Classical, Quantum & Statistical Mechanics

Gavin Conran

Introduction

This is a brief revision/background paper on Classical, Quantum and Statistical mechanics. It introduces some of the giants of Physics and Mathematics: Newton; Leibniz, Lagrange; Hamilton; Markov; Boltzmann; Feynman and Schrödinger plus others.

The document starts with Classical Mechanics, which mathematically describes the movement of the macroscopic, e.g. golf balls and even heavenly bodies, where both Newton's, especially his 2nd law resulting in a differential equation of motion, and Hamilton's formulations are discussed. Newton actually invented Calculus to mathematically describe his mechanics. Leibnez also lays claim to be the father of calculus, which was fiercely contested by Newton and his supporters, but whose notation is more widely used today and is used through out this document.

Leibniz	***Lagrange***	***Euler***	***Newton***
$\frac{dy}{dx}, \frac{d^2y}{dx^2}, \frac{d^ny}{dx^n}$	$f', f'', f^{(n)}$	$\frac{df}{dx}, \partial_{xx} f = \frac{\partial^2 f}{\partial x^2}$	$\dot{y}, \ddot{y}, \dddot{y}$

Table 1: Notation for Differentiation

Together with what we learnt from Gauss and Fourier in earlier papers, the Hamiltonian is brought forward to the world of Quantum Mechanics and wave-particle duality; the study of the mechanics of the atomic, where the Schrödinger Equation, another differential equation of motion, describes the internal dynamics of particle physics, is introduced.

The paper then moves onto a discussion on Statistical Mechanics, where statistically ('probablistically' is, probably, a better word to describe the process) simulating the atomic or molecular stimulates macroscopic phenomena, for both classical and quantum mechanics. Statistical methods also provides a convenient way to conduct multi-dimensional integration and an optimisation algorithm through simulated annealing.

The science finishes with Lagrange Multipliers, a mathematical technique used in optimisation, especially in Economics, and Jacobian and Hessian matrices; tools used generally when combining linear algebra and calculus. It is worth pointing out that Hamilton derived his mechanics from Lagrange rather than directly from Newton but as they are not widely used, Lagragian Mechanics are not discussed here. It is also worth pointing out that Optimisation problems can be solved using Graph algorithms, such as those created by Dijkstra, or Linear Programming.

As always, the document finishes with a discussion on applications enabled by the mathematics and physics discussed in the paper and, again, as usual, relies on Finite Difference Methods, this time a simple time-step technique from Euler, to provide us with a numerical solution to the Schrödinger equation which actually shows quantum tunnelling. This time though we have the addition of Monte Carlo algorithms to provde us with a statistical, approximation technique that allows us to compute Richard Feynman's Path Integrals. This means that over the three revision papers we have covered the main numerical approximation techniques: Fourier; Taylor (FDM) and Monte Carlo.

Classical (Newtonian) Mechanics

Classical Mechanics models real world objects as Point Particles. It is worth noting that when modelling very samll particles, such as electrons, Quantum Mechanics is used. The motion of the Point Particle is characteristed by a small number of parameters:

- Position of the particle
- Mass of the particle
- Forces applied to the particle

With the following assumptions:

- Matter and energy have definite, knowable attributes, such as, location and speed
- Forces act instantaneously

Position

The position of a Point particle is defined in relation to a Coordinate System centred on an arbitrary fixed reference point in space, 0, the origin.

1. When particle, p, is moving relative to 0, the vector, r, is defined as a function of time, t. This is referred to as Absolute Time which is the same for all observers, i.e. it is non-relativistic.
2. Assumes Euclidean Geometry for the structure of space.

Velocity and Speed

Velocity, or the rate of change of the position wrt time, is defined as the derivative of the position wrt time:

$$v = \frac{dr}{dt}$$

Velocities are directly additive as Vector Quantities and so they must be dealt with using Vector Analysis.

Acceleration

Acceleration is the rate of change of Velocity wrt time

$$a = \frac{dv}{dt} = \frac{d^2 r}{dt^2}$$

Newton's Laws of Motion

1st Law

In an internal frame of reference, an object either remains at rest or continues to move at constant velocity unless acted upon by a Force.

2nd Law

Newton's 2nd Law is an expression between Force and Momentum

$$F = \frac{dp}{dt} = \frac{d(mv)}{dt}$$

where mv is called the (canonical) momentum

The net force on a particle is thus equal to the rate of change of the momentum of the particle wrt time. Since

$$a = \frac{dv}{dt}$$

the 2nd law can be written:

$$F = ma$$

where

F is the Force acting on the particle
m is the mass of the particle
a is the acceleration of the particle

As long as the force acting on a particle is known, Newton's 2nd law is sufficient to describe the motion of the particle. Once independent relations between each force acting on a particle are available, they can be substituted into Newton's 2nd law to obtain an Ordinary Differential Equation which is called the ***Equation of Motion***.

Example: Assume friction is the only force acting on the particle and that it may be modelled as a function of the velocity of the particle. If

$$F_R = -\lambda v$$

where

λ is apositive constanr
and the -ive sign indicates the force is opposite to the velocity

Then the ***Equation of Motion*** (ODE) is

$$-\lambda v = ma = m\frac{dv}{dt}$$

This can be integrated to give:

$$v = v_0 e^{-\frac{\lambda t}{m}}$$

where v_0 is the Initial Velocity

This means that the velocity decays exponentialy to zero ar $t \to \infty$. The Conservation of Energy interpretation is that the KE of the particle is absorbed by the Friction (which converts to Heat Energy).

This equation can be further integrated to obtain the position, r, of the particle as a function of time, t:

3rd Law

When a body exerts a force on a 2nd body, the 2nd body simultaneously exerts an opposite and equal force on the 1st body.

Frames of Reference

- Classical Mechanics assumes the existence of a family of Reference Frames in which the mechanical laws of nature take a comparatively simple form, called ***Inertial Frames***.
- An Inertial Frame is a frame of reference within which an object interacting with no forces appears either at rest or moving uniformly in a straight line. These are characterised by the

requirement that all forces entering the observer's physical laws originate from identifiable sources caused by Fields, such as:

- Electro-Static Field (caused by Static Electrical charges)
- Electro-Magnetic Field (caused by moving Electrical charges)
- Gravitational Field (caused by mass)

Work & Energy

Work Done

If a constant force, F, is applied to a particle that makes a displacement, Δr, the ***Work Done*** by the Force is defined by the scalar product of the Force and Displacement vectors:

$$W = F \cdot \Delta r$$

More generally, if the ***force varies as a function of position*** as the particle moves from r_1 to r_2 along a path, C, the Work Done on the particle is given by the Line Integral:

$$W = \int_C F(r) \cdot dr$$

If the Work Done is the same, no matter the path taken, the ***Force is said to be Conservative***, e.g. Gravity is a Conservative Force but Friction is non-conservative.

Kinetic Energy

The Kinetic Energy, E_k, of a particle of mass, m, travelling at a speed, v, is give by:

$$E_k = \frac{1}{2} m v^2$$

The KE of a composite body of many particles is the sum of the KEs of the particles.

Work Energy Theorem

For a particle of constant mass, m, the Total Work, W, done on the particle as it moves from r_1 to r_2 is equal to the change in Kinetic Energy, E_k, of the particle.

$$W = \Delta E_k = E_{k,2} - E_{k,1} = \frac{1}{2} m \left(v_2^2 - v_1^2 \right)$$

Potential Energy

Conservative Forces can be expressed as the ***Gradient*** of a scalar function, known as the Potential Energy, E_p :

$$F = -\nabla E_p$$

The decrease in the Potential Energy is equal to the increase in Kinetic Energy:

$$-\Delta E_p = \Delta E_k = \Delta \left(E_p + E_k \right) = 0$$

The result is known as the ***Conservation of Energy*** and states that

$$\sum E = E_k + E_p$$

is a constant in time.

Classical (Hamiltonian) Mechanics

Hamiltonian Mechanics is a reformulation of Newtonian Classical Mechanics which predicts the same outcomes as non-Hamiltonian Mechanics. It later contributed to the formulation of Statistical and Quantum Mechanics.

In Hamiltonian Mechanics a classical physical system is described as a set of canonical coordinates, $r=(q,p)$, where each component of the coordinate q_i , p_i is indexed to the Frame of Reference of the system.

The Time Evolution of the system is uniquely defined by ***Hamilton's Equations***:

$$\frac{dp}{dt}=-\frac{\partial H}{\partial q} \qquad \frac{dq}{dt}=+\frac{\partial H}{\partial p}$$

where

$H=H(q,p,t)$ is the ***Hamiltonian*** which corresponds to the ***Total Energy of the system***. For a closed system it is the sum of the Kinetic and Potential Energies.

Example: 1-D system consisting of one particle of mass, m.

The Hamiltonian represents the Total Energy of the system, which is the sum of the Kinetic and Potential Energies, T and V, respectively. Here:

- q is the space coordinate
- p is the momentum, mv

Then

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

Note:

- T is a function of p alone
- V is a function of q alone

In this example:

First Hamilton Equation (Momentum):
The Time derivative of the Momentum, p, equals the Newtonian Force and so the Force equals the -ive gradient of the Potential Energy

$$\frac{dp}{dt}=-\frac{\partial H}{\partial q} \qquad F=-\nabla E_p$$

Hamilton ***Newton (equivalent)***

Second Hamilton Equation (Velocity):
The Time derivative of q is the velocity which equals the Kinetic Energy wrt momentum

$$\frac{dq}{dt}=+\frac{\partial H}{\partial p} \qquad E_k=\frac{1}{2}mv^2$$

Hamilton ***Newton (equivalent)***

Newtonian vs. Hamiltonian Mechanics

Newton	***Hamilton***
• Time evolution is obtained by computing the total force being exerted on each particle of the system • From Newton's 2nd law the time-evolution of both position and velocity are computed	• Time evolution is obtained by computing the Hamiltonian of the system in the Generalised Coordinates and inserting it into Hamilton's Equations • Real strength is in complex dynamical systems, such as planatery orbits in celestical mechanics. The more degrees of freedom the system has, the more complicated its time evolution → Chaotic Systems!!!

Quantum Mechanics

Quantum Mechanics describes nature at the smallest scales of energy levels of atoms and subatomic particles. Classical Physics describes nature at ordinary (macroscopic) scale. Quantum Mechanics differs from Classical Physics in that

- Energy, Momentum, Angular Momentum and other quantities of a bound system are restricted to discrete values, i.e. ***Quantisation***
- Objects have characteristics of both particles and waves, i.e. ***Wave-Particle Duality***
- There are limits to the precision with which quantities can be measured, i.e. ***Uncertainty Principle***

Overview of Postulates of Quantum Mechanics

1st Postulate: Wave Function Postulate

- The ***state of the system*** is completely characterised by a wave function, $\Psi(\vec{r},t)$
- The ***probability of finding either a particle*** or, if we have a many particle system, in a small volume element, d^3r , centred around the position, r, at time, t, is given by:
  - taking the complex conjugate of the wave function,
  - multiplying it by the wave function,
  - and then multiplying by the volume element

$$\rho(\vec{r},t)=\Psi^*(\vec{r},t)\Psi(\vec{r},t)d^3r$$

where

$$\Psi^*(\vec{r},t)=\Re[\Psi(\vec{r},t)]-i\Im[\Psi(\vec{r},t)]$$

2nd Postulate: Observables

- How do we define quantities that we want to observe, such as, momentum of a particle, the position, or the energy?
- The 2nd postulate says that to every observable quantity there corresponds a linear, self-adjoint operator
- We use the generic symbol, $\hat{L}$, for observables
- We are unable to determine all the observables at the same time with infinite precision → ***Heisenberg Uncertianty Principle***
- To every classical observable there corresponds a quantum observable:
  - Classical: $x(t)$; Quantum: $\hat{x}\Psi(\vec{r},t)$
  - Classical: $p(t)$; Quantum: $\hat{p}\Psi(\vec{r},t)$
  - Classical: E ; Quantum: $\hat{H}\Psi(\vec{r},t)$

- Observables are EigenFunctions and EigenValues
- The EigenFunctions are Harmonics (same as Fourier Series) of the Energy function, i.e the Hamiltonian
- The EigenValues are the Energy Levels of the Quantum Mechanical System

3rd Postulate: Measurements

- Concerns how we make measurements of these observables
- In any measurement of the observable, $\hat{L}$, the only possible values we will ever observe are the eigenvales, λ , of the operator, $\hat{L}$, which satisfy the eigenvalue equation, $\hat{L}U_\lambda(x)=\lambda_\lambda U_\lambda(x)$, where U_λ is an EigenFunction.
- As with classical waves, an arbitrary Wave Function can be written as an expression over the set of Complete and Orthogonal EigenFunctions:

$$\Psi(\vec{r},t)=\sum_\lambda C_\lambda(t)\,U_\lambda(x)$$

- So arbitrary wave functions actually conatin a Superposition of many EigenFunctions. Henceforth, their observable will also conatin a Superposition of many different EigenValues.
- Upon Measurement, the Wave Function of the system will ***Collapse***, i.e. it will select an EigenValue and EigenFunction out of its Superposition, an EigenState. This means that Measurement affects the State. Recall that in Classical Mechanics, measurement is non-invasive.

4th Postulate: Expectation Values

- Given that the system is characterised by a state, Ψ , the average of many measurements of the observable, $\hat{L}$, at time, t, is given by:

$$\langle\hat{L}\rangle_t=\int_D \Psi^*(\vec{r},t)\,\hat{L}\,\Psi(\vec{r},t)\,d^3r$$

5th Postulate: Time-Dependent Schrodinger Equations

- State of the system, $\Psi(\vec{r},t)$, is obtain ed by solving the following 1st order ODE:

$$i\hbar\frac{\partial}{\partial t}\Psi(\vec{r},t)=\hat{H}\Psi(\vec{r},t)$$

where
i: square root of -1
$\hbar$: Plank's Constant divided by 2π
$\hat{H}$: Hamiltonian Operator

The solution to the Schrödinger Equation, $\Psi(\vec{r},t)$, is given by:

$$\Psi(\vec{r},t)=\sum_E C_E(0)\,e^{-\frac{iEt}{\hbar}}U_E(x)$$

where

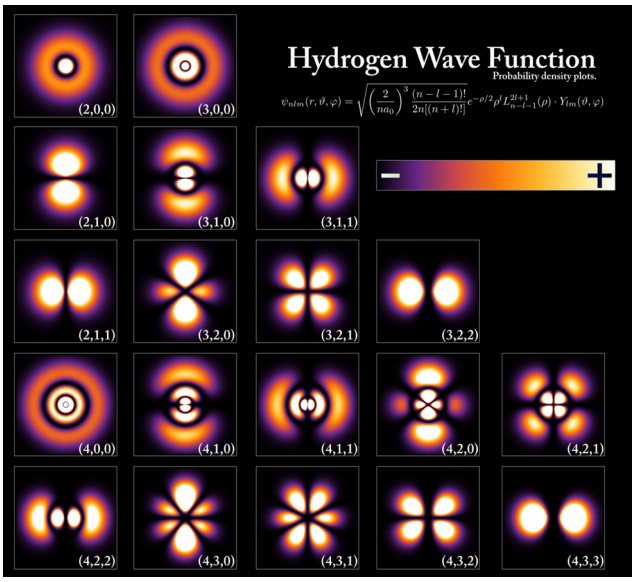
E : Energy Level
C_E : Time Dependent Coefficient, giving the EigenValue
U_E : EigenFunction (Basis Function)

- ***Schrödinger Equation*** is the Equation of Motion for the Wave Function in the same way that ***Newton's 2nd Law*** was the Motion of Equation for the position or the state of the system of a particle in Classical Mechanics
- A special observable quantity is the Energy and is given by the Hamiltonian Operator, $\hat{H}$

Classical vs. Quantum Mechanics

	Classical	Quantum
Initial Conditions	$\{\vec{r}(0), \vec{p}(0)\}$	$\Psi(\vec{r}, 0)$
Dynamical Variables	$\vec{r}(t), \vec{p}(t)$	$\Psi(\vec{r}, t)$
Equations of Motion	$\{\frac{d\vec{p}}{dt} = -\nabla_{\vec{r}} H, \frac{d\vec{r}}{dt} = \nabla_{\vec{p}} H\}$	$i\hbar \frac{\partial}{\partial t} \Psi(\vec{r}, t) = \hat{H} \Psi(\vec{r}, t)$

Table 2: Equations of Motion for Classical and Quantum Mechanics



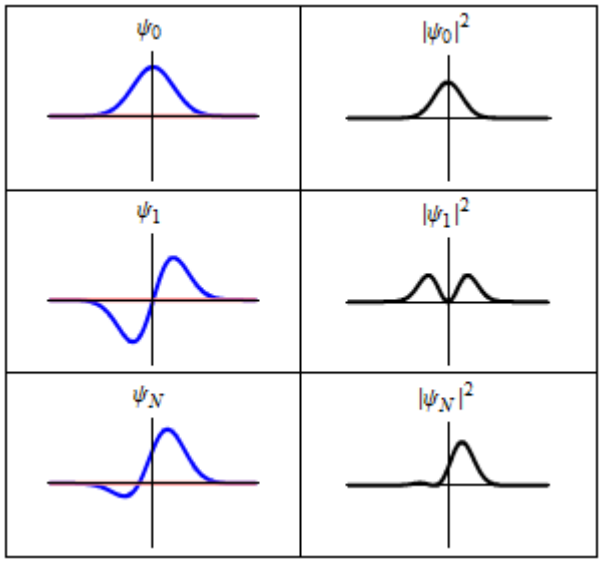


Table 3: LHS: Wave functions of the electron of a hydrogen atom at different energies. The brightness at each point represents the probability of observing the electron at that point.

RHS: Left: The real part (blue) and imaginary part (red) of the wave function. Right: The probability distribution of finding the particle with this wave function at a given position

Classical	Quantum
Describing the State of the System	
• In Classical Mechanics the state of the system is fully characterised by the ***position, x(t)***, or the ***momentum, p(x)*** • Occasionally, we get probabilities from incomplete information	• QM & 1st Postulate says that the state of the system is completely characterised by a ***Wave Function,*** $\Psi(\vec{r}, t)$, & the corresponding ***Probability Density,*** $\rho(\vec{r}, t)$
Quantum Tunnelling	
• When a ball is rolled slowly up a hill, it will come to a stop and roll back, because it does not have enough energy to get over the top of the hill to other side	• The Schrödinger equation predicts that there is a small probability that the ball will get to the other side of the hill, even if it has too little energy to reach the top.
Measurement	
• Non-invasive	• Affects the state

Statistical Mechanics

Statistical Mechanics shows how the concepts from macroscopic observations (such as temperature and pressure) are related to the description of microscopic state that fluctuates around an average state. To help summarise the difference between classical and quantum particles; classical particles are related to points but, analogously, quantum particles are related to one-dimensional objects, Path Integrals.

Classical Statistical Mechanics

Explicit numerical integration of Newton's equations and the Markov-chain Monte Carlo algorithm are used to illustrate mechanics and statistical mechanics, and to introduce the fundamental concepts of statistical mechanics:

- the Equiprobability Principle
  - Equivalence between Newton and Boltzmann Mechanics, called Ergodicity Hypothesis
  - Principle that configurations with the same enegy are equally probable
- the Boltzman Distribution of Energies
  - relates the probabilities, Π(a) and Π(b), of configurations, a and b, with different energies
  - $\Pi(a) \propto \exp(-\frac{E(a)}{kBT})$
- the Thermodynamic Temperature and Pressure
  - Consider constant temperature & pressure rather than constant temperature & volume

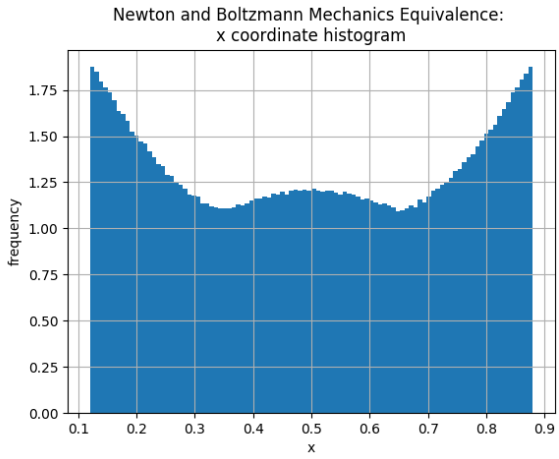
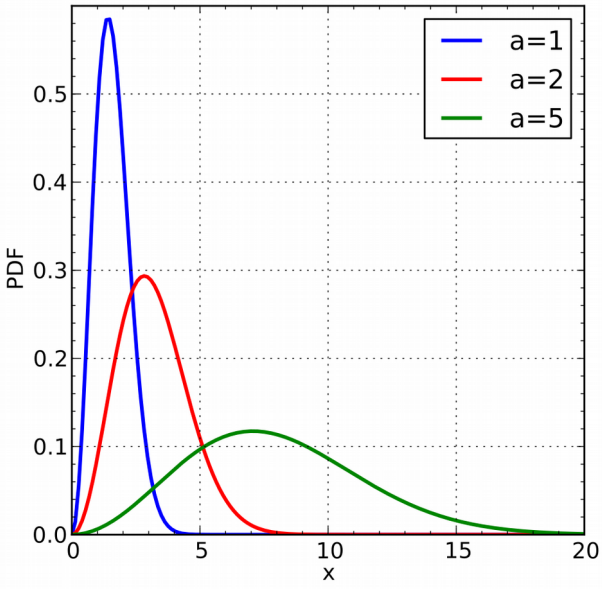
Equiprobability Principle	Boltzman Distribution
	

Table 4: Equiprobability Principle and Boltzmann Distribution

Quantum Statistical Mechanics

We now examine the field of Quantum Mechanics, where the ***Density Matrix*** generalises the classical ***Boltzmann Distribution***. The Density Matrix constructs the quantum statistical probabilities from their two origins:

- the Quantum Mechanical wave functions and
- the Boltzmann probabilities of the energy levels

The Density Matrix can thus be defined in terms of the complete solution of the quantum problem (wave functions and energies).

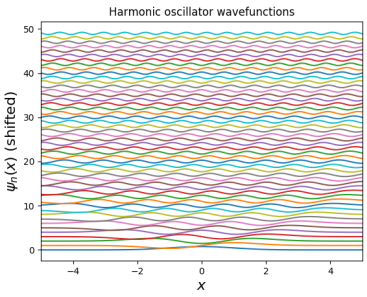
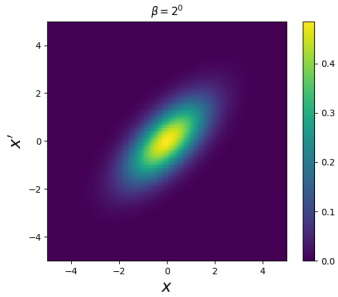
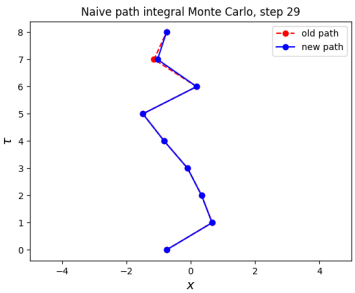
Wavefunctions (Schrödinger View)	***Density Matrix*** (Quantum Statistical Probabilities)	***Path Integrals*** (Feynman View)
		

Table 5: Wavefunctions, the Density Matrix and Path Integrals

To make the connection between wavefunctions on the one hand and path integrals on the other, we must provide the missing link, the ***Density Matrix*** that holds the path together. A convolution procedure, called ***Matrix Squaring***, is used to compute the quantum statistical probabilities for particles in an external potential. The convolution procedure, and the connection it establishes between the classical and quantum systems, is the basis of the ***Feynman Path Integral***.

Density Matrix

The ***Density Matrix*** , a matrix of the probabilities for a particle moving from position x to position x', takes the mathematical form:

$$\rho(x,x',\beta)=\sum_n e^{-\beta E_n}\psi(x)\psi^*(x')$$

where $\beta=\frac{1}{K_B T}$

Properties of the Density Matrix

1) Convolution: allows us to compute the density matrix at any given temperature from a product of two density matrices at higher temperatures

$$\rho(x,x'',2\beta)=\int dx_1\,\rho(x,x',\beta)\rho(x',x'',\beta)$$

2) Free Density Matrix: As β incresaes, the variance, σ^2 , of the Gaussian becomes larger and the system becomes more 'Quantum', i.e. more spread out.

$$\rho^{free}(x,x',\beta)=\frac{1}{\sqrt{(2\pi\beta)}}\exp\left(-\frac{(x-x')^2}{2\beta}\right)$$

3) Trotter Decomposition: This connects the Density Matrix to the Free Density Matrix

$$\rho(x,x',\beta)=e^{-\frac{\beta}{2}V(x)}\rho^{free}(x,x',\beta)e^{-\frac{\beta}{2}V(x')} \quad \text{for } \beta \to 0$$

We now have an EXPLICIT formula for $\rho(x,x',\beta)$ without solving the Schrodinger equation for any potential. The associated Density Matrix is shown in Illustration 2, LHS.

Feynman's Path Integral

Density matrices can be expressed as multiple integrals over path variables, so called Path Integrals. By using Markov-Chain Monte Carlo simulations we can obtain the Density Matrix, i.e. the description of quantum statistical probabilities, without solving the Schrödinger Equation.

In classical motion, only the start and end points that indicate the action are of importance whereas in Quantum Mechanics all the paths that connect the starting and the end points are important. Every possible path going from 'a' to 'b' contributes to the amplitude. The probability, ***P(a,b)***, to go from x_a at time, t_a, to x_b al later time, t_b, is the absolute square of the amplitude $P(a,b)=|K(b,a)|^2$.

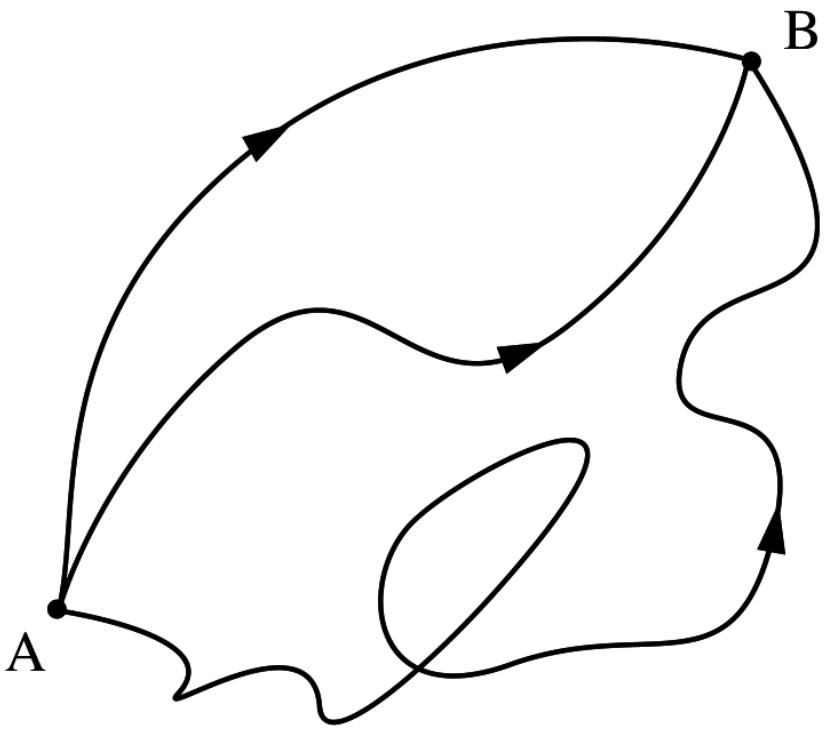


Illustration 1: The diagrammatic representation of some of the possible trajectories

The amplitude, ***K(b, a)***, is the sum of contribution ***ψ[x(t)]*** from each path

$$K(b,a)=\sum_{a\to b}\psi[x(t)]$$

The contribution of a given path has a phase which is propositional to action S.

$$\psi[x(t)]=A\cdot\exp(\frac{i}{\hbar}S[x(t)])$$

where A is a constant.

Metropolis-Hastings Algorithm

As we learnt in the Gauss paper, a Markov process (or chain) is a sequence of random variables such that the next value or state of the sequence depends only on the previous one. Monte Carlo integration is a method of numerically evaluating integrals. It simply consists of integrating over a random sampling of points instead of a regular array of points.

A Markov Chain Monte Carlo (MCMC) is a class of ***random walk*** algorithms commonly used for carrying out complicated numerical integration. This is useful for computing expectations of complicated high-dimension probability distributions.

The Metropolis-Hastings algorithm is a Markov Chain Monte Carlo (MCMC) method for obtaining a sequence of random samples from a probability distribution for which direct sampling is difficult. This sequence can be used to approximate the distribution, i.e. create a histogram, or to compute an integral, such as an expected value.

The algorithm allows for the construction of Markov Chains based on the proposal distribution $q(\cdot|X_t)$. It is a rejection sampling algorithm that only requires being able to evaluate the density of the target distribition $\Pi(\cdot)$.

It draws a sample from the proposal and accepts it with probability α. If the sample is rejected the chain remains unchanged at state t+1, otherwise it moves. A multivariate normal distribution usually provides a good proposal:

$$\alpha(\boldsymbol{X}_t, \boldsymbol{Y}) = min(1, \frac{\Pi(\boldsymbol{Y})q(\boldsymbol{X}_t|\boldsymbol{Y})}{\Pi(\boldsymbol{X}_t)q(\boldsymbol{Y}|\boldsymbol{X}_t)})$$

The associated Density Matrix is shown in Illustration 2, RHS.

With reference to the Path Integrals in Table 6 below, it is possible to move from one path configuration to the next by choosing one position, x_t , and making a little displacement, δx , that can be +ive or -ive. We compute the statistical weight after the move using the Metropolis Acceptance Probability. We can also move x_0 so the path can move as a whole.
It is worth noting that Path Integrals are also used to solve the Black Scholes PDE via the Feynman Kac equation, which is described below.

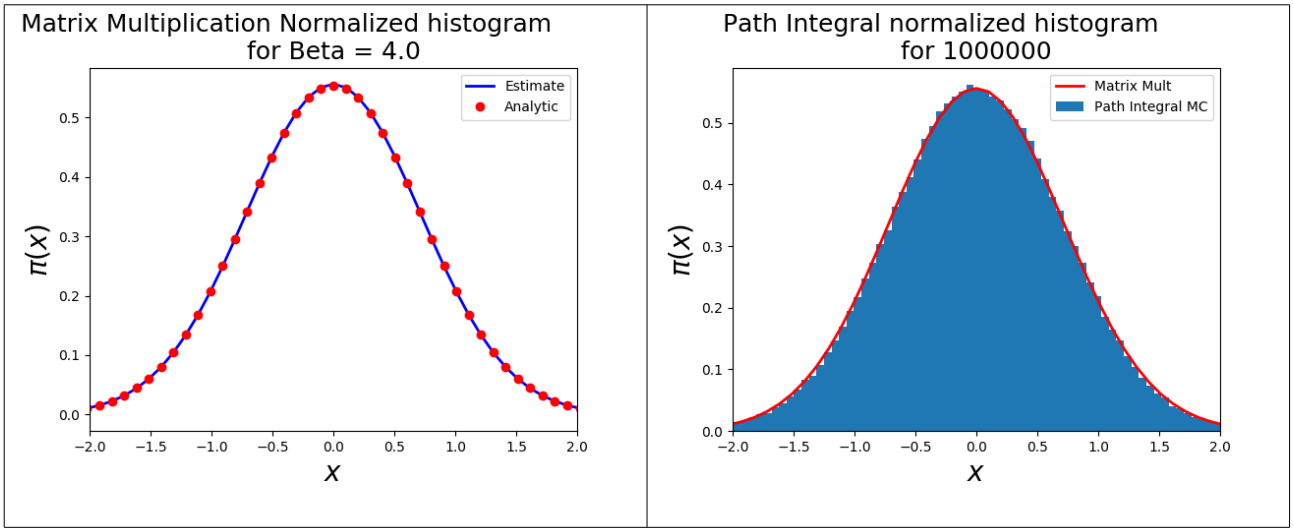


Illustration 2: Density Matrix computed using TrotterDecomposition and Path Integrals (MCMC)

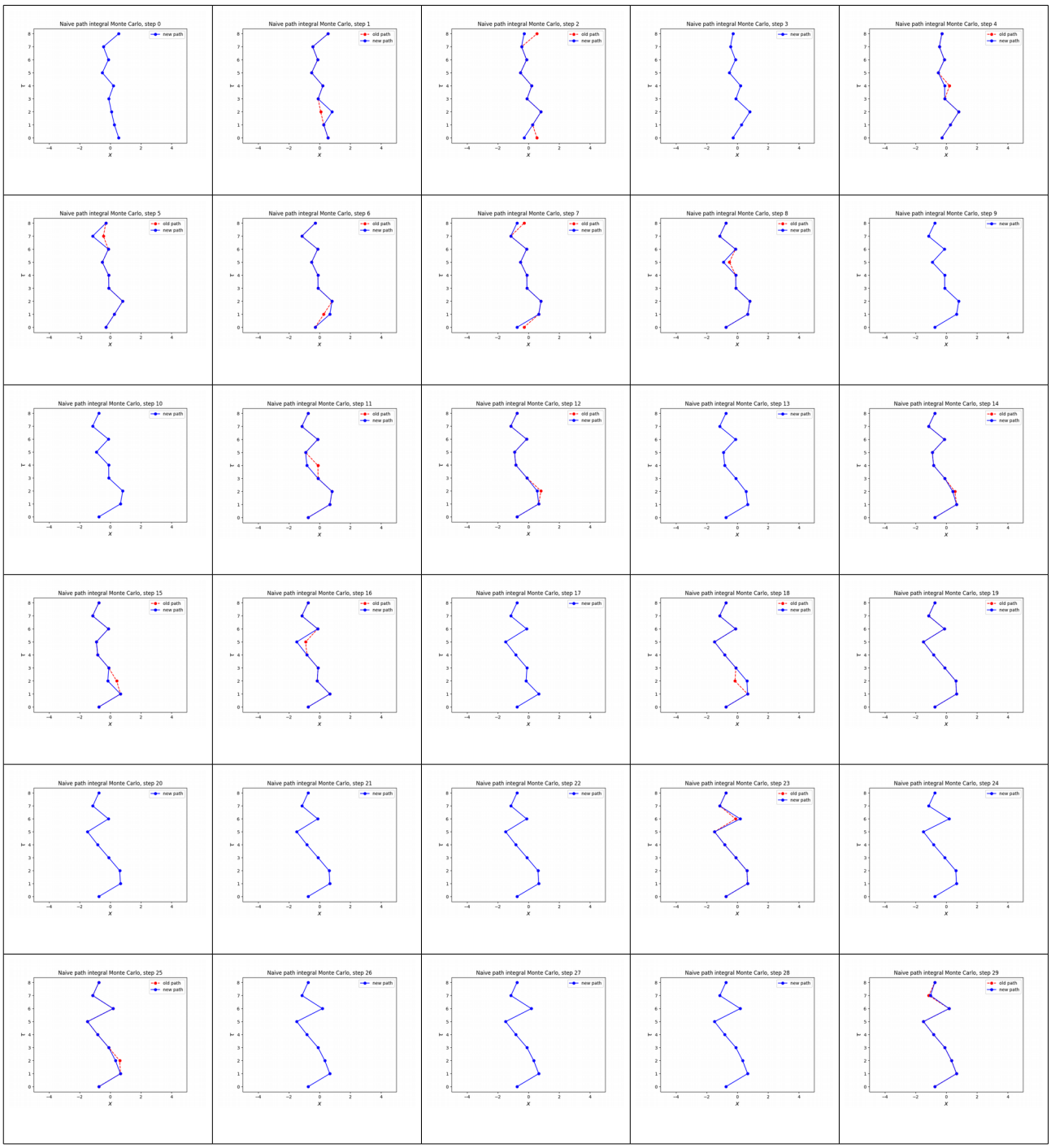


Table 6: Path Integrals: Markov-Chain Monte Carlo simulations used to obtain the Density Matrix

Simulating the solution to the Schrödinger Equation

The simulation shown in Table 7 is a time evolution of the solution to the Schrödinger equation computed as an Initial Value Problem (IVP). The code computes the density matrix, using convolution (via the Discrete Fourier Transform), and uses simple Euler time-stepping discretisation. If you look carefuly you can see quantum tunelling on the RHS of the potential well.

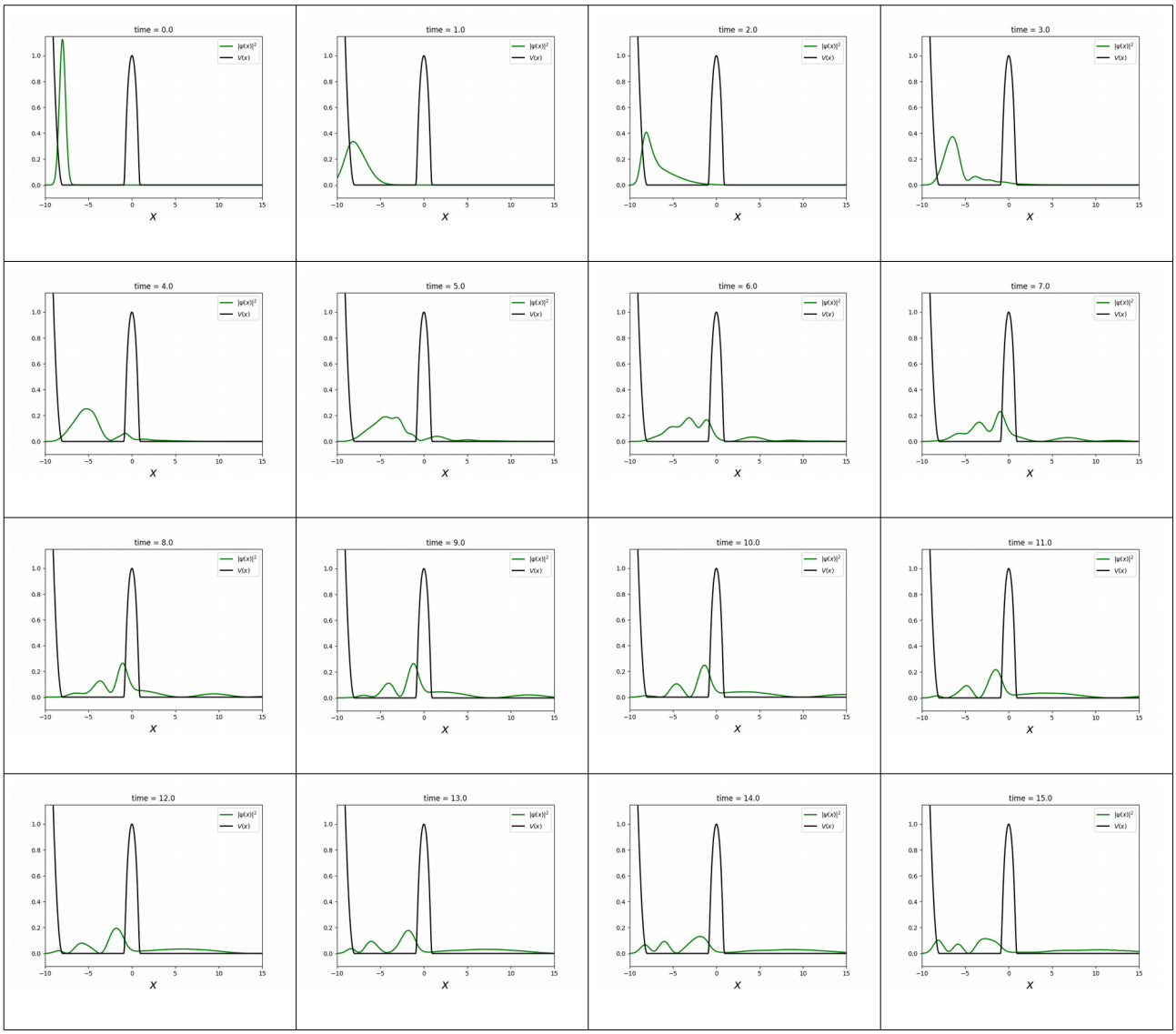


Table 7: Time evolution solution of Schrodinger's equation using Trotter Decomposition & simple time-step discretisation

Feynman-Kac Formula

Establishes a link between Parabolic PDEs and Stochastic Processes. The equation proves rigorously the ‘real’ case of Feynman’s Path Integrals; the ‘complex’ case, which occurs when a particle’s spin is included, is still unproven.

It offers a method of solving certain PDEs by simulating random paths of a stochastic process. Conversely, an important class of expectations of random processes can be computed by deterministic methods.

Consider the ***Partial Differential Equation (PDE)***:

$$\frac{\partial u}{\partial t}(x,t)+\mu(x,t)\frac{\partial u}{\partial x}+\frac{1}{2}\sigma^2(x,t)\frac{\partial^2 u}{\partial x^2}(x,t)-V(x,t)u(x,t)+f(x,t)=0$$

defined for all $x\in\mathbb{R}$ and $t\in[0,T]$, subject to the terminal condition:

$$u(x,T)=\psi(x)$$

where μ,σ,ψ,V,f are known functions,
T is a parameter and
$\mu:\mathbb{R}x[0,T]\rightarrow\mathbb{R}$ is the unknown.

Then the ***Feynman-Kac formula*** says the solution can be written as a ***Conditional Expectation:***

$$\mu(x,t)=\boldsymbol{E}^{\boldsymbol{Q}}[\int_t^T e^{-\int_t^r V(X_\tau,\tau(d\tau))}f(X_r,r)\,dr+e^{-\int_t^T V(X_\tau,\tau)d\tau}\psi(X_T)/X_t=x]$$

under the probability Q such that X is an ***Ito process*** driven by the equation:

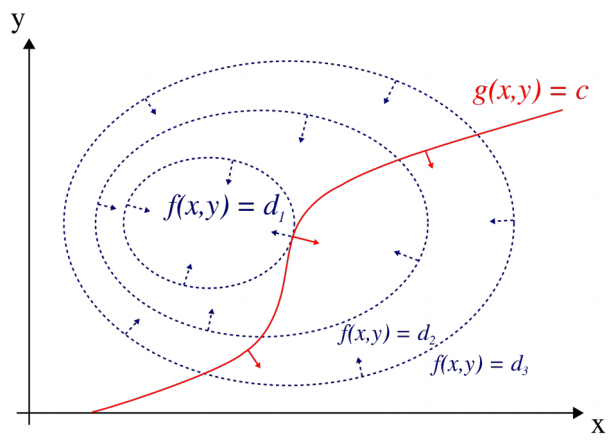
$$dX=\mu(X,t)dt+\sigma(X,t)dW^Q$$

with $dW^Q(t)$ being a ***Wiener process*** (also called ***Brownian motion***) under Q and the initial condition for $X(t)$ is $X(t)=x$.

Lagrange Multipliers

Strategy for finding local Maxima and Minima of a function subject to Equality Constraints.

Intuition



- At a maximum, f(x,y) cannot be increasing in the direction of any neighbouring point where g=0.
- If it were, we could walk along g=0 to get higher, meaning that the starting point wasn't actually the maximum.
- We want points (x,y) where g(x,y)=0 and $\nabla_{x,y} f = \lambda \nabla_{x,y} g$ for some λ.
- λ is required because, although the two gradient vectors are parallel, the magnitudes of the gradient vectors are generally not equal.
- To incorporate these conditions into one equation, we introduce an auxillary function: $L(x,y,\lambda) = f(x,y) - \lambda \cdot g(x,y)$ and solve $\nabla_{x,y,\lambda} L(x,y,\lambda) = 0$, as below.

Table 8: The red curve shows the constraint g(x, y) = c. The blue curves are contours of f(x, y). The point where the red constraint tangentially touches a blue contour is the maximum of f(x, y) along the constraint, since d1 > d2.

Single Constraint

- One constraint and only two variables
- functions, f and g, have continuous first partial derivatives

Maximise f(x, y)
Subject to g(x, y) = 0

- We introduce a new variable, λ, called a Lagrange Multiplier
- Then study the Lagrange Function defined by:

$$L(x,y,\lambda) = f(x,y) - \lambda \cdot g(x,y)$$

and solve

$$\nabla_{x,y,\lambda} L(x,y,\lambda) = 0$$

which is equivalent to 3 equations with 3 unknowns.

The method generalises readily to functions on n variables:

$$\nabla_{x_1,\ldots,x_n,\lambda} L(x_1,\ldots,x_n,\lambda) = 0$$

which amounts to solving n+1 equations with n+1 unknowns

Example:

Maximise f(x, y) = x + y
Subject to $x^2+y^2=1$

Step 1: The Lagrange Function is:

$$L(x,y,\lambda)=f(x,y)-\lambda\cdot g(x,y)$$

after substuiting the equations, we get:

$$L(x,y,\lambda)=x+y+\lambda\cdot(x^2+y^2-1)$$

Step 2: calculate the Gradients

$$\nabla_{x,y,\lambda}L(x,y,\lambda)=(\frac{\partial L}{\partial x},\frac{\partial L}{\partial y},\frac{\partial L}{\partial \lambda})$$

which gives

$$\nabla_{x,y,\lambda}L(x,y,\lambda)=(1+2x\lambda,1+2y\lambda,x^2+y^2-1)$$

and therefore gives us three equations to find three variables:

$$\nabla_{x,y,\lambda}L(x,y,\lambda)=0 \quad \Leftrightarrow \quad \begin{aligned} 1+2x\lambda&=0\\ 1+2y\lambda&=0\\ x^2+y^2-1&=0\end{aligned}$$

Step 3: The first two equations yield:

$$x=y=-\frac{1}{2\lambda} \qquad \lambda\neq 0$$

By substuiting into the last equation we have:

$$\frac{1}{4\lambda^2}+\frac{1}{4\lambda^2}-1=0 \quad \text{so} \quad \lambda=\pm\frac{1}{\sqrt{(2)}}$$

Step 4: Which implies that the stationary points of L are:

$$(\frac{\sqrt{(2)}}{2},\frac{\sqrt{(2)}}{2,}-\frac{1}{\sqrt{(2)}}) \ , \qquad (-\frac{\sqrt{(2)}}{2},-\frac{\sqrt{(2)}}{2},\frac{1}{\sqrt{(2)}})$$

Step 5: Evaluating the Objective Function, f, at these points yields:

$$f(\frac{\sqrt{(2)}}{2},\frac{\sqrt{(2)}}{2})=\sqrt{(2)} \qquad f(-\frac{\sqrt{(2)}}{2},-\frac{\sqrt{(2)}}{2})=-\sqrt{(2)}$$

Thus the ***Constrained Maximum*** is $\sqrt{(2)}$
And the ***Constrained Minimum*** is $-\sqrt{(2)}$

Jacobian & Hessian Matrices

Jacobian

(Matrix of 1st Order Partial Derivatives of a ***Vector***-valued function)

Suppose $f:\mathbb{R}^n \to \mathbb{R}^m$ is a function which takes as input the vector $x \in \mathbb{R}^n$ and produces, as output, the vector $f(x) \in \mathbb{R}^m$ then the Jacobian Matrix, J, of f is an m * n matrix:

$$J = \left[\frac{\partial f}{\partial x_1} \ldots\ldots \frac{\partial f}{\partial x_n}\right] = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \ldots\ldots & \frac{\partial f_1}{\partial x_n} \\ & \vdots & \\ \frac{\partial f_m}{\partial x_1} & \ldots\ldots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

or, component wise

$$J_{ij} = \frac{\partial f_i}{\partial x_j}$$

- If the function is differentiable at point ***X***
- Then the Jacobian Matrix defines a linear map, $\mathbb{R}^n \to \mathbb{R}^m$, which is the best point wise linear approximation of the function, f, near the point **X**. This linear map is thus the generalisation of the usual notion of a derivative and is called the ***Derivative*** or the ***Differential of f at X***.
- if m = n the Jacobiam Matrix is a square matrix and its Determinant is the ***Jacobian Determinant of f***. It carries importatant information about the local behaviour of f.
- f in the neighbourhood of point **X** has an Inverse Function that is differentiable iff the ***Jacobian Determinant*** is non-zero at **X**.

Example: Consider the function, $f:\mathbb{R}^2 \to \mathbb{R}^2$, given by:

$$f(x,y) = \begin{pmatrix} x^2 y \\ 5x + \sin(y) \end{pmatrix}$$

Then we have

$$f_1(x,y) = x^2 y \quad \text{and} \quad f_2(x,y) = 5x + \sin(y)$$

And the Jacobian Matrix, J, of f is

$$J_f(x,y) = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix} = \begin{pmatrix} 2xy & x^2 \\ 5 & \cos(y) \end{pmatrix}$$

And the Jacobian Determinant is

$$det(J_f(x,y)) = 2xy\cos(y)5x^2$$

Hessian

(Square Matrix of 2nd Order Partial Derivatives of a ***Scalar***-valued function)

Describes the Local Curvature (remember it is a 2nd order derivative) of a function of many variables.

Suppose $f:\mathbb{R}^n \to \mathbb{R}^m$ is a function taking as input a vector, $x \in \mathbb{R}^n$, and outputing a scalar, $f(x) \in \mathbb{R}$; if all 2nd Partial Derivatives of f exist and are continuous over the domain of the function, then the Hessian Matrix, H, of f, H_f, is a square n * n matrix, usually defined and arranged as follows:

$$H_f(X) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ . & . & & . \\ . & . & . & . \\ . & . & & . \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

or component wise

$$H_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

The Hessian Matrix is related to the Jacobian Matrix by:

$$H_f(X) = J(\nabla f(X))^T$$

Note: H_f is a matrix with functions as entries and it is meant to be evaluated at some point, ***X***, this is sometimes called a Matrix-valued Function.

Applications

Application 1: 2nd Derivative Test

IF Hessian is +ive definite (+ive for every non-zero column vector) at x

THEN f attains an isolated Minimum at x.

IF Hessian is -ive definite (-ive for every non-zero column vector) at x

THEN f attains an isolated Maximum at x.

IF Hessian has both +ive and -ive Eigenvalues

THEN x is a Saddle Point for f.

Application 2: Quadratic Approximations of Multi-Variate Functions

Hessian Matrices are used in large scale Optimisation problems within Newton type methods because they are the Coefficient of the Quadratic Term of a local Taylor Expansion of afunction, that is:

$$y = f(x+\Delta x) \approx f(x) + \nabla f(x)^T \Delta x + \frac{1}{2!} \Delta x^T H(x) \Delta x$$

where ∇f is the gradient $(\frac{\partial f}{\partial x_1} \cdots \frac{\partial f}{\partial x_N})$

Application 3: Image Processing

The Hessian Matrix is commonly used for expressing Image Processing Operators in field of Computer Vision.

Bordered Hessian

Used for the 2nd derivative test in certain Constrained Optimisation problems. Given a function, f, but adding a constraint function, g, such as g(x) = c, the bordered Hessian of the Lagrange Function, $L(x,\lambda) = f(x) + \lambda[g(x) - c]$ is:

$$H(L) = \begin{pmatrix} \frac{\partial^2 \lambda}{\partial \lambda^2} & \frac{\partial^2 \lambda}{\partial \lambda \partial x} \\ (\frac{\partial^2 \lambda}{\partial \lambda \partial x})^T & \frac{\partial^2 \lambda}{\partial x^2} \end{pmatrix} = \begin{pmatrix} 0 & \frac{\partial g}{\partial X} \\ (\frac{\partial g}{\partial X})^T & \frac{\partial^2 L}{\partial X^2} \end{pmatrix}$$

$$= \begin{pmatrix} 0 & \frac{\partial g}{\partial x_1} & \frac{\partial g}{\partial x_2} & \frac{\partial g}{\partial x_n} \\ \frac{\partial g}{\partial x_1} & \frac{\partial^2 L}{\partial x_1^2} & \frac{\partial^2 L}{\partial x_1 \partial x_2} \cdots\cdots & \frac{\partial^2 L}{\partial x_1 \partial x_n} \\ \frac{\partial g}{\partial x_2} & \frac{\partial^2 L}{\partial x_2 \partial x_1} & \frac{\partial^2 L}{\partial x_2^2} \cdots\cdots\cdots & \frac{\partial^2 L}{\partial x_1 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g}{\partial x_n} & \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 L}{\partial x_n \partial x_2} \cdots\cdots\cdots & \frac{\partial^2 L}{\partial x_n^2} \end{pmatrix}$$

- If there are m Constraints Then the Zero in the upper-left corner is an m * m block of zeros
- 2nd derivative test consists here of sign restrictions of the determinants of a certain set of n-m submatrices of the bordered Hessian
    - Normal 2nd order derivative test rules cannot apply here since a bordered Hessian can neither be -ive / +ive definite

Applications

Classical Mechanics:

- Mechanical / Aeronatical Engineering
- Rocket / Satellite / Space Engineering
- Physics Engines for Video Games, CGI and Computer Animation
- Astronomy: Celestial Mechanics
- Chemistry: Dynamics of molecular collisions

Quantum Mechanics:

- Explains many of the features of our universe
  - Sub-atomic particles, string theories
- Quantum Chemistry
  - Understanding covalent bonding to form molecules
  - Combined with Artificial Intelligence (AI) to create new drugs
- Electronics
  - Lasers, semi-conductors (diodes, transistors)
  - Electron Microscope, Magnetic Resonance Imaging (MRI)

Statistical Mechanics:

- Thermodynamics
- Machine Learning
  - Boltzmann Machine (stochastic recurrent neural network)

Lagrange Multipliers:

- Managerial Economics

References

Papers / Books

Although not directly referenced in the paper, the following texts were used:

Newton, I. (1687). Philosophiae Naturalis Principia Mathematica

Leibniz, G. (1684). Nova Methodus pro Maximis et Minimis

Lagrange, J. (1811). Mécanique Analytique

Hamilton, W. (1835). Second Essay on a General Method in Dynamics

Blanchard, P. and Devaney, R. (2011). Differential Equations

Krauth, W. (2006). Statistical Mechanics: Algorithms and Computation

Rangel, A. (2016). Principles of Economics with Calculus

Aspuru-Guzik, A. (2017). The Quantum World

Sedgewick, R. and Wayne, K. (2011). Algorithms

Goyal, V. (2016). Data, Models and Decisions in Business Analytics

Grinspun, E. (2016). Animation and CGI Motion

Konis, K. (2015). Mathematical Methods for Quantitative Finance

Jha, R.G. (2012). On the Numerical Simulations of Feynman's Path Integrals using Markov Chain Monte-Carlo with Metropolis-Hastings Algorithms

Wikipedia (multiple links)

Codes

NewtonBoltzmannEquivalence.py

quantum_time_evolution.py

harmonic_wavefunction_movie.py

matrix_square_harmonic_movie.py

naive_harmonic_path_movie.py

DensityMatrix_Trotter.py

DensityMatrix_PathIntegrals.py