, P\psi \rangle = |\langle \psi, \phi \rangle|^2.$$

If $\psi$ is some initial state evolving as $\psi_t = e^{itH}\psi$, the probability that $\psi_t$ remains in the initial state after time $t$ is $|\langle \psi, \psi_t \rangle|^2 = |\langle \psi, e^{itH}\psi \rangle|^2$. For example, if $\psi$ is an eigenfunction then this quantity

is 1 for all time. On the other hand, if the spectral measure $\mu_\psi^H$ is absolutely continuous, then by the spectral theorem

$$\langle \psi, e^{itH}\psi \rangle = \int_{\sigma(H)} e^{it\lambda} d\mu_\psi^H(\lambda).$$

The Riemann Lebesgue lemma, a result in measure theory, says that this tends to zero as $t \to \infty$.

But is such a state really a scattering state? Does it leave bounded regions of space in the sense that for a bounded region $I$ in configuration space $\langle \psi_t, \chi_I(x)\psi_t \rangle \to 0$? And if so, supposing that the expected value of $x^2$ given by $\langle \psi, x^2\psi \rangle$ is initially finite, how does it change in time?

There are many interesting questions like this and answering them for systems of interest can be challenging. There isn't time in this course to make a systematic study. Instead we will pick out a few examples for the simple systems we considered in the classical case. To simplify the notation, we will stick to one dimensional systems.

#### Example 1: Free motion

The Hilbert space is $\mathcal{H} = L^2(\mathbb{R}, dx)$ and

$$H = \frac{p^2}{2m} = -\frac{1}{2m}\frac{d^2}{dx^2}.$$

This operator is easy to study since it is diagonalized by the Fourier transform. This means that $\mathcal{F}H\mathcal{F}^{-1}$ is multiplication by $\dfrac{k^2}{2m}$, which implies that the spectrum is $[0, \infty)$ and is purely absolutely continuous. Moreover the time evolution $\psi_t$ is given by

$$\psi_t = e^{-itH}\psi = \mathcal{F}^{-1} e^{-itk^2/(2m)} \mathcal{F}\psi.$$

Notice that the momemtum distribution $|\widehat{\psi}_t(k)|^2 dk$ doesn't change in time.

To begin lets take our initial state to be a Gaussian $\varphi_0 = \pi^{-1/4} e^{-x^2/2}$. Then $\widehat{\varphi}_0(k) = \varphi_0(k)$ and the constant momentum distribution is $\pi^{-1/2} e^{-k^2} dk$. The expected values for both $x$ and $p$ in this state are zero. So this corresponds to a classical particle sitting at the origin.

How does this state evolve? We can calculate

$$\begin{aligned} \mathcal{F}(e^{-itH}\varphi_0)(k) &= \pi^{-1/4} e^{-itk^2/(2m)} e^{-k^2/2} \\ &= \pi^{-1/4} e^{-(1+it/m)k^2/2} \\ &= \pi^{-1/4} e^{-\alpha k^2/2} \end{aligned}$$

where $\alpha = 1 + it/m$ is complex. Thus the time evolution is

$$e^{-itH}\varphi_0 = \pi^{-1/4} \mathcal{F}^{-1} e^{-\alpha k^2/2}$$

*Problem 3.1:* Even though the complex dilation $D(\alpha^{-1/2})\psi$ has no meaning for a general $\psi \in L^2(\mathbb{R}, dk)$, show that $D(\alpha^{-1/2})\varphi_0$ makes sense and verifies $\mathcal{F}^{-1} D(\alpha^{-1/2})\varphi_0 = D(\alpha^{1/2})\mathcal{F}^{-1}\varphi_0$.

Using this result we continue the calculation.

$$\begin{aligned}(e^{-itH}\varphi_0)(x) &= \alpha^{-1/4}\mathcal{F}^{-1}D(\alpha^{-1/2})\varphi_0(k)\\ &= \alpha^{-1/4}D(\alpha^{1/2})\mathcal{F}^{-1}\varphi_0(k)\\ &= \alpha^{-1/2}\varphi_0(x/\sqrt{\alpha})\end{aligned}$$

This yields the distribution

$$\frac{1}{\sqrt{\pi}\sqrt{1+(t/m)^2}}e^{\frac{-x^2}{(1+(t/m)^2)}}dx$$

for the position $x$ in the state $e^{-itH}\varphi_0$. What is happening here? The state is not stationary as in the classical case, but it is also not moving in any direction. It is just slowly spreading out as time goes on.

To see some motion, lets start with an initial state where the momentum distribution has been shifted up. Let $\psi = M[k_0]\varphi_0$ so that $\widehat{\psi} = \mathcal{F}M[k_0]\varphi_0 = T[k_0]\mathcal{F}\varphi_0$. Then the initial momentum distribution $|\widehat{\psi}(k)|^2dk = |\varphi_0(k-k_0)|^2dk$ is centered at $k_0$. Now

$$e^{-itH}\psi = \pi^{-1/4}\mathcal{F}^{-1}e^{-itk^2/(2m)}e^{-(k-k_0)^2/2}$$

*Problem 3.2:* Using the expansion

$$\frac{tk^2}{2m} = \frac{t(k_0+k-k_0)^2}{2m} = \frac{t(k_0)^2}{2m} + \frac{tk_0(k-k_0)}{m} + \frac{t(k-k_0)^2}{2m}$$

and the commutation properties of $T[k_0]$, $M[tk_0/m]$ with $\mathcal{F}^{-1}$ show that

$$|\psi_t(x)|^2 = \frac{1}{\sqrt{\pi}\sqrt{1+(t/m)^2}}e^{\frac{-(x-tk_0/m)^2}{(1+(t/m)^2)}}$$

Now we see the particle moving to the right with velocity $k_0/m$ while still spreading at the same rate.

Even when the initial state is not a Gaussian we can still use the Fourier transform to compute the time evolution. To begin we assume that $\psi \in \mathcal{S}$. Then one can compute

$$(e^{-itH}\psi)(x) = (2\pi it/m)^{-1/2}\int_{\mathbb{R}} e^{im\frac{(x-y)^2}{2t}}\psi(y)dy$$

Then, as with the Fourier transform, we can extend the definition to $L^2(\mathbb{R}, dx)$ since we know the operator is unitary.

Now think of expanding $(x-y)^2 = x^2 - 2xy + y^2$ in the exponent and throwing out the $y^2$ term. The resulting operator is

$$\tilde{U}(t)\psi = \frac{1}{(t/m)^{1/2}}e^{i\frac{mx^2}{2t}}\widehat{\psi}(xm/t)$$

If the momentum distribution $\widehat{\psi}$ is sharply peaked at $k_0$ then $\widehat{\psi}(xm/t)$ is concentrated near where $xm/t = k_0$, or $x = k_0t/m$ which is the classical trajectory starting at $x_0 = 0$. Moreover the exponent looks like the classical action we computed before.

*Problem 3.3:* Show that $\|(\tilde{U}(t) - e^{-itH})\psi\| \to 0$ as $t \to \infty$.

### Using the Hamilton Jacobi equation

In this section we will construct approximate scattering solutions to the Schrödinger equation using the Hamilton Jacobi equation. For the free motion this may seem pointless since we can already compute the exact solution using the Fourier transform! But there are two reasons to proceed. So far we have just recognized the classical behaviour in solutions constructed another way. But now we will use the classical flow to actually build the solution. Secondly, this procedure works in situations where the Fourier transform is not available. In this section we will go back to general dimension $n$.

To get an approximate solution to the time dependent Schrödinger equation $\left(H - i\frac{\partial}{\partial t}\right)\psi = 0$ of the form $\psi(x,t) = e^{iS(x,t)}u(x,t)$, we substitute this expression into the left side of the equation. This is the key calculation, which we do for $V \neq 0$.

$$\left(\frac{-1}{2m}\Delta + V(x) - i\frac{\partial}{\partial t}\right)e^{iS}u$$
$$= e^{iS}\left(\frac{|\nabla S|^2}{2m} + V(x) + \frac{\partial S}{\partial t}\right) - ie^{iS}\left(\frac{1}{m}\nabla S\cdot\nabla u + \frac{\partial u}{\partial t} + \frac{1}{2m}(\Delta S)u\right) - e^{iS}\frac{1}{2m}\Delta u$$

Then the first term on the right side vanishes if $S(x,t)$ solves the Hamilton Jacobi equation

$$\frac{|\nabla S|^2}{2m} + V(x) + \frac{\partial S}{\partial t} = 0.$$

From our previous work, we know that given a Lagrangian submanifold $\mathcal{L}$ in extended phase space that (i) is contained in the constant energy surface $\{(x,p,\tau,E) : H(x,p) - E = 0\}$ and (ii) is a graph over $(x,t)$, then the generating function (action) $S(x,t)$ is a solution to the Hamilton Jacobi equation. Moreover, we can find $S$ by integrating the 1–form $\sum_i p_i dx_i - Ed\tau$ over paths in $\mathcal{L}$. We constructed such a Lagrangian using a flow out. In this case $(x,p,\tau,E) \in \mathcal{L}$ if there is an point $(x_0,p_0)$ in the initial constant energy Lagrangian manifold $\mathcal{L}_0$ in the original phase space (this implies $E = H(x_0,p_0)$), and a $\tau$ such that the Hamiltonian trajectory $(x(t),p(t))$ with initial condition $(x_0,p_0)$ arrives at $(x,p)$ at time $\tau$.

Now suppose we have solved the Hamilton Jacobi equation so that we know $S(x,t)$. Then our next task is to solve the equation

$$\frac{1}{m}\nabla S\cdot\nabla u + \frac{\partial u}{\partial t} + \frac{1}{2m}(\Delta S)u = 0$$

for $u(x,t)$. This is a first order equation, called a transport equation, that we can solve using flows. The standard trick for such equations is to assume we have a solution and determine its

values along the flow associated to the vector field in the equation. Of course, here these flows will be Hamiltonian trajectories. So assume that we have a solution $u(x,t)$. Fix a starting point $(x_0, p_0)$ the original Lagrangian of initial conditions. For notational simplicity lets suppose this Lagrangian is a graph over $p$ so a unique starting point is determined by $p_0$. Now let us determine the values of $u$ along this trajectory. Setting $v(t) = u(x(t), t)$ we find that

$$\dot{v}(t) = \dot{x}(t) \cdot \nabla u(x(t), t) + \frac{\partial u}{\partial t}(x(t), t)$$

But

$$\dot{x}(t) = p(t)/m = \nabla S(x(t), t)/m.$$

The first equality is Hamilton's equation and the second follows from the fact that $S$ is a generating function. So

$$\begin{aligned}\dot{v}(t) &= \nabla S \cdot \nabla u + \frac{\partial u}{\partial t} \\ &= \nabla S \cdot \nabla u + \frac{\partial u}{\partial t} + \frac{1}{2m}(\Delta S)u - \frac{1}{2m}(\Delta S)u \\ &= -\frac{1}{2m}(\Delta S)(x(t), t)v(t).\end{aligned}$$

Since we know $S(x,t)$, the function $\frac{1}{2m}(\Delta S)(x(t),t)$ is a known function of $t$, once we specify $p_0$, which tells us the orbit we are on. Thus we know $\alpha(p_0, t) = \frac{1}{2m}(\Delta S)(x(t), t)$ and

$$\dot{v}(t) = -\alpha(p_0, t)v(t)$$

Integrating, and keeping in mind that the constant of integration may depend on the orbit we are on, we find

$$v(\tau) = f(p_0)e^{-\int_0^\tau \alpha(p_0,t)dt}.$$

Now we can verify that any smooth function whose values along the trajectory $(x(t), t)$ labelled by $p_0$ are given by

$$u(x(t), t) = f(p_0)e^{-\int_0^\tau \alpha(p_0,t)dt}$$

where $\alpha(p_0, t)$ is given above, solves the equation.

If we construct $S(x,t)$ and $u(x,t)$ according to this scheme we find

$$\left(\frac{-1}{2m}\Delta + V(x) - i\frac{\partial}{\partial t}\right)e^{iS}u = -e^{iS}\frac{1}{2m}\Delta u$$

The term remaining on the left is an error term which may become negligible for large $t$.

To illustrate this procedure, let us sketch the constructing of scattering solutions for the free motion. In this case we found that

$$S(x,t) = \frac{m|x - x_0|^2}{2t}$$

Thus

$$\frac{1}{2m}\Delta S(x,t) = \frac{n}{2t}$$

which implies that in this case $\alpha(p_0,t) = n/(2t)$ is independent of the trajectory and so

$$u(x,t) = f(p_0)e^{-n\int_1^t \log(\tau)d\tau/2} = f(p_0)t^{-n/2}.$$

To complete the construction we must determine $p_0$, the initial momentum, if we know $(x,t)$. The Hamiltonian trajectory is $x = x_0 + p_0 t/m$, so $p_0 = m(x-x_0)/t$. Thus our approximate solutions have the form

$$\psi(x,t) = e^{im(x-x_0)^2/(2t)} f\left(\frac{m(x-x_0}{t}\right) t^{-n/2}$$

Now suppose we want to show that

$$\lim_{t\to\infty} e^{itH}\psi(x,t)$$

exists. This is like a wave operator in scattering theory. We can show existence by writing

$$e^{itH}\psi(x,t) = \psi(x,1) + \int_1^t \frac{d}{dt}e^{itH}\psi(x,t)$$

Using what is known as Cook's method, it suffices to prove

$$\int_1^\infty \left\| \frac{d}{dt}e^{itH}\psi(x,t)\right\| < \infty.$$

By the calculation above

$$\begin{aligned}\left\| \frac{d}{dt}e^{itH}\psi(x,t)\right\| &= \left\| e^{itH}\left(iH + \frac{\partial}{\partial t}\right)\psi(x,t)\right\| \\ &= \left\| \left(H - i\frac{\partial}{\partial t}\right)\psi(x,t)\right\| \\ &= \left\| \frac{1}{2m}\Delta u\right\|\end{aligned}$$

where $u(x,t) = f\left(\frac{m(x-x_0}{t}\right)t^{-n/2}$. So

$$\begin{aligned}\|\Delta u(x,t)\|^2 &\le C\int_{\mathbb{R}^n}\left| t^{-n/2-2} f\left(\frac{m(x-x_0)}{t}\right)\right|^2 d^n x \\ &= t^{-4}\left|f\left(m(y)\right)\right|^2 d^n y\end{aligned}$$

This shows $\|\Delta u(x,t)\| \le Ct^{-2}$ and so the integral is finite.

Now given a choice of $f$ set

$$\phi = \lim_{t\to\infty} e^{itH}\psi(x,t).$$

Then

$$\lim_{t\to\infty}\left\| e^{-itH}\phi - \psi(x,t)\right\| = \lim_{t\to\infty}\left\|\phi - e^{itH}\psi(x,t)\right\| = 0$$

and we have constructed an approximate scattering solution. A harder question is whether every scattering solution can be obtained by such a construction. This the question of asymptotic completeness.

I learned about this construction from Ira Herbst who used related more complicated constructions in his work with Erik Skibsted. For the case of Schrödinger operators see [DG].

### Example 2: The Harmonic Oscillator

We will again restrict ourselves to one dimension. Then the Hilbert space is $L^2(\mathbb{R}, dx)$ and the Hamiltonian operator is

$$H = \frac{1}{2}(p^2 + x^2)$$

(We have set $m = 1$). The classical motion consist of only bound states. My original intention was to present a proof of a theorem here that states that if $\lim_{|x|\to\infty} V(x) = \infty$ then the spectrum of $H$ only contains isolated eigenvalues with finite multiplicity. This can be found, for example, in Reed and Simon IV [RSIV] Theorem XIII.67.

The Harmonic Oscillator is special in many ways, and we can proceed differently using creation and anihilation operators. This is explained in almost every book on quantum mechanics! Define the (differential) operators

$$\begin{aligned} a &= \frac{1}{\sqrt{2}}\left(x + \frac{d}{dx}\right) && \text{anihilation operator} \\ a^* &= \frac{1}{\sqrt{2}}\left(x - \frac{d}{dx}\right) && \text{creation operator} \\ N &= a^* a && \text{number operator} \end{aligned}$$

and let $\varphi_0 = \pi^{-1/4} e^{-x^2/2}$ be the normalized Gaussian. Then

$$\begin{aligned} H &= N + \frac{1}{2} \\ [a, a^*] &= 1 \\ a\varphi_0 &= 0 \end{aligned}$$

and one defines $\psi_0 = \varphi_0$ and for $n = 1, 2, 3, \ldots$,

$$\psi_n = \frac{1}{\sqrt{n!}}(a^*)^n \psi_0.$$

These can be shown to form an orthonormal set of eigenfunctions for $H$ with eigenvalues $E_n = n + \frac{1}{2}$. The have the explicit form

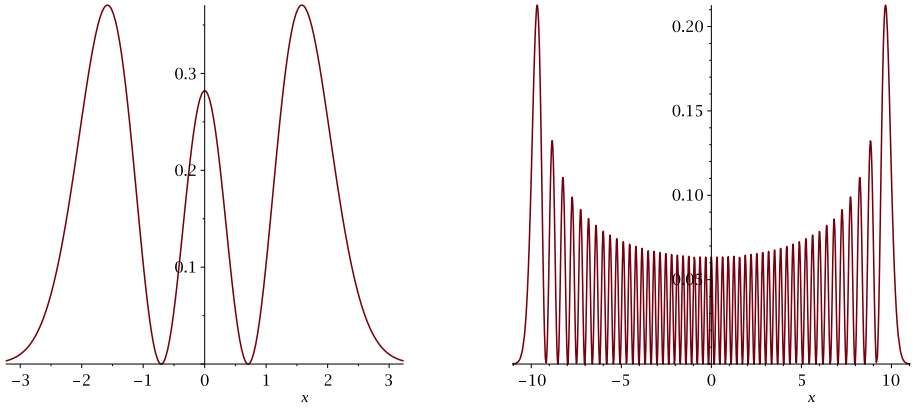
$$\psi_n(x) = \frac{1}{\sqrt{2^n n!}} \pi^{-1/4} H_n(x)$$

where

$$H_n(z) = (-1)^n e^{z^2} \frac{d^n}{dz^n} e^{-z^2}$$

are the classical Hermite polynomials. There is a theorem that states that the set $\{p(x)\varphi_0(x) : p(x) \text{ is a polynomial}\}$ is dense in $L^2(\mathbb{R}, dx)$. This implies that $\{\psi_n : n \in \mathbb{N}\}$ is an orthonormal basis of eigenvectors for $H$.

The first thing I want to do is to set fix an energy $E_n = n + \frac{1}{2}$ and compare the classical and quantum states with this energy. The classical state is a a circular trajectory in phase space with radius $\sqrt{2E_n}$. The quantum bound state is stationary. It doesn't move in time and the probability distribution for the position of the particle is the constant measure $|\psi_n(x)|^2 dx$ which looks like this (for $n = 2$ and $n = 50$):



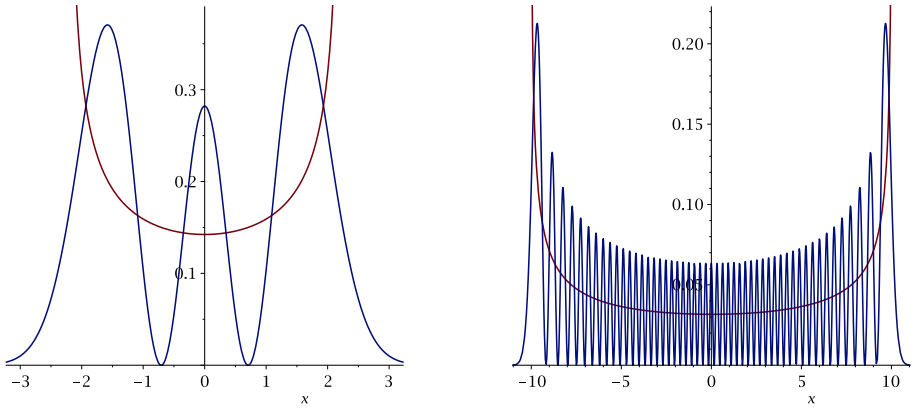

What classical quantity could one compare to these distributions? Maybe the quantum distribution represents a sort of average of the classical motion. To test this we must compute measure the on $[-\sqrt{2E_n}, \sqrt{2E_n}]$ defined by

$$m_n(I) = \lim_{T\to\infty} T^{-1} \int_{-T}^{T} \chi_I(t)dt.$$

This gives the asymptotic proportion of the time that the classical particle spends in the set $I$. A calculation yields

$$dm_n = \frac{1}{\pi\sqrt{2E_n - x^2}} dx$$

Here is what we get if we superimpose the classical and quantum densities.




When $n = 2$ they are pretty different, but when $n = 50$ the classical density looks like some sort of average.

*Problem 3.4:* Is it true that for $x \in [-\sqrt{2E_n}, \sqrt{2E_n}]$

$$\lim_{n\to\infty} \left| \int_{-\infty}^{x} |\psi_n(y)|^2 dy - \int_{-\sqrt{2E_n}}^{x} m_n(y) dy \right| = 0?$$

The above discussion might leave you with the impression that there is no motion in the quantum Harmonic Oscillator. But when you take a linear combination, or superposition, of eigenstates with different eigenvalues the result is no longer a stationary state. Amazingly, one can construct states which exhibit exactly the same motion as the classical particle! These states are called coherent states. They were introduced by Klauder and have been greatly generalized. For an exposition see, for example, the second volume of Simon's Course in Analysis [S2]

A simple way to define them is as eigenstates of the annihilation operator. We already know that $a\varphi_0 = 0$ so $0$ is an eigenvalue of $a$ with eigenvector $\varphi_0$.

**Proposition 3.1** *Every complex number $\alpha$ is an eigenvalue of the annihilation operator $a$. If $\alpha = \frac{1}{\sqrt{2}}(x_0 + ip_0)$, then the normalized eigenvector $\varphi_\alpha$ is given by*

$$\varphi_\alpha(x) = M[p_0]T[x_0]\varphi_0 = e^{ip_0 x}\varphi_0(x - x_0).$$

We can think of $\alpha$ as a point in classical phase space. Then $|\varphi_\alpha(x)|^2 = \pi^{-1/2}e^{-(x-x_0)^2}$ while $|\widehat{\varphi}_\alpha(k)|^2 = \pi^{-1/2}e^{-(k-p_0)^2}$. So the position and momentum distributions for this state are Gaussians centered at the classical values $x_0$ and $p_0$.

*Proof:* Writing out the equation $a\varphi_\alpha = \alpha\varphi_\alpha$ yields $x\varphi_\alpha + \varphi'_\alpha = \sqrt{2}\alpha\varphi_\alpha$. whose solution is

$$\begin{aligned}\varphi_\alpha(x) &= Ce^{-x^2/2+\sqrt{2}\alpha x} \\ &= Ce^{-x^2/2+(x_0+ip_0)x} \\ &= \tilde{C}e^{-(x-x_0)^2/2}e^{ip_0x}.\end{aligned}$$

The normalization condition then implies $\tilde{C} = 1$ (up to an inessential phase). □

Recall that the Harmonic Oscillator time evolution of a point $\alpha$ in classical phase space, thought of as a point in $\mathbb{C}$, is $\alpha(t) = e^{-it}\alpha$. The quantum time evolution of $\varphi_\alpha$ completely classical in the following sense.

**Proposition 3.2**

$$e^{-itH}\varphi_\alpha = e^{-it/2}\varphi_{\alpha(t)}$$

Of course the phase factor $e^{-it/2}$ doesn't change the state.

*Proof:* Expand $\varphi_\alpha$ in the orthonormal basis of eigenvectors. We have

$$\varphi_\alpha = \sum_n \langle \psi_n, \varphi_\alpha \rangle \psi_n$$

where

$$\langle \psi_n, \varphi_\alpha \rangle = \frac{1}{\sqrt{n!}} \langle (a^*)^n \varphi_0, \varphi_\alpha \rangle = \frac{1}{\sqrt{n!}} \langle \varphi_0, a^n \varphi_\alpha \rangle = \frac{\alpha^n}{\sqrt{n!}} \langle \varphi_0, \varphi_\alpha \rangle$$

To determine $\langle \varphi_0, \varphi_\alpha \rangle$ use

$$1 = \|\varphi_\alpha\|^2 = \sum_n |\langle \psi_n, \varphi_\alpha \rangle|^2 = \sum_n \frac{|\alpha|^{2n}}{n!} |\langle \varphi_0, \varphi_\alpha \rangle|^2 = e^{|\alpha|^2} |\langle \varphi_0, \varphi_\alpha \rangle|^2.$$

So $|\langle \varphi_0, \varphi_\alpha \rangle| = e^{-|\alpha|^2/2}$ and (up to a phase factor)

$$\varphi_\alpha = e^{-|\alpha|^2/2} \sum_n \frac{\alpha^n}{\sqrt{n!}} \psi_n$$

Since each $\psi_n$ is an eigenvector of $H$ with eigenvalue $E_n = n + \frac{1}{2}$ we know that $e^{-itH}\psi_n = e^{-it(n+\frac{1}{2})}\psi_n$. Therefore

$$\begin{aligned}e^{-itH}\varphi_\alpha &= e^{-|\alpha|^2/2} \sum_n \frac{\alpha^n}{\sqrt{n!}} e^{-itH}\psi_n \\ &= e^{-|\alpha|^2/2} \sum_n \frac{\alpha^n}{\sqrt{n!}} e^{-it(n+\frac{1}{2})}\psi_n \\ &= e^{-it/2}e^{-|\alpha|^2/2} \sum_n \frac{(e^{-it}\alpha)^n}{\sqrt{n!}} \psi_n \\ &= e^{-it/2}e^{-|\alpha(t)|^2/2} \sum_n \frac{\alpha(t)^n}{\sqrt{n!}} \psi_n \\ &= e^{-it/2}\varphi_{\alpha(t)}\end{aligned}$$

□

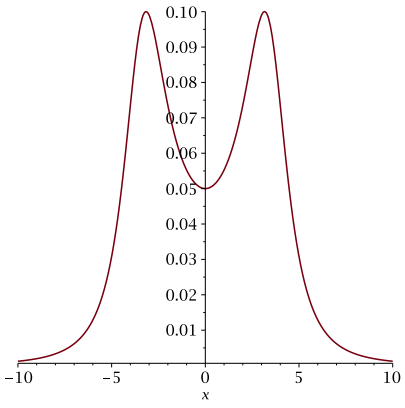
The final topic I intended to cover for the Harmonic Oscillator is the Maslov-WKB construction of approximate eigenfunctions using the Hamilton-Jacobi equation (1.11). Again, this construction is only important as an illustration of the method, as we know the eigenfunctions exactly. In this case, one needs to take large $E$ to get a good approximation.

The difficulty is that the Lagrangian manifold is now a circle, which is not a graph over $(x, E)$. We can define $S = \int p dx$ as a smooth function on the circle, but the approximate eigenfunction we get as a function of $x$ blows up at the turning points $x = \pm\sqrt{2E}$ where the manifold fails to be a graph. The solution is to work locally. Near the turning points project onto $p$ instead. This is particularly easy for the Harmonic Oscillator since $p$ and $x$ appear symmetrically in Hamiltonian. The resulting function of $p$ is then changed to a function of $x$ by taking the Fourier transform. Combining these local function into a smooth function requires the introduction of the Maslov index at each turning point. Moreover the final definition of the approximate eigenfunction gives zero unless the Bohr-Sommerfeld quantization condition holds.

There was no time for this, but fortunately there is a beautiful pedagogical paper by Eckmann and Seneor [ES] that explains the procedure.

#### Example 3: The two bump potential

The two bump potential we considered before illustrates some interesting differences between classical and quantum mechanics. Recall that the classical particle could be trapped by the potential well. What can we say about bound and scattering states for the quantum Hamiltonian $H = p^2 + V(x)$ acting in $L^2(\mathbb{R})$?




Two bump potential

There was no time for this either in the course, but here are some results about this operator that we could have considered. For simplicity let us assume that $V(x)$ is smooth and compactly supported.

**Theorem 3.3** *The spectrum of $H$ is $\sigma(H) = [0, \infty)$ and is absolutely continuous.*

So in contrast to the classical case, where there are bound states with positive energy corresponding to particles trapped by the well, there are no bound states in the quantum case. The physical explanation for this is that the particle can tunnel through the barrier. The mathematical theory of the absence of eigenvalues imbedded in the continuous spectrum is harder. There are examples of potentials (e.g., the Wigner–von Neumann potential) which tend to zero (slowly) for large $|x|$ but do have a positive eigenvalue imbedded in the continuous spectrum. See, for example, Chapter 4 of [CFKS].

What would happen if we moved the bottom of the central well so that it dipped below zero? Since we are working in one dimension, $H$ would immediately have a negative eigenvalue (bound state) in addition to the absolutely continuous spectrum on the postive half line. The proof of this is a nice application of min-max methods (see [RSIV]).

What would happen if we moved the bottom of the central well so that it's minimum was exactly zero, and then multiplied $V(x)$ by a large positive coupling constant $\mu$? If the minimum was non-degenerate, then the potential would look more and more like a harmonic oscillator as $\mu$ goees to infinity. Yet for each finite $\mu$ the potential would satisfy the hypothesis of the theorem. In this case the spectrum would be absolutely continuous in $[0, \infty)$ for every $\mu$. However there would be *resonances* in the lower half plane converging to the harmonic oscillator eigenvalues. For a definition and introduction to resonances, see e.g., [CKFS] [HS]. Resonances could be the subject of another course like this!

## 4. Hidden variables and non-locality

### Introduction

In general, quantum mechanics only provides probability distributions for the measured values of a quantum observables., and cannot predict the outcomes with certainty. In the early days of quantum mechanics, Einstein strongly resisted the idea that a physical theory could be essentially probabilistic in this way. Experiments like the one described by Einstein, Podolsky and Rosen in 1935 were proposed to show that measurement of spatially separated but quantum mechanically entangled particles leads to action at a distance. In the intervening eighty years there has been much philosophical debate on the foundations of quantum mechanics and the nature of physical reality. A readable account is *The Infamous Boundary* by David Wick [W]. Remarkably, recent experiments support the idea that action at a distance is real [FTZWF]. In this lecture, we will consider an example of Lucien Hardy [H] that is an outgrowth of the fundamental work of J.S. Bell [B]. This account is based on Faris [F1] [F2] and conversations with David Brydges.

It is possible (and usual) to work on the mathematics of quantum mechanics without worrying about the philosophical underpinnings. Still, I think it is worthwhile to spend some time to appreciate the essential weirdness of this theory.

Recall that for a system in the pure state $\psi$, given a Borel set $I \subseteq \mathbb{R}$, the probability that the measured value of a quantum observable $A$ will lie in $I$ is

$$\mathbf{P}_\psi[A \in I] = \langle \psi, \chi_I(A)\psi \rangle \,, \tag{4.1}$$

where $\chi_I(A)$ is the spectral projection for the operator $A$ corresponding to the set $I$.

Two commuting $A$ and $B$ can in principle be measured simultaneously. In this case $\chi_I(A)$ and $\chi_J(B)$ commute and $\chi_I(A)\chi_J(B)$ is again a projection. The probability that the simultaneously measured values of $A$ and $B$ lie in $I$ and $J$ respectively is

$$\mathbf{P}_\psi[A \in I \textbf{ and } B \in J] = \langle \psi, \chi_I(A)\chi_J(B)\psi \rangle \tag{4.2}$$

Equations (4.1) and (4.2) can be taken as definitions of the quantum probabilities $\mathbf{P}_\psi[A \in I]$ and $\mathbf{P}_\psi[A \in I \textbf{ and } B \in J]$ on the left sides. Alternatively, we can think of $\mathbf{P}_\psi[A \in I]$ and $\mathbf{P}_\psi[A \in I \textbf{ and } B \in J]$ as quantities that can be determined experimentally by running the same experiment many times. Then (4.1) and (4.2) express the agreement of theory and experiments.

It is natural to ask whether a more complete theory could predict outcomes with certainty. Perhaps quantum observables have exact values, and saying that a system is in a quantum state $\psi$ only gives partial information about them. In other words, maybe the situation is like a mixed state in Classical mechanics, where there is a probability measure on some underlying state space.

Could there be a quantum state space $\Omega$ with quantum observables $A, B, \ldots$ corresponding to functions $a, b, \ldots : \Omega \to \mathbb{R}$ with $a(\omega), b(\omega), \ldots$ giving the exact value of the observables if the system is in the state $\omega$. Then each quantum state $\psi$ would correspond to a probability measure $\mathbb{P}_\psi$ on $\Omega$ such that the spectral measures for $A, B, \ldots$ are the image measures of $\mathbb{P}_\psi$ under $a(\omega), b(\omega), \ldots$. Or perhaps this is true for some subset $\mathcal{A}$ of observables. Thus we are led to the following question.

*Question 1:* Given a quantum system and a collection $\mathcal{A}$ of observables does there exist

- a measureable space $(\Omega, \mathcal{F})$ and
- for each $A \in \mathcal{A}$ a random variable $a : \Omega \to \mathbb{R}$

such that every pure state $\psi$ is associated with a probability measure $\mathbb{P}_\psi$ that reproduces the quantum distributions for observables, that is,

$$\langle \psi, P_I(A)\psi \rangle = \mathbb{P}_\psi[\{a \in I\}] \tag{4.3}$$

and if $A$ and $B$ commute,

$$\langle \psi, P_I(A)P_j(B)\psi \rangle = \mathbb{P}_\psi[\{a \in I\} \cap \{b \in J\}]? \tag{4.4}$$

*Example 1:* If $\mathcal{A} = \{A\}$ contains only one observable, then the answer to Question 1 is obviously yes. We can take $\Omega = \mathbb{R}$ and $\mathbb{P}_\psi = \mu^A_\psi$.

*Example 2:* If $\mathcal{A}$ a collection of non-commuting observables, then the answer to Question 1 is also yes. We don't have to verify (4.4) and can satisfy (4.3) by assuming that measurements of different observables are independent.

In general the answer to Question 1 is no. Our first goal is to construct an example to illustrate this.

### Spin observables $A$ and $B$

Our examples are constructed from observables $A$, $B$ given by $2 \times 2$ Hermitian matrices acting on the Hilbert space $\mathbb{C}^2$. We will assume that they have the same two distinct eigenvalues $\lambda_1$ and $\lambda_2$ but different eigenvectors.

Such matrices occur in Physics when describing the spin of a spin $\frac{1}{2}$ particle. In units where $\frac{\hbar}{2} = 1$ these have the form

$$S_e = \sum_{i=1}^{3} e_i \sigma_i$$

where $e = [e_1, e_2, e_3]$ is a unit vector (the spin direction) in $\mathbb{R}^3$, and $\sigma_i$ are Pauli spin matrices

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

For any choice or direction $e$ the matrix $S_e$ has eigenvalues $-1$ and $1$. These are the possible outcomes of measuring the spin in this direction. To make the examples easier to follow, we will think of $A$ and $B$ as spin matrices.

Let $\{\phi_1, \phi_2\}$ be an orthonormal basis of eigenvectors for $A$ and denote by $P_i$, $i = 1, 2$ the corresponding spectral projections. Similarly, let $\{\psi_1, \psi_2\}$ be an orthonormal basis of eigenvectors for $B$ and denote by $Q_i$, $i = 1, 2$ the corresponding spectral projections. We will assume that the bases are related by the unitary change of basis matrix $U = [\langle \psi_i, \phi_j \rangle]$ given by the rotation matrix

$$U = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix},$$

with $\alpha, \beta \in [-1, 1]$, $\alpha^2 + \beta^2 = 1$. We will assume that neither $\alpha$ nor $\beta$ is zero.

To begin, let us illustrate explicitly the construction of independent random variables in Example 2 above. This requires a product measure. We can use $\mathcal{A} = \{A, B\}$. Consider a state $\psi = \frac{1}{\sqrt{2}}\phi_1 + \frac{1}{\sqrt{2}}\phi_2$. The quantum probabilities are

$$\begin{aligned}
\mathbf{P}[A = \lambda_1] &= |\langle \psi, P_1 \psi \rangle|^2 = \frac{1}{2} \\
\mathbf{P}[A = \lambda_2] &= |\langle \psi, P_2 \psi \rangle|^2 = \frac{1}{2} \\
\mathbf{P}[B = \lambda_1] &= |\langle \psi, Q_1 \psi \rangle|^2 = \frac{(\alpha - \beta)^2}{2} \\
\mathbf{P}[B = \lambda_2] &= |\langle \psi, Q_2 \psi \rangle|^2 = \frac{(\alpha + \beta)^2}{2}
\end{aligned}$$

Let $\Omega$ be the finite probability space

$$\Omega = \{\lambda_1, \lambda_2\} \times \{\lambda_1, \lambda_2\} = \{(\lambda_1, \lambda_1), (\lambda_1, \lambda_2), (\lambda_2, \lambda_1), (\lambda_2, \lambda_2)\},$$

$\mathcal{F}$ be all subsets of $\Omega$ and define the random variables $a$ and $b$ by

$$a((\lambda_i, \lambda_j)) = \lambda_i, \quad b((\lambda_i, \lambda_j)) = \lambda_j.$$

Then the probability measure generated by

$$\begin{aligned}
\mathbb{P}_\psi[\{(\lambda_1, \lambda_1)\}] &= \frac{1}{2} \cdot \frac{(\alpha - \beta)^2}{2} \\
\mathbb{P}_\psi[\{(\lambda_1, \lambda_2)\}] &= \frac{1}{2} \cdot \frac{(\alpha + \beta)^2}{2} \\
\mathbb{P}_\psi[\{(\lambda_2, \lambda_1)\}] &= \frac{1}{2} \cdot \frac{(\alpha - \beta)^2}{2} \\
\mathbb{P}_\psi[\{(\lambda_2, \lambda_2)\}] &= \frac{1}{2} \cdot \frac{(\alpha + \beta)^2}{2}
\end{aligned}$$

will work.

### Tensor Product

We will need to use the tensor product $\mathbb{C}^2 \otimes \mathbb{C}^2$ in the following examples. Tensor products are important in describing multi-particle systems in physics and situations where two systems interact. In our examples they describe two spin measurements, done on the left and the right.

One way to describe the tensor product is with bases. If $\{e_1, e_2\}$ is a basis for the copy of $\mathbb{C}^2$ on the left and $\{f_1, f_2\}$ is a basis for the copy of $\mathbb{C}^2$ on the right, then $\mathbb{C}^2 \otimes \mathbb{C}^2$ is a $4$ dimensional space with basis $\{e_1 \otimes f_1, e_2 \otimes f_1, e_2 \otimes f_1, e_2 \otimes f_2\}$. The inner product is defined on basis vectors as $\langle e_i \otimes f_j, e_k \otimes f_l \rangle = \langle e_i, e_k \rangle \langle f_j, f_l \rangle$ and is then extended by linearity to the whole space. Similarly, if $A$, $B$ are operators on $\mathbb{C}^2$ (i.e., $2 \times 2$ matrices) then $A \otimes B$ acts on basis vectors via $(A \otimes B)(e_i \otimes f_j) = Ae_i \otimes Bf_j$ and can be extended by linearity to the whole space.

One way to think about the tensor product is to think of the elements of $\mathbb{C}^2$ on the left as column vectors $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ and elements of $\mathbb{C}^2$ on the right as row vectors $[b_1, b_2]$. Then $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \otimes [b_1, b_2]$ corresponds to the $2 \times 2$ matrix $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} [b_1, b_2] = \begin{bmatrix} a_1 b_1 & a_1 b_2 \\ a_2 b_1 & a_2 b_2 \end{bmatrix}$ and $\mathbb{C}^2 \otimes \mathbb{C}^2$ can be identified with the space of $2 \times 2$ matrices. Product states, that is, states of the form $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \otimes [b_1, b_2]$ correspond to matrices of rank 1. However a generic $2 \times 2$ has rank 2 and cannot be factored in this way. Such a state is called entangled.

Recall that if $A$, $B$ are operators on $\mathbb{C}^2$ then $A \otimes B$ acts on $\mathbb{C}^2 \otimes \mathbb{C}^2$. If we represent a vector in $\mathbb{C}^2 \otimes \mathbb{C}^2$ as a $2 \times 2$ matrix $M$, then the action of $A \otimes B$ on $M$ is $AMB^T$.

### Hardy's example

Now we present Hardy's example of a system where a random variable description is not possible. This system models spin measurement for two particles in an entangled state. The Hilbert space for this system is $\mathbb{C}^2 \otimes \mathbb{C}^2$. We can imagine the two particles as being separated in space, one on the left and one on the right, having been sent there after the state is prepared in some central location. Let $A$ and $B$ be the spin matrices defined above. We consider the four observables given by

$$A^L = A \otimes I$$

$$A^R = I \otimes A$$

$$B^L = B \otimes I$$

$$B^R = I \otimes B$$

and notice that $A^L$ commutes with $A^R$ and $B^R$. This agrees with the idea that measurements on the left do not interfere with measurements on the right. Thus we may simultaneously measure $A^L$ and $A^R$, and (in a different experiment) simultaneously measure $A^L$ and $B^R$. Similarly $A^R$ commutes with $A^L$ and $B^L$, etc.

We have four bases for $\mathbb{C}^2 \otimes \mathbb{C}^2$ at our disposal, namely $\{\phi_i \otimes \phi_j\}$, $\{\psi_i \otimes \phi_j\}$, $\{\phi_i \otimes \psi_j\}$, and $\{\psi_i \otimes \psi_j\}$, for $i, j = 1, 2$. If we expand a state $\Psi \in \mathbb{C}^2 \otimes \mathbb{C}^2$ in the first of these as

$$\Psi = \sum_{i,j} c_{i,j} \phi_i \otimes \phi_j$$

and write the coefficients as a matrix

$$C = \begin{bmatrix} c_{1,1} & c_{1,2} \\ c_{2,1} & c_{2,2} \end{bmatrix},$$

then the corresponding coefficient matrices for the other basis are $UC$ for $\{\psi_i \otimes \phi_j\}$, $CU^T$ for $\{\phi_i \otimes \psi_j\}$ and $UCU^T$ for $\{\psi_i \otimes \psi_j\}$.

To calculate the quantum probability that a simultaneous measurement of $A^L$ and $A^R$ for two particles in the state $\Psi$ yields $A^L = \lambda_1$ and $A^R = \lambda_1$ we must take the product of the corresponding spectral projections for $A^L$ and $A^R$ (which is again a projection since $A^L$ and $A^R$ commute)

$$(P_1 \otimes I)(I \otimes P_1) = P_1 \otimes P_1.$$

Then the quantum probability is given by

$$\mathbf{P}[A^L = \lambda_1 \textbf{ and } A^R = \lambda_1] = |\langle \Psi, P_1 \otimes P_1 \Psi \rangle|^2 = |c_{1,1}|^2.$$

For our example we take the coefficient matrix for the first basis to be

$$C = \frac{1}{N} \begin{bmatrix} 0 & -\alpha \\ -\alpha & \beta \end{bmatrix},$$

where $N$ is the normalization constant $N = \sqrt{2\alpha^2 + \beta^2} = \sqrt{1 + \alpha^2}$. Then the corresponding coefficient matrices for the other bases are

$$\begin{aligned} UC &= \frac{1}{N} \begin{bmatrix} \alpha\beta & -1 \\ -\alpha^2 & 0 \end{bmatrix} \\ CU^T &= \frac{1}{N} \begin{bmatrix} \alpha\beta & -\alpha^2 \\ -1 & 0 \end{bmatrix} \\ UCU^T &= \frac{1}{N} \begin{bmatrix} \alpha^2\beta + \beta & \alpha\beta^2 - \alpha \\ -\alpha^2 & -\beta\alpha^2 \end{bmatrix} \end{aligned}$$

The key point is the position of the zero entries, and the fact that the $2, 2$ entry of the final matrix is non-zero. These lead to the quantum probabilities

$$\begin{aligned} &\mathbf{P}[A^L = \lambda_1 \textbf{ and } A^R = \lambda_1] = 0 \\ &\mathbf{P}[B^L = \lambda_2 \textbf{ and } A^R = \lambda_2] = 0 \\ &\mathbf{P}[A^L = \lambda_2 \textbf{ and } B^R = \lambda_2] = 0 \\ &\mathbf{P}[B^L = \lambda_2 \textbf{ and } B^R = \lambda_2] = \frac{\alpha^4\beta^2}{N^2} > 0 \end{aligned}$$

We now suppose that there is an underlying probability space such that

$$\begin{aligned}
\mathbb{P}[\{a^L = \lambda_1\} \cap \{a^R = \lambda_1\}] &= 0 \\
\mathbb{P}[\{b^L = \lambda_2\} \cap \{a^R = \lambda_2\}] &= 0 \\
\mathbb{P}\{[a^L = \lambda_2\} \cap \{b^R = \lambda_2\}] &= 0 \\
\mathbb{P}[\{b^L = \lambda_2\} \cap \{b^R = \lambda_2\}] &= \frac{\alpha^4\beta^2}{N^2} > 0
\end{aligned}$$

and derive a contradiction. Here is the key point. When we assume that a probability space exists, we can write the set $\{a^L = \lambda_1\} \cap \{a^R = \lambda_1\}$ as a disjoint union

$$\begin{aligned}
&\{a^L = \lambda_1\} \cap \{a^R = \lambda_1\} = \\
&\qquad \{a^L = \lambda_1\} \cap \{b^L = \lambda_1\} \cap \{a^R = \lambda_1\} \cup \{a^L = \lambda_1\} \cap \{b^L = \lambda_2\} \cap \{a^R = \lambda_1\}
\end{aligned}$$

This is saying is that even though we don't know the value of $b^L$, it has some value, and we can decompose an event into a disjoint union of events depending on that value. We can do this one more time with the value of $b^R$ to produce the disjoint union

$$\begin{aligned}
&\{a^L = \lambda_1\} \cap \{a^R = \lambda_1\} = \\
&\qquad \{a^L = \lambda_1\} \cap \{b^L = \lambda_1\} \cap \{a^R = \lambda_1\} \cap \{b^R = \lambda_1\} \cup \\
&\qquad \{a^L = \lambda_1\} \cap \{b^L = \lambda_1\} \cap \{a^R = \lambda_1\} \cap \{b^R = \lambda_2\} \cup \\
&\qquad \{a^L = \lambda_1\} \cap \{b^L = \lambda_2\} \cap \{a^R = \lambda_1\} \cap \{b^R = \lambda_1\} \cup \\
&\qquad \{a^L = \lambda_1\} \cap \{b^L = \lambda_2\} \cap \{a^R = \lambda_1\} \cap \{b^R = \lambda_2\} \cup
\end{aligned}$$

Now let

$$p(i,j,k,l) = \mathbb{P}\left[\{a^L = \lambda_i\} \cap \{b^L = \lambda_j\} \cap \{a^R = \lambda_k\} \cap \{b^R = \lambda_l\}\right]$$

Then we can rewrite our equations as

$$\begin{array}{ccccccccc}
p(1,1,1,1) & + & p(1,1,1,2) & + & p(1,2,1,1) & + & p(1,2,1,2) & = 0 \\
p(1,2,2,1) & + & p(1,2,2,2) & + & p(2,2,2,1) & + & p(2,2,2,2) & = 0 \\
p(2,1,1,2) & + & p(2,1,2,2) & + & p(2,2,1,2) & + & p(2,2,2,2) & = 0 \\
p(1,2,1,2) & + & p(1,2,2,2) & + & p(2,2,1,2) & + & p(2,2,2,2) & = \alpha^4\beta^2/N^2 > 0
\end{array}$$

But the $p(i,j,k,l)$ are non-negative numbers. So all the ones that appear in the first three equations are zero. But this includes all all the terms in the last equation. So they cannot add up to something positive. This contradiction shows that the probability space cannot exist.

### Bell's second theorem: non-locality

This section follows Faris' appendix in the book by Wicks. He includes a more careful explanation of the needed concepts from probability aimed at non-mathematicians.

Since we have to set up different experiments for each pair of measurements in the example above, maybe the probability measure which reproduces the quantum results depends on which experiment we are performing. In other words we might have a probability space $\Omega$ and $\sigma$-algebra $\mathcal{F}$ with four different probability measures $\mathbb{P}_{AA}$, $\mathbb{P}_{AB}$, $\mathbb{P}_{BA}$ and $\mathbb{P}_{BB}$ depending on which spin direction we are measuring on the left and on the right.

In this case there are some natural restrictions if we want to disallow action at a distance. Measurements which only involve the equipment on the left should not be influenced by what happens on the right, and vice versa. Furthermore, simultaneous measurements on the left and right should be conditionally independent, if we take into account events in the preparation stage. The idea is that events in the preparation stage in the experiment, for example preparing a singlet state where the spins on the left and right always have opposite values, can result in correlations for the measurements on the left and the right. However, if we factor out these effects, the residual randomness leads to independent outcomes on the left and right.

To put this in a mathematical form we need some more ideas from probability theory. We assume that there are three sub $\sigma$-algebras of $\mathcal{F}$, denoted $\mathcal{F}^L$, $\mathcal{F}^R$ and $\tilde{\mathcal{F}}$. The events in $\mathcal{F}^L$ describe what happens on the left. Knowing whether a experimental outcome is contained in an event from $\mathcal{F}^L$ only tells us about measurements performed on the left. An example would be $\{a^L = 1\}$. Similarly, the events in $\mathcal{F}^R$ are associated with what happens on the right. The events in $\tilde{\mathcal{F}}$ are associated what happened in the preparation stage of the experiment. Assume we are given four probability measures $\mathbb{P}_{AA}$, $\mathbb{P}_{AB}$, $\mathbb{P}_{BA}$ and $\mathbb{P}_{BB}$ such that

1. If $E \in \mathcal{F}^L \cup \tilde{\mathcal{F}}$ then $\mathbb{P}_{AA}(E) = \mathbb{P}_{AB}(E)$ and $\mathbb{P}_{BA}(E) = \mathbb{P}_{BB}(E)$.

In other words, events on the left or from the preparation stage have probabilities which don't change when we move around the equipment on the right.

2. Similarly, if $E \in \mathcal{F}^R \cup \tilde{\mathcal{F}}$ then $\mathbb{P}_{AA}(E) = \mathbb{P}_{BA}(E)$ and $\mathbb{P}_{AB}(E) = \mathbb{P}_{BB}(E)$.

The third assumption expresses the conditional independence of events on the right and left. Recall that if $E$ is an event and $G$ is another event with $\mathbb{P}(G) \neq 0$ then the conditional probability of $F$ given $G$ is the number $\mathbb{P}[E|G] = \dfrac{\mathbb{P}(E \cap G)}{\mathbb{P}(G)}$. If $\tilde{\mathcal{F}}$ is a sub $\sigma$-algebra of $\mathcal{F}$ then the conditional probability $\mathbb{P}[E|\tilde{\mathcal{F}}]$ is not a number, but a random variable. The idea (which is actually correct if $\Omega$ is finite and there are no non-empty events with zero probability) is that

$$\mathbb{P}[E|\tilde{\mathcal{F}}](\omega) = \mathbb{P}[E|G_\omega] \quad \text{where} \quad G_\omega = \bigcap_{G \in \tilde{\mathcal{F}} : \omega \in G} G.$$

In other words, $\mathbb{P}[E|\tilde{\mathcal{F}}](\omega)$ is the probability of $E$ given all the information in $\tilde{\mathcal{F}}$ about $\omega$. The situation is more complicated when there are set of measure zero and we can't talk about the value of a random variable at a point. We won't worry about this. What we use the fact that $\mathbb{E}[\mathbb{P}[E|\tilde{\mathcal{F}}]] = \mathbb{P}[E]$. The third assumption is

3. If $E \in \mathcal{F}^L$ and $F \in \mathcal{F}^R$ then $P_{xy}[E \cap F|\tilde{\mathcal{F}}] = P_{xy}[E|\tilde{\mathcal{F}}]\ P_{xy}[F|\tilde{\mathcal{F}}]$. Here $x$ and $y$ are any choice of $A, B$.

These assumptions are reasonable if we don't believe in action at a distance. Yet we can show that there can be no hidden variable description of the experiment unless they are violated.

**Proposition 4.1** *Let the matrices $A$ and $B$ and the coefficients $\alpha$, $\beta$ and $N$ be as defined in the previous section. Suppose that there are probability measures $\mathbb{P}_{AA}$, $\mathbb{P}_{AB}$, $\mathbb{P}_{BA}$ and $\mathbb{P}_{BB}$ and sub $\sigma$–algebras $\mathcal{F}^L$, $\mathcal{F}^R$ and $\tilde{\mathcal{F}}$ satisfying the assumptions above. Then there are no random variables $a^L$, $a^R$, $b^L$, $b^R$ such that*

$$\begin{aligned}
\mathbb{P}_{AA}[\{a^L = \lambda_1\}\ \cap\ \{a^R = \lambda_1\}] &= 0 \\
\mathbb{P}_{BA}[\{b^L = \lambda_2\}\ \cap\ \{a^R = \lambda_2\}] &= 0 \\
\mathbb{P}_{AB}[\{a^L = \lambda_2\}\ \cap\ \{b^R = \lambda_2\}] &= 0 \\
\mathbb{P}_{BB}[\{b^L = \lambda_2\}\ \cap\ \{b^R = \lambda_2\}] &= \frac{\alpha^4\beta^2}{N^2} > 0
\end{aligned}$$

*Proof:* If the probability of an event is zero, then so is the conditional probability. Thus the first three equations, together with the conditional independence, imply

$$\begin{aligned}
\left(\mathbb{P}_{AA}[\{a^L = \lambda_1\}|\tilde{\mathcal{F}}]\right)\left(\mathbb{P}_{AA}[\{a^R = \lambda_1\}|\tilde{\mathcal{F}}]\right) &= 0 \\
\left(\mathbb{P}_{BA}[\{b^L = \lambda_2\}|\tilde{\mathcal{F}}]\right)\left(\mathbb{P}_{BA}[\{a^R = \lambda_2\}|\tilde{\mathcal{F}}]\right) &= 0 \\
\left(\mathbb{P}_{AB}[\{a^L = \lambda_2\}|\tilde{\mathcal{F}}]\right)\left(\mathbb{P}_{AB}[\{b^R = \lambda_2\}|\tilde{\mathcal{F}}]\right) &= 0
\end{aligned}$$

almost everywhere. Then

$$\begin{aligned}
&\mathbb{P}_{BB}[\{b^L = \lambda_2\}\ \cap\ \{b^R = \lambda_2\}|\tilde{\mathcal{F}}] \\
&\quad= \left(\mathbb{P}_{BB}[\{b^L = \lambda_2\}|\tilde{\mathcal{F}}]\right)\left(\mathbb{P}_{BB}[\{b^R = \lambda_2\}|\tilde{\mathcal{F}}]\right) \\
&\quad= \left(\mathbb{P}_{BA}[\{b^L = \lambda_2\}|\tilde{\mathcal{F}}]\right)\left(\mathbb{P}_{AB}[\{b^R = \lambda_2\}|\tilde{\mathcal{F}}]\right) \\
&\quad= \left(\mathbb{P}_{BA}[\{b^L = \lambda_2\} \cap \{a^R = \lambda_1\}|\tilde{\mathcal{F}}] + \mathbb{P}_{BA}[\{b^L = \lambda_2\} \cap \{a^R = \lambda_2\}|\tilde{\mathcal{F}}]\right) \\
&\qquad\left(\mathbb{P}_{AB}[\{a^L = \lambda_1\} \cap \{b^R = \lambda_2\}|\tilde{\mathcal{F}}] + \mathbb{P}_{AB}[\{a^L = \lambda_2\} \cap \{b^R = \lambda_2\}|\tilde{\mathcal{F}}]\right) \\
&\quad= \left(\mathbb{P}_{BA}[\{b^L = \lambda_2\} \cap \{a^R = \lambda_1\}|\tilde{\mathcal{F}}]\right)\left(\mathbb{P}_{AB}[\{a^L = \lambda_1\} \cap \{b^R = \lambda_2\}|\tilde{\mathcal{F}}]\right) \\
&\quad\leq \mathbb{P}_{BA}[\{a^R = \lambda_1\}|\tilde{\mathcal{F}}]\mathbb{P}_{AB}[\{a^L = \lambda_1\}|\tilde{\mathcal{F}}] \\
&\quad= \mathbb{P}_{AA}[\{a^R = \lambda_1\}|\tilde{\mathcal{F}}]\mathbb{P}_{AA}[\{a^L = \lambda_1\}|\tilde{\mathcal{F}}] \\
&\quad= 0
\end{aligned}$$

almost everywhere. This implies

$$\mathbb{P}_{BB}[\{b^L = \lambda_2\}\ \cap\ \{b^R = \lambda_2\}] = \mathbb{E}[\mathbb{P}_{BB}[\{b^L = \lambda_2\}\ \cap\ \{b^R = \lambda_2\}|\tilde{\mathcal{F}}]] = 0,$$

which contradicts the last line in our original set of equations. □

## References


[A] V.I. Arnold, *Mathematical Methods of Classical Mechanics, 2nd Ed.*, Springer 1997

[B] J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics,* Cambridge University Press, 2004.

[CFKS] H.L. Cycon, R.G. Froese, W. Kirsch and B. Simon, *Schrödinger Operators, with Applications to Quantum Mechanics and Global Geometry*, 2nd ed., Springer, 2007

[DG] J. Derezinski and C. Gerard, *Scattering Theory of Classical and Quantum N-Particle Systems*, Springer, 1997

[ES] J.P. Eckmann and R. Seneor, *The Maslov-WKB method for the (an-)Harmonic Oscillator*, Archive for Rational Mechanics and Analysis, June 1976, Volume 61, Issue 2, pp 153-173.

[F1] William Faris, *Outline of Quantum Mechanics*, Contemporary Math. Vol 529, 2010.

[F2] William Faris, Appendix to Wick's book [W].

[FTZWF] Maria Fuwa, Shuntaro Takeda, Marcin Zwierz, Howard M. Wiseman & Akira Furusawa, *Experimental proof of nonlocal wavefunction collapse for a single particle using homodyne measurements,* Nature Communications, 2015 DOI: 10.1038/ncomms7665

[H] Lucien Hardy, *Nonlocality for two particles without inequlalities for almost all entangled states,* Phys. Rev. Letters 71 (1993) 1665–1668

[HS] P.D. Hislop and I.M. Sigal, *Introduction to Spectral Theory: With Applications to Schrödinger Operators,* Springer, 1997

[MR] J.E. Marsden and T.S. Ratiu, *Introduction to Mechanics and symmetry*, Springer, 1999

[RSIV] M. Reed and B. SImon, *Methods of Modern Mathematical Physics, vol. IV, Analysis of operators,* Academic Press, 1978

[S2] B. Simon, *A Comprehensive Course in Analysis, Volume 2,* AMS, 2015

[W] David Wick, *The Infamous Boundary: Seven Decades of Heresy in Quantum Physics*, Copernicus, 1995.