MOMENTUM AND SPIN IN ENTROPIC QUANTUM DYNAMICS

by

Shahid Nawaz

A Dissertation

Submitted to the University at Albany, State University of New York

in Partial Fulfillment of

the Requirements for the Degree of

Doctor of Philosophy

College of Arts & Sciences

Department of Physics

2014

Abstract

We study quantum theory as an example of entropic inference. Our goal is to remove conceptual difficulties that arise in quantum mechanics. Since probability is a common feature of quantum theory and of any inference problem, we briefly introduce probability theory and the entropic methods to update probabilities when new information becomes available. Nelson's stochastic mechanics and Caticha's derivation of quantum theory are discussed in the subsequent chapters.

Our first goal is to understand momentum and angular momentum within an entropic dynamics framework and to derive the corresponding uncertainty relations. In this framework momentum is an epistemic concept – it is not an attribute of the particle but of the probability distributions. We also show that the Heisenberg's uncertainty relation is an osmotic effect. Next we explore the entropic analog of angular momentum. Just like linear momentum, angular momentum is also expressed in purely informational terms.

We then extend entropic dynamics to curved spaces. An important new feature is that the displacement of a particle does not transform like a vector. It involves second order terms that account for the effects of curvature . This leads to a modified Schrödinger equation for curved spaces that also take into account the curvature effects. We also derive Schrödinger equation for a charged particle interacting with external electromagnetic field on general Riemannian manifolds.

Finally we develop the entropic dynamics of a particle of spin 1/2. The particle is modeled as a rigid point rotator interacting with an external EM field. The configuration space of such a rotator is $R^3 \times S^3$ (S^3 is the 3-sphere). The model describes the regular representation of $SU(2)$ which includes all the irreducible representations (spin 0, 1/2, 1, 3/2,...) including spin 1/2.

Acknowledgments

I would like to express my deepest gratitudes to my advisor Prof. Ariel Caticha who introduced me to Information Physics. I was very lucky to join his group during a time when his interests were in unifying information theory with that of quantum mechanics. His deep understanding of the subject matter significantly changed my way of thinking. I'm committed to collaborate with him in the years to come.

I owe special thanks to my PhD dissertation committee members Dr. Carlo Cafaro, Prof. Akira Inomata, Prof. Kevin H. Knuth and Prof. Philip Goyal . Especially I'm gratefully indebted to Prof. Akira Inomata for his constant encouragements. I learned a lot from him during casual discussions with him from time to time.

My thanks also go to the graduate students, faculty and staff of the Department of Physics, University at Albany. I'm also greatly influenced by the friendly attitude of Prof. Carolyn MacDonald who has been my teaching supervisor.

Last but not the least I'm very thankful to my wife Ayisha for providing me very kind support during my studies.

Contents

Chapter 1

Introduction

General relativity and quantum mechanics are two foundational pillars of modern physics. The mathematical frameworks of both theories are well-established. In contrast to general relativity quantum mechanics lacks a conceptual foundational principle. Although quantum mechanics describes the properties of microscopic particles to a high level of accuracy, its formalism is very abstract and non-intuitive. The fundamental object in quantum mechanics is the wave function. Whether the wave function corresponds to some physical wave has been the subject of debate since the beginning of the theory in the 1930's. Several interpretations of quantum mechanics have been proposed, mainly the orthodox Copenhagen interpretation [1], the causal interpretation [2], the statistical interpretation [3], the many-worlds interpretation [4], the transactional interpretation [5], the consistent history interpretation [6], and many more.

Although the interpretational difficulties need not affect the main conclusions of a theory, the lack of conceptual understanding can give rise to paradoxes and controversies. Take for example the interpretation of special relativity. Some relativistic equations appearing in special relativity such as Lorentz contraction were already known before Einstein, and were controversial. Once Einstein established the theory based on two simple postulates the interpretational problems were resolved. A similar treatment is needed for quantum mechanics.

Since quantum mechanics involves probabilities in a fundamental way its conceptual foundation may lie in the theory of inference. Recently A. Caticha [7, 8] derived quantum mechanics as an example of entropic inference. In this approach QM is a dynamics driven by entropy — it is therefore called Entropic Dynamics (ED). In the original formulation the basic assumption was that in addition to the particles of interest the world contains other 'extra' variables whose entropy depends on the positions of the particles. In ED, not only the amplitude but also the phase of the wave function receives a statistical interpretation: the phase keeps track of the entropy of those extra variables. In a more recent formulation [9] this assumption has been dropped.

Entropic dynamics is a formulation that is much broader than QM. Depending on the choice of microstates or variables, and the choice of constraints we can represent forms of dynamics that have quantum properties, or deterministic classical properties, or stochastic classical properties. The goal of this thesis is to explore some particular versions of ED in order to eventually attain a deeper understanding of QM.

Entropic dynamics has close formal similarities with Nelson's stochastic mechanics. Both theories are modeled in configuration space and Brownian motion is their common feature. But there is an important difference. Brownian motion is merely postulated in Nelson's theory while it is derived in entropic dynamics. Also Nelson's theory operates at an ontological level, while ED operates completely at the epistemological level.

Entropic dynamics differs from other information-based approaches to quantum theory in that the position observable assumes a privileged role: particles have well-defined, albeit unknown positions. This opens the possibility of explaining all other observables in purely informational terms. In this thesis one of our goals is to identify what concepts, within entropic framework, play the role of momentum, and angular momentum. Another goal is to generalize ED formalism to curved spaces.

This thesis is organized as follows. In chapter 2 we introduce probability theory. Special focus is given to the comparative study of Kolmogorov and Cox frameworks of probability. The Kolmogorov framework (KF) heavily rests on *measure theory* while Cox framework (CF) makes use of Boolean logic to derive the rules of probability. The mathematical structure of probability theory remains the same but the meaning of probability changes while going from KF to CF framework.

In chapter 3 we review the methods to update probabilities when new information becomes available. Information either comes in the form of data or constraints. One normally uses Bayes' rule to update probabilities when information is available in the form of data and when information comes in the form of constraints, one uses the method of maximum entropy (MaxEnt) originally designed by E. T. Jaynes [10, 11].

In this chapter we review the extended method of maximum entropy (ME) due to A. Caticha [12]. It is shown that both Bayes' rule and MaxEnt are special cases of ME.

Chapter 4 is devoted to Nelson's theory of stochastic mechanics. In 1966, E. Nelson derived Schrödinger equation from an entirely classical notion of Brownian motion [13–15]. In this chapter we define Brownian motion and then review Nelson's formalism of quantum theory.

Chapter 5 deals with Caticha's derivation of quantum theory. Specifying the relevant statistical manifold and choosing the appropriate constraints, and then using the method of maximum entropy, the transition probability is derived for short steps which leads to Brownian motion. Finally requiring that the diffusion be non-dissipative – that there exists a conserved energy – the Schrödinger equation is derived.

The new contributions of this thesis are contained in chapters 6 through 9. Chapter 6 deals with the concept of linear momentum within the framework of entropic dynamics. We also review the concept of momentum within classical and quantum mechanics. Next we define momentum within the entropic framework. We note that since the particle follows a non-differentiable trajectory it is clear that the classical momentum, $m\,d\vec{x}/dt$, along the trajectory cannot be defined. Nevertheless three different notions of momentum can be usefully defined. They are the *drift*, *osmotic* and *current* momenta. It turns out that these momenta are not associated with the particles but to the probability distributions. The drift momentum reflects probability flow along the entropy gradient, the osmotic momentum indicates diffusion of probability flow, while the current momentum reflects the flow of total probability. It is shown that these momenta share properties with the quantum momentum $\vec{p}_q = -i\hbar\vec{\nabla}$, and in the appropriate classical limit the drift and current momenta converge to the classical momentum, $m\vec{v}$, while the osmotic momentum tends to zero. Finally we derive the uncertainty relations for all momenta that appear in entropic dynamics. In the same chapter we also explore a special case of entropic dynamics that involves hybrid classical-quantum features. It obeys the classical Hamilton-Jacobi equation and also the usual uncertainty principle.

Another important concept in physics is angular momentum. In chapter 7 we develop the entropic analog of angular momentum. Just like linear momentum, angular momentum is also statistical in nature—angular momentum is an attribute of the probability distributions. We introduce four different notions of angular momenta. They are the drift angular momentum, osmotic angular momentum, current angular momentum and the standard quantum angular momentum operator. We show that the current/drift angular momentum represents the entropic analog of angular momentum. The expected values of current/drift angular momentum is the same as the expectation of quantum angular momentum while osmotic momentum has vanishing expectation. Having defined angular momenta, we also establish their uncertainty relations.

In chapter 8 we extend entropic dynamics to curved spaces. An important new feature is that the displacement of a particle does not transform like a vector. It involves second order terms that account for the curvature effects. This leads to a modified Schrödinger equation for curved spaces that take into account the curvature effects. In the same chapter, we also derive Schrödinger equation for a charged particle interacting with external electromagnetic field on general Riemannian manifolds.

Chapter 9 is an application of the theory developed in chapter 8. We develop entropic analog of the models of spin developed by Dankel [16]; and Dohrn, Guerra and Ruggiero [17] in connection with Nelson's stochastic mechanics. Consider a single particle, we assume that the particle has the usual spatial coordinates as well as some internal degree of freedom. The configuration space is enlarged from R^3 to $R^3 \times S^3$, where R^3 corresponds to the usual three dimensional Euclidean space that accounts for the translational degrees of freedom while S^3 is the 3-sphere that takes into account the rotational degrees of freedom. We also discuss the possible shortcomings of the rotator models of Dankel [16]; and Dohrn, Guerra and Ruggiero [17]. These models reproduce the regular Pauli equation that corresponds to the regular representation of $SU(2)$ which includes the irreducible spinor representation.

In chapter 10 we collect our conclusions.

Chapter 2

Probability Theory

2.1 Introduction

The mathematical theory of probability arose from correspondence between Blaise Pascal (1623-1662) and Pierre Fermat (1601-1665) to solve some problems in the games of chance. The earlier notion of probability was intuitive and only in twentieth century A. N. Kolmogorov [18] provided the axiomatic foundation of probability. The Kolmogorov's framework rests heavily on measure theory of sets, it does not address the question of interpretation of probability. A meaningful interpretation of probability as a tool for inference was provided by Richard T. Cox [19,20] and others. This chapter is devoted to the comparative study of the two frameworks.

2.2 Kolmogorov's Framework

The axiomatic development makes it possible to encompass the totality of the objects studied by a given mathematical theory [21]. Like any other mathematical theory such as geometry and algebra, probability theory can also be developed from certain axioms. Here we state the axioms of Kolmogorov in connection with probability and briefly describe them. The axioms or definitions stated here are given in modern terms, see for example [22, 23].

Definition 1. (Borel Sigma-Field) *Let Ω be a set and $\mathcal{F}$ be a collection of subsets of Ω. Then $\mathcal{F}$ is called a Borel sigma-field or Borel sigma-algebra if the following conditions are met:*

1. *The empty set $\varphi \in \mathcal{F}$*
2. *If $A \in \mathcal{F}$, then the complement $A^c \in \mathcal{F}$*
3. *If $A_i \in \mathcal{F}$ for all $i = 1, 2, \ldots$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$*

The set Ω is called a *sample space*, the elements of Ω are called *outcomes* and the elements A of $\mathcal{F}$ are called *events*. It is clear that Borel sigma-field is closed under the operation of complement and countable union of its members.

Definition 2. (Probability Measure) *A set function $P : \mathcal{F} \to [0, 1]$ is said to be a probability measure if the following conditions hold:*

1. $P(\Omega) = 1$
2. $P(A) \geq 0,\ A \in \mathcal{F}$
3. $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum\limits_{i=1}^{\infty} P(A_i),\ A_i \in \mathcal{F},\ A_i \cap A_j = \phi,\ i \neq j$

Here the set function P assigns real values in the unit interval $[0, 1]$ to the events in $\mathcal{F}$. We say that $A \subseteq \Omega$ is measurable if $A \in \mathcal{F}$. One can observe that probability measure is finitely additive, that is, if A_1 and A_2 are disjoint then $P(A_1 \cup A_2) = P(A_1) + P(A_2)$. Moreover it is also monotonic whenever $A \subseteq B$ then $P(A) \leq P(B)$ for any A, B in $\mathcal{F}$. Finally the triple $(\Omega, \mathcal{F}, P)$ is called a *probability space*.

Definition 3. (Conditional Probability) *Let $(\Omega, \mathcal{F}, P)$ be a probability space. Then the conditional probability of the event A given that the event B has occurred, denoted by $P(A|B)$ is defined as*

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0\,. \tag{2.1}$$

The basic idea of conditioning is that whenever it is given that an event say B has occurred then the original probability space $(\Omega, \mathcal{F}, P)$ becomes irrelevant. A new probability space $(B, \mathcal{F}_B, P_B)$ takes over the original space. Here $B \subseteq \Omega$, $\mathcal{F}_B = \{A \cap B : A \in \mathcal{F}\}$, and $P_B \overset{\text{def}}{=} P(A|B)$. The space $(B, \mathcal{F}_B, P_B)$ is called a conditional probability space or simply subprobability space [24]. The events A and B are said to be *independent* if $P(A|B) = P(A)$, that is the probability of A has not affected when the event B is observed.

Definition 4. (Random Variables) *Let $\omega \in \Omega$. Let $B \subseteq R^* = [-\infty, \infty]$. Then the real-valued function $X : \Omega \to \mathcal{R}^*$ is called random variable if the image of B under the inverse mapping*

$$X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathcal{F}\,, \tag{2.2}$$

where $\mathcal{F}$ is a sigma-field on Ω.

One might think that a random variable is just a 'variable' but this is misleading. As evident from the definition just stated X is not a variable but a function. Here the inverse mapping $X^{-1}(B)$ is very crucial. Consider a collection $\mathcal{G}$ of all those subsets B of R^* for which $X^{-1}(B) \in \mathcal{F}$. Since inverse mapping preserves unions, intersections and differences of sets, if $\mathcal{F}$ happens to be a sigma-field then $\mathcal{G}$ is also a sigma-field. The proof is trivial, see for example, [25]. A random variable induces a probability measure on $\mathcal{G}$. If $(\Omega, \mathcal{F}, P)$ is a probability space, then the triple $(R^*, \mathcal{G}, \mu)$ is also a probability space, defined by, $\mu(B) = P(X^{-1}(B))$. Computationally this correspondence is very useful because the element ω is not necessarily a number but $X(\omega)$ is a real number.

These are the basic elements of Kolmogorov's framework of probability theory. The Kolmogorov formulation does not address the issue of the meaning of probability. Its great virtue is mathematical rigour.

2.3 Cox's Framework

The meaning of probability is central to Cox framework. It is worth noting that several interpretations of probability exist even today. The *frequency* interpretation and the *Bayesian* interpretation are the two major interpretations of probability. According to frequency theory the probability of a random event is defined as the limiting frequency with which that event occurs in a sufficiently large number of identical and independent trials.

The Bayesian interpretation is further classified into *subjective* and *objective* theories. According to subjective view probability is identified with the degree of belief of a particular individual and so different individuals may have different degrees of belief in the truth of the same proposition. This is sometimes called the personalistic view. At the other end of the spectrum, in the objective view, probability is identified with the degree of belief of an ideally rational agent. It is assumed that all rational agents reasoning on the basis of the same information will arrive at the same degree of belief about the truth of a proposition. The objective view places probability as an extension of logic.

Richard Cox [19] made use of the objective view of probability to derive the sum and product rules of probability. The advantage of this approach is that degrees belief can be quantified. The idea is that beliefs can be compared with respect to the intensity with which they are held. This means that degrees of belief can be represented by real numbers. But first some notation is in order.

Let $a, b, c, \ldots$ represent *statements* or *propositions* that obey Boolean logic. The negation not-a of a proposition a, is denoted by a'. If a is true, then a' is false and vice versa. The *logical product* or *conjunction* is represented by ab. The conjunction ab is true when both a and b are true. The *logical sum* or *disjunction* is represented by $a+b$. The disjunction $a + b$ is true when either a or b or both are true. The propositions $a, b, c, \ldots$ do not represent real numbers, however, they can be quantified using the following two requirements [19, 26, 27]:

1. *The degree to which we believe that proposition a is true when proposition b is known to be true, is represented by a real number.*

2. *The assignment of degrees of belief must be consistent.*

Mathematically the first condition will be written as $P(a|b)$ which will be later related with the probability measure. However at the present moment we shall simply call it *plausibility.* We seek the degree of plausibility of a proposition a given a proposition b. The second requirement, called the consistency condition, is very crucial. For if the plausibility of a can be found in two different ways, then the two ways should agree.

Our goal is to show that after suitable regraduation, a change of scale, plausibility turns out to be probability. This must be the case if we are able to show that our function P satisfies the following criteria:

1. If $0 \leq P \leq 1$

2. If P satisfies probability sum rule.

3. If P satisfies probability product rule.

Cox derived sum and product rules of probability by focusing on negation and conjunction properties of Boolean logic. Here are his axioms [19, 26, 27]:

Axiom 1. *The plausibility of negation a' is monotonic function of the plausibility of a,*

$$P(a'|b) = f\left(P(a|b)\right) . \tag{2.3}$$

The function f just relates the plausibility of a' with the plausibility of a.

Axiom 2. *The plausibility* $P\left(a_1 a_2|b\right)$ *of a conjunction* $a_1 a_2$*, is a function of plausibility of* $P\left(a_1|b\right)$ *of* a_1*, and* $P\left(a_2|a_1 b\right)$ *of* a_2 *given* a_1*,*

$$P\left(a_1 a_2|b\right) = g\left(P\left(a_1|b\right), P\left(a_2|a_1 b\right)\right) \tag{2.4}$$

where g is some function.

The functions f and g are not specified.

2.3.1 Product Rule

To find the functions f and g we first obtain a constraint that follows from the associativity of the Boolean algebra, $abc = (ab)c = a(bc)$. Treating ab and bc as single statements, then

$$P((ab)c|d) = P(a(bc)|d)\,. \tag{2.5}$$

Applying eq. (2.4) repeatedly, we arrive at

$$g\left[g\left[P(a|d), P(b|ad)\right], P(c|abd)\right] = g\left[P(a|d), g\left[P(b|ad), P(c|bad)\right]\right]\,. \tag{2.6}$$

This equation is called 'the associativity equation' which has a unique solution. The general solution is given by [19]

$$G[P(ab|c)] = G[P(a|b)]G[P(b|ac)]\,, \tag{2.7}$$

where G is another function. Originally our goal was to find the function P. But there is nothing special about P. If $P(a|b)$ is a plausibility, then any monotonic function of $P(a|b)$ is also an acceptable degree or measure of plausibility. It is just a matter of rescaling. Therefore $G[P(a|b)]$ is also a plausibility. It is convenient to regraduate $P(a|b)$ to a new set of positive numbers $G[P(a|b)]$,

$$p(a|c) \stackrel{\text{def}}{=} G[P(a|b)], \quad p(b|ac) \stackrel{\text{def}}{=} G[P(b|ac)], \ldots \tag{2.8}$$

we thus arrive at the product rule of probability,

$$p(ab|c) = p(a|c)p(b|ac) \tag{2.9}$$

For now we shall still continue to call it 'plausibility' instead of 'probability'. Let us derive the other two conditions—namely $0 \leq p \leq 1$, and the sum rule.

2.3.2 The range of plausibility

To find the range of p, set $a = b$ in eq. (2.9)

$$p(aa|c) = p(a|c)p(a|ac)\,. \tag{2.10}$$

Since $a, b, c, \ldots$ are propositions, they obey Boolean algebra. We can write $p(aa|c) = p(a|c)$ because $aa = a$. We obtain

$$p(a|c) = p(a|c)p(a|ac) \quad \Rightarrow \quad p(a|ac) = 1\,, \tag{2.11}$$

where $p(a|ac)$ is the plausibility of a given a which happens in the situation of total certainty. Let p_T reflect total certainty. Therefore

$$p_T = 1\,. \tag{2.12}$$

This is one of the extreme value of p. To find the other extreme value use the product rule again with a different setting,

$$p(ab'|b) = p(a|b)p(b'|ab)\,. \tag{2.13}$$

Since both $ab'|b$ and $b'|ab$ are absurdities, call $p(ab'|b) = p_F = p(b'|ab)$, therefore

$$p_F = p(a|b)p_F\,. \tag{2.14}$$

Since $p(a|b)$ is arbitrary, therefore $p_F = 0, \infty$ or $-\infty$. The negative values of p are not allowed because it violates consistency with the product rule. If $p_F = \infty$, then p is a decreasing function which is changing from $p = \infty$ for absurdity down to $p = 1$ for certainty – that is $1 \leq p \leq \infty$. In the later case, the plausibility p can be regraduateed to a new plausibility $\mathcal{P}$ such that $\mathcal{P}(a|b) = 1/p(a|b)$ [26]. This means that $0 \leq \mathcal{P} \leq 1$. Therefore without loss of generality we can set: $0 \leq p \leq 1$. Hence the second condition of probability is also derived.

2.3.3 Probability sum rule

Now let us solve eq. (2.3). Start with the product rule, eq. (2.9).

$$P(ab|c) = P(a|c)P(b|ac) = P(a|c)f\left(P(b'|ac)\right)\,, \tag{2.15}$$

but

$$P(ab'|c) = P(a|c)P(b'|ac)\,, \tag{2.16}$$

then

$$P(ab|c) = P(a|c)f\left(\frac{P(ab'|c)}{P(a|c)}\right)\,. \tag{2.17}$$

Since $P(ab|c)$ is symmetric in $ab = ba$, therefore

$$P(a|c)f\left(\frac{P(ab'|c)}{P(a|c)}\right) = P(b|c)f\left(\frac{P(a'b|c)}{P(b|c)}\right) . \tag{2.18}$$

Since a, b, and c are arbitrary, we can choose $b' = ad$. In the l. h. s. $ab' = aad = b'$, therefore $P(ab'|c) = P(b'|c) = f(P(b|c))$. In the r. h. s. $b = (ad)' = a' + d'$ so that $a'b = a'a' + a'ab = a'$, therefore $P(a'b|c) = P(a'|c) = f(P(a|c))$

$$P(a|c)f\left(\frac{f(P(b|c))}{P(a|c)}\right) = P(b|c)f\left(\frac{f(P(a|c))}{P(b|c)}\right) . \tag{2.19}$$

Writing $P(a|c) = u$, and $P(b|c) = v$, we have

$$uf\left(\frac{f(v)}{u}\right) = vf\left(\frac{f(u)}{v}\right) . \tag{2.20}$$

The general solution of eq. (2.20) is given by [19, 26],

$$f(u) = (1 - u^\alpha)^{1/\alpha} \quad \text{or} \quad u^\alpha + (f(u))^\alpha = 1\,, \tag{2.21}$$

which means that

$$[P(a|c)]^\alpha + [P(a'|c)]^\alpha = 1\,. \tag{2.22}$$

Performing regraduation again

$$p(a|c) \stackrel{\text{def}}{=} [P(a|c)]^\alpha \ , \tag{2.23}$$

leads to

$$p(a|c) + p(a'|c) = 1\,. \tag{2.24}$$

This is the 'sum rule of probability'. For any statements a and b the 'extended sum rule' is given by

$$p(a + b|c) = p(a|c) + p(b|c) - p(ab|c)\,. \tag{2.25}$$

In summary, since all conditions of probability — the product and sum rules and that probability is normalized are derived, we have proved that 'plausibility' is indeed 'probability'. From now on we shall call it probability.

2.4 Conclusions

The very motive of this chapter was to put forward probability theory either in purely mathematical settings, the Kolmogorov's framework (KF), or on the interpretational grounds, the Cox's framework (CF). In KF probability is just measure based on set theory. For mathematicians this maybe the standard way of dealing with probability. Of course, it is still needed to attach some interpretation to probability in order to apply it to practical problems. One naturally recovers KF if one associates probability with the relative frequency of *equally likely events*. To demonstrate this let A and B be disjoint sets, and $\Omega = A \cup B$, then $|\Omega| = |A| + |B|$, where $|.|$ represents total number of elements (cardinality) of a set. Now define probability of event A as $P(A) = |A|/|\Omega|$, and similarly $P(B) = |B|/|\Omega|$. One observers that $P(\Omega) = P(A) + P(B) = 1$. Unfortunately one arrives at a contradiction if A, B and Ω are all infinite sets having cardinality $\aleph_0$, where $\aleph_0$ is the cardinality of the set of integers, then $P(\Omega) = P(A) + P(B) = 1$ is not guaranteed.

Formulating a theory purely in a mathematical way (theorem-proof style) and leaving the interpretation to the end may lead to difficulties if the interpretation is disputable. It may satisfy the curiosity of mathematicians but may not be much appreciated by other disciplines. Returning to probability theory, Cox framework (CF) does not have this problem. In CF both the mathematics and interpretation of probability are in complete agreement. The fundamental object in CF is the conditional probability $P(a|b)$. The credibility of a proposition a is relative to some known evidence b. This makes all probabilities conditional. One assigns probability on the basis of partial knowledge. Furthermore a probability can be assigned to single events without the need for large ensemble of identical trials and that the 'randomness' in KF is nothing but incomplete information.

Chapter 3

The Problem of Inference

3.1 Introduction

The process of making judgments on the basis of partial knowledge is called inductive inference. The tool to handle incomplete information or uncertainty is probability. Since probability is defined as the degree of belief of a rational agent, our beliefs change when new information becomes available. Likewise probabilities must also be updated. A common method of updating probabilities from the prior given information to the posterior probability distribution is the Bayes' rule.

Although Bayes' rule is very successful, it also has some limitations. It is restricted to situations where the information is in the form of data,, it cannot handle other forms of information such as constraints. When information is in the form of constraints, one uses a different method, namely the method of maximum entropy (MaxEnt), originally formulated by E. T. Jaynes [10, 11] to reconcile the statistical mechanics of J. W. Gibbs [28] and communication theory of C. E. Shannon [29].

MaxEnt also has some limitations. It allows arbitrary constraints but it does not allow information contained in arbitrary priors. A more general method of updating that allows arbitrary constraints and arbitrary priors involves maximizing relative entropy. Such a method of entropic inference is an extended method of maximum entropy (abbreviated ME).

Processing information on the basis of available information is the subject of this chapter. We begin with Bayesian inference and then present entropic inference.

3.2 Bayesian Inference

Bayesian inference deals with updating probabilities from old beliefs about one or several parameters $\theta \in \Theta$ when the new evidence is available in the form of data $x \in \mathcal{X}$. First we wish to describe the old probabilities before the data has been observed. At this stage the relevant space is neither Θ nor $\mathcal{X}$ but the product space $\Theta \times \mathcal{X}$ whose probability is represented by the joint distribution

$$q(\theta, x) = q(\theta) q(x|\theta)\,, \tag{3.1}$$

where $q(\theta)$ is called the prior probability distribution that represent our knowledge about θ before the data has been collected. The relation between x and θ is encoded into the conditional probability distribution $q(x|\theta)$ called the *likelihood*.

We note that since the joint distribution $q(\theta, x)$ is symmetric in its arguments, we can also write as

$$q(x, \theta) = q(x) q(\theta|x)\,. \tag{3.2}$$

Equating eqs. (3.1) and (3.2), we obtain

$$q(\theta|x) = \frac{q(\theta)q(x|\theta)}{q(x)}\,, \tag{3.3}$$

which is called Bayes' theorem. It is named after Reverend Thomas Bayes who addressed the problem of *inverse probabilities* sometime during 1740s. Its modern mathematical form is due to Pierre-Simon Laplace who discovered the theorem independently.

It should be noted that Bayes' theorem just relates two prior conditional probabilities regardless of what has been observed in the data. One uses *Bayes' rule* in order to take into account the actual data. Bayes' rule is a method of update from prior probabilities to posterior probability distribution when the new evidence is available in the form of actual data. Prior probability should be revised only the extent required by the new data —the so-called Principle of Minimal Updating (PMU), see for example [30].

Our goal is to obtain the new joint posterior probability distribution $p(\theta, x)$. Using the product rule

$$p(\theta, x) = p(x)p(\theta|x)\,. \tag{3.4}$$

When the experiment is performed, the actual value of x is found to be x', therefore

$$p(x) = \delta(x - x')\,. \tag{3.5}$$

Next PMU takes over and we require that the prior conditional probability $q(\theta|x')$ not be updated,

$$p(\theta|x') = q(\theta|x')\,. \tag{3.6}$$

Combining PMU with Bayes' theorem lead to Bayes' rule

$$p(\theta) = q(\theta)\frac{q(x'|\theta)}{q(x')}\,, \tag{3.7}$$

where $p(\theta)$ is the desired posterior distribution. The factor in the denominator is the normalization constant given by

$$q(x') = \int q(\theta)q(x'|\theta)d\theta\,. \tag{3.8}$$

3.3 Entropic Inference

The goal of inductive inference is to update from the prior to the posterior probability distribution when new information, either in the form of data or constraints, becomes available. In general there could be several candidate distributions p_1, p_2, p_3,... that satisfy the constraints and qualify for the desired posterior. To select the posterior one has to rank all p's in increasing order of "preference" and then select the one that maximizes "preference". In order to work the ranking scheme, the measure of preference must be transitive: if p_1 is better than p_2, and p_2 is better than p_3, then p_1 is better than p_3. To implement the ranking one introduces certain quantity $S[p,q]$. The quantity $S[p,q]$ is called the entropy of p relative to the prior q, and it is designed in such a way that if p_1 is preferred over p_2, then $S[p_1,q] > S[p_2,q]$. The preferred distribution is obtained by maximizing the entropy functional.

It is desirable to first find the functional form of $S[p,q]$ that follows from certain axioms based on the process of *eliminative induction*. The method presented here is due to Caticha [31, 32] see also [30], and is based on previous work by Shore and Johnson [33] and Skilling [34].

Axiom 1. (Locality) *Local information has local effects.*

Let $x \in \mathcal{X}$, and let $\mathcal{X}$ be partitioned into two non-overlapping domains $\mathcal{D}$ and $\tilde{\mathcal{D}}$ with $\mathcal{X} = \mathcal{D} \cup \tilde{\mathcal{D}}$. Suppose that the information to be processed does not refer to $\mathcal{D}$, then according to the principle of minimal updating (PMU) the prior conditional probability $q(x|\mathcal{D})$ is not updated. It should be noted that PMU is now generalized to include any kind of information. The generalized version of PMU states that beliefs should be updated only to the extent required by the new information [30]. In the earlier section PMU was only restricted to the new information in the form of data but now it includes information either in the form of data or constraints.

The consequence of the locality criterion is that non-overlapping domains of x contribute additively to the entropy functional

$$S[p,q] = \int dx\, F(p(x), q(x), x)\,, \tag{3.9}$$

where F is a function yet to be determined.

Axiom 2. (Coordinate Invariance) *The system of coordinates carries no information.*

The process of updating remains unchanged by a mere change of coordinates. The consequence of axiom 2 is that eq. (3.9) can be written in a coordinate invariant way

$$S[p,q] = \int dx\, m(x)\, \Phi\left(\frac{p(x)}{m(x)}, \frac{q(x)}{m(x)}\right) \,, \tag{3.10}$$

where $m(x)$ is a density and Φ is another function of only two arguments in contrast to F which has three arguments.

The function Φ is still undetermined. Combining the locality and coordinate invariance criteria reduces Φ to be a function of only argument. A second use of the locality axiom 1 allows us to determine the unknown density $m(x)$. Recall the locality criterion and require that the domain $\mathcal{D}$ covers the whole space, that is $\mathcal{D} = \mathcal{X}$. This corresponds to a situation when no new information is available. The immediate consequence of this is that up to a normalization $m(x)$ turns out to be the same as the prior $q(x)$. Therefore the entropy functional (3.10) reduces to the following form

$$S[p,q] = \int dx\, q(x)\, \Phi\left(\frac{p(x)}{q(x)}\right) \,. \tag{3.11}$$

Axiom 3. (Independence) *When two systems are a priori believed to be independent and we receive independent information about each then it should not matter if one is included in the analysis of the other or not (and vice versa).*

Suppose that the system is composed of subsystems, $(x_1, x_2) \in \mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2$. If the subsystems happen to be independent then the probability of composite system should update to $p(x_1, x_2) = p(x_1)p(x_2)$ whether they are treated jointly or separately. The consequences of this is that the function Φ is determined to be $\Phi(z) = -z \log z$, and therefore

$$S[p,q] = -\int dx\, p(x)\, \log \frac{p(x)}{q(x)} \,, \tag{3.12}$$

where $q(x)$ is the prior and $p(x)$ is the posterior probability distribution.

3.4 Special Cases of ME

The entropic inference framework or ME is of general applicability. Whether the information is available in the form of data or constraints, it can be used for updating when new information becomes available. It turns out that both the MaxEnt of Jaynes and Bayes' rule are special cases of ME. Here we derive them as special cases of ME.

3.4.1 Jaynes' method of maximum entropy (MaxEnt)

Originally the method of maximum entropy (MaxEnt) was designed by Jaynes to assign probabilities on the basis of constraint information [10]. To obtain MaxEnt, we write eq. (3.12) in the discrete form (for a uniform prior $q_i =$ constant),

$$S[p] = -\sum_i p_i \log p_i \, , \tag{3.13}$$

which can be recognized as the Shannon's entropy.

Now specifying the constraints: Beside the normalization condition, $\sum_i p_i = 1$, it is assumed that the information is available in the form of expected values of certain quantities

$$\langle f^k \rangle = \sum_i p_i f_i^k = F^k \, , \tag{3.14}$$

where F^k are numerical values of the functions f^k, for $k = 1, 2, \ldots$

Next maximizing $S[p]$ subject to the constraints

$$0 = \delta\big(S[p] - \alpha \textstyle\sum_i p_i - \sum_k \lambda_k \sum_i p_i f_i^k\big) \ , \tag{3.15}$$

where λ's are Lagrange multipliers. The solution of eq. (3.15) is the generalized canonical distribution

$$p_i = \frac{\exp\left[-\sum_k \lambda_k f_i^k\right]}{Z} \tag{3.16}$$

where Z is the partition function

$$Z = \sum_i \exp\left[-\textstyle\sum_k \lambda_k f_i^k\right] \tag{3.17}$$

For example, the Maxwell-Boltzmann distribution of statistical mechanics follows immediately if the only information available is the expected energy, $\langle E\rangle = \sum_i p_i E_i$, then

$$p_i = \frac{e^{-\beta E_i}}{Z} \ . \tag{3.18}$$

This concludes that statistical physics which is commonly regarded as a "physical" theory is nothing but an example of inference.

3.4.2 Bayes' rule

Bayes' rule and MaxEnt were regarded as two parallel methods for update. Apparently they operate at different levels. The difficulties in their unification is that Bayes' rule allows for the information contained in an arbitrary prior and in data, it cannot handle arbitrary constraints. On the other hand, MaxEnt can cope with arbitrary constraints but fixed prior. In MaxEnt the prior is the underlying measure.

The unification of Bayes' rule and MaxEnt was finally achieved by Caticha and Giffin in 2006 [12, 35]. It turned out that Bayes' rule is also a special case of ME. Here we outline their work. For detailed analysis see [30].

In Bayes' rule one aims to infer one or more parameters $\theta \in \Theta$ on the basis of information available in the form of data $x \in \mathcal{X}$. But before the data is available we do not know both θ and x, therefore the relevant space is the product space $\Theta \times \mathcal{X}$. We want to update from joint prior distribution $q(x, \theta)$ to the posterior $p(x, \theta)$. To find it maximize the appropriate entropy

$$S[p, q] = -\int dx d\theta\, p(x, \theta) \log \frac{p(x, \theta)}{q(x, \theta)} \,. \tag{3.19}$$

Constraints: When the data x' is collected, the data imposes a constraint on the posterior: $p(x, \theta)$ must reflect complete knowledge of the value of x. In addition to the normalization, the relevant constraint that takes into account the observed data is given by

$$p(x) = \textstyle\int d\theta\, p(\theta, x) = \delta(x - x') \,. \tag{3.20}$$

Now use the machinery of ME

$$\delta\{S - \alpha[\textstyle\int dx d\theta\, p(x, \theta) - 1] + \textstyle\int dx\, \lambda(x)[\,\textstyle\int d\theta\, p(x, \theta) - \delta(x - x')\,]\} = 0 \,, \tag{3.21}$$

which yields

$$p(x, \theta) = q(x, \theta) \frac{e^{\lambda(x)}}{Z} \,, \tag{3.22}$$

where Z is a normalization constant and $\lambda(x)$ is the Lagrange multiplier. The Lagrange multiplier can be found by using eq. (3.20).

$$\int d\theta \, p(x,\theta) = \int d\theta \, q(x,\theta) \frac{e^{\lambda(x)}}{Z} = \delta(x-x')\,, \tag{3.23}$$

this gives

$$q(x)\frac{e^{\lambda(x)}}{Z} = \delta(x-x')\,, \tag{3.24}$$

and therefore eq. (3.22) becomes

$$p(x,\theta) = q(x,\theta)\frac{\delta(x-x')}{q(x)}\,. \tag{3.25}$$

The joint prior can be written as

$$q(x,\theta) = q(\theta)q(x|\theta)\,, \tag{3.26}$$

so that

$$p(x,\theta) = q(\theta)q(x|\theta)\frac{\delta(x-x')}{q(x)}\,. \tag{3.27}$$

Finally marginalizing over x yields

$$\int dx \, p(x,\theta) = \int dx \, q(\theta)q(x|\theta)\frac{\delta(x-x')}{q(x)}\,, \tag{3.28}$$

and thus we obtain the Bayes' rule

$$p(\theta) = q(\theta)\frac{q(x'|\theta)}{q(x')}\,. \tag{3.29}$$

3.5 Conclusions

We have shown that the extended method of maximum Entropy (ME) is capable of processing any form of information. It extends beyond the original scope of Bayes' method and MaxEnt. The ME method described here and MaxEnt are formally very similar but there is an important difference. In the original formalism of MaxEnt by Jaynes, the constraints always take the form of expected values of certain functions. ME is not restricted in this way: any type of constraint is acceptable. This will be important in the derivation of QM. For example, in chapter 5 a constraint, namely eq. (5.9), is used which is not in the form of expected values.

Another difference between MaxEnt and ME is that in both Shannon's and Jaynes' formalism, S[p] measures an amount of information. This is what people call Shannon's information — it is an "amount" (usually measured in *bits*). In the entropic inference described here "information" is not an amount. Its meaning is close to the colloquial meaning: information is what changes our beliefs. Information is the constraints that induce the updating from prior to posterior.

Entropic inference has universal applicability. It provides a framework to derive not only statistical physics but also quantum theory as an application of ME method [8].

Chapter 4

Nelson's Stochastic Mechanics

4.1 Introduction

Classical mechanics and quantum mechanics deal with physical systems very differently. The fundamental objects in the former are positions and velocities or momenta while in the latter the state of a physical system is described by the wave function. The mathematical objects in classical mechanics can be intuitively understood but the interpretation of the wave function in quantum mechanics remains controversial. Our common sense would accept it more readily if quantum phenomena were expressible in purely classical terms. Indeed there have been several attempts along these lines. One approach is due to E. Nelson. In 1966, he derived the Schrödinger equation as an entirely classical but unusual type of Brownian motion [13–15]. The particle obeys a stochastic version of $F = ma$.

This chapter is devoted to Nelson's theory of quantum mechanics. Since Brownian motion plays a crucial rule in this theory, we shall explore it in section 4.2, Nelson's formalism is reviewed in section 4.3 and the chapter is concluded in section 4.4.

4.2 Brownian Motion

Brownian motion was first noticed by the Scottish botanist R. Brown in 1827 while investigating pollen in water through a microscope. He observed that microscopic particles immersed in water exhibit a continuously zigzagging motion. The explanation of the phenomenon was unclear until 1905 when A. Einstein [36] and independently M. von Smoluchowski [37] explained that the incessant motion of small suspended particle in fluid is actually caused by the bombardment of molecules of the fluid.

Mathematically Brownian motion is a *stochastic process* which can be formally defined as below, see e.g., [38];

Definition 1. (Brownian Motion) *A real-valued stochastic process* $\{w(t) : t \geq 0\}$ *is called a Wiener process or Brownian motion if the following conditions hold:*

1. $w(0) = x$ *where* $x \in \mathcal{R}$ *is the starting point.*

2. *The process has independent increments, i.e., for all times* $0 \leq t_1 \leq t_2 \leq \ldots \leq t_n$ *the increments* $w(t_n) - w(t_{n-1}), \ldots, w(t_2) - w(t_1)$ *are independent random variables.*

3. *For all* $t \geq 0$ *and* $\Delta t > 0$*, the increments* $w(t + \Delta t) - w(t)$ *are normally distributed.*

4. *The process* $w(t)$ *is almost surely continuous.*

The process has a starting point but never ceases. The transition probability of Brownian motion keeps track of the future given the present while it is independent of the past,

$$p(x',t'|x,t) = \frac{1}{(2\pi\sigma^2\Delta t)^{d/2}} e^{-\frac{(x'-x)^2}{2\sigma^2\Delta t}}\,, \tag{4.1}$$

where $\Delta t = t' - t$ and d is the dimension of the space. The expectation and variance of the increment $\Delta w(t)$ can be computed as follows

$$\langle \Delta w^i \rangle = \int (x'^i - x^i)\, p(x',t'|x,t) dx' = 0\,, \tag{4.2}$$

and

$$\langle \Delta w^i \Delta w^j \rangle = \int (x'^i - x^i)(x'^j - x^j)\, p(x',t'|x,t) dx' = \sigma^2 \Delta t \delta^{ij}\,. \tag{4.3}$$

This shows that $\Delta w \sim O(\Delta t^{1/2})$, which means that the trajectory of the particle is continuous but it is nowhere differentiable (see eq. (4.6) below).

4.3 Nelson's Formalism

Nelson's stochastic mechanics is based on the assumption that particles in vacuum undergo a continuous Brownian motion in real space. The description presented here follows Nelson's original paper [13]. Nelson assumes that the displacement of the particle at any time t is given by the stochastic differential equation

$$\Delta x^i(t) = b^i(x(t),t)\Delta t + \Delta w^i(t)\,, \tag{4.4}$$

where $w^i(t)$ is a Wiener process with

$$\langle \Delta w^i \rangle = 0\,, \quad \text{and} \quad \langle \Delta w^i \Delta w^j \rangle = \sigma^2 \Delta t \delta^{ij}\,, \tag{4.5}$$

where $\sigma^2/2$ is the diffusion coefficient which will be later identified with $\hbar/2m$.

One can check that the velocity dx^i/dt does not exist because

$$\frac{\Delta x^i}{\Delta t} = b^i + \frac{\Delta w^i}{\Delta t} = b^i + O\left(\frac{1}{\Delta t^{1/2}}\right) \to \infty\,, \tag{4.6}$$

but its expected value does because $\langle \Delta w^i \rangle = 0$. First take $\langle \cdot\cdot \rangle$ and then the limit,

$$\lim_{\Delta t \to 0} \left\langle \frac{\Delta x^i}{\Delta t} \right\rangle = b^i\,. \tag{4.7}$$

Formally the function b^i is called the mean forward velocity defined by

$$b^i(x(t),t) = Dx^i(t) = \lim_{\Delta t \to 0^+} \frac{\langle x^i(t+\Delta t) - x^i(t) \rangle_{x(t)}}{\Delta t}\,, \tag{4.8}$$

where $Dx^i(t)$ is the mean forward derivative and $\langle . \rangle_{x(t)}$ is the conditional expectation given the state of the particle at the earlier position $x(t)$.

Similarly the mean backward velocity is defined by

$$b^i_*(x(t),t) = D_* x^i(t) = \lim_{\Delta t \to 0^+} \frac{\langle x^i(t) - x^i(t-\Delta t) \rangle_{x(t)}}{\Delta t}\,, \tag{4.9}$$

so that the displacement from the past is

$$\Delta x^i_*(t) = x^i(t) - x^i(t-\Delta t) = b^i_*(t)\Delta t + \Delta w^i_*(t)\,, \tag{4.10}$$

where w_* is again a Wiener process with

$$\left\langle \Delta w^i_* \right\rangle = 0\,, \quad \text{and} \quad \left\langle \Delta w^i_* \Delta w^j_* \right\rangle = \sigma^2 \Delta t \delta^{ij}\,. \tag{4.11}$$

Let $\rho(x,t)$ be the probability density at location x and time t, then Nelson showed that ρ evolves according to the forward Fokker-Planck equation

$$\frac{\partial \rho}{\partial t} = -\partial_i \left(b^i \rho\right) + \frac{\sigma^2}{2}\partial^2 \rho\,, \tag{4.12}$$

and according to the backward Fokker-Planck equation

$$\frac{\partial \rho}{\partial t} = -\partial_i \left(b_*^i \rho \right) - \frac{\sigma^2}{2} \partial^2 \rho \,, \tag{4.13}$$

where $\partial_i = \partial / \partial x^i$, and $\partial^2 = \delta^{ij} \partial_i \partial_j$.

Adding eqs. (4.12) and (4.13), and using

$$v^i = \frac{1}{2} \left(b^i + b_*^i \right) \,, \tag{4.14}$$

we obtain the continuity equation

$$\frac{\partial \rho}{\partial t} = -\partial_i \left(v^i \rho \right) \,, \tag{4.15}$$

where v^i is the velocity with which the probability flows. Accordingly, v^i is called the current velocity.

Subtracting eq. (4.12) from eq. (4.13) we get

$$u^i = \frac{1}{2} \left(b^i - b_*^i \right) = \frac{\sigma^2}{2} \delta^{ij} \frac{\partial_j \rho}{\rho} \,, \tag{4.16}$$

where u^i is called the osmotic velocity. It reflects the velocity of the Brownian particle in equilibrium when the external force is balanced by the osmotic force [13].

Next take the time derivative of eq. (4.16) and using eq. (4.15), we get the following useful equation

$$\frac{\partial u^i}{\partial t} = -\delta^{ij} \partial_j \left(\frac{\sigma^2}{2} \partial_k v^k + \delta_{kl} v^k u^l \right) \,. \tag{4.17}$$

We also want to compute $\partial v^i / \partial t$. To do this we proceed as follows: Let $f(x(t), t)$ be a function, then by Taylor expansion

$$\Delta f = \frac{\partial f}{\partial t} \Delta t + \Delta x^i \partial_i f + \frac{1}{2} \Delta x^i \Delta x^j \partial_i \partial_j f + \dots \,, \tag{4.18}$$

where $\Delta f = f(x(t+\Delta t), t+\Delta t) - f(x(t), t)$, and $\Delta x = x(t+\Delta t) - x(t)$. Next find the expected value of f given the state of the function at the earlier position $x(t)$ and divide it by Δt and also use eqs. (4.4), and (4.5), we get

$$Df = \left\langle \frac{\Delta f}{\Delta t} \right\rangle = \frac{\partial f}{\partial t} + b^i \partial_i f + \frac{\sigma^2}{2} \partial^2 f + \dots , \tag{4.19}$$

where D is the mean forward derivative defined earlier, eq. (4.8).

In the similar way we can also find the mean backward derivative of f,

$$D_* f = \frac{\partial f}{\partial t} + b_*^i \partial_i f - \frac{\sigma^2}{2} \partial^2 f + \dots . \tag{4.20}$$

Now set $f = D_* x^i(t) = b_*^i$ in eq. (4.19), and $f = Dx^i(t) = b^i$ in eq. (4.20), and add we get

$$a^i = \frac{\partial}{\partial t} \left(\frac{b^i + b_*^i}{2} \right) + \frac{1}{2} b^j \partial_j b_*^i + \frac{1}{2} b_*^j \partial_j b^i - \frac{\sigma^2}{2} \partial^2 \left(\frac{b^i - b_*^i}{2} \right) , \tag{4.21}$$

where a^i is the mean acceleration defined by

$$a^i = \frac{1}{2} (DD_* x^i(t) + D_* D x^i(t)) . \tag{4.22}$$

Finally using $b^i = v^i + u^i$, and $b_*^i = v^i - u^i$ in eq. (4.21) and rearrange,

$$\frac{\partial v^i}{\partial t} = a^i - v^j \partial_j v^i + u^j \partial_j u^i + \frac{\sigma^2}{2} \partial^2 u^i . \tag{4.23}$$

We have derived the stochastic dynamical equations given by eqs. (4.17) and (4.23) but further assumptions are needed to derive Schrödinger equation.

The first assumption involves the mean acceleration a^i, Nelson requires that it is given by Newton's law such that

$$F^i = m a^i = -\delta^{ij} \partial_j V , \tag{4.24}$$

where V is an external potential. Thus, in Nelson's stochastic mechanics the particle obeys a stochastic form of Newton's $F = ma$.

The second assumption of Nelson deals with the current velocity, he requires that it is a gradient of a scalar function such that

$$v^i = \delta^{ij}\partial_j \phi \,. \tag{4.25}$$

Note that osmotic velocity is also a gradient given by (4.16).

Next substitute eqs. (4.16), (4.24), and (4.25) in eqs. (4.17) and (4.23) and furthermore identify the diffusion coefficient $\sigma^2/2$ with $\hbar/2m$. After making this substitutions eqs. (4.17), and (4.23) can be combined into a single equation by introducing a complex function $\Psi = \rho^{1/2} e^{i\phi}$, we arrive at

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m}\partial^2 \Psi + V\Psi \,, \tag{4.26}$$

which is the Schrödinger equation.

In summary Nelson's derivation of Schödinger equation is based on three assumptions — the background field hypothesis, the requirements that the mean acceleration is given by the Newton law, and that the current velocity is a gradient. However in a later work [39], Nelson avoided the Newton's law and required that the diffusion process can be non-dissipative such that the expected energy is constant in time

$$\frac{dE}{dt} = \frac{d}{dt}\int dx^3 \rho \left(\frac{1}{2}mv^2 + \frac{1}{2}mu^2 + V\right) = 0 \,. \tag{4.27}$$

4.4 Conclusions

The goal of stochastic mechanics is to give a physically realistic interpretation of quantum mechanics. According to Nelson, "It is an attempt to derive and explain nonrelativistic quantum mechanics as an emergent theory in which particle trajectories are physically real and governed by stochastic laws of motion" [40].

In this picture of quantum mechanics, the wave function Ψ is not fundamental but the drift field $\vec{b}$ is the fundamental object. It is a classical theory in the sense that the fundamental equation is the Newton-Nelson Law, (4.24), which governs the evolution of the drift field. The Schrödinger equation is equivalent to eqs. (4.17) and (4.23) that determines the drift field.

The quantum theory derived in this way smells like a classical theory but not in the fullest. The theory rests on the background field hypothesis, the Brownian motion. There is no classical analog of such a field to exist in the vacuum. Also, for several particles the background field induces a Brownian motion that is highly non-local in real space R^3 [15]. It is not at all clear how a physical field that lives in real space could perform this role.

Nelson's second assumption is also ad hoc. It is not clear why the mean acceleration should take the form of eq. (4.24). However it can be more elegantly avoided by requiring the diffusion process can be non-dissipative [39].

Another serious objection was raised by T. C. Wallstrom [41, 42]. It deals with the third assumption of Nelson's theory. The current velocity is required to be a local gradient of a scalar function ϕ, eq. (4.25), it does not guarantee whether ϕ is single-valued or multi-valued. If it is single-valued then Schrödinger equation immediately follows from eqs. (4.17) and (4.23). Since ϕ is the phase of the wave function, to include non-zero angular momentum one has to allow that the phase to be multi-valued while still requiring that the wave function is single-valued. It is so because the wave functions of the angular momentum contain factors of the form $e^{im\varphi}$, where φ is the azimuthal angle and m is integer. If that is the case then the phase is multi-valued ($\phi = m\varphi$) and therefore eqs. (4.17) and (4.23) have other solutions that do not correspond to Schrödinger equation. There are also other variants of stochastic mechanics [43], which allows the current velocity be a global gradient but then the phase is single-valued and one cannot have non zero angular momentum. Whether the current velocity is a local or global gradient the solutions are either too many or too few, therefore stochastic mechanics does not reproduce the Schrödinger equation in full generality.

Chapter 5

Entropic Quantum Dynamics

5.1 Introduction

Quantum theory introduced in the last century is a highly successful physical theory. It describes the properties of atoms, nuclei, elementary particles and photons to a high level of accuracy. Despite of its successful applications at microscopic level, the formalism of standard quantum mechanics is highly abstract and non-intuitive. The basic equation is the Schrödinger equation that describes the evolution of wave function rather than the motion of particles. This gives rise to serious interpretational issues. Does the wave function correspond to reality itself or does it concern with our knowledge of reality? In philosophical terminology, the question is whether the wave function represents the ontic state (Ψ represents what the system is really, objectively doing) or whether the wave function represents an epistemic state (Ψ represents information that is known about the system — its previous history, how it was prepared, and so on).

Since the wave function is defined over an abstract configuration space rather than the real three dimensional space, the ontic interpretation makes Ψ more mysterious. Though the ontic view has a long history dating back to Schrödinger and many others, the most popular one is the de Broglie-Bohm pilot wave theory which is referred to as a *causal interpretation* [2, 44]. On the other hand the epistemic view also has a long history. Einstein's argument of quantum theory being incomplete [45] and Ballentine's *statistical interpretation* [3] favor an epistemic view of the wave function.

The advent of information theory, which handles incomplete information in a natural way, has further enhanced the epistemic view. Still in the earlier works it was only the amplitude of the wave function, $|\Psi|^2$, that represents a state of knowledge. A completely epistemological interpretation was provided by A. Caticha in 2009 [7, 8]. In this approach not only the amplitude but also the phase of the wave function is also expressed in purely informational terms. The basic assumption is that in addition to the particle of interest the world contains other variables whose entropy is reflected in the phase of the wave function. It is shown that quantum theory is also an application of the method of maximum entropy. In a more recent development [9] ED is formulated without appealing to extra variables.

This chapter is devoted to Caticha's original approach to quantum theory.

5.2 Entropic Dynamics

Here we review Caticha's approach to foundations of quantum mechanics within the framework of entropic dynamics (ED). For a detailed analysis see [8, 30]. The theory is defined on the configuration space. It is assumed that the particles have definite positions x. For a single particle the configuration space $\mathcal{X}$ is Euclidean with the metric

$$\gamma_{ab} = \delta_{ab}/\sigma^2, \quad a, b = 1, 2, 3. \tag{5.1}$$

where σ^2 is a scale factor. The full significance of the scale factor only becomes apparent when discussing several particles with different masses [8].

In addition to the particle of interest there exists other variables which we call y and live in a space $\mathcal{Y}$. We do not need to be very specific about the y variables. We will assume that their value is uncertain and that this uncertainty depends on the location x of the particle and is expressed by some probability distribution $p(y|x)$. We do not need to be very specific about $p(y|x)$ either. As we shall see it is their entropy that matters. The entropy of the y variables is given by

$$S[p, q] = -\int dy\, p(y|x) \log \frac{p(y|x)}{q(y)} = S(x)\,. \tag{5.2}$$

where $q(y)$ is some underlying measure that need not be specified further. Since x enters as a parameter in $p(y|x)$ the entropy is a function of x: $S[p, q] = S(x)$.

When the particle moves from an initial position x to a neighboring position x' and the y variables change from y to y', then we want to find the joint distribution $P(x', y'|x)$. Thus the relevant space is $\mathcal{X} \times \mathcal{Y}$, in which case, the appropriate entropy is

$$\mathcal{S}[P, Q] = -\int dx' dy' P(x', y'|x) \log \frac{P(x', y'|x)}{Q(x', y'|x)}\,, \tag{5.3}$$

where $Q(x', y'|x)$ is the prior probability distribution. The acceptable posteriors $P(x', y'|x)$ can be obtained by making use of the prior information and specifying the relevant constraints.

The prior

The prior probability distribution codifies relation between x' and y' given x before the actual information contained in the constraints has been processed. At this point we are ignorant about any relation between x' and y'. When the knowledge of x' tells us nothing about y' and vice versa, then the joint prior can be written as a product

$$Q(x', y'|x) = Q(x'|x)Q(y'|x)\,. \tag{5.4}$$

We will furthermore assume complete ignorance so that Q reflects a uniform distribution, that is, it assigns equal probabilities to equal volumes,

$$Q(x'|x)d^3x' \propto \gamma^{1/2}d^3x'\,, \tag{5.5}$$

where $\gamma = \det \gamma_{ab}$, and

$$Q(y'|x)dy' \propto q(y')dy'\,. \tag{5.6}$$

Therefore up to a proportionality constant, the joint prior becomes

$$Q(x', y'|x) = \gamma^{1/2}q(y')\,, \tag{5.7}$$

The constraints

To specify the constraints, we write the posterior as

$$P(x', y'|x) = P(x'|x)P(y'|x', x) \tag{5.8}$$

The first constraint is introduced through the second factor in eq. (5.8) which codifies information about the uncertainty in y' given x, and x'. We will assume that the uncertainty in y' depends only the present value of x', and not on the earlier value x. This means that

$$P(y'|x',x) = p(y'|x')\,, \tag{5.9}$$

where $p(y'|x')$ is the probability distribution of y variables.

The second constraint concerns the factor $P(x'|x)$ in eq. (5.8) which represents the transition probability from x to x'. We require that actual physical changes happen continuously, there is no discontinuity while moving from x to x'. To allow the continuity condition we require that x' is infinitesimally close to x. This information is incorporated in to the following constraint: Let $\Delta x = x' - x$, then we require that the expectation

$$\left\langle \Delta\ell^2 \right\rangle = \left\langle \gamma_{ab} \Delta x^a \Delta x^b \right\rangle \ , \tag{5.10}$$

be some small numerical value, which we take to be independent of x in order to reflect the translational symmetry of the space $\mathcal{X}$.

The last constraint involves the normalization condition

$$\int d^3x'\, P(x'|x) = 1 \tag{5.11}$$

Finally substituting eq. (5.7) in eq. (5.3), and incorporating the constraints eqs. (5.9), (5.10) and (5.11), the machinery of Maximum Entropy method (ME) leads to

$$P(x'|x) = \frac{1}{\zeta} e^{S(x') - \frac{1}{2}\alpha \Delta\ell^2}\,, \tag{5.12}$$

where ζ is a normalization constant and α is a Lagrange multiplier.

The transition probability $P(x'|x)$ is meant to hold for short steps, eq. (5.10). This happens when α is very large. For large α, eq. (5.12) can be approximated to a Gaussian

$$P(x'|x) \approx \frac{1}{Z} \exp\left[-\frac{\alpha}{2\sigma^2}\delta_{ab}\left(\Delta x^a - \Delta\bar{x}^a\right)\left(\Delta x^b - \Delta\bar{x}^b\right)\right] . \tag{5.13}$$

where Z is a new normalization constant. The displacement Δx^a can be expressed as an expected drift plus a fluctuation,

$$\Delta x^a = \Delta\bar{x}^a + \Delta w^a , \tag{5.14}$$

where

$$\langle \Delta x^a \rangle = \Delta\bar{x}^a = \frac{\sigma^2}{\alpha}\delta^{ab}\partial_b S(x) , \tag{5.15}$$

$$\langle \Delta w^a \rangle = 0 \quad \text{and} \quad \langle \Delta w^a \Delta w^b \rangle = \frac{\sigma^2}{\alpha}\delta^{ab} . \tag{5.16}$$

As can be seen from eq. (5.15), the particle tends to drift along the entropy gradient. For large α the step size becomes very small but the fluctuations become dominant, because $\Delta\bar{x} \sim O(\alpha^{-1})$ while $\Delta w \sim O(\alpha^{-1/2})$. It means that as $\alpha \to \infty$ the trajectory is continuous but not differentiable—just like Brownian motion.

5.3 The Construction of Entropic Time

The concept of time is closely connected with motion and change [46]. In entropic dynamics (ED) motion is described by the transition probability, eq. (5.13), that describes small changes in short steps. Larger changes will be obtained as the accumulation of very many small short steps.

To construct time in ED we must define: (a) an instant of time, (b) the temporal order of instants, (c) the duration of time [47]. We begin with constructing an instant of time. Consider the particle is initially at position x and it moves to a final position x'. In general both x and x' are unknown. This means that we must deal with the joint probability $P(x, x')$, and then using the product rule

$$P(x', x) = P(x'|x)P(x)\,. \tag{5.17}$$

We note that $P(x'|x)$ is the probability of x' given x, but x is also unknown so we marginalize over x

$$P(x') = \int P(x', x)dx = \int P(x'|x)P(x)dx\,, \tag{5.18}$$

where $P(x)$ is the probability of the particle being located at position of x and $P(x')$ is the probability of the particle being found at x'. Since x is the initial position which occurs at an initial time t and x' occurs at a later time $t' > t$, therefore we write $P(x) = \rho(x, t)$ and $P(x') = \rho(x', t')$ so that

$$\rho(x', t') = \int P(x'|x)\rho(x, t)dx\,, \tag{5.19}$$

where t and t' are different instants of time which are ordered according to earlier and later ($t' > t$).

Having introduced the notion of ordered instants in entropic dynamics the next important issue is of the duration or interval of time. Since we want to reconstruct non relativistic quantum mechanics, we need to construct Newtonian time. In Newtonian time, the time interval is independent of position x and of time t. To achieve this we assume that the Lagrange multiplier α is a constant such that

$$\alpha = \frac{\tau}{\Delta t} = \text{constant}\,, \tag{5.20}$$

where τ is a constant that sets the unit of time interval Δt.

Finally the transition probability, eq. (5.13), becomes

$$P\left(x'|x\right) \approx \frac{1}{Z}\exp\left[-\frac{\tau}{2\sigma^2\Delta t}\delta_{ab}\left(\Delta x^a - \Delta\bar{x}^a\right)\left(\Delta x^b - \Delta\bar{x}^b\right)\right]\;. \tag{5.21}$$

which can be recognized as standard Wiener process where now eq. (5.14) can be expressed in a familiar form

$$\Delta x^a = b^a\left(x\right)\Delta t + \Delta w^a, \tag{5.22}$$

where

$$b^a\left(x\right) = \frac{\sigma^2}{\tau}\delta^{ab}\partial_b S\left(x\right)\;, \tag{5.23}$$

is the drift velocity, and Δw^a is a fluctuation with

$$\left\langle \Delta w^a \right\rangle = 0 \quad \text{and} \quad \left\langle \Delta w^a \Delta w^b \right\rangle = \frac{\sigma^2}{\tau}\Delta t\delta^{ab}\,, \tag{5.24}$$

where $\sigma^2/2\tau$ is the diffusion constant.

5.4 Derivation of the Schrödinger Equation

The set of equations (5.21-5.24) describe small changes. Standard methods show that the successive iteration of eq. (5.19) yields a probability distribution $\rho\left(x,t\right)$ that evolves according to Fokker-Planck equation [15, 30, 48]

$$\frac{\partial\rho}{\partial t} = -\partial_a\left(b^a\rho\right) + \frac{\sigma^2}{2\tau}\nabla^2\rho\,, \tag{5.25}$$

which can be written as an equation for conservation of probability

$$\partial_t\rho = -\partial_a\left(\rho v^a\right) \tag{5.26}$$

Clearly v^a is interpreted as the velocity of flow of probability — it is called the current velocity. The current velocity can also be written as

$$v^a = b^a + u^a\,, \tag{5.27}$$

where b^a is the drift velocity given by eq. (5.23), and

$$u^a = -\frac{\sigma^2}{\tau}\delta^{ab}\partial_b\log\rho^{1/2}\,, \tag{5.28}$$

To interpret eq. (5.28) we write it as

$$\rho u^a = -\frac{\sigma^2}{2\tau}\delta^{ab}\partial_b\rho\,, \tag{5.29}$$

which we recognize as Fick's Law and shows that ρu^a is the probability flux due to diffusion. The velocity u^a is called the osmotic velocity.

Since both b^a (in eq. (5.23)) and u^a (in eq. (5.28)) are gradients, therefore the current velocity is also a gradient,

$$v^a = \frac{\sigma^2}{\tau}\delta^{ab}\partial_b\phi\ , \quad \text{with} \quad \phi\left(x,t\right) = S\left(x\right) - \log\rho^{1/2}\left(x,t\right) \tag{5.30}$$

In Nelson theory the current velocity was postulated to be a gradient (see eq. (4.25)). In this version of ED, this fact is derived!

The dynamics just described is standard irreversible diffusion. It is not QM. In particular there is no conservation of energy. To fix this problem we recognize that an additional constraint must be imposed. Here we borrow Nelson's brilliant idea that diffusion can be non-dissipative if the expected energy is conserved [39]. In entropic dynamics, this is achieved by allowing $p(y|x)$ and $S\left(x\right)$ to be functions of time, $S=S\left(x,t\right)$.

To this end introduce an energy functional [8, 30],

$$E[\rho,S]=\int d^{3}x\rho\left(x,t\right)\left(\frac{1}{2}mv^{2}+\frac{1}{2}\mu u^{2}+V\left(x\right)\right)\,, \tag{5.31}$$

where m and μ are constants that will be called the mass and the osmotic mass respectively.

For static potential $\dot{V}=0$, it is assumed that the energy is constant

$$\frac{dE}{dt}=0\,. \tag{5.32}$$

Otherwise the energy increases at a rate given by eq. (5.36) below.

For arbitrary initial choices of ρ and ϕ the energy conservation leads to the quantum Hamilton-Jacobi equation,

$$\eta\dot{\phi}+\frac{\eta^{2}}{2m}\left(\partial_{a}\phi\right)^{2}+V-\frac{\mu\eta^{2}}{2m^{2}}\frac{\nabla^{2}\rho^{1/2}}{\rho^{1/2}}=0\,, \tag{5.33}$$

where we have defined a new constant η so that

$$\eta\overset{\text{def}}{=}m\sigma^{2}/\tau\,, \tag{5.34}$$

In terms of η, the Focker-Planck equation (5.26) becomes

$$\dot{\rho}=-\frac{\eta}{m}\partial^{a}(\rho\partial_{a}\phi)\,. \tag{5.35}$$

Eqs. (5.33) and (5.35) are the entropic dynamical equations that determine the evolution of the dynamical variables $\phi(x,t)$ and $\rho(x,t)$.

It should be noted that eq. (5.33) can be obtained without loss of generality even when the potential is time dependent in which case the energy increases at the rate of

$$\dot{E} = \int d^3x \rho \dot{V} \, . \tag{5.36}$$

The two coupled equations (5.33) and (5.35), which involve real quantities, can be combined into a single complex equation by introducing a complex quantity

$$\Psi = \rho^{1/2} e^{i\phi} \, , \tag{5.37}$$

then

$$i\eta\dot{\Psi} = -\frac{\eta^2}{2m}\nabla\Psi + V\Psi + \frac{\eta^2}{2m}\left(1 - \frac{\mu}{m}\right)\frac{\nabla^2\left(\Psi\Psi^*\right)^{1/2}}{\left(\Psi\Psi^*\right)^{1/2}}\Psi \, . \tag{5.38}$$

This reproduces the Schrödinger equation provided $\mu = m$,

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi \, , \tag{5.39}$$

where we have also identified η with $\hbar$.

There are essentially two possibilities if $m \neq \mu$: either $\mu > 0$ or $\mu = 0$. Here we analyze both cases separately. First we consider the former case. It turns out [8, 30] that all theories with $\mu > 0$ are physically equivalent in that they can be regraduated to a theory with $\mu_{\text{new}} = m$. To show this we note that the units η and τ can always be rescaled into $\eta = \kappa\eta'$ and $\tau = \kappa\tau'$ while simultaneously rescaling ϕ into $\phi = \phi'/\kappa$ where κ is some constant. Making these substitutions in eqs. (5.35) and (5.31) we get

$$\frac{\partial\rho}{\partial t} = -\frac{\eta'}{m}\partial_a\left(\rho\partial^a\phi'\right) , \tag{5.40}$$

and

$$E[\rho, S] = \int d^3x\rho\left(\frac{\eta'^2}{2m}\left(\partial_a\phi'\right)^2 + \frac{\mu\kappa^2\eta'^2}{8m^2}\left(\partial_a\log\rho\right)^2 + V\right) . \tag{5.41}$$

Again follow the same procedure that led to eq. (5.38) we get

$$i\eta'\dot{\Psi}' = -\frac{\eta'^2}{2m}\nabla^2\Psi' + V\Psi + \frac{\eta'^2}{2m}\left(1 - \frac{\mu\kappa^2}{m}\right)\frac{\nabla^2\left(\Psi'\Psi'^*\right)^{1/2}}{\left(\Psi'\Psi'^*\right)^{1/2}}\Psi' , \tag{5.42}$$

where now $\Psi' = \rho^{1/2}e^{i\phi'}$. Since κ is just a rescaling factor which has no physical implications we can tune it so that $\mu_{\text{new}} = \mu\kappa^2 = m$, and thus we again recover the Schrödinger equation provided $\mu_{\text{new}} = m$,

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi , \tag{5.43}$$

where we dropped primes over Ψ and identified η' with $\hbar$.

The other possibility occurs for $\mu = 0$ which allows no regraduation and leads to a non-linear Schrödinger equation,

$$i\hbar\dot{\Psi} = -\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi + \frac{\hbar^2}{2m}\frac{\nabla^2\left(\Psi\Psi^*\right)^{1/2}}{\left(\Psi\Psi^*\right)^{1/2}}\Psi . \tag{5.44}$$

This case will be further discussed in chapter 6.

5.5 External Electromagnetic Field

Entropic dynamics can handle an external electromagnetic field in a natural way. If the particle is placed in an external field, it constrains the possible trajectories of the particle. To encode this additional information in the transition probability, the following constraint is to be used

$$\langle \Delta x^a A_a(x) \rangle = C \,, \tag{5.45}$$

where $A_a(x)$ are the components of the vector potential and C is a constant. This condition only constrains the expected components of displacements $\Delta\vec{x}$ along the direction of $\vec{A}$.

Carrying out the calculations as in the previous sections, the transition probability turns out to be [8, 30],

$$P(x'|x) \propto \exp\left[-\frac{m}{2\hbar\Delta t}\delta_{ab}\left(\Delta x^a - \Delta\bar{x}^a\right)\left(\Delta x^b - \Delta\bar{x}^b\right)\right] \,, \tag{5.46}$$

where the displacement Δx^a is given by

$$\Delta x^a = \Delta\bar{x}^a + \Delta w^a \,, \tag{5.47}$$

with

$$\Delta\bar{x}^a = b^a \Delta t \quad \text{where} \quad b^a = \frac{\hbar}{m}\delta^{ab}[\partial_b S - \lambda A_b] \,, \tag{5.48}$$

where λ is a Lagrange multiplier that arises due to the additional constraint, eq. (5.45). The fluctuations remain unaffected

$$\langle \Delta w^a \rangle = 0 \quad \text{and} \quad \langle \Delta w^a \Delta w^b \rangle = \frac{\hbar}{m}\Delta t \delta^{ab} \,. \tag{5.49}$$

The Fokker-Planck equation takes the form

$$\dot{\rho} = -\partial_a(\rho v^a) \,, \tag{5.50}$$

where now the current velocity is given by

$$v^a = \frac{\hbar}{m}\delta^{ab}(\partial_b\phi - \lambda A_b)\,. \tag{5.51}$$

The phase ϕ and the osmotic velocity u^a do not change,

$$\phi(x,t) = S(x,t) - \log\rho^{1/2}(x,t)\,, \tag{5.52}$$

and

$$u^a = -\frac{\hbar}{m}\delta^{ab}\partial_b\log\rho^{1/2}\,. \tag{5.53}$$

The energy functional is the same as in eq. (5.31), but now the current velocity is given by eq. (5.51),

$$E = \int d^3x\rho\left(\frac{\hbar^2}{2m}(\partial_a\phi - \lambda A_a)^2 + \frac{\hbar^2}{8m}(\partial_a\log\rho)^2 + V\right)\,. \tag{5.54}$$

Just as before, the energy conservation ($\dot{E} = 0$) can be imposed if the external potentials are time-independent ($\dot{V} = 0$ and $\dot{A} = 0$), otherwise we require that the energy increases at the rate

$$\dot{E} = \int d^3x\rho(\dot{V} + \hbar\lambda v^a\dot{A}_a) \tag{5.55}$$

In general,

$$\dot{E} - \int d^3x\rho(\dot{V} + \hbar\lambda v^a\dot{A}_a) = \int d^3x\dot{\rho}\left(\hbar\dot{\phi} + \frac{\hbar^2}{2m}(\partial_a - \lambda A_a)^2 + V - \frac{\hbar^2}{2m}\frac{\nabla^2\rho^{1/2}}{\rho^{1/2}}\right) = 0\,. \tag{5.56}$$

For arbitrary choice of $\dot{\rho}$, we have

$$\hbar\dot{\phi} + \frac{\hbar^2}{2m}(\partial_a - \lambda A_a)^2 + V - \frac{\hbar^2}{2m}\frac{\nabla^2\rho^{1/2}}{\rho^{1/2}} = 0\,. \tag{5.57}$$

This is the Hamilton-Jacobi equation in an external electromagnetic field.

Now again let $\Psi = \rho^{1/2}e^{i\phi}$, then eqs. (5.50) and (5.57) lead to the Schrödinger equation in an external electromagnetic field,

$$i\hbar\frac{\partial\Psi}{\partial t} = \frac{\hbar^2}{2m}(i\partial_a - \lambda A_a)^2\Psi + V\Psi \tag{5.58}$$

It turns out that the Lagrange multiplier λ plays the rule of electric charge e by making the identification $\lambda = e/\hbar c$.

5.6 Conclusions

This chapter was devoted to deriving quantum theory as an example of entropic inference. Entropic dynamics (ED) allows one to overcome the conceptual difficulties that arise due to the interpretation of the wave function. In ED, the wave function is fully epistemic — both the amplitude and phase of the wave function are expressed in purely informational terms.

Entropic dynamics developed here has formal similarities with Nelson's stochastic mechanics. They both are position-based theories and that the Schrödinger equation is derived as a non-dissipative diffusion. However there is an important difference. Nelson's stochastic theory is meant to be realistic, an extension of classical mechanics. On the other hand, ED operates at an epistemological level. The second difference is that the basic assumptions, the Brownian motion and that the current velocity be a gradient which are postulated in Nelson's theory, are derived in ED.

However Wallstrom's objection [41, 42] of multivaluedness of the phase of the wave function also applies to ED as the dynamical equations (5.33) and (5.35) involve single-valued functions ρ and ϕ, where ρ is amplitude and ϕ is phase of the wave function. This means that eqs.(5.33) and (5.35) are not fully equivalent to the Schrödinger equation (5.39). This apparent non-equivalence of ED and the Schrödinger equation is due to the existence of y variables as the phase ϕ involves entropy of y variables, $\phi = S - \log \rho^{1/2}$. In a more recent formulation of ED [9], the y variables are eliminated. In the newer version of ED the phase ϕ has the properties of an angular variable and satisfies a quantization condition [49]

$$\oint \vec{\nabla}\phi \cdot d\vec{\ell} = 2\pi n \,, \tag{5.59}$$

which guarantees that the wave function will remain single-valued even for multi-valued phases. Therefore ED and the Schrödinger equation are in full equivalence.

Chapter 6

Linear Momentum

In this chapter we present the first original contribution of this thesis —the theory of linear momentum and the uncertainty relations. The presentation of the material is found in the publication [50].

6.1 Introduction

In the previous chapter we have shown how the method of maximum entropy (ME) can be used to derive quantum theory from a purely informational perspective. Since entropic dynamics (ED) is formulated in configuration space, it distinguishes ED from other information-based approaches to quantum theory in that the position observable assumes a privileged role: particles have well-defined, albeit unknown, positions. This opens the possibility of explaining all other observables in purely informational terms.

The notion of momentum has undergone a remarkable evolution from Descartes' early imperfect notion of a scalar "quantity of motion" to Newton's vectorial quantity of motion, then through Lagrange's generalized momenta and Hamilton's canonical momenta, to the modern quantum version of momentum as the generator of infinitesimal translations. Each theory of motion demands its own concept of momentum. Our goal is to identify what concept, within the entropic dynamics (ED) framework, plays the role of momentum.

6.2 Momentum in Classical Mechanics

In classical mechanics linear momentum is defined as mass times velocity

$$p_i = m\frac{dx_i}{dt}\,, \tag{6.1}$$

and, more generally, momentum associated with generalized coordinate q^i in the Lagrange formalism is defined as

$$p_i = \frac{\partial L}{\partial \dot{q}^i}\,, \tag{6.2}$$

where L is the Lagrangian and $\dot{q}^i$ is the generalized velocity. Since $\dot{q}$ is not necessarily the linear velocity, the generalized momentum defined by eq. (6.2) does not necessarily represent linear momentum.

6.3 Momentum in Quantum Mechanics

In quantum mechanics momentum is the generator of a translation group. If $\psi(\vec{x})$ is a wave function, then under space translation, $\vec{x} \to \vec{x}\,' = \vec{x} + \vec{a}$, the shape of the new wave function $\psi'(\vec{x} + \vec{a})$ remains the same [51, 52]

$$\psi'(\vec{x} + \vec{a}) = \psi(\vec{x})\,, \tag{6.3}$$

or equivalently

$$\psi'(\vec{x}) = \psi(\vec{x} - \vec{a}) = \hat{U}(\vec{a})\psi(\vec{x})\,, \tag{6.4}$$

where $\hat{U}$ is a unitary operator. For infinitesimal transformation

$$\hat{U}(\delta\vec{a}) = 1 - i\delta\vec{a}\cdot\vec{p}_q/\hbar\,, \tag{6.5}$$

where $\vec{p}_q = -i\hbar\vec{\nabla}$ is the generator of translation.

A finite transformation can be obtained by using group property

$$\hat{U}(\delta a_1)\hat{U}(\delta a_2) = \hat{U}(\delta a_1 + \delta a_2)\,, \tag{6.6}$$

so that

$$\hat{U}(\vec{a}) = \lim_{N\to\infty}\left(1 - i\frac{\vec{a}\cdot\vec{p}_q}{N\hbar}\right)^N = e^{-i\vec{a}\cdot\vec{p}_q/\hbar} \tag{6.7}$$

Therefore

$$\psi'(\vec{x}) = e^{-i\vec{a}\cdot\vec{p}_q/\hbar}\psi(\vec{x})\,. \tag{6.8}$$

6.4 Momentum in ED

When quantum mechanics was invented a central problem was to identify the quantum concept of momentum that would, in the appropriate limit, correspond to the classical momentum. We face an analogous (but easier) problem: our goal is to identify what concept, within the entropic framework, may reasonably be called momentum.

Since the particle follows a Brownian non-differentiable trajectory it is clear that the classical momentum $md\vec{x}/dt$ along the trajectory cannot be defined. Nevertheless three different notions of momentum can be usefully defined. They are associated to the *drift*, *osmotic* and *current* velocities.

6.4.1 The drift momentum

The entropic dynamics developed in chapter 5 does not allow us to define the instantaneous velocity. This can be seen by recalling eq. (5.22),

$$\Delta x^a = b^a\,(x)\,\Delta t + \Delta w^a\,. \tag{6.9}$$

Since $\Delta w \sim O(\Delta t^{1/2})$, therefore

$$\lim_{\Delta t\to 0^+} \frac{\Delta x^a}{\Delta t} = b^a + O(\Delta t^{-1/2}) \to \infty\,. \tag{6.10}$$

However

$$\lim_{\Delta t\to 0^+} \left\langle \frac{\Delta x^a}{\Delta t} \right\rangle = b^a \quad \text{exists because} \quad \langle \Delta w^a \rangle = 0\,, \tag{6.11}$$

where b^a is called the drift velocity defined by

$$b^a = \lim_{\Delta t\to 0^+} \frac{1}{\Delta t} \int d^3x'\, P(x'|x) \Delta x^a\,, \tag{6.12}$$

where $x = x(t)$, $x' = x(t+\Delta t)$, and $\Delta x^a = x'^a - x^a$.

The drift velocity is a fundamental object both in Nelson's stochastic mechanics as well as in ED. However in Nelson's theory it is postulated that the drift velocity is the gradient of some scalar function. In the ED discussed in Ch. 5 this fact is derived. ED provides us that the drift velocity is the gradient of the entropy of y variables,

$$b^a(x) = \frac{\hbar}{m} \partial^a S(x)\,. \tag{6.13}$$

We are now in a position to define our first type of momentum by re-arranging eq. (6.13)

$$p_d^a = m b^a = \hbar \delta^{ab} \partial_b S(x)\,. \tag{6.14}$$

The drift momentum reflects probability flow along the entropy gradient. Since entropy is a statistical concept, therefore the drift momentum is not an attribute of the particle but of the probability distributions.

6.4.2 The osmotic momentum

The osmotic (or diffusion) effects are central to quantum behavior — As reflected by Fick's law for the osmotic current flow

$$\rho u^a = -\frac{\hbar}{2m}\delta^{ab}\partial_b \rho\,, \tag{6.15}$$

which can be rewritten as

$$u^a = -\frac{\hbar}{m}\delta^{ab}\partial_b \log \rho^{1/2}\,, \tag{6.16}$$

where u^a is called the osmotic velocity.

We can now define the second type of momentum

$$p_o^a = m u^a = -\hbar \delta^{ab}\partial_b \log \rho^{1/2}\,, \tag{6.17}$$

which is the osmotic momentum. It reflects the flow of probability by diffusion. Just like the drift momentum, the osmotic momentum is also statistical in nature.

6.4.3 The current momentum

The probability density $\rho(x,t)$ is one of the entropic dynamical variables, its time evolution is given by

$$\frac{\partial \rho}{\partial t} = -\partial_a(\rho v^a)\,, \tag{6.18}$$

where ρv^a is the total probability flux, and v^a is the velocity of the probability current. Furthermore, we have also derived in chapter 5 that v^a is a gradient of a scalar function

$$v^a = \frac{\hbar}{m}\delta^{ab}\partial_b\phi\,, \quad \text{with} \quad \phi = S - \log\rho^{1/2}\,. \tag{6.19}$$

This leads in to define the third type of momentum

$$p_c^a = mv^a = \hbar\delta^{ab}\partial_b\phi\,, \tag{6.20}$$

which is the current momentum. It reflects the total flow of probability. Since the current momentum is the gradient of phase ϕ that involves statistical attributes S and ρ, therefore the current momentum is also statistical in nature.

Thus, we have constructed three different momenta that are associated with the particle. The fourth notion of momentum that one can introduce in ED is the differential operator that generates infinitesimal translation—it coincides with the standard quantum momentum $\vec{p}_q = -i\hbar\vec{\nabla}$.

Notice that the three momenta $\vec{p}_d$, $\vec{p}_o$, and $\vec{p}_c$ are local functions of $\vec{x}$ and this makes them conceptually very different from the differential operator $\vec{p}_q$. In the next section we explore what properties they share with $\vec{p}_q$. We calculate their first and second moments and their corresponding uncertainty relations. The results below are formally similar to analogous relations derived in the context of Nelson's stochastic mechanics [53–56] and the Hall-Reginatto exact uncertainty formalism of [57].

6.5 Expected Values

The three momenta that appear in ED are not independent. Recall eq. (6.20),

$$\begin{aligned} p_c^a &= \hbar\delta^{ab}\partial_b\phi \\ &= \hbar\delta^{ab}\partial_b S - \hbar\delta^{ab}\partial_b \log\rho^{1/2} \\ &= p_d^a + p_o^a \,. \end{aligned} \tag{6.21}$$

We wish to calculate their expected values. The important theorem here is the vanishing expectation of the osmotic momentum. Using eq. (6.17) and since ρ vanishes at infinity,

$$\langle p_o^a\rangle = -\hbar\int d^3x\,\rho\,\partial^a\log\rho^{1/2} = -\frac{\hbar}{2}\int d^3x\,\partial^a\rho = 0\ . \tag{6.22}$$

The immediate consequence is that

$$\langle p_c^a\rangle = \langle p_d^a\rangle\,. \tag{6.23}$$

To study the connection to the quantum mechanical momentum we calculate

$$\langle p_q^a\rangle = \int d^3x\Psi^*\frac{\hbar}{i}\partial^a\,\Psi\ . \tag{6.24}$$

Using $\Psi = \rho^{1/2}e^{i\left(S-\log\rho^{1/2}\right)}$ and (6.22) and (6.20) one gets

$$\langle p_q^a\rangle = -i\hbar\int d^3x\rho\left(\partial^a\log\rho^{1/2} + i\partial^a S - i\partial^a\log\rho^{1/2}\right) = \hbar\left\langle\partial^a S\right\rangle\ . \tag{6.25}$$

Therefore

$$\langle\vec{p}_q\rangle = \langle\vec{p}_c\rangle = \langle\vec{p}_d\rangle\ , \tag{6.26}$$

the expectations of quantum momentum, current momentum and drift momentum coincide.

6.6 Uncertainty Relations

We start by stating a couple of definitions and an inequality. The variance of a quantity A is

$$\mathrm{Var}A = \langle (A - \langle A \rangle)^2 \rangle = \left\langle A^2 \right\rangle - \langle A \rangle^2 \ , \tag{6.27}$$

and its covariance with B is

$$\mathrm{Cov}\,(A, B) = \langle (A - \langle A \rangle)\,(B - \langle B \rangle) \rangle = \langle AB \rangle - \langle A \rangle \langle B \rangle \ . \tag{6.28}$$

The general form of uncertainty relation to be used below follows from the Schwarz inequality

$$\langle a^2 \rangle \langle b^2 \rangle \geq |\langle ab \rangle|^2 \,, \tag{6.29}$$

or

$$\begin{aligned} (\mathrm{Var}A)\,(\mathrm{Var}B) &= \left\langle (A - \langle A \rangle)^2 \right\rangle \left\langle (B - \langle B \rangle)^2 \right\rangle \\ &\geq |\langle (A - \langle A \rangle)\,(B - \langle B \rangle) \rangle|^2 = \mathrm{Cov}^2\,(A, B) \ . \end{aligned} \tag{6.30}$$

Next we apply these notions to the various momenta. An analogous calculation in the context of stochastic mechanics is given in [55].

6.6.1 Uncertainty relation for osmotic momentum

Starting from the covariance inequality

$$\begin{aligned} (\mathrm{Var}\ x^a)\left(\mathrm{Var}\ p_o^b\right) &\geq \mathrm{Cov}^2\left(x^a, p_o^b\right) \\ &= \langle x^a p_o^b \rangle - \langle x^a \rangle \langle p_o^b \rangle \,. \end{aligned} \tag{6.31}$$

The second term vanishes by (6.22) and while integrating by parts the first term

$$\begin{aligned}\langle x^a p_o^b \rangle &= -\hbar \int d^3x \, \rho x^a \delta^{bc} \partial_c \log \rho^{1/2} \\ &= -\frac{\hbar}{2} \left\{ \int dx^2 \rho x^a \delta^{bc} \rho|_{x^c=-\infty}^{x^c=\infty} - \int d^3x \, \rho \, \delta^{bc} \frac{\partial x^a}{\partial x^c} \right\} ,\end{aligned} \tag{6.32}$$

the surface term vanishes, and the second term involves the Kronecker delta

$$\frac{\partial x^a}{\partial x^c} = \delta^a_c \, , \tag{6.33}$$

so that

$$\langle x^a p_o^b \rangle = \frac{\hbar}{2} \delta^{ab} \, . \tag{6.34}$$

Therefore

$$\left(\text{Var } x^a \right) \left(\text{Var } p_o^b \right) \geq \left(\frac{\hbar}{2} \delta^{ab} \right)^2 , \tag{6.35}$$

or

$$\Delta x^a \, \Delta p_o^b \geq \frac{\hbar}{2} \delta^{ab} \, . \tag{6.36}$$

The osmotic uncertainty relation coincides with the Heisenberg uncertainty relation.

6.6.2 Uncertainty relation for drift momentum

The uncertainty relation is

$$\left(\text{Var } x^a \right) \left(\text{Var } p_d^b \right) \geq \text{ Cov}^2 \left(x^a, p_d^b \right) \, . \tag{6.37}$$

Evaluating the r.h.s. by making use of (6.28) and (6.20), we obtain

$$\begin{aligned}\text{Cov} \left(x^a, p_d^b \right) &= \langle x^a p_d^b \rangle - \langle x^a \rangle \langle p_d^b \rangle \\ &= \hbar \textstyle\int d^3x \, \rho x^a \partial^b S - \left(\int d^3x \rho x^a \right) \left(\hbar \int d^3x \, \rho \partial^b S \right) \, .\end{aligned} \tag{6.38}$$

The integrands involve ρ and ∂S which can be chosen independently. We can choose as narrow a probability distribution as we like, for example $\rho \to \delta\left(x^a - x^a_0\right)$, which trivially leads to $\mathrm{Cov}\big(x^a, p^b_d\big) \to 0$. Therefore, the uncertainty relation for drift momentum is

$$(\mathrm{Var}\ x^a)\left(\mathrm{Var}\ p^b_d\right) \geq 0 \quad \text{or} \quad \Delta x^a\, \Delta p^b_d \geq 0\,. \tag{6.39}$$

6.6.3 The Schrödinger and the Heisenberg uncertainty relations

To derive uncertainty relation for quantum momentum p_q, we calculate the second moment of quantum momentum. Using $\Psi = \rho^{1/2} e^{i\phi}$, (6.17) and (6.20) we have, after an integration by parts

$$\langle p^2_q \rangle = \int d^3x \Psi^* (\frac{\hbar}{i}\partial)^2\, \Psi = \langle p^2_c \rangle + \langle p^2_o \rangle\,. \tag{6.40}$$

Together with eqs.(6.22) and (6.26) this leads to

$$\mathrm{Var}\ p^b_q = \langle p^2_q \rangle - \langle p_q \rangle^2 = \mathrm{Var}\ p^b_c + \mathrm{Var}\ p^b_o\,, \tag{6.41}$$

then

$$\begin{aligned}
(\mathrm{Var}\ x^a)\left(\mathrm{Var}\ p^b_q\right) &= (\mathrm{Var}\ x^a)\left(\mathrm{Var}\ p^b_c\right) + (\mathrm{Var}\ x^a)\left(\mathrm{Var}\ p^b_o\right) \\
&\geq \mathrm{Cov}\left(x^a, p^b_c\right) + \mathrm{Cov}\left(x^a, p^b_o\right) \\
(\mathrm{Var}\ x^a)\left(\mathrm{Var}\ p^b_q\right) &\geq \mathrm{Cov}\left(x^a, p^b_c\right) + \left(\frac{\hbar}{2}\delta^{ab}\right)^2 .
\end{aligned} \tag{6.42}$$

Next we also want to calculate $\mathrm{Cov}\left(x^a, p^b_q\right)$. Since it involves non commuting operators, one has to modify (6.28) accordingly [55]

$$\mathrm{Cov}(\hat{A}, \hat{B}) = \frac{1}{2}\Big\langle (\hat{A} - \langle \hat{A} \rangle)(\hat{B} - \langle \hat{B} \rangle) + (\hat{B} - \langle \hat{B} \rangle)(\hat{A} - \langle \hat{A} \rangle) \Big\rangle$$

$$= \frac{1}{2}\langle \hat{A}\hat{B} + \hat{B}\hat{A}\rangle - \langle \hat{A}\rangle\langle \hat{B}\rangle \,, \tag{6.43}$$

so that

$$\text{Cov}\left(x^a, p_q^b\right) = \frac{1}{2}\langle x^a p_q^b + p_q^b x^a\rangle - \langle x^a\rangle\langle p_q^b\rangle \,. \tag{6.44}$$

Consider the first term

$$\frac{1}{2}\langle x^a p_q^b + p_q^b x^a\rangle = -\frac{i\hbar}{2}\int dx^3 \Psi^* x^a \partial^b \Psi - \frac{i\hbar}{2}\int d^3x \Psi^* \partial^b (x^a \Psi) \,. \tag{6.45}$$

Now $\Psi = \rho^{1/2}\, e^{i\phi}$, then after an integration by parts we get

$$\frac{1}{2}\langle x^a p_q^b + p_q^b x^a\rangle = \hbar \int d^3x \rho \, x^a \partial^b \phi = \langle x^a p_c^b\rangle \,. \tag{6.46}$$

Previously we have $\langle p_q^b\rangle = \langle p_c^b\rangle$. Finally eq. (6.44) becomes

$$\text{Cov}\left(x^a, p_q^b\right) = \langle x^a p_c^b\rangle - \langle x^a\rangle\langle p_c^b\rangle = \text{Cov}\left(x^a, p_c^b\right) \,. \tag{6.47}$$

Substitute eq. (6.47) in eq. (6.42), we obtain

$$\left(\text{Var } x^a\right)\left(\text{Var } p_q^b\right) \geq \text{Cov}\left(x^a, p_q^b\right) + \left(\frac{\hbar}{2}\delta^{ab}\right)^2 \,, \tag{6.48}$$

which is a version of the quantum uncertainty relation originally proposed by Schrödinger [58]. Since $\text{Cov}^2\left(x^a, p_q^b\right) \geq 0$ the somewhat weaker Heisenberg uncertainty relation follows immediately

$$\left(\text{Var } x^a\right)\left(\text{Var } p_q^b\right) \geq (\frac{\hbar}{2}\delta^{ab})^2 \quad \text{or} \quad \Delta x^a \,\Delta p_q^b \geq \frac{\hbar}{2}\delta^{ab} \,. \tag{6.49}$$

6.6.4 Uncertainty relation for current momentum

The current momentum uncertainty relation is given by

$$\left(\text{Var } x^a\right)\left(\text{Var } p_c^b\right) \geq \text{Cov}^2\left(x^a, p_c^b\right) \,. \tag{6.50}$$

We wish to place a bound on the r.h.s. of this equation. This can be done by making use of eq. (6.47) which involves quantum momentum p_q.

Recall eq. (6.44)

$$\text{Cov}\left(x^a, p_q^b\right) = \frac{1}{2}\langle x^a p_q^b + p_q^b x^a\rangle - \langle x^a\rangle\langle p_q^b\rangle \,. \tag{6.44}$$

The minimum value of the l.h.s. can be found by using, for example, the ground state of harmonic oscillator. For simplicity we consider the one dimensional harmonic oscillator. In the ground state, the wave function of harmonic oscillator is given by

$$\Psi(x) = \left(\frac{1}{\pi\sigma^2}\right)^{1/4} e^{-x^2/2\sigma^2} \,. \tag{6.51}$$

Now use $p_q = -i\hbar\partial/\partial x$, and

$$\langle\hat{A}\rangle = \int_{-\infty}^{\infty} dx\, \Psi^* \hat{A} \Psi \,. \tag{6.52}$$

We find that

$$\text{Cov}\,(x, p_q) = \frac{1}{2}\langle x p_q + p_q x\rangle - \langle x\rangle\langle p_q\rangle = 0 \,. \tag{6.53}$$

In general

$$\text{Cov}^2(x^a, p_q^b) \geq 0 \,. \tag{6.54}$$

It implies that

$$\text{Cov}^2(x^a, p_c^b) \geq 0 \,, \tag{6.55}$$

and therefore

$$(\text{Var } x^a)\left(\text{Var } p_c^b\right) \geq 0 \quad \Rightarrow \quad \Delta x^a \Delta p_c^b \geq 0 \,. \tag{6.56}$$

Results:

We now collect the results. We have constructed momentum in the framework of ED. Since the trajectory of the particle is non differentiable, several momenta are associated with the particle. In the classical limit, $\hbar \to 0$, the drift/current momenta converge to the classical momentum while the osmotic momentum tends to zero. In this limit

$$\vec{p}_d = \vec{p}_c = m\vec{v} \quad \text{and} \quad \vec{p}_o = m\vec{u} = 0\,, \tag{6.57}$$

where $\vec{v}$ is the velocity of the particle. The classical limit will be further elaborated in section 6.7.

The expected values of drift and current momenta coincide with the expectation of quantum momentum while osmotic momentum has vanishing expectation

$$\langle \vec{p}_d \rangle = \langle \vec{p}_c \rangle = \langle \vec{p}_q \rangle \ , \quad \langle \vec{p}_o \rangle = 0\,. \tag{6.58}$$

We have found that the Heisenberg uncertainty relation

$$\Delta x^a \Delta p_q^b \geq \frac{\hbar}{2}\delta^{ab}\,, \tag{6.59}$$

is the same as the osmotic momentum uncertainty relation

$$\Delta x^a \Delta p_o^b \geq \frac{\hbar}{2}\delta^{ab}\,. \tag{6.60}$$

The Heisenberg uncertainty relation is a diffusion effect, which in ED, arises due to the constraint given by eq. (5.10) (i.e. $\left\langle \gamma_{ab}\Delta x^a \Delta x^b \right\rangle = \langle \Delta\ell^2 \rangle$). This constraint is responsible for the non differentiability of the Brownian paths.

Furthermore, we have also found the drift momentum uncertainty relation

$$\Delta x^a \Delta p_d^b \geq 0\,, \tag{6.61}$$

and the current momentum uncertainty relation

$$\Delta x^a \Delta p_c^b \geq 0\,. \tag{6.62}$$

6.7 A Hybrid Theory

Non-dissipative ED is defined by the quantum Hamilton-Jacobi equation (5.33) and the Fokker-Planck equation (5.35). We can consider a different theory obtained by setting the osmotic mass, $\mu = 0$, which is neither a classical nor a quantum theory. To better understand this case we first discuss the classical limit.

Classical limit: Define $S_{HJ} = \eta\phi$ in eq. (5.33) and letting $\eta = \hbar \to 0$ with S_{HJ}, m, and μ fixed, gives the classical Hamilton-Jacobi equation

$$\dot{S}_{HJ} + \frac{1}{2m}(\partial^a S_{HJ})^2 + V = 0\,, \tag{6.63}$$

where

$$mv^a = \partial^a S_{HJ}\,, \tag{6.64}$$

is the classical momentum. In classical limits, the drift and current momentum coincide with the classical momentum while the osmotic momentum is zero.

$$p_d^a = p_c^a = \partial^a S_{HJ} \quad \text{and} \quad p_o^a = 0 \tag{6.65}$$

This suggests that up to a proportionality constant the Hamilton-Jacobi function S_{HJ} is the entropy S of the y variables

$$S_{HJ} = \hbar S \tag{6.66}$$

The fluctuations vanish in the classical limits

$$\langle \Delta w^a \Delta w^b \rangle = \frac{\hbar}{m} \Delta t \delta^{ab} \to 0\,. \tag{6.67}$$

But our main concern in this section is to focus on a different limit which also reproduces the classical Hamilton Jacobi theory but with non vanishing fluctuations.

We can also arrive at the classical Hamilton-Jacobi equation (6.63) in the limit where the osmotic mass is set to zero, $\mu = 0$, while keeping $\hbar$ and m fixed. In this case the fluctuations do not vanish

$$\langle \Delta w^a \Delta w^b \rangle = \frac{\hbar}{m} \Delta t \delta^{ab}\,. \tag{6.68}$$

The expected trajectory still lies along a classical path but the osmotic momentum does not vanish,

$$p_c^a = m v^a = \partial^a S_{HJ}, \quad \text{and} \quad p_o^a = m u^a = -\hbar \partial^a \log \rho^{1/2}\,. \tag{6.69}$$

All the considerations about momentum described in the previous sections apply to this $\mu = 0$ model. In particular, the momentum operator $\vec{p}_q = -i\hbar\vec{\nabla}$ can be introduced—for exactly the same reasons that one would introduce it in quantum theory—as a generator of translations, and this means that the $\mu = 0$ model obeys uncertainty relations identical to quantum theory

$$\Delta x^a \Delta p_o^b \geq \frac{\hbar}{2} \delta^{ab}, \quad \text{and} \quad \Delta x^a \Delta p_q^b \geq \frac{\hbar}{2} \delta^{ab}\,. \tag{6.70}$$

And yet, this is not quantum theory: the corresponding Schrödinger equation, eq.(5.44), is nonlinear and therefore there is no superposition principle.

6.8 Conclusions

We have explored the notion of momentum in entropic dynamics. We find that both the drift and current momenta converge to the classical momentum in the appropriate limit, and their expected values coincide with the expectation of quantum momentum. The Heisenberg uncertainty relation can be explained as a diffusion or osmotic effect arises due to a diffusion effect; it can be traced back to the osmotic momentum.

In the ED framework the various momenta we considered are not attributes of the particle; they are the properties of probability distributions. In ED, unlike the standard interpretation of quantum mechanics, particles have positions with definite values just as they would have in classical physics, but they do not have a momentum.

Finally, we have also briefly explored ED for the $\mu = 0$ model that exhibits both classical and quantum features. Whether it can be used to model any actual system remains to be seen.

Chapter 7

Angular Momentum

7.1 Introduction

The concept of angular momentum is important in classical as well as in quantum mechanics. It deals with the rotational symmetry of physical systems. In CM, angular momentum is the vector product of position vector and linear momentum $\vec{L} = \vec{r} \times \vec{p}$. In QM, the operator of orbital angular momentum is the generator of the rotation group $\vec{L}_q = -i\hbar\, \vec{r} \times \vec{\nabla}$. Our main concern in this chapter is to establish an entropic analog of angular momentum and to derive its corresponding uncertainty relations.

7.2 Angular Momentum in ED

As we have noted in chapter 6 there are three linear momenta, $\vec{p}_d$, $\vec{p}_o$, and $\vec{p}_c$, associated with the particle. Their corresponding angular momenta can be defined as follows

The drift angular momentum is associated with the drift momentum

$$\vec{L}_d = \vec{r} \times \vec{p}_d = \hbar\, \vec{r} \times \vec{\nabla} S\,. \tag{7.1}$$

The osmotic angular momentum is associated with the osmotic momentum

$$\vec{L}_o = \vec{r} \times \vec{p}_o = -\hbar\, \vec{r} \times \vec{\nabla} \log \rho^{1/2} \,. \tag{7.2}$$

The current angular momentum is associated with the current momentum

$$\vec{L}_c = \vec{r} \times \vec{p}_c = \hbar\, \vec{r} \times \vec{\nabla} \phi \,. \tag{7.3}$$

All these angular momenta involve various attributes of probabilities, therefore just as linear momentum, angular momentum is also statistical in nature.

We are interested to explore how these angular momenta are related to the quantum angular momentum operator $\vec{L}_q = -i\hbar \vec{r} \times \vec{\nabla}$. We calculate the expected values of $\vec{L}_q$, and solve the eigenvalue equation for the eigenstates of $\vec{L}_q$ in the following subsections.

7.2.1 Expected values

We calculate the expectation value of quantum angular momentum starting with the third component of angular momentum L_q^3

$$\left\langle L_q^3 \right\rangle = \int d^3x \Psi^* \frac{\hbar}{i} (x^1 \partial^2 - x^2 \partial^1) \Psi \,. \tag{7.4}$$

Using $\Psi = \rho^{1/2} e^{i\phi}$, we have

$$\left\langle L_q^3 \right\rangle = \hbar \int d^3x \rho (x^1 \partial^2 \phi - x^2 \partial^1 \phi) - i\hbar \int d^3x \rho \left(x^1 \frac{\partial^2 \rho}{2\rho} - x^2 \frac{\partial^1 \rho}{2\rho} \right) \,. \tag{7.5}$$

The second integral vanishes on integrating by parts, and since $\phi = S - \log \rho^{1/2}$ we have

$$\left\langle L_q^3 \right\rangle = \left\langle L_c^3 \right\rangle = \left\langle L_d^3 \right\rangle \;, \quad \text{and} \quad \left\langle L_o^3 \right\rangle = 0 \,. \tag{7.6}$$

The same equations hold for all 3 components,

$$\left\langle L_q^i \right\rangle = \left\langle L_c^i \right\rangle = \left\langle L_d^i \right\rangle \ , \quad \text{and} \quad \left\langle L_o^i \right\rangle = 0\,. \tag{7.7}$$

The expectation value of quantum angular momentum coincides with the expectations of current angular momentum and drift angular momentum while the osmotic angular momentum has vanishing expectation.

7.2.2 The eigenvalue equation

If Ψ is an eigenfunction of L_q^3 , then

$$L_q^3 \Psi = m\hbar\Psi\,. \tag{7.8}$$

Now using $L_q^3 = -i\hbar\left(x^1\partial^2 - x^2\partial^1\right)$ and $\Psi = \rho^{1/2}e^{i\phi}$, eq. (7.8) becomes

$$L_q^3\Psi = \hbar\left(x^1\partial^2\phi - x^2\partial^1\phi\right)\Psi - i\hbar\left(x^1\frac{\partial^2\rho}{2\rho} - x^2\frac{\partial^1\rho}{2\rho}\right)\Psi\,. \tag{7.9}$$

Comparing eqs. (7.8) and (7.9), and then separating into real and imaginary parts. We have

$$L_c^3 = \hbar\left(x^1\partial^2\phi - x^2\partial^1\phi\right) = m\hbar\,, \tag{7.10}$$

and

$$L_o^3 = -\hbar\left(x^1\frac{\partial^2\rho}{2\rho} - x^2\frac{\partial^1\rho}{2\rho}\right) = 0\,. \tag{7.11}$$

Since $\phi = S - \log\rho^{1/2}$, eq. (7.10) simplifies to

$$L_d^3 = \hbar\left(x^1\partial^2 S - x^2\partial^1 S\right) = m\hbar\,. \tag{7.12}$$

Results: When Ψ is the eigenstate of the $z-$component of angular momentum, then the osmotic angular momentum vanishes while the drift/current angular momenta become quantized and the eigenvalue equation, eq. (7.8) reduces to

$$L_q^3\Psi = L_c^3\Psi = L_d^3\Psi = m\hbar\Psi\,. \tag{7.13}$$

In Nelson's stochastic theory, it was shown previously that the current angular momentum is the classical analog of angular momentum [59]. It is important to emphasize that the stochastic analog of angular momentum only becomes possible by requiring that the current velocity is a gradient of some scalar function ϕ that represents phase of the wave function. On the hand hand, this fact is derived in ED and therefore both L_c and L_d are good candidates.

7.3 Angular Momentum Uncertainty Relations in QM

In standard quantum mechanics observable are represented by the Hermitian operators. If $\hat{A}$ and $\hat{B}$ are two Hermitian operators, then the uncertainties in $\hat{A}$ and $\hat{B}$ are given by

$$\Delta A\Delta B \geq \frac{1}{2}|\langle[\hat{A},\hat{B}]\rangle|\,, \tag{7.14}$$

where $(\Delta A)^2 = \langle(A-\langle A\rangle)^2\rangle$ is the uncertainty of $\hat{A}$, $(\Delta B)^2 = \langle(B-\langle B\rangle)^2\rangle$ is the uncertainty of B, and $[\hat{A},\hat{B}] = \hat{A}\hat{B}-\hat{B}\hat{A}$. For instance, when $\hat{A}=\hat{x}$, and $\hat{B}=\hat{p}_x=-i\hbar\partial/\partial x$, then

$$[\hat{x},\hat{p}_x] = i\hbar\,, \tag{7.15}$$

therefore

$$\Delta x \Delta p_x \geq \frac{\hbar}{2} \tag{7.16}$$

It should be noted that there is also a shortcoming of eq. (7.14). Take, for example, the $z-$component of angular momentum,

$$L_q^3 = -i\hbar\partial/\partial\varphi\,, \tag{7.17}$$

where φ is the azimuthal angle. This implies that

$$[\varphi, L_q^3] = i\hbar\,. \tag{7.18}$$

This suggests that

$$\Delta\varphi\Delta L_q^3 \geq \frac{\hbar}{2}\,, \tag{7.19}$$

but this is wrong because ΔL_q^3 can be arbitrarily small but $\Delta\varphi \leq 2\pi$. In fact ΔL_q^3 can even be chosen to be zero. If Ψ is an eigenfunction of L_q^3, then $\Delta L_q^3 = 0$ which implies that $\Delta\varphi = \infty$ but $\Delta\varphi \leq 2\pi$ which is a contradiction. In fact, eq. (7.18) can only be assumed if L_q^3 hermitian. Furthermore, L_q^3 is hermitian only when acting on a single-valued functions while φ is not.

The correct uncertainty relation for ΔL_q^3 follows for the functions periodic in φ such as $\cos\varphi$ and $\sin\varphi$. These are the functions that are single valued. One can compute [3, 60, 61]

$$(\Delta L_q^3)^2\{(\Delta\cos\varphi)^2 + (\Delta\sin\varphi)^2\} \geq \frac{\hbar^2}{4}(\langle\cos\varphi\rangle^2 + \langle\sin\varphi\rangle^2) \tag{7.20}$$

7.4 Angular Momentum Uncertainty Relations in ED

There are several angular momenta in ED. They are the drift, osmotic and current angular momenta. As we have seen in the previous section the angular momentum uncertainty relation involves the azimuthal angle φ. The azimuthal angle can be introduced in ED by treating $\varphi = \varphi(x, y, z)$. Here we follow Golin [62] to derive the corresponding uncertainty relations for angular momentum.

Starting with the covariance inequality (6.30). First we find uncertainty relations for the osmotic angular momentum. Let $A = \sin\varphi$, and $B = L_o^3$, then

$$(\text{Var } \sin\varphi)\left(\text{Var } L_o^3\right) \geq \text{Cov}^2\left(\sin\varphi, L_o^3\right) . \tag{7.21}$$

Similarly

$$(\text{Var } \cos\varphi)\left(\text{Var } L_o^3\right) \geq \text{Cov}^2\left(\cos\varphi, L_o^3\right) , \tag{7.22}$$

Computing the r.h.s. of eq. (7.21),

$$\text{Cov}\left(\sin\varphi, L_o^3\right) = \left\langle \sin\varphi L_o^3\right\rangle - \langle\sin\varphi\rangle\left\langle L_o^3\right\rangle , \tag{7.23}$$

the second term vanishes by using eq. (7.7) while the first term can be written as

$$\begin{aligned}\left\langle \sin\varphi L_o^3\right\rangle &= -\hbar\int d^3x\,\rho\sin\varphi\left(x^1\frac{\partial^2\rho}{2\rho} - x^2\frac{\partial^1\rho}{2\rho}\right)\\ &= \frac{\hbar}{2}\int d^3x\,\rho\cos\varphi\left(x^1\partial^2\rho - x^2\partial^1\rho\right) .\end{aligned} \tag{7.24}$$

It is convenient to express the r.h.s. in spherical polar coordinates,

$$x^1 = r\sin\theta\cos\varphi ,$$

$$x^2 = r\sin\theta\sin\varphi\,,$$

$$x^3 = r\cos\theta\,, \tag{7.25}$$

then a straightforward calculation yields,

$$\mathrm{Cov}\left(\sin\varphi, L_o^3\right) = \frac{\hbar}{2}\left\langle\cos\varphi\right\rangle\ , \tag{7.26}$$

and thus eq. (7.21) becomes

$$(\mathrm{Var}\ \sin\varphi)\left(\mathrm{Var}\ L_o^3\right) \geq \frac{\hbar^2}{4}\left\langle\cos\varphi\right\rangle^2\ . \tag{7.27}$$

Similarly

$$(\mathrm{Var}\ \cos\varphi)\left(\mathrm{Var}\ L_o^3\right) \geq \frac{\hbar^2}{4}\left\langle\sin\varphi\right\rangle^2\ . \tag{7.28}$$

The corresponding uncertainty relations for current angular momentum are given by

$$(\mathrm{Var}\ \sin\varphi)\left(\mathrm{Var}\, L_c^3\right) \geq \mathrm{Cov}^2\left(\sin\varphi, L_c^3\right)\ , \tag{7.29}$$

and

$$(\mathrm{Var}\ \cos\varphi)\left(\mathrm{Var}\, L_c^3\right) \geq \mathrm{Cov}^2\left(\cos\varphi, L_c^3\right)\ , \tag{7.30}$$

which cannot be written in a more familiar form, however the combination of eqs. (7.27)–(7.30) lead to useful results

$$(\mathrm{Var}\ \sin\varphi)\left(\mathrm{Var}\, L_c^3 + \mathrm{Var}\, L_o^3\right) \geq \mathrm{Cov}^2\left(\sin\varphi, L_c^3\right) + \frac{\hbar^2}{4}\left\langle\cos\varphi\right\rangle^2 \tag{7.31}$$

$$(\mathrm{Var}\ \cos\varphi)\left(\mathrm{Var}\, L_c^3 + \mathrm{Var}\, L_o^3\right) \geq \mathrm{Cov}^2\left(\cos\varphi, L_c^3\right) + \frac{\hbar^2}{4}\left\langle\sin\varphi\right\rangle^2\ . \tag{7.32}$$

On further combining we have

$$\left\{\mathrm{Var}\, L_c^3 + \mathrm{Var}\, L_o^3\right\}\frac{(\mathrm{Var}\ \sin\varphi) + (\mathrm{Var}\ \cos\varphi)}{\left\langle\sin\varphi\right\rangle^2 + \left\langle\cos\varphi\right\rangle^2}$$

$$\geq \frac{\mathrm{Cov}^2\left(\sin\varphi, L_c^3\right) + \mathrm{Cov}^2\left(\cos\varphi, L_c^3\right)}{\langle \sin\varphi\rangle^2 + \langle\cos\varphi\rangle^2} + \frac{\hbar^2}{4}\,. \tag{7.33}$$

Let us calculate the variance of quantum angular momentum operator L_q^3:

$$\mathrm{Var}\, L_q^3 = \left\langle \left(L_q^3\right)^2\right\rangle - \left\langle L_q^3\right\rangle^2 \tag{7.34}$$

Since

$$\begin{aligned} \left\langle \left(L_q^3\right)^2\right\rangle &= -\hbar^2 \int d^3x \Psi^* \left(x^1\partial^2 - x^2\partial^1\right)^2 \Psi \\ &= \hbar^2 \int d^3x \left|\left(x^1\partial^2 - x^2\partial^1\right)\Psi\right|^2 \\ &= \left\langle \left(L_c^3\right)^2\right\rangle + \left\langle \left(L_o^3\right)^2\right\rangle\,, \end{aligned} \tag{7.35}$$

while $\langle L_q^3\rangle = \langle L_c^3\rangle$ by eq. (7.7). Therefore

$$\mathrm{Var}\, L_q^3 = \mathrm{Var}\, L_c^3 + \mathrm{Var}\, L_o^3\,, \tag{7.36}$$

and

$$\mathrm{Cov}\left(\sin\varphi, L_q^3\right) = \mathrm{Cov}\left(\sin\varphi, L_c^3\right) \tag{7.37}$$

$$\mathrm{Cov}\left(\cos\varphi, L_q^3\right) = \mathrm{Cov}\left(\cos\varphi, L_c^3\right)\,. \tag{7.38}$$

Finally eq. (7.33) becomes

$$\mathrm{Var}\, L_q^3 \frac{\mathrm{Var}\ \sin\varphi + \mathrm{Var}\ \cos\varphi}{\langle \sin\varphi\rangle^2 + \langle\cos\varphi\rangle^2} \geq \frac{\mathrm{Cov}^2\left(\sin\varphi, L_q^3\right) + \mathrm{Cov}^2\left(\cos\varphi, L_q^3\right)}{\langle \sin\varphi\rangle^2 + \langle\cos\varphi\rangle^2} + \frac{\hbar^2}{4}\,, \tag{7.39}$$

which is a quantum angular momentum uncertainty relation (UR). This is the analogue of the Schrödinger's UR. It places a stronger bound on the uncertainty for angular momentum than that given by eq. (7.20).

7.5 Conclusions

In this chapter we have defined angular momentum within the framework of ED. We observed that there are several angular momenta. They are the drift, osmotic and current angular momenta. Just as linear momentum, angular momentum is also statistical in nature — angular momentum is an attribute of the probability distributions. We also observed that it is the current/drift angular momentum that represents entropic analog of angular momentum. We have also explored the connection of entropic angular momenta with that of quantum angular momentum. Finally, we have derived the corresponding uncertainty relations for angular momentum.

Chapter 8

Entropic Dynamics on a Curved Space

8.1 Introduction

Entropic dynamics (ED) developed in chapter 5 successfully explains quantum mechanics in the non relativistic limit. The Schrödinger equation is obtained if the particle without spin is assumed to lie in flat Euclidean space, R^3. It is desirable to generalize the formulation of ED on a curved space. The immediate consequences of this is that we can incorporate spin into ED. In this chapter, we derive a modified Schrödinger equation on a curved space. The theory of spin will be discussed in the subsequent chapter.

8.2 The Statistical Model

We consider a particle lying in an n-dimensional curved space. The generalization for several particles is immediate. For a single particle the configuration space $\mathcal{X}$ is Riemannian with metric $g_{ab}(x)$. As in flat space, in this version of ED, we assume that there are extra variables, lying in a space $\mathcal{Y}$, whose entropy is given by

$$S(x) = -\int dy\, p(y|x) \log \frac{p(y|x)}{q(y)} \ , \tag{8.1}$$

where all terms have their usual meaning as in chapter 5 (cf. eq. (5.2)).

In order to predict the new position x' we proceed as in chapter 5. We want to find $P(x'|x)$ but the actual relevant space is $\mathcal{X} \times \mathcal{Y}$. Therefore the appropriate entropy is

$$\mathcal{S}\left[P, Q\right] = -\int d^n x' dy' P\left(x', y'|x\right) \log \frac{P\left(x', y'|x\right)}{Q\left(x', y'|x\right)} \, , \tag{8.2}$$

Having specified the appropriate entropy, we want to maximize it to find the acceptable posterior $P(x', y'|x)$. The relevant information is introduced through the prior $Q(x', y'|x)$ and the constraints.

The prior

Before the actual information contained in the constraints is processed we are ignorant about any relation between x' and y', therefore the prior is a product. We also assume that the prior to be uniform which is expressed by equal probabilities for equal volumes

$$Q\left(x', y'|x\right) = g^{1/2}(x') q\left(y'\right) \ , \tag{8.3}$$

where $g(x') = \det g_{ab}(x')$.

The constraints

First we write the posterior $P(x', y'|x)$ in the following form

$$P\left(x', y'|x\right) = P\left(x'|x\right) P\left(y'|x', x\right) \ . \tag{8.4}$$

We require that x' and y' are related such that $P\left(y'|x',x\right)=p(y'|x')$, where $p(y'|x')$ is the probability distribution of y-variables. This is our first constraint and so

$$P\left(x',y'|x\right)=P\left(x'|x\right)p\left(y'|x'\right)\ . \tag{8.5}$$

The second constraint concerns the factor $P\left(x'|x\right)$ and represents the fact that actual physical changes happen continuously. We assume that x' is short step away from x, that is, $x'^a=x^a+\Delta x^a$. We require the expectation

$$\left\langle \Delta\ell^2\left(x',x\right)\right\rangle=\left\langle g_{ab}(x)\Delta x^a\Delta x^b\right\rangle=\Delta\bar{\ell}^2\left(x\right)\ , \tag{8.6}$$

be small but for now unspecified numerical value $\Delta\bar{\ell}^2\left(x\right)$ which might depend on x.

The last constraint is the normalization

$$\int d^nx'P(x'|x)=1\,. \tag{8.7}$$

Finally maximize eq. (8.2) subject to constraints (8.5), (8.6), and (8.7) to get

$$P(x'|x)=\frac{1}{\zeta(x,\alpha)}g^{1/2}(x')\exp\left[S(x')-\tfrac{1}{2}\alpha(x)g_{ab}(x)\Delta x^a\Delta x^b\right], \tag{8.8}$$

where

$$\zeta(x,\alpha)=\int d^nx'g^{1/2}(x')e^{S(x')-\frac{1}{2}\alpha(x)g_{ab}(x)\Delta x^a\Delta x^b}\,. \tag{8.9}$$

The Lagrange multiplier α can be determined from the constraint (8.6)

$$\frac{\partial}{\partial\alpha}\log\zeta(x,\alpha)=-\frac{1}{2}\Delta\bar{\ell}^2\,. \tag{8.10}$$

The Jacobian factor $g^{1/2}(x')$ in eq. (8.8) makes $P(x'|x)d^nx'$ invariant under changes of the coordinates x' in the space $\mathcal{X}$. Sometimes it is convenient to write eq. (8.8) as

$$P(x'|x)=g^{1/2}(x')\mathcal{P}(x'|x)\,, \tag{8.11}$$

with

$$\mathcal{P}(x'|x) = \frac{1}{\zeta(x,\alpha)} \exp\left[S(x') - \tfrac{1}{2}\alpha(x) g_{ab}(x) \Delta x^a \Delta x^b\right] , \tag{8.12}$$

where $\mathcal{P}(x'|x)$ is a scalar density, an invariant. Its transformation does not involve a Jacobian. On the other hand $P(x'|x)$ is a tensor density of rank zero and weight 1. Its transformation involves a Jacobian.

The transition probability $\mathcal{P}(x'|x)$ holds for short steps which can be guaranteed if the Lagrange multiplier α is large. To examine this limit we write eq. (8.12) in locally Cartesian coordinates, called the *normal coordinates*. In normal coordinates (NC) at a point p the metric tensor in the vicinity of p is approximately that of flat Euclidean space. That is, at the point p we can choose

$$g_{ab}(x_p) = \gamma_{ab} \,, \quad \text{with} \quad \gamma_{ab} = \frac{\delta_{ab}}{\sigma^2} \,, \tag{8.13}$$

and

$$\left.\frac{\partial g_{ab}}{\partial x^c}\right|_{x_p} = 0 \,. \tag{8.14}$$

However the second derivatives do not vanish

$$\left.\frac{\partial^2 g_{ab}}{\partial x^c \partial x^d}\right|_{x_p} \neq 0 \,, \tag{8.15}$$

they are the effects of curvature if the manifold is not exactly flat.

For large α or short step, expanding the exponent of eq. (8.12)) about its maximum.

$$\mathcal{P}\left(x'|x\right) \approx \frac{1}{Z\left(x\right)} \exp\left[-\frac{\alpha\left(x\right)}{2\sigma^2} \delta_{ab} \left(\Delta x^a - \Delta \bar{x}^a\right) \left(\Delta x^b - \Delta \bar{x}^b\right)\right] . \tag{8.16}$$

This is the expression for transition probability in normal coordinates which is obviously a Gaussian. The factors independent of x' are absorbed into a new normalization $Z(x)$. The displacement Δx^a and the expected drift $\Delta\bar{x}^a$ are given by

$$\Delta x^a = \Delta\bar{x}^a + \Delta w^a \ , \tag{8.17}$$

$$\Delta\bar{x}^a = \frac{\sigma^2}{\alpha(x)}\delta^{ab}\partial_b S\,(x) \ , \tag{8.18}$$

and the fluctuations

$$\langle\Delta w^a\rangle = 0\,, \quad \text{and} \quad \langle\Delta w^a \Delta w^b\rangle = \frac{\sigma^2}{\alpha(x)}\delta^{ab}\,. \tag{8.19}$$

Having established diffusion process in normal coordinates its transformation to any other coordinates is immediate. If $x^{a'}$ is another coordinate in the neighborhood of $x^a(x^{a'})$, then by Taylor expansion

$$\Delta x^a = \frac{\partial x^a}{\partial x^{a'}}\Delta x^{a'} + \frac{1}{2}\Delta x^{a'}\Delta x^{b'}\frac{\partial^2 x^a}{\partial x^{a'}\partial x^{b'}}\,. \tag{8.20}$$

It should be noted that in ordinary differential geometry, one only keeps the first term on the right. But in our case the displacement Δx^a involves fluctuations which is of the order of $\mathcal{O}(t^{1/2})$, therefore the second order term must be included in the Taylor expansion, which is $\Delta x^a \Delta x^b \sim \mathcal{O}(t)$. Notice also that Δx^a does not transform like a vector because the second order term on the right spoils its vectorial nature. It is convenient to define a new quantity,

$$\tilde{\Delta} x^a = \Delta x^a \quad \text{in NC,} \tag{8.21}$$

and require that it transforms like a vector [15]. Then

$$\tilde{\Delta} x^a = \frac{\partial x^a}{\partial x^{a'}} \Delta x^{a'} , \tag{8.22}$$

Now for $\Delta x^{a'}$ on the l.h.s. of eq. (8.22) use eq. (8.20) by reversing the roles of the prime indices

$$\begin{aligned} \tilde{\Delta} x^{a'} &= \frac{\partial x^{a'}}{\partial x^a} \Delta x^a \\ &= \frac{\partial x^{a'}}{\partial x^a} \left(\frac{\partial x^a}{\partial x^{b'}} \Delta x^{b'} + \frac{1}{2} \Delta x^{b'} \Delta x^{c'} \frac{\partial^2 x^a}{\partial x^{b'} \partial x^{c'}} \right) \\ &= \frac{\partial x^{a'}}{\partial x^a} \frac{\partial x^a}{\partial x^{b'}} \Delta x^{b'} + \frac{1}{2} \Delta x^{b'} \Delta x^{c'} \frac{\partial x^{a'}}{\partial x^a} \frac{\partial^2 x^a}{\partial x^{b'} \partial x^{c'}} . \end{aligned} \tag{8.23}$$

The second derivative on the r.h.s. can be eliminated by introducing the *Christoffel symbols*. The Christoffel symbols are defined by

$$\Gamma^a_{bc} = \frac{1}{2} g^{ad} \left(\frac{\partial g_{bd}}{\partial x^c} + \frac{\partial g_{cd}}{\partial x^b} - \frac{\partial g_{bc}}{\partial x^d} \right) . \tag{8.24}$$

The transformation rule of Γ^a_{bc} is

$$\Gamma^{a'}_{b'c'} = \frac{\partial x^{a'}}{\partial x^a} \frac{\partial x^b}{\partial x^{b'}} \frac{\partial x^c}{\partial x^{c'}} \Gamma^a_{bc} + \frac{\partial x^{a'}}{\partial x^a} \frac{\partial^2 x^a}{\partial x^{b'} \partial x^{c'}} , \tag{8.25}$$

Since the unprimed coordinates in eq. (8.23) are normal coordinates we have $\Gamma^a_{bc} = 0$ because they involve first derivative of metric tensor. Therefore

$$\tilde{\Delta} x^{a'} = \Delta x^{a'} + \frac{1}{2} \Delta x^{b'} \Delta x^{c'} \Gamma^{a'}_{b'c'} . \tag{8.26}$$

Re-arranging eq. (8.26)

$$\Delta x^{a'} = \tilde{\Delta} x^{a'} - \frac{1}{2} \Delta x^{b'} \Delta x^{c'} \Gamma^{a'}_{b'c'} , \tag{8.27}$$

which is the displacement in general coordinates. On taking the expectation

$$\Delta\bar{x}^{a'} = \tilde{\Delta}\bar{x}^{a'} - \frac{1}{2}\langle\Delta x^{b'}\Delta x^{c'}\rangle\Gamma^{a'}_{b'c'} \,. \tag{8.28}$$

Recall eq. (8.18), in normal coordinates $\tilde{\Delta}\bar{x}^{a'}$ is given by eq. (8.21) and since it is assumed that it transforms like a vector, therefore in general coordinates

$$\tilde{\Delta}\bar{x}^{a'} = \frac{1}{\alpha(x)}g^{a'b'}\partial_{b'}S \,. \tag{8.29}$$

Similarly $\langle\Delta x^{a'}\Delta x^{b'}\rangle$ in normal coordinates is given by eqs. (8.17), (8.18), and (8.19)

$$\langle\tilde{\Delta}x^{a'}\tilde{\Delta}x^{b'}\rangle = \frac{\gamma^{a'b'}}{\alpha(x)} \quad \text{with} \quad \gamma^{a'b'} = \sigma^2\delta^{a'b'} \tag{8.30}$$

Therefore in general coordinates

$$\langle\Delta x^{a'}\Delta x^{b'}\rangle = \frac{g^{a'b'}(x)}{\alpha(x)} \,, \tag{8.31}$$

where

$$g^{a'b'}(x) = \frac{\partial x^{a'}}{\partial x^{a''}}\frac{\partial x^{b'}}{\partial x^{b''}}\gamma^{a''b''} \,. \tag{8.32}$$

Finally eq. (8.27) can also be written as an expectation plus a fluctuation

$$\Delta x^{a'} = \Delta\bar{x}^{a'} + \Delta w^{a'} \,, \tag{8.33}$$

where now

$$\Delta\bar{x}^{a'} = \tilde{\Delta}\bar{x}^{a'} - \frac{1}{2\alpha(x)}\Gamma^{a'} = \frac{1}{\alpha(x)}\left(g^{a'b'}\partial_{b'}S - \frac{1}{2}\Gamma^{a'}\right) \,, \text{ with } \Gamma^{a'} = g^{b'c'}\,\Gamma^{a'}_{b'c'} \,, \tag{8.34}$$

$$\langle\Delta w^{a'}\rangle = 0 \quad \text{and} \quad \langle\Delta w^{a'}\Delta w^{b'}\rangle = \frac{g^{a'b'}(x)}{\alpha(x)} \,, \tag{8.35}$$

which describes diffusion process in general coordinates. In what follows, we drop primes over the indices.

8.3 Entropic Time in Riemannian Manifolds

The notion of time introduced in chapter 5 can be easily extended to Riemannian manifolds simply by defining

$$\rho(x', t') = \int d^n x \, g^{1/2}(x)\rho(x,t)\mathcal{P}(x'|x)\,. \tag{8.36}$$

Note that the ρ is a scalar function.

Next is to introduce the duration of time (time interval). In chapter 5 we introduced the time interval Δt though the Lagrange multiplier α (see eq. (5.20)). The same form of α continues to hold in curved spaces within non-relativistic regime. Therefore

$$\alpha = \frac{\tau}{\Delta t} = \text{constant}\,. \tag{8.37}$$

Having introduced time we can now define drift velocities in a similar manner

$$b^a(x) = \lim_{\Delta t \to 0^+} \frac{\Delta \bar{x}^a}{\Delta t} = \lim_{\Delta t \to 0^+} \frac{1}{\Delta t} \int d^n x' g^{1/2}(x')\mathcal{P}(x'|x)\Delta x^a \ , \tag{8.38}$$

which is the *future drift velocity.* The displacement is by eqs. (8.33-8.35)

$$\Delta x^a = b^a(x)\Delta t + \Delta w^a \,, \tag{8.39}$$

where

$$\langle \Delta w^a \rangle = 0 \quad \text{and} \quad \langle \Delta w^a \Delta w^b \rangle = \frac{1}{\tau} g^{ab} \Delta t \,, \tag{8.40}$$

$$b^a(x) = \tilde{b}^a(x) - \frac{1}{2\tau}\Gamma^a \;, \quad \text{with} \quad \tilde{b}^a(x) = \frac{1}{\tau} g^{ab}\partial_b S(x) \,, \tag{8.41}$$

Note that b^a does not transform like a vector but $\tilde{b}^a$ does.

Similarly the *past drift velocity* is defined by

$$b^a_*(x') = \lim_{\Delta t \to 0^+} \frac{1}{\Delta t} \int d^n x g^{1/2}(x) \mathcal{P}(x|x') \Delta x^a \tag{8.42}$$

The connection between the two drift velocities is given by

$$b^a_*(x) = b^a(x) - \frac{1}{\tau} g^{ab}\partial_b \log \rho(x,t) \;, \tag{8.43}$$

which follows as a straightforward application of Bayes' theorem

$$\mathcal{P}(x|x') = \frac{\mathcal{P}(x)}{\mathcal{P}(x')} \mathcal{P}(x'|x) \,. \tag{8.44}$$

8.4 Fokker-Planck Equation

The transition probability $\mathcal{P}(x'|x)$ holds for small changes. The result of building up finite change from initial time t_0 up to final time t leads to the density

$$\rho\,(x,t) = \textstyle\int d^n x_0 \, g^{1/2}(x_0) \rho\,(x_0,t_0) \, \mathcal{P}\,(x,t|x_0,t_0) \;, \tag{8.45}$$

where the finite-time transition probability, $\mathcal{P}\,(x,t|x_0,t_0)$, can be constructed by iterating the infinitesimal changes

$$\mathcal{P}\,(x,t+\Delta t|x_0,t_0) = \textstyle\int d^n z g^{1/2}(z) \mathcal{P}\,(z,t|x_0,t_0) \, \mathcal{P}\,(x,t+\Delta t|z,t) \;. \tag{8.46}$$

The integral equation can be written as differential equation by a Taylor expanding in Δt. Since $\mathcal{P}\,(x,t+\Delta t|z,t) \to \delta(x-z)$ as $\Delta t \to 0$. To avoid this singular behavior we multiply eq. (8.46) by a smooth function $f(x)$ and integrate over x

$$\int d^n x\, g^{1/2}(x) \mathcal{P}(x, t+\Delta t|x_0, t_0)\, f(x) = \int d^n z\, g^{1/2}(z) \mathcal{P}(z, t|x_0, t_0)\, F(z)\,, \tag{8.47}$$

where

$$F(z) = \int d^n x\, g^{1/2}(x) \mathcal{P}(x, t+\Delta t|z, t)\, f(x)\ . \tag{8.48}$$

Expand $f(x)$ about z

$$\begin{aligned} F(z) &= \int d^n x g^{1/2}(x) \mathcal{P}(x, t+\Delta t|z, t) \left(f(z) + \frac{\partial f}{\partial z^a}\Delta z^a + \frac{1}{2}\frac{\partial^2 f}{\partial z^a \partial z^b}\Delta z^a \Delta z^b + \ldots \right) \\ &= f(z) + \Delta t b^a(z) \frac{\partial f}{\partial z^a} + \frac{1}{2\tau}\Delta t g^{ab} \frac{\partial^2 f}{\partial z^a \partial z^b} + \ldots\,, \end{aligned} \tag{8.49}$$

where $\Delta z^a = x^a - z^a$, and by eqs. (8.38) and (8.40)

$$\lim_{\Delta t \to 0^+} \frac{1}{\Delta t} \int d^n x g^{1/2}(x) \mathcal{P}(x, t+\Delta t|z, t)\, \Delta z^a = b^a(z)\,, \tag{8.50}$$

$$\lim_{\Delta t \to 0^+} \frac{1}{\Delta t} \int d^n x g^{1/2}(x) \mathcal{P}(x, t+\Delta t|z, t)\, \Delta z^a \Delta z^b = \frac{1}{\tau} g^{ab}\,. \tag{8.51}$$

Substituting eq. (8.49) in to eq. (8.47) and then integrating by parts

$$\begin{aligned} \int d^n x\, g^{1/2}(x) \frac{\Delta \mathcal{P}}{\Delta t} f(x) &= \int d^n z [-\partial_a \left(g^{1/2} b^a \mathcal{P}(z, t|x_0, t_0) \right) \\ &\quad + \frac{1}{2\tau}\partial_a \partial_b \left(g^{1/2} g^{ab}\, P(z, t|x_0, t_0) \right)] f(z)\,, \end{aligned} \tag{8.52}$$

where $\partial_a = \partial/\partial z^a$, and $\Delta\mathcal{P} = \mathcal{P}(x, t+\Delta t|x_0, t_0) - \mathcal{P}(x, t|x_0, t_0)$. Finally let $\Delta t \to 0$, and since $f(x)$ is an arbitrary test function we get

$$\frac{\partial}{\partial t}\mathcal{P}(x, t|x_0, t_0) = -\frac{1}{\sqrt{g}}\partial_a \left(\sqrt{g} b^a \mathcal{P}(x, t|x_0, t_0) \right) + \frac{1}{2\tau}\frac{1}{\sqrt{g}}\partial_a \partial_b \left(\sqrt{g} g^{ab}\, P(x, t|x_0, t_0) \right)\,, \tag{8.53}$$

where now $\partial_a = \partial/\partial x^a$. Next multiply by $g^{1/2}(x_0)\rho(x_0,t_0)$, integrate eq. (8.53) with respect to x_0, and make use of eq. (8.45) to get the FP equation,

$$\frac{\partial}{\partial t}\rho(x,t) = -\frac{1}{\sqrt{g}}\partial_a\left(\sqrt{g}b^a\rho(x,t)\right) + \frac{1}{2\tau}\frac{1}{\sqrt{g}}\partial_a\partial_b\left(\sqrt{g}g^{ab}\rho(x,t)\right) \ . \tag{8.54}$$

Recall eq. (8.41)

$$b^a(x) = \tilde{b}^a(x) - \frac{1}{2\tau}\Gamma^a \,. \tag{8.41}$$

Use the identity

$$\Gamma^a = -\frac{1}{\sqrt{g}}\partial_b\left(\sqrt{g}g^{ab}\right) \ . \tag{8.55}$$

and finally eq. (8.54) becomes

$$\frac{\partial\rho}{\partial t} = -\frac{1}{\sqrt{g}}\partial_a\left(\sqrt{g}\,\tilde{b}^a\rho\right) + \frac{1}{2\tau}\Delta_g\rho \,, \tag{8.56}$$

which is the forward Fokker-Planck equation, where Δ_g is the Laplace-Beltrami operator given by

$$\Delta_g = \frac{1}{\sqrt{g}}\partial_a\left(\sqrt{g}g^{ab}\partial_b\right) \tag{8.57}$$

The Fokker-Planck equation can be written as a continuity equation

$$\frac{\partial\rho}{\partial t} = -\frac{1}{\sqrt{g}}\partial_a\left(\sqrt{g}v^a\rho\right) \ , \tag{8.58}$$

where v^a is the current velocity

$$v^a = \tilde{b}^a + u^a \,, \tag{8.59}$$

the osmotic velocity u^a is given by

$$u^a = -\frac{1}{2\tau}g^{ab}\frac{\partial_b\rho}{\rho} \,. \tag{8.60}$$

The continuity equation can also be written as

$$\frac{\partial \rho}{\partial t} = -\frac{1}{\tau\sqrt{g}}\partial_a\left(\sqrt{g}\rho g^{ab}\partial_b\phi\right) \ , \tag{8.61}$$

where

$$v^a = \frac{1}{\tau}g^{ab}\partial_b\phi \quad \text{with} \quad \phi(x,t) = S(x,t) - \log\rho^{1/2}(x,t) \, . \tag{8.62}$$

Note that since ρ and S are scalars (not densities) then v^a and u^a are vectors, and ϕ is a scalar.

8.5 The Schrödinger Equation in Riemannian Manifolds

We now derive the Schrödinger equation in curved space. To derive it, we recall that what we have learned so far as the joint evolution of ρ and S: ρ evolves as constrained by S, and then S is updated by the new ρ, in such a way the energy E is conserved. We start with the appropriate energy functional,

$$E[S,\rho] = \textstyle\int d^n x\sqrt{g}\,\rho[Ag_{ab}v^a v^b + Bg_{ab}u^a u^b + V(x) + V_c(x)] \, , \tag{8.63}$$

where $V(x)$ is the usual external potential, and $V_c(x)$ is potential due to a possible curvature effect. In a flat space, or a space of constant curvature V_c becomes a constant.

Having specified the energy functional next we put the values of v^a and u^a from eqs. (8.60) and (8.62) respectively

$$E[S,\rho] = \int d^n x\sqrt{g}\,\rho\left(\frac{A}{\tau^2}g^{ab}\partial_a\phi\partial_b\phi + \frac{B}{4\tau^2}g^{ab}\frac{\partial_a\rho\partial_b\rho}{\rho^2} + V(x) + V_c(x)\right) \, . \tag{8.64}$$

We assume that the Riemannian manifold is static (i.e. $g_{ab}(x)$ is independent of time). Furthermore the potential terms are also assumed to be time independent, then the time derivative of eq. (8.64) gives

$$\begin{aligned}\dot{E} &= \int d^n x \sqrt{g}\, \dot{\rho} \left(\frac{A}{\tau^2} g^{ab} \partial_a \phi \partial_b \phi + \frac{B}{4\tau^2} g^{ab} \frac{\partial_a \rho \partial_b \rho}{\rho^2} + V(x) + V_c(x) \right) \\ &+ \frac{2A}{\tau^2} \int d^n x \sqrt{g} \rho g^{ab} \partial_a \phi \partial_b \dot{\phi} \\ &+ \frac{B}{2\tau^2} \int d^n x \sqrt{g} \rho g^{ab} \left(\frac{\partial_a \rho \partial_b \dot{\rho}}{\rho^2} - \frac{\partial_a \rho \partial_b \rho}{\rho^3} \dot{\rho} \right) . \end{aligned} \tag{8.65}$$

Consider the second and third terms

$$I_1 = \frac{2A}{\tau^2} \int d^n x \sqrt{g} \rho g^{ab} \partial_a \phi \partial_b \dot{\phi} \,, \tag{8.66}$$

$$I_2 = \frac{B}{2\tau^2} \int d^n x \sqrt{g} \rho g^{ab} \left(\frac{\partial_a \rho \partial_b \dot{\rho}}{\rho^2} - \frac{\partial_a \rho \partial_b \rho}{\rho^3} \dot{\rho} \right) . \tag{8.67}$$

Integrate by parts the r.h.s. of eq. (8.66) and use eq. (8.61) to get

$$I_1 = \frac{2A}{\tau} \int d^n x \sqrt{g} \dot{\rho} \dot{\phi} \,. \tag{8.68}$$

Similarly eq. (8.67) gives

$$I_2 = -\frac{B}{2\tau^2} \int d^n x \sqrt{g} \left(\frac{1}{\sqrt{g}} \frac{\partial_a (\sqrt{g} g^{ab} \partial_b \rho)}{\rho} \right) \dot{\rho} \,. \tag{8.69}$$

Thus eq. (8.65) becomes

$$\begin{aligned}\dot{E} &= \int d^n x \sqrt{g}\, \dot{\rho} \left(\frac{2A}{\tau} \dot{\phi} + \frac{A}{\tau^2} g^{ab} \partial_a \phi\, \partial_b \phi \right. \\ &\qquad \left. + V(x) + V_c(x) - \frac{B}{\tau^2} \frac{\Delta_g \sqrt{\rho}}{\sqrt{\rho}} \right) , \end{aligned} \tag{8.70}$$

where we have also used the identity

$$\frac{\Delta_g\sqrt{\rho}}{\sqrt{\rho}} = \frac{1}{\sqrt{g}}\frac{\partial_a(\sqrt{g}\,g^{ab}\partial_b\rho)}{2\rho} - g^{ab}\frac{\partial_a\rho\,\partial_b\rho}{4\rho^2} \tag{8.71}$$

Any time t can be considered the initial time for evolution into the future. If we require that $\dot{E}=0$ for arbitrary choices of the initial $\dot{\rho}$ which from eq. (8.61), is equivalent to arbitrary choices of initial ρ and initial ϕ then the integrand of eq. (8.70) is zero

$$\frac{2A}{\tau}\dot{\phi} + \frac{A}{\tau^2}g^{ab}\partial_a\phi\,\partial_b\phi + V(x) + V_c(x) - \frac{B}{\tau^2}\frac{\Delta_g\sqrt{\rho}}{\sqrt{\rho}} = 0\ , \tag{8.72}$$

which can be recognized as the Hamilton-Jacobi equation in a curved space. The Hamilton-Jacobi equation and the continuity equation (8.61) can be combined in to a single equation by introducing $\Psi = \sqrt{\rho}\,e^{i\phi}$. The result is

$$i\frac{2A}{\tau}\frac{\partial\Psi}{\partial t} = -\frac{A}{\tau^2}\Delta_g\Psi + V(x)\Psi + V_c(x)\Psi + \frac{A}{\tau^2}\left(1-\frac{B}{A}\right)\frac{\Delta_g\sqrt{\rho}}{\sqrt{\rho}}\Psi\,. \tag{8.73}$$

It should be noted that the Laplace-Beltrami operator Δ_g involves the metric g^{ab} of the configuration space. It is convenient to write g^{ab} in terms of the metric h^{ab} of the curved space,

$$g^{ab}(x) = \sigma^2 h^{ab}(x)\,. \tag{8.74}$$

where σ^2 is scale factor.

We now introduce new constants m, μ and η such that

$$m = \frac{2A}{\sigma^2} \quad \text{and} \quad \mu = \frac{2B}{\sigma^2}\ , \tag{8.75}$$

and

$$\eta = \frac{2A}{\tau} \quad \text{so that} \quad \frac{\sigma^2}{\tau} = \frac{\eta}{m}\ , \tag{8.76}$$

In terms of the new constants eq. (8.73) becomes

$$i\eta\frac{\partial\Psi}{\partial t} = -\frac{\eta^2}{2m}\Delta_h\Psi + V(x)\Psi + V_c(x)\Psi + \frac{\eta^2}{2m}\left(1-\frac{\mu}{m}\right)\frac{\Delta_h\sqrt{\rho}}{\sqrt{\rho}}\Psi\,, \tag{8.77}$$

where now

$$\Delta_h = \frac{1}{\sqrt{h}}\partial_a\left(\sqrt{h}h^{ab}\partial_b\right)\,. \tag{8.78}$$

The last term in eq. (8.77) can be dropped out by rescaling the constants η and τ as $\eta = \kappa\eta'$, $\tau = \tau'/\kappa$, and introducing a new $\Psi' = \rho^{1/2}e^{i\phi'}$ where $\phi = \phi'/\kappa$, so that

$$i\eta\frac{\partial\Psi'}{\partial t} = -\frac{\eta^2}{2m}\Delta_h\Psi' + V(x)\Psi' + V_c(x)\Psi' + \frac{\eta^2}{2m}\left(1-\frac{\mu\kappa^2}{m}\right)\frac{\Delta_h\sqrt{\rho}}{\sqrt{\rho}}\Psi'\,, \tag{8.79}$$

We can choose κ such that $\mu\kappa^2 = m$, and by setting $\eta = \hbar$, and also dropping primes over Ψ for brevity, then

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\Delta_h\Psi + V(x)\Psi + V_c(x)\Psi\,. \tag{8.80}$$

This is the familiar Schrödinger equation in a curved space. The modified Schrödinger equation for a curved space is different from its counterpart in flat space for two reasons: Firstly, the Laplacian $\nabla^2 = \delta^{ab}\partial_a\partial_b$ is replaced by the Laplace-Beltrami operator Δ_h and secondly there is a possible additional potential term $V_c(x)$ due to curvature.

The curvature term has a long history. B. Podolsky [63] in 1928 derived the Schrödinger equation in curved space without the curvature potential. Using *path integral formulation* B. DeWitt [64] in 1957 proposed that there is a curvature potential proportional to *Ricci scalar curvature* R, that is $V_c(x) \propto R$. However for a particle constrained to move in a two dimensional curved surface, Jensen and Koppe [65], da Costa [66] and more recently Inomata and Junker [67] have noted that the curvature term is of the form of

$$V_c(x) = \frac{\hbar^2}{2m}\left(K(x) - H^2(x)\right) , \tag{8.81}$$

where $K(x)$ is the *Gaussian curvature* and $H(x)$ is the *mean curvature* of the surface.

The Schrödinger equation in a curved space has been also a subject of interest in the framework of Nelson's stochastic mechanics. In 1978, Dohrn and Guerra [68] derived the Schrödinger equation in the framework of Nelson's stochastic mechanics on Riemannian manifolds without the curvature term. However, here we remark that the curvature term does not arise in Ref. [68] because it is not included in the Lagrangian from the start.

In entropic dynamics, on the other hand, the curvature term is quite arbitrary it only enters through the energy constraint, namely eq. (8.63). However this term can be thrown away for a space of constant curvature because in that case it only introduces an overall irrelevant phase factor.

8.6 Interaction with External Electromagnetic Field

The interaction of charged particle with an external EM field is represented by imposing an additional constraint,

$$\langle \Delta x^a A_a(x)\rangle = C\,. \tag{8.82}$$

The contraction of indices is performed with respect to the metric $g_{ab}(x)$. Here $A_a(x)$ is the vector potential in a curved space. This constraint together with the old constraints in section 8.2 lead to the transition probability

$$P(x'|x) = g^{1/2}(x')\mathcal{P}(x'|x)\,, \tag{8.83}$$

with

$$\mathcal{P}(x'|x) = \frac{1}{\zeta(x,\alpha,\lambda)} e^{S(x')-\frac{1}{2}\alpha\Delta\ell^2-\lambda\Delta x^a A_a(x)}\,, \tag{8.84}$$

$$\zeta(x,\alpha,\lambda) = \int d^n x' g^{1/2}(x') e^{S(x')-\frac{1}{2}\alpha\Delta\ell^2-\lambda\Delta x^a A_a(x)}\,, \tag{8.85}$$

and the Lagrange multiplier λ can be determined from the constraint (8.82),

$$\frac{\partial}{\partial\lambda}\log\zeta(x,\alpha,\lambda) = -C\,. \tag{8.86}$$

For large α the diffusion process takes the form

$$\Delta x^a = \tilde{b}^a(x)\Delta t - \frac{1}{2\tau}\Gamma^a\Delta t + \Delta w^a\,, \tag{8.87}$$

where now

$$\tilde{b}^a = \frac{1}{\tau}g^{ab}(\partial_b S - \lambda A_b)\,, \tag{8.88}$$

and

$$\langle\Delta w^a\rangle = 0 \quad \text{and} \quad \langle\Delta w^a\Delta w^b\rangle = \frac{1}{\tau}g^{ab}\Delta t\,. \tag{8.89}$$

The Fokker-Planck equation is obtained in a similar way. The only difference is that ∂S is replaced by a gauge invariant term $\partial S - \lambda A$,

$$\frac{\partial\rho}{\partial t} = -\frac{1}{\sqrt{g}}\partial_a\left(\sqrt{g}\,\rho v^a\right)\,, \tag{8.90}$$

where

$$v^a = \frac{1}{\tau} g^{ab}(\partial_b \phi - \lambda A_b) \quad \text{with} \quad \phi(x,t) = S(x,t) - \log \rho^{1/2}(x,t)\,. \tag{8.91}$$

The energy constraint is the same as in eq. (8.63)

$$E(S,\rho) = \textstyle\int d^n x \sqrt{g}\, \rho [A g_{ab} v^a v^b + B g_{ab} u^a u^b + V(x) + V_c(x)]\,, \tag{8.92}$$

where now v^a is given by eq. (8.91). Impose the energy conservation, $\dot{E} = 0$, and the result is

$$\frac{2A}{\tau}\dot{\phi} + \frac{A}{\tau^2} g^{ab}(\partial_a \phi - \lambda A_a)(\partial_b \phi - \lambda A_b) + V(x) + V_c(x) - \frac{B}{\tau^2}\frac{\Delta_g \sqrt{\rho}}{\sqrt{\rho}} = 0\,, \tag{8.93}$$

The Lagrange multiplier λ is related to charge of the particle by the equation

$$\lambda = \frac{e}{\hbar c}\,. \tag{8.94}$$

Let

$$\Psi = \sqrt{\rho}\, e^{i\phi}\,, \tag{8.95}$$

then the couple of equations (8.90) and (8.93) can be combined in to a single equation

$$\begin{aligned} i\frac{2A}{\tau}\frac{\partial \Psi}{\partial t} &= \frac{A}{\tau^2 \sqrt{g}} \left(i\partial_a - \frac{e}{\hbar c} A_a \right) \sqrt{g} g^{ab} \left(i\partial_b - \frac{e}{\hbar c} A_b \right) \Psi \\ &\quad + V(x)\Psi + V_c(x)\Psi + \frac{A}{\tau^2}\left(1 - \frac{B}{A}\right)\frac{\Delta_g \sqrt{\rho}}{\sqrt{\rho}}\Psi\,. \end{aligned} \tag{8.96}$$

We can redefine and rescale the constants A, B, τ, σ^2, and the metric g^{ab} as in eqs. (8.75), (8.76), and (8.74) so that

$$i\hbar \frac{\partial \Psi}{\partial t} = \frac{\hbar^2}{2m\sqrt{h}} \left(i\partial_a - \frac{e}{\hbar c} A_a \right) \sqrt{h} h^{ab} \left(i\partial_b - \frac{e}{\hbar c} A_b \right) \Psi$$

$$+V(x)\Psi + V_c(x)\Psi + \frac{\hbar^2}{2m}\left(1-\frac{\mu}{m}\right)\frac{\Delta_h\sqrt{\rho}}{\sqrt{\rho}}\Psi\,. \tag{8.97}$$

The last term also vanishes when $\mu = m$, therefore

$$i\hbar\frac{\partial\Psi}{\partial t} = \frac{\hbar^2}{2m\sqrt{h}}\left(i\partial_a - \frac{e}{\hbar c}A_a\right)\sqrt{h}h^{ab}\left(i\partial_b - \frac{e}{\hbar c}A_b\right)\Psi + V(x)\Psi + V_c(x)\Psi\,. \tag{8.98}$$

This is the Schrödinger equation in a curved space for a charged particle interacting with an EM field.

8.7 Conclusions

We have derived the modified Schrödinger equation on a Riemannian manifold in the framework of entropic dynamics (ED). The modified equation replaces the Laplacian $\nabla^2 = \delta^{ab}\partial_a\partial_b$ by the Laplace-Beltrami operator Δ_h. Furthermore, the modified equation may contain an additional potential term $V_c(x)$ due to curvature provided it is included in the energy functional (8.63). However, this term can be thrown away for a space of constant curvature because it only introduces an overall irrelevant phase factor.

An additional constraint, eq. (8.82), is introduced for a charged particle interacting with electromagnetic field and thus the Schrödinger equation takes the form of eq. (8.98). In the next chapter we will show that the motion of a non relativistic particle with spin can be modeled by an appropriate choice of the curved Riemannian manifold.

Chapter 9

Entropic Dynamics of Spin-Half Particles

9.1 Introduction

The idea of spin was proposed by Uhlenbeck and Goudsmith in 1925 [69], to explain the splitting of spectral lines of atoms placed in a magnetic field, the so-called anomalous Zeeman effect. In 1927, W. Pauli [70] formulated the theory of electron spin within non relativistic quantum mechanics while P. Dirac in 1928 [71, 72] established a relativistic version of quantum mechanics that describes all *fermions*..

In the model discussed here, unlike angular momentum, the spin angular momentum does not depend on spatial coordinates of a particle. This makes it difficult to realize it classically, although many authors (Kramers [73], Takabayasi [74], Schiller [75] and others) have developed classical descriptions of spinning particles. The goal of the present chapter is to incorporate spin within the framework of entropic dynamics. Our specific concern is *spin-half* particles. The model presented here has formal similarities with Dankel [16]; and Dohrn, Guerra and Ruggiero [17] in Nelson's stochastic mechanics.

9.2 The Model

We consider a single particle. We assume that the particle has the usual spatial coordinates as well as some internal degrees of freedom. To incorporate those internal degrees of freedom in entropic dynamics we extend the configuration space from R^3 to $R^3 \times S^3$, where R^3 corresponds to the three dimensional Euclidean space and S^3 is a 3-sphere (see below).

The metric g_{ab} of $R^3 \times S^3$ is written in block matrix form

$$g_{ab} = \begin{pmatrix} \delta_{ij}/\sigma_1^2 & 0 \\ 0 & h_{\mu\nu}/\sigma_2^2 \end{pmatrix} , \tag{9.1}$$

where the indices i, j and μ, ν correspond to directions in R^3 and S^3 respectively, and σ_1^2 and σ_2^2 are scale factors that are introduced for later convenience.

Now recall eq. (8.96)

$$\begin{aligned} i\frac{2A}{\tau}\frac{\partial\Psi}{\partial t} &= \frac{A}{\tau^2\sqrt{g}}\left(i\partial_a - \frac{e}{\hbar c}A_a\right)\sqrt{g}g^{ab}\left(i\partial_b - \frac{e}{\hbar c}A_b\right)\Psi \\ &\quad +V(x)\Psi + V_c(x)\Psi + \frac{A}{\tau^2}\left(1 - \frac{B}{A}\right)\frac{\Delta_g\sqrt{\rho}}{\sqrt{\rho}}\Psi \, . \end{aligned} \tag{9.2}$$

Since $R^3 \times S^3$ has a constant curvature therefore $V_c(x)$ is constant, which can be omitted. Furthermore we are only interested in the case when the osmotic mass μ is equal to the current mass m or $A = B$. Then eq. (9.2) simplifies to

$$i\frac{2A}{\tau}\frac{\partial\Psi}{\partial t} = \frac{A}{\tau^2\sqrt{g}}\Big(i\partial_a - \frac{e}{\hbar c}A_a\Big)\sqrt{g}g^{ab}\Big(i\partial_b - \frac{e}{\hbar c}A_b\Big)\Psi + V(x)\Psi \tag{9.3}$$

Put $A_a = (A_i, \mathcal{B}_\mu)$, where $A_i = (A_1, A_2, A_3)$ is the vector potential and $\mathcal{B}_\mu$ is the spin potential. In ED, the potential A_i reflects a constraint on the motion:

$$\langle \Delta x^i \rangle A_i = C_1 \,. \tag{9.4}$$

The same happens with $\mathcal{B}_\mu$, the spin potential but as the choice of A_i is not arbitrary (determined by the external environment) we will find that the choice of $\mathcal{B}_\mu$ is not arbitrary either. As we will later see that the potentials A_i and $\mathcal{B}_\mu$ are not independent. $\mathcal{B}_\mu$ is determined by A_i.

Next, use eq. (9.1) for the metric g_{ab} which decomposes the first term on r.h.s. into two parts,

$$\begin{aligned} i\frac{2A}{\tau}\frac{\partial\Psi}{\partial t} &= \frac{A\sigma_1^2\delta^{ij}}{\tau^2}\Big(i\partial_i - \frac{e}{\hbar c}A_i\Big)\Big(i\partial_j - \frac{e}{\hbar c}A_j\Big)\Psi + V(x)\Psi \\ &+\frac{A\sigma_2^2}{\tau^2\sqrt{h}}\Big(i\partial_\mu - \frac{e}{\hbar c}\mathcal{B}_\mu\Big)\sqrt{h}h^{\mu\nu}\Big(i\partial_\nu - \frac{e}{\hbar c}\mathcal{B}_\nu\Big)\Psi \,. \end{aligned} \tag{9.5}$$

Introducing new constants

$$\hbar = \frac{2A}{\tau}, \quad \frac{\sigma_1^2}{\tau} = \frac{\hbar}{m}, \quad \frac{\sigma_2^2}{\tau} = \frac{\hbar}{I}\,, \tag{9.6}$$

where m is the mass and I is the moment of inertia, then eq. (9.5) becomes

$$\begin{aligned} i\hbar\frac{\partial\Psi}{\partial t} &= \frac{\hbar^2\delta^{ij}}{2m}\Big(i\partial_i - \frac{e}{\hbar c}A_i\Big)\Big(i\partial_j - \frac{e}{\hbar c}A_j\Big)\Psi + V(x)\Psi \\ &+\frac{\hbar^2}{2I\sqrt{h}}\Big(i\partial_\mu - \frac{e}{\hbar c}\mathcal{B}_\mu\Big)\sqrt{h}h^{\mu\nu}\Big(i\partial_\nu - \frac{e}{\hbar c}\mathcal{B}_\nu\Big)\Psi \,. \end{aligned} \tag{9.7}$$

For $I > 0$, this equation represents a rigid charged sphere also known as the Bopp-Haag equation [76]. For a point particle we will let $I \to 0$ at the end. The metric $h^{\mu\nu}$ is the inverse of $h_{\mu\nu}$. To find $h^{\mu\nu}$ we study the connection of S^3 with $SU(2)$ group.

9.3 The Geometry of the 3–Sphere and the SU(2) Group

A three sphere S^3 is a unit sphere embedded in R^4. Let $q^k \in R$ (with $k = 0, 1, 2, 3.$), then

$$S^3 = \{(q^0, q^1, q^2, q^3) \in R^4 : (q^0)^2 + (q^1)^2 + (q^2)^2 + (q^3)^2 = 1\}\,. \tag{9.8}$$

The components q^k can be parameterized in several ways but a convenient one is that of Euler angles (α, β, γ). To find out Euler angles parameterization we note that S^3 is isomorphic to $SU(2)$. The special unitary group $SU(2)$ is a set of 2×2 matrices

$$SU(2) = \left\{ \begin{pmatrix} a & b \\ -b^\star & a^\star \end{pmatrix} : a, b \in C,\ |a|^2 + |b|^2 = 1 \right\}. \tag{9.9}$$

The isomorphism of S^3 and $SU(2)$ follows immediately if we define $a = q^0 + iq^3$, and $b = iq^1 + q^2$, then $|a|^2 + |b|^2 = 1$ implies $(q^0)^2 + (q^1)^2 + (q^2)^2 + (q^3)^2 = 1$. Hence $S^3 \cong SU(2)$. The advantage of this isomorphism is that any element of $SU(2)$ can be expressed in terms of Euler angles. For example, the elements in the spin half representation are [77, 78]

$$\begin{aligned} D^{(1/2)} &= e^{-i\alpha\hat{\sigma}_3/2} e^{-i\beta\hat{\sigma}_2/2} e^{-i\gamma\hat{\sigma}_3/2} \\ &= \begin{pmatrix} e^{-\frac{1}{2}i(\alpha+\gamma)} \cos\frac{\beta}{2} & -e^{-\frac{1}{2}i(\alpha-\gamma)} \sin\frac{\beta}{2} \\ e^{\frac{1}{2}i(\alpha-\gamma)} \sin\frac{\beta}{2} & e^{\frac{1}{2}i(\alpha+\gamma)} \cos\frac{\beta}{2} \end{pmatrix}. \end{aligned} \tag{9.10}$$

where $\hat{\sigma}$'s on the right in the first equality are Pauli matrices. Compare eq. (9.9) and eq. (9.10) we obtain

$$a = e^{-\frac{i}{2}(\alpha+\gamma)} \cos\frac{\beta}{2}$$

$$b = -e^{-\frac{i}{2}(\alpha-\gamma)} \sin\frac{\beta}{2} \tag{9.11}$$

Separating the real and imaginary parts we obtain

$$\begin{aligned} q^0 &= \cos\frac{\beta}{2}\cos\frac{\alpha+\gamma}{2}\,, \\ q^1 &= \sin\frac{\beta}{2}\sin\frac{\alpha-\gamma}{2}\,, \\ q^2 &= -\sin\frac{\beta}{2}\cos\frac{\alpha-\gamma}{2}\,, \\ q^3 &= -\cos\frac{\beta}{2}\sin\frac{\alpha+\gamma}{2}\,. \end{aligned} \tag{9.12}$$

Note that the entire three sphere can be covered only once if the angles are taken in the range $0 \le \alpha \le 2\pi$, $0 \le \beta \le \pi$, and $0 \le \gamma \le 4\pi$.

The metric δ_{kl} of R^4 induces a metric $h_{\mu\nu}$ on a hypersurface (in this case the hypersurface is S^3)

$$h_{\mu\nu} = \frac{\partial q^k}{\partial x^\mu}\frac{\partial q^l}{\partial x^\nu}\delta_{kl}\,, \tag{9.13}$$

where q^k are given by eq. (9.12), and $x^\mu = (\alpha, \beta, \gamma)$. Carrying out the calculations we obtain the matrix elements of $h_{\mu\nu}$

$$h_{\mu\nu} = \begin{pmatrix} h_{\alpha\alpha} & h_{\alpha\beta} & h_{\alpha\gamma} \\ h_{\beta\alpha} & h_{\beta\beta} & h_{\beta\gamma} \\ h_{\gamma\alpha} & h_{\gamma\beta} & h_{\gamma\gamma} \end{pmatrix} = \begin{pmatrix} \frac{1}{4} & 0 & \frac{1}{4}\cos\beta \\ 0 & \frac{1}{4} & 0 \\ \frac{1}{4}\cos\beta & 0 & \frac{1}{4} \end{pmatrix}. \tag{9.14}$$

The volume of $SU(2)$ or S^3 can be computed likewise

$$V_{SU(2)} = \int_0^{2\pi} d\alpha \int_0^{\pi} \frac{1}{8}\sin\beta\, d\beta \int_0^{4\pi} d\gamma = 2\pi^2\,. \tag{9.15}$$

It is customary to introduce a normalized invariant measure

$$d\tau = \frac{1}{2\pi^2}\sqrt{h}\, d\alpha\, d\beta\, d\gamma\ , \quad \text{with} \quad \sqrt{h} = \frac{1}{8}\sin\beta\,. \tag{9.16}$$

We also note that for the $j-$representation of $SU(2)$ equation (9.10) can be written as

$$D^{(j)}_{mm'}(\alpha,\beta,\gamma) = \langle jm\left| R(\alpha,\beta,\gamma)\right| jm'\rangle \ , \tag{9.17}$$

where $R(\alpha,\beta,\gamma)$ is the rotation matrix on $SU(2)$

$$R(\alpha,\beta,\gamma) = e^{-i\frac{\alpha}{\hbar}J_3}e^{-i\frac{\beta}{\hbar}J_2}e^{-i\frac{\gamma}{\hbar}J_3} \ , \tag{9.18}$$

where J's are the generators of the group. For future considerations we also need to calculate the derivatives of $D^{(j)}_{mm'}$ with respect to the Euler angles

$$\frac{\partial}{\partial\alpha}D^{(j)}_{mm'} = \left\langle jm\left|\frac{\partial}{\partial\alpha}R(\alpha,\beta,\gamma)\right| jm'\right\rangle \ , \tag{9.19}$$

$$\frac{\partial}{\partial\beta}D^{(j)}_{mm'} = \left\langle jm\left|\frac{\partial}{\partial\beta}R(\alpha,\beta,\gamma)\right| jm'\right\rangle \ , \tag{9.20}$$

$$\frac{\partial}{\partial\gamma}D^{(j)}_{mm'} = \left\langle jm\left|\frac{\partial}{\partial\gamma}R(\alpha,\beta,\gamma)\right| jm'\right\rangle \ , \tag{9.21}$$

where

$$\frac{\partial}{\partial\alpha}R(\alpha,\beta,\gamma) = -\frac{i}{\hbar}J_3R = -\frac{i}{\hbar}R\left[R^{-1}J_3R\right] \ , \tag{9.22}$$

$$\frac{\partial}{\partial\beta}R(\alpha,\beta,\gamma) = -\frac{i}{\hbar}R\left[e^{i\frac{\gamma}{\hbar}J_3}J_2e^{-i\frac{\gamma}{\hbar}J_3}\right] \ , \tag{9.23}$$

$$\frac{\partial}{\partial\gamma}R(\alpha,\beta,\gamma) = -\frac{i}{\hbar}RJ_3 \, . \tag{9.24}$$

We want to compute the quantities in the square brackets in the r.h.s. of eqs. (9.22) and (9.23). To compute them use the identity (see for example, [79])

$$a^{\hat{A}}\hat{B}a^{-\hat{A}} = \hat{B} + [\hat{A},\hat{B}] + \frac{1}{2!}[\hat{A},[\hat{A},\hat{B}]] + \frac{1}{3!}[\hat{A},[\hat{A},[\hat{A},\hat{B}]]] + \dots\,, \tag{9.25}$$

and the commutation relations

$$[J_i, J_j] = i\hbar J_k\,, \quad \text{(in cyclic order.)} \tag{9.26}$$

For instance

$$e^{i\frac{\gamma}{\hbar}J_3} J_2 e^{-i\frac{\gamma}{\hbar}J_3} = J_1 \sin\gamma + J_2\cos\gamma\,. \tag{9.27}$$

Here we have used $\hat{A} = i\gamma J_3/\hbar$, and $\hat{B} = J_2$. In a similar way we obtain

$$R^{-1}J_3R = -J_1\sin\beta\cos\gamma + J_2\sin\beta\sin\gamma + J_3\cos\beta\,. \tag{9.28}$$

Collecting the results, eqs. (9.19–9.21) become,

$$\begin{aligned}\frac{\partial}{\partial\alpha}D^{(j)}_{mm'} &= -\frac{i}{\hbar}\left\langle jm\left|R[R^{-1}J_3R]\right|jm'\right\rangle \\ &= -\frac{i}{\hbar}\sum_{m''}\left\langle jm\left|R\right|jm''\right\rangle\left\langle jm''\right|R^{-1}J_3R\left|jm'\right\rangle\,,\end{aligned} \tag{9.29}$$

where we have introduced an identity matrix $\sum_{m''}|jm''\rangle\langle jm''| = 1$. On substituting eq. (9.28) we have

$$\frac{\partial}{\partial\alpha}D^{(j)}_{mm'} = -\frac{i}{\hbar}\sum_{m''}D^{(j)}_{mm''}[-\sin\beta\cos\gamma(\Sigma_1)^{(j)}_{m''m'}+\sin\beta\sin\gamma(\Sigma_2)^{(j)}_{m''m'}+\cos\beta(\Sigma_3)^{(j)}_{m''m'}]\,, \tag{9.30}$$

where

$$(\Sigma_i)^{(j)}_{m''m'} = \langle jm''\left|J_i\right|jm'\rangle\,, \quad i = 1,2,3. \tag{9.31}$$

Similarly

$$\frac{\partial}{\partial \beta} D^{(j)}_{mm'} = -\frac{i}{\hbar} \sum_{m''} D^{(j)}_{mm''} [\sin\gamma (\Sigma_1)^{(j)}_{m''m'} + \cos\gamma (\Sigma_2)^{(j)}_{m''m'}] \,, \tag{9.32}$$

and

$$\frac{\partial}{\partial \gamma} D^{(j)}_{mm'} = -\frac{i}{\hbar} \sum_{m''} D^{(j)}_{mm''} (\Sigma_3)^{(j)}_{m''m'} \,. \tag{9.33}$$

Furthermore it is also straightforward to show that

$$\left\{ \frac{1}{\sin\beta} \frac{\partial}{\partial\beta} \left(\sin\beta \frac{\partial}{\partial\beta} \right) + \frac{1}{\sin^2\beta} \left(\frac{\partial^2}{\partial\alpha^2} + \frac{\partial^2}{\partial\gamma^2} - 2\cos\beta \frac{\partial^2}{\partial\alpha\partial\gamma} \right) + j(j+1) \right\} D^{(j)}_{mm'} = 0 \,, \tag{9.34}$$

and the orthnormalization condition,

$$\int d\tau D^{(j)}_{mm'} D^{*(j')}_{m_1 m'_1} = \frac{\delta_{jj'} \delta_{mm_1} \delta_{m'm'_1}}{2j+1} \,, \tag{9.35}$$

where the measure $d\tau$ is given by eq. (9.16).

9.4 Derivation of the Pauli Equation

As we saw in chapters 5 and 8, the wave function ψ was constructed from two real quantities ρ and ϕ, that is $\psi = \sqrt{\rho}\, e^{i\phi}$. As a result we obtained Schrödinger equation in which ψ is a one component scalar. In contrast Ψ is a two component *spinor* in Pauli equation. The Pauli equation for spin half can be derived from eq. (9.7) by using the Peter-Weyl decomposition [78]

$$\Psi(x; \alpha, \beta, \gamma; t) = \sum_{j,m,m'} D^{(j)}_{mm'}(\alpha, \beta, \gamma) \psi^{(j)}_{m'm}(x, t) \,, \tag{9.36}$$

where $D^{(j)}_{mm'}$, given by eq. (9.17), is the $j-$ representation of $SU(2)$. The advantage of Peter-Weyl decomposition is that it separates spatial coordinates and Euler angles as can be seen $D^{(j)}_{mm'}$ depends only on the Euler angles while $\psi^{(j)}_{m'm}$ depends on x and t.

At this stage the evolution according to eq. (9.7) is such that different j's are coupled. The different spins are uncoupled only if $\mathcal{B}_\mu$ are chosen accordingly (see eq. (**??**) below). Next we carry on the calculations and Substitute eq. (9.36) into eq. (9.7)

$$i\hbar \sum_{j,m,m'} D^{(j)}_{mm'} \frac{\partial \psi^{(j)}_{m'm}}{\partial t} = \sum_{j,m,m'} \left(D^{(j)}_{mm'} C^{(j)}_{m'm} + K^{(j)}_{mm'} \psi^{(j)}_{m'm} \right) , \tag{9.37}$$

where

$$C^{(j)}_{m'm} = \left[\frac{\hbar^2 \delta^{ij}}{2m} \left(i\partial_i - \frac{e}{\hbar c} A_i \right) \left(i\partial_i - \frac{e}{\hbar c} A_i \right) + V(x) \right] \psi^{(j)}_{m'm} , \tag{9.38}$$

and

$$K^{(j)}_{mm'} = \frac{\hbar^2}{2I\sqrt{h}} \left(i\partial_\mu - \frac{e}{\hbar c} \mathcal{B}_\mu \right) \sqrt{h} h^{\mu\nu} \left(i\partial_\nu - \frac{e}{\hbar c} \mathcal{B}_\nu \right) D^{(j)}_{mm'} . \tag{9.39}$$

Note that the time derivative $\partial/\partial t$, and spatial derivatives $\partial_i = (\partial/\partial x, \partial/\partial y, \partial/\partial z)$ act on $\psi^{(j)}_{mm'}$, while the Euler angles derivatives $\partial_\mu = (\partial/\partial\alpha, \partial/\partial\beta, \partial/\partial\gamma)$ act on $D^{(j)}_{mm'}$ only. We first focus on eq. (9.39),

$$K^{(j)}_{mm'} = T_1 + T_2 + T_3 + T_4 , \tag{9.40}$$

where

$$T_1 = -\frac{\hbar^2}{2I} \frac{1}{\sqrt{h}} \partial_\mu \left(\sqrt{h} h^{\mu\nu} \partial_\nu D^{(j)}_{mm'} \right) , \tag{9.41}$$

$$T_2 = -\frac{ie\hbar}{2Ic} h^{\mu\nu} \mathcal{B}_\mu \partial_\nu D^{(j)}_{mm'} , \tag{9.42}$$

$$T_3 = -\frac{ie\hbar}{2Ic}\frac{1}{\sqrt{h}}\partial_\mu\left(\sqrt{h}h^{\mu\nu}\mathcal{B}_\nu\right)D^{(j)}_{mm'}\,, \tag{9.43}$$

$$T_4 = \frac{e^2}{2Ic^2}h^{\mu\nu}\mathcal{B}_\mu\mathcal{B}_\nu D^{(j)}_{mm'}\,, \tag{9.44}$$

while $h^{\mu\nu}$ is the inverse of $h_{\mu\nu}$

$$h^{\mu\nu} = \begin{pmatrix} h^{\alpha\alpha} & h^{\alpha\beta} & h^{\alpha\gamma} \\ h^{\beta\alpha} & h^{\beta\beta} & h^{\beta\gamma} \\ h^{\gamma\alpha} & h^{\gamma\beta} & h^{\gamma\gamma} \end{pmatrix} = \begin{pmatrix} \frac{4}{\sin^2\beta} & 0 & -\frac{4\cos\beta}{\sin^2\beta} \\ 0 & 4 & 0 \\ -\frac{4\cos\beta}{\sin^2\beta} & 0 & \frac{4}{\sin^2\beta} \end{pmatrix}, \tag{9.45}$$

and $\sqrt{h} = \frac{1}{8}\sin\beta$.

We want to compute each term in eq. (9.40). Starting with T_1 which after rearrangements becomes,

$$\begin{aligned} T_1 &= -\frac{2\hbar^2}{I}\left\{\frac{1}{\sin\beta}\frac{\partial}{\partial\beta}\left(\sin\beta\frac{\partial}{\partial\beta}\right) + \frac{1}{\sin^2\beta}\left(\frac{\partial^2}{\partial\alpha^2} + \frac{\partial^2}{\partial\gamma^2} - 2\cos\beta\frac{\partial^2}{\partial\alpha\partial\gamma}\right)\right\}D^{(j)}_{mm'} \\ &= \frac{2\hbar^2}{I}j\,(j+1)\,D^{(j)}_{mm'}\,, \end{aligned} \tag{9.46}$$

where we have also used eq. (9.34) to obtain the second equality.

Now computing eq. (9.42)

$$T_2 = -\frac{ie\hbar}{2Ic}\left(\mathcal{B}^\alpha\partial_\alpha D^{(j)}_{mm'} + \mathcal{B}^\beta\partial_\beta D^{(j)}_{mm'} + \mathcal{B}^\gamma\partial_\gamma D^{(j)}_{mm'}\right)\,, \tag{9.47}$$

where $\mathcal{B}^\mu = h^{\mu\nu}\mathcal{B}_\nu$. Use eqs. (9.30–9.33),

$$\begin{aligned} T_2 = &-\frac{e}{2Ic}\sum_{m''}D^{(j)}_{mm''}\Big[\left\{-\mathcal{B}^\alpha\sin\beta\cos\gamma + \mathcal{B}^\beta\sin\gamma\right\}(\Sigma_1)^{(j)}_{m''m'} \\ &+\left\{\mathcal{B}^\alpha\sin\beta\sin\gamma + \mathcal{B}^\beta\cos\gamma\right\}(\Sigma_2)^{(j)}_{m''m'} + \left\{\mathcal{B}^\alpha\cos\beta + \mathcal{B}^\gamma\right\}(\Sigma_3)^{(j)}_{m''m'}\Big]\,. \end{aligned} \tag{9.48}$$

Next we follow [17] and define the quantities in the curly brackets {..} on the right as

$$\frac{I}{m}B^1 \stackrel{def}{=} -\mathcal{B}^\alpha \sin\beta\cos\gamma + \mathcal{B}^\beta \sin\gamma\,, \tag{9.49}$$

$$\frac{I}{m}B^2 \stackrel{def}{=} \mathcal{B}^\alpha \sin\beta\sin\gamma + \mathcal{B}^\beta \cos\gamma\,, \tag{9.50}$$

$$\frac{I}{m}B^3 \stackrel{def}{=} \mathcal{B}^\alpha \cos\beta + \mathcal{B}^\gamma\,, \tag{9.51}$$

so that

$$T_2 = -\frac{e}{2mc}\sum_{m''} D^{(j)}_{mm''}\vec{B}\cdot(\vec{\Sigma})^{(j)}_{m''m'}\,, \tag{9.52}$$

where $\vec{B}$ will eventually be identified with the magnetic field. The significance of eqs. (9.49–9.51) become clear when it is inverted for the spin potential $\mathcal{B}$'s

$$\mathcal{B}^\alpha = \frac{I}{m}\left(-B^1\frac{\cos\gamma}{\sin\beta} + B^2\frac{\sin\gamma}{\sin\beta}\right)\,, \tag{9.53}$$

$$\mathcal{B}^\beta = \frac{I}{m}\left(B^1\sin\gamma + B^2\cos\gamma\right)\,, \tag{9.54}$$

$$\mathcal{B}^\gamma = \frac{I}{m}\left(B^1\cos\gamma\cot\beta - B^2\sin\gamma\cot\beta + B^3\right)\,. \tag{9.55}$$

Note that the spin potential is a function of both the spatial coordinates and the internal Euler angles: $\vec{\mathcal{B}} = \vec{\mathcal{B}}(x;\alpha,\beta,\gamma)$. This means that, in general, the quantities B^i's on the right are functions of both the spatial coordinates and the Euler angles: $\vec{B} = \vec{B}(x;\alpha,\beta,\gamma)$. This implies that in general $\vec{B}(x;\alpha,\beta,\gamma)$ is not a legitimate magnetic field because it allows a transition between different spins ($0 \leftrightarrow 1/2 \leftrightarrow 1 \leftrightarrow 3/2\ldots$). This transition between spins occurs if $T_3 \neq 0$. T_3 is given by (9.43). On the other hand there is no spins transition if $B^i = B^i(x)$ which gives

$$T_3 = -\frac{ie\hbar}{2Ic}\frac{1}{\sqrt{h}}\partial_\mu\left(\sqrt{h}h^{\mu\nu}\mathcal{B}_\nu\right)D^{(j)}_{mm'} = 0\,, \tag{9.56}$$

because ∂_μ only acts on the Euler angles. The vanishing value of T_3 means that different values of j's are uncoupled.

Finally the calculations of T_4 give

$$\begin{aligned} T_4 &= \frac{e^2}{2Ic^2} h_{\mu\nu} \mathcal{B}^\mu \mathcal{B}^\nu D^{(j)}_{mm'} \\ &= \frac{e^2 I}{8m^2c^2} \vec{B}^2 D^{(j)}_{mm'} \, . \end{aligned} \tag{9.57}$$

Substituting eqs. (9.46), (9.52), (9.56) and (9.57) in eq. (9.40), we have

$$K^{(j)}_{mm'} = \frac{2\hbar^2}{I} j\,(j+1)\, D^{(j)}_{mm'} - \frac{e}{2mc} \sum_{m''} D^{(j)}_{mm''} \vec{B} \cdot (\vec{\Sigma})^{(j)}_{m''m'} + \frac{e^2 I}{8m^2c^2} \vec{B}^2 D^{(j)}_{mm'} \, . \tag{9.58}$$

Hence eq. (9.37) becomes

$$\begin{aligned} i\hbar \sum_{j,m,m'} D^{(j)}_{mm'} \frac{\partial \psi^{(j)}_{m'm}}{\partial t} &= \sum_{j,m,m'} D^{(j)}_{mm'} \left[\frac{\hbar^2}{2m} \left(i\vec{\nabla} - \frac{e}{\hbar c} \vec{A} \right)^2 + V(x) \right] \psi^{(j)}_{m'm} \\ &+ \frac{2\hbar^2}{I} \sum_{j,m,m'} j\,(j+1)\, D^{(j)}_{mm'} \psi^{(j)}_{m'm} \\ &- \frac{e}{2mc} \sum_{j,m,m',m''} D^{(j)}_{mm''} \vec{B} \cdot (\vec{\Sigma})^{(j)}_{m''m'} \psi^{(j)}_{m'm} \\ &+ \frac{e^2 I}{8m^2c^2} \sum_{j,m,m'} \vec{B}^2 D^{(j)}_{mm'} \psi^{(j)}_{m'm} \, . \end{aligned} \tag{9.59}$$

Multiplying eq. (9.59) by $D^{*(j')}_{m_1 m_1'}$ and integrate over the Euler angles and making use of the orthogonality condition (9.35), we get

$$\begin{aligned} i\hbar \frac{\partial \psi^{(j)}_{mm'}}{\partial t} &= \left[\frac{\hbar^2}{2m} \left(i\vec{\nabla} - \frac{e}{\hbar c} \vec{A} \right)^2 + V(x) \right] \psi^{(j)}_{mm'} \\ &+ \frac{2\hbar^2}{I} j\,(j+1)\, \psi^{(j)}_{mm'} \\ &- \frac{e}{2mc} \sum_{m''} \vec{B} \cdot (\vec{\Sigma})^{(j)}_{mm''} \psi^{(j)}_{m''m'} \\ &+ \frac{e^2 I}{8m^2c^2} \vec{B}^2 \psi^{(j)}_{mm'} \, . \end{aligned} \tag{9.60}$$

The second term is interesting. As $I \to 0$ (point particle), the different spins correspond to different subspaces that are largely separated in energy ($\Delta E \sim 1/I \to \infty$) and therefore become decoupled: typical finite energy interactions cannot induce transitions. However, in this limit the fourth term vanishes while the second term is an infinite constant which contributes an overall irrelevant phase factor that can be omitted and so we are left with,

$$i\hbar \frac{\partial \psi^{(j)}_{mm'}}{\partial t} = \left[\frac{\hbar^2}{2m} \left(i\vec{\nabla} - \frac{e}{\hbar c} \vec{A} \right)^2 + V(x) \right] \psi^{(j)}_{mm'} - \frac{e}{2mc} \sum_{m''} \vec{B} \cdot (\vec{\Sigma})^{(j)}_{mm''} \psi^{(j)}_{m''m'} . \quad (9.61)$$

This equation holds for any value of j. In particular for $j = 1/2$ we can obtain the Pauli equation for spin-half particle. To write the Pauli equation in a familiar form we recall eq. (9.31)

$$\vec{\Sigma}^{(j)}_{m'm} = \left\langle jm' \middle| \vec{J} \middle| jm \right\rangle , \quad (9.62)$$

For $j = 1/2$, we have $m, m' = \pm 1/2$, and

$$\vec{J} = \frac{\hbar}{2} \vec{\sigma} , \quad (9.63)$$

where $\vec{\sigma}$ are the Pauli matrices,

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \quad (9.64)$$

Therefore eq. (9.61) becomes,

$$i\hbar \frac{\partial \psi^{(1/2)}_{mm'}}{\partial t} = \left[\frac{\hbar^2}{2m} \left(i\vec{\nabla} - \frac{e}{\hbar c} \vec{A} \right)^2 + V(x) \right] \psi^{(1/2)}_{mm'} - \frac{e\hbar}{2mc} \vec{B} \cdot \vec{\sigma}_{mm''} \psi^{(1/2)}_{m''m'} , \quad (9.65)$$

Note that the summation over the repeated index m'' is understood. Equation (9.65) is also known as Dankel equation [16], it describes two identical copies of Pauli equation because $\psi^{(1/2)}$ is a 2×2 matrix where each column represents a spinor

$$\psi^{1/2} = \begin{pmatrix} \psi^{1/2}_{1/2\,1/2} & \psi^{1/2}_{1/2\,-1/2} \\ \\ \psi^{1/2}_{-1/2\,1/2} & \psi^{1/2}_{-1/2\,-1/2} \end{pmatrix} . \tag{9.66}$$

We further note that eq. (9.65) is the $j = 1/2$ sector of the regular representation of $SU(2)$. The regular representation is reducible; it is the direct sum of the $(2j+1)$ copies of the j^{th} representation for all values of $j = 0, 1/2, 1, 3/2, \ldots$

We can decompose eq. (9.66) into the minimal left ideals by which we mean subspaces that are invariant under left multiplication. This is done as follows. Write eq. (9.66) as

$$\psi^{1/2} = \begin{pmatrix} \psi^{1/2}_{1/2\,1/2} & 0 \\ \\ \psi^{1/2}_{-1/2\,1/2} & 0 \end{pmatrix} + \begin{pmatrix} 0 & \psi^{1/2}_{1/2\,-1/2} \\ \\ 0 & \psi^{1/2}_{-1/2\,-1/2} \end{pmatrix} , \tag{9.67}$$

For brevity, we write eq. (9.67) as

$$\psi^{1/2} = \psi_1 + \psi_2 \,, \tag{9.68}$$

where ψ_1 and ψ_2 are Pauli spinors. If eq. (9.68) is multiplied from left by a matrix U, then it does not mix the spinors, that is,

$$U\psi^{1/2} = U\psi_1 + U\psi_2 = \psi_1' + \psi_2' \,, \tag{9.69}$$

It is in this this sense that eq. (9.65) describes two copies of the Pauli equation.

Remarks: Since ED has close similarities with Nelson's stochastic mechanics, some of the questionable features that are present in Nelson/Dankel theory also show up in this version of ED.

The entropic model presented in this chapter gives rise to several questions;

1. Gauge invariance is one of the fundamental symmetries in physics. It implies that electric charge is conserved. The Schrödinger equation is also invariant under the gauge transformation,

$$\Psi \to \Psi' = e^{i\chi(x)}\Psi\,, \quad \text{and} \quad A_i \to A_i' = A_i + \frac{i\hbar c}{e}\partial_i \chi(x)\,. \tag{9.70}$$

   In our case, eq. (9.7) is also invariant under (9.70). However there is an unwanted gauge symmetry inherent in eq. (9.7): We can see that eq. (9.7) is also invariant if we use the following gauge transformation

$$\Psi \to \Psi' = e^{i\chi(\alpha,\beta,\gamma)}\Psi\ , \quad \text{and} \quad \mathcal{B}_\mu \to \mathcal{B}_\mu' = \mathcal{B}_\mu + \frac{i\hbar c}{e}\partial_\mu \chi(\alpha,\beta,\gamma) \tag{9.71}$$

   where $\mathcal{B}$ is the field on $SU(2)$, and the arguments of χ are Euler angles. Nevertheless, this apparent ambiguity disappears if we let the moment of inertia $I \to 0$ in which case the final equation takes the form of eq. (9.65) which is invariant under the standard gauge transformation (9.70).

2. Another shortcoming of the Dohrn, Guerra and Ruggiero [17] derivation in connection with stochastic mechanics is that the $\mathcal{B}$ field in eqs. (9.53), (9.54) and (9.55) is constrained to satisfy $T_3 = 0$ in eq. (9.56), and then the B field on the right of eqs. (9.53), (9.54) and (9.55) is identified with the magnetic field.

However, the problem remains that the connection between $\vec{B}$ and $\vec{\nabla} \times \vec{A}$ is merely stated, not established.

3. Another objection was raised by Wallstrom [80], it deals with eq. (9.65). Since eq. (9.65) generates two identical copies of Pauli equation, it represents the *regular* representation of $SU(2)$ rather than the *irreducible* spinor representation.

9.5 Conclusions

In this chapter our goal was to construct spin 1/2 theory as an example of entropic dynamics. To construct a statistical model it is necessary to first identify the appropriate configuration space. To model spin 1/2 theory we chose $R^3 \times S^3$ manifold which is isomorphic to $R^3 \times SU(2)$. But the difficulty with this model is that it enlarges the configuration space from R^3 to $R^3 \times SU(2)$; the rotator can have any spin, as in eq. (9.61). We derived the regular Pauli equation that corresponds to the regular representation of $SU(2)$which include the irreducible spinor representation.

The difficulty in constructing any statistical model is to identify the subject matter. In our case here, what is the configuration space appropriate to the description of spin? What is a spin? Is it a "point" rigid rotator as in the Bopp-Hagg model [16, 76, 81]? or is it a precessing dipole as described by Kramers [73] and Schiller [75]? Is spin a property of the particle or is it a property of the probability distributions that describe its motion? Then we must identify the relevant information—that is, the constraints—that when taken into account through entropic inference lead to the Pauli equation. And even if we succeed in describing a single spin 1/2 particle the problem remains of describing many particles and the spin-statistics connection.

We conclude that the formulation developed in this chapter represents an interesting preliminary exercise but that a more definitive approach must involve developing the ED of fermion fields.

Chapter 10

Conclusions

Entropic dynamics (ED) is a framework for inference on the basis of incomplete information. Probabilities play a key role when dealing with incomplete information and the goal is to update probabilities when new information becomes available. Since probability is also a common feature of quantum mechanics, in the ED framework quantum mechanics is also an example of inference. ED does not challenge the formalism of quantum theory but focuses on removing the conceptual difficulties that arise in understanding of quantum mechanics. Conceptually the most important advance is the understanding of the fundamental object in quantum mechanics — the wave function. The work of A. Caticha [8,9,30] clarifies that the wave function is fully epistemic — both the amplitude and the phase of the wave function receive a statistical interpretation, and the modes of evolution are dictated by updating according to entropic methods.

Entropic quantum dynamics (EQD) developed in Chapter 5 allows us to define almost all observables in quantum theory in purely informational terms. Along these lines, the problem of measurement in quantum mechanics is addressed in [82, 83].

ED of relativistic scalar quantum fields is developed in [84, 85]. In this thesis we explored the concepts of momentum, angular momentum and spin within the entropic dynamics framework, and extended the theory to curved spaces.

In chapter 6 we established the entropic analog of linear momentum. We noted that the particle follows a non-differentiable trajectory so that the classical momentum $m\, d\vec{x}/dt$ along the trajectory cannot be defined. We introduced three different types of momentum. They are the *drift*, *osmotic* and *current* momenta— the drift momentum reflects flow along the entropy gradient, the osmotic momentum indicates diffusion of probability flow, while the current momentum reflects the flow of total probability. It is shown that these momenta share properties with the quantum momentum $\vec{p}_q = -i\hbar\vec{\nabla}$, and in the appropriate classical limit the drift/current momenta converge to classical momentum while the osmotic momentum tends to zero.

The conclusion of this chapter is that momentum is a statistical concept. Momentum is not an attribute of the particles but of the probability distributions. Finally we derived the uncertainty relations for all momenta that show up in entropic dynamics and we showed that the Heisenberg's uncertainty relation is an effect that can be attributed to diffusion – it is an osmotic effect. In the same chapter we also explored a special case of entropic dynamics that involves hybrid features. It obeys the classical Hamilton-Jacobi equation and also the usual uncertainty principle.

In chapter 7, just like linear momentum, angular momentum was also expressed in purely informational terms. In addition to the quantum angular momentum, there are three other local angular momenta. They are the drift angular momentum, osmotic angular momentum, and current angular momentum. Having defined angular momenta, we also established their uncertainty relations.

In chapter 8 we extended entropic dynamics to curved spaces. Here the configuration space was assumed to be a general Riemannian manifold. No additional assumptions were introduced. Time was introduced exactly in the same way as in flat case. The diffusion process in general coordinates was derived. Taking into account the second order terms we noted that the displacement vector Δx^a as well as the drift velocity b^a do not transform like a vector. This led to a modified Schrödinger equation for curved spaces which also takes into account the curvature effects. We also derived Schrödinger equation for a charged particle interacting with external electromagnetic field on general Riemannian manifold.

In chapter 9 we discussed an application of the theory developed in chapter 8. We developed the entropic analog of the spin models of Dankel [16]; and Dohrn, Guerra and Ruggiero [17] models of spin in connection with Nelson's stochastic mechanics. To model spin 1/2 theory we chose $R^3 \times S^3$ manifold which is isomorphic to $R^3 \times SU(2)$. But the difficulty with this model is that it enlarges the configuration space from R^3 to $R^3 \times SU(2)$; the rotator can have any spin. In the limit $I \to 0$ the various spins decouples and we arrived at the regular Pauli equation that corresponds to the regular representation of $SU(2)$ which includes the irreducible spinor representation.

In the future studies I have several interesting problems in mind. Especially I wish to explore further applications of the theory developed in chapter 8. Can we derive classical Einstein field equations within the framework of ED? What is the ED of quantum gravity? Is it possible to develop a theory of incorporating spin without enlarging the configuration space from R^3? These questions and other interesting problems can be addressed once we know the appropriate statistical manifold and the relevant constraints.

Bibliography

[1] H. P. Stapp, "The Copenhagen Interpretation," *Am. J. Phys.* **40** (1972) 1098.

[2] P. R. Holland, *The Quantum Theory of Motion: An Account of the de Broglie-Bohm Interpretation of Quantum Mechanics.* (Cambridge U.P., 1993).

[3] L. E. Ballentine, "The Statistical Interpretation of Quantum Mechanics," *Rev. Mod. Phys.* **42** (1970) 358. `http://link.aps.org/doi/10.1103/RevModPhys.42.358`.

[4] H. Everett, ""Relative State" Formulation of Quantum Mechanics," *Rev. Mod. Phys.* **29** (1957) 454. `http://link.aps.org/doi/10.1103/RevModPhys.29.454`.

[5] J. G. Cramer, "The Transactional Interpretation of Quantum Mechanics," *Rev. Mod. Phys.* **58** (1986) 647. `http://link.aps.org/doi/10.1103/RevModPhys.58.647`.

[6] J. J. Halliwell, "A Review of the Decoherent Histories Approach to Quantum Mechanics.," in *Fundamental Problems in Quantum Theory*, D. M. Greenberger

and A. Zelinger, eds., vol. 755 of *Annals of the New York Academy of Sciences*, p. 726. 1995. arXiv:gr-qc/9407040.

[7] A. Caticha, "From Entropic Dynamics to Quantum Theory," in *American Institute of Physics Conference Series*, P. M. Goggans and C.-Y. Chan, eds., vol. 1193 of *American Institute of Physics Conference Series*, p. 48. 2009. arXiv:0907.4335 [quant-ph].

[8] A. Caticha, "Entropic Dynamics, Time and Quantum Theory," *Journal of Physics A Mathematical General* **44** (2011) 225303, arXiv:1005.2357 [quant-ph].

[9] A. Caticha, "Entropic Dynamics: an Inference Approach to Quantum Theory, Time and Measurement," *Journal of Physics Conference Series* **504** no. 1, (Apr., 2014) 012009, arXiv:1403.3822 [quant-ph].

[10] E. T. Jaynes, "Information Theory and Statistical Mechanics," *Physical Review* **106** (1957) 620.

[11] E. T. Jaynes, "Information Theory and Statistical Mechanics. II," *Physical Review* **108** (1957) 171.

[12] A. Caticha and A. Giffin, "Updating Probabilities," in *Bayesian Inference and Maximum Entropy Methods In Science and Engineering*, A. Mohammad-Djafari, ed., vol. 872 of *American Institute of Physics Conference Series*, p. 31. 2006. arXiv:physics/0608185.

[13] E. Nelson, "Derivation of the Schrödinger Equation from Newtonian Mechanics," *Physical Review* **150** (1966) 1079.

[14] E. Nelson, *Dynamical Theories of Brownian Motion.* (Princeton U. Press, Princeton, 1967).

[15] E. Nelson, *Quantum Fluctuations.* (Princeton U. Press, Princeton, 1985).

[16] T. G. Dankel, "Mechanics on Manifolds and the Incorporation of Spin into Nelson's Stochastic Mechanics," *Archive for Rational Mechanics and Analysis* **37** (1970) 192.

[17] D. Dohrn, F. Guerra, and P. Ruggiero, "Spinning Particles and Relativistic Particles in the Framework of Nelson's Stochastic Mechanics," in *Feynman Path Integrals*, S. Albeverio, P. Combe, R. Høegh-Krohn, G. Rideau, M. Sirugue-Collin, M. Sirugue, and R. Stora, eds., vol. 106 of *Lecture Notes in Physics, Berlin Springer Verlag*, p. 165. 1979.

[18] A. N. Kolmogorov, *Foundations of the Theory of Probability.* (Chelsea Publishing Company, New York, 1956).

[19] R. T. Cox, "Probability, Frequency and Reasonable Expectation," *American Journal of Physics* **14** (1946) 1.

[20] R. T. Cox, *The Algebra of Probable Inference.* (Johns Hopkins, Baltimore, 1961).

[21] L. E. Maistrov, *Probability Theory: A Historical Sketch.* (Academic Press, New York, 1974).

[22] N. Mukhopadhyay, *Probability and Statistical Inference.* (Marcel Dekker, Inc. New York, 2000).

[23] M. Grigoriu, *Stochastic Calculus: Applications in Science and Engineering.* (Birkhauser, Boston, 2002).

[24] M. Lynn, "Categories of Probability Spaces," 2010. `http://www.math.uchicago.edu/~may/VIGRE/VIGRE2010/REUPapers/Lynn.pdf`.

[25] K. L. Chung, *A Course in Probability Theory.* (Academic Press, New York, 1974).

[26] E. T. Jaynes, *Probability Theory: The Logic of Science.* (Cambridge U. Press, 2003).

[27] A. Caticha, "Lectures on Probability, Entropy, and Statistical Physics," *ArXiv e-prints* (2008) , `arXiv:0808.0012 [physics.data-an]`.

[28] J. W. Gibbs, *Elementary Principles in Statistical Mechanics: Developed with Especial Reference to the Rational Foundation of Thermodynamics.* (Dover Publication, New York, 1960).

[29] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication.* (The University of Illinois Press, Urbana, 1964).

[30] A. Caticha, *Entropic Inference and the Foundations of Physics.* (São Paulo, Brazil, 2012). http://www.albany.edu/physics/ACaticha-EIFP-book.pdf.

[31] A. Caticha, "Relative Entropy and Inductive Inference," in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, G. J. Erickson and Y. Zhai, eds., vol. 707 of *American Institute of Physics Conference Series*, p. 75. 2004. arXiv:physics/0311093.

[32] A. Caticha, "Information and Entropy," in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, K. H. Knuth, A. Caticha, A. Giffin, and C. C. Rodríguez, eds., vol. 954 of *American Institute of Physics Conference Series*, p. 11. 2007. arXiv:0710.1068 [physics.data-an].

[33] J. E. Shore and R. W. Johnson, "Axiomatic Derivation of the Principle of Maximum Entropy and the Principle of Minimum Cross-Entropy," *IEEE Transactions on Information Theory* **26** (1980) 26.

[34] J. Skilling, "The Axioms of Maximum Entropy," in *Maximum-Entropy and Bayesian Methods in Science and Engineering*. 1988.

[35] A. Giffin and A. Caticha, "Updating Probabilities with Data and Moments," in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, K. H. Knuth, A. Caticha, A. Giffin, and C. C. Rodríguez, eds., vol. 954 of *American Institute of Physics Conference Series*, p. 74. 2007. arXiv:0708.1593 [physics.data-an].

[36] A. Einstein, "On the Motion of Small Particles Suspended in Liquids at Rest Required by the Molecular-Kinetic Theory of Heat," in *Einstein's Miraculous Year: Five Papers That Changed the Face of Physics*. 1998.

[37] M. Smoluchowski, "Kinetische Theorie der Brownsche Bewegung und der Suspensionen," *Ann. D. Phys.* **21** (1993) 756.

[38] P. Mörters and Y. Peres, *Brownian Motion.* (Cambridge U. Press, 2010).

[39] E. Nelson, "Connection between Brownian Motion and Quantum Mechanics," in *Einstein Symposion Berlin*, H. Nelkowski, A. Hermann, H. Poser, R. Schrader, and R. Seiler, eds., vol. 100 of *Lecture Notes in Physics, Berlin Springer Verlag*, p. 168. 1979.

[40] E. Nelson, "Review of Stochastic Mechanics," *Journal of Physics Conference Series* **361** (2012) 012011.

[41] T. C. Wallstrom, "On the Derivation of the Schrödinger Equation from Stochastic Mechanics," *Found. Phys. Lett.* **2** (1989) 113.

[42] T. C. Wallstrom, "The Inequivalence between the Schrödinger Equation and the Madelung Hydrodynamic Equations," *Phys. Rev. A* **49** (1994) 1613.

[43] F. Guerra and L. Morato, "Quantization of Dynamical Systems and Stochastic Control Theory," *Phys. Rev. D* **27** (1983) 1774.

[44] D. Bohm and B. J. Hiley, *The Undivided Universe: An Ontological Interpretation of Quantum Theory.* (Routledge, London, 1993).

[45] A. Einstein, B. Podolsky, and N. Rosan, "Can Quantum-Mechanical Description of Physical Reality be Considered Complete?," *Phys. Rev.* **47** (1935) 777.

[46] A. Caticha, "Change, Time and Information Geometry," in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, A. Mohammad-Djafari, ed., vol. 568 of *American Institute of Physics Conference Series*, p. 72. 2001. `arXiv:math-ph/0008018`.

[47] A. Caticha, "Entropic Time," in *American Institute of Physics Conference Series*, A. Mohammad-Djafari, J.-F. Bercher, and P. Bessiére, eds., vol. 1305 of *American Institute of Physics Conference Series*, p. 200. 2011. `arXiv:1011.0746 [quant-ph]`.

[48] S. Chandrasekhar, "Stochastic Problems in Physics and Astronomy," *Rev. Mod. Phys.* **15** (1943) 1. `http://link.aps.org/doi/10.1103/RevModPhys.15.1`.

[49] A. Caticha, D. Bartolomeo, and M. Reginatto, "Entropic Dynamics: from Entropy and Information Geometry to Hamiltonians and Quantum Mechanics," in *34th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering (21 – 26 September 2014), Château Clos Lucé, Parc Leonardo Da Vinci, Amboise, France.*

[50] S. Nawaz and A. Caticha, “Momentum and Uncertainty Relations in the Entropic Approach to Quantum Theory,” in *American Institute of Physics Conference Series*, P. Goyal, A. Giffin, K. H. Knuth, and E. Vrscay, eds., vol. 1443 of *American Institute of Physics Conference Series*, p. 112. 2012. arXiv:1108.2629 [quant-ph].

[51] L. I. Schiff, *Quantum Mechanics.* (McGrah-Hill Book Company, New York, 1968).

[52] W. Greiner and B. Müller, *Quantum Mechanics: Symmetry.* (Springer-Verlag, Berlin, 1994).

[53] D. de Falco, S. De Martino, and S. De Siena, “Position-Momentum Uncertainty Relations in Stochastic Mechanics,” *Phys. Rev. Lett.* **49** (1982) 181. http://link.aps.org/doi/10.1103/PhysRevLett.49.181.

[54] S. D. Martino and S. D. Siena, “Quantum Uncertainty Relations in Stochastic Mechanics,” *Il Nuovo Cimento* **79B** (1984) 175.

[55] S. Golin, “Uncertainty Relations in Stochastic Mechanics,” *J. Math. Phys.* **26** (1985) 2781.

[56] S. Golin, “Comment on Momentum in Stochastic Mechanics,” *J. Math. Phys.* **27** (1986) 1549.

[57] M. J. W. Hall and M. Reginatto, "Schrödinger Equation from an Exact Uncertainty Principle," *Journal of Physics A Mathematical General* **35** (2002) 3289, `arXiv:quant-ph/0102069`.

[58] E. Schrödinger, "About the Heisenberg Uncertainty Relation", Sitzungberichten der Preussischen Akademie der Wissenschaften," *Phys. Math. Kasse* **19** (1930) 296.

[59] B. Apolloni and D. de Falco, "Orbital Angular Momentum in Nelson's Stochastic Mechanics," *J. Phys. A: Math. Gen.* **33** (2000) 3225.

[60] W. H. Louisell, "Amplitude and Phase Uncertainty Relations," *Phys. Lett.* **7** (1963) 60.

[61] P. Carruthers, "Phase and Angle Variables in Quantum Mechanics," *Rev. Mod. Phys.* **40** (1968) 411.

[62] S. Golin, "Indeterminacy Relations in Stochastic Mechanics," in *Stochasitic Processes in Classical and Quantum System.* 1985.

[63] B. Podolsky, "Quantum-Mechanically Correct Form of Hamiltonian Function for Conservative Systems," *Physical Review* **32** (1928) 812.

[64] B. S. Dewitt, "Dynamical Theory in Curved Spaces. I. A Review of the Classical and Quantum Action Principles," *Reviews of Modern Physics* **29** (1957) 377.

[65] H. Jensen and H. Koppe, "Quantum Mechanics with Constraints," *Ann. Phys.* **63** (1971) 586.

[66] R. C. T. da Costa, "Quantum mechanics of a Constrained Particle," **23** (1981) 1982.

[67] A. Inomata and G. Junker, "Path Integration in Conical Space," *Physics Letters A* **376** (2012) 305, `arXiv:1110.2279 [math-ph]`.

[68] D. Dohrn and F. Guerra, "Nelson's Stochastic Mechanics on Riemannian Manifolds," *Lettere Al Nouvo Cimento* **22** (1978) 121.

[69] G. E. Uhlenbeck and S. Goudsmit, "Spinning Electrons and the Structure of Spectra," **117** (1926) 264.

[70] W. Pauli, "Zur Quantenmechanik des Magnetischen Elektrons," *Zeitschrift fur Physik* **43** (1927) 601.

[71] P. A. M. Dirac, "The Quantum Theory of the Electron," *Royal Society of London Proceedings Series A* **117** (1928) 610.

[72] P. A. M. Dirac, "The Quantum Theory of the Electron. Part II," *Royal Society of London Proceedings Series A* **118** (1928) 351.

[73] H. A. Kramers, *Quantum Mechanics.* (Dover Publications, Inc., New York, 1964).

[74] T. Takabayasi, "The Vector Representation of Spinning Particle in the Quantum Theory, I," *Progress of Theoretical Physics* **14** (1955) 283.

[75] R. Schiller, "Quasi-Classical Theory of the Spinning Electron," *Physical Review* **125** (1962) 1116.

[76] F. Bopp and R. Haag, "Über die Möglichkeit von Spinmodellen," *Zeitschrift Naturforschung Teil A* **5** (1950) 644.

[77] K. Gottfried and T. M. Yan, *Quantum Mechanics: Fundamentals.* (Springer, New York, second ed., 2004).

[78] W. K. Tung, *Group Theory in Physics.* (World Scientific, Philadelphia, 1985).

[79] R. L. Liboff, *Introductory Quantum Mechanics.* Addison-Wisley Publishing Company, New York, second ed., 1992.

[80] T. C. Wallstrom, "The Stochastic Mechanics of the Pauli Equation," *Transactions of the American Mathematical Society* **318** no. 2, (1990) 749. http://www.jstor.org/stable/2001329.

[81] F. Guerra and R. Marra, "Configuration Spaces for Quantum Spinning Particles," *Physical Review Letters* **50** (1983) 1715.

[82] D. T. Johnson and A. Caticha, "Entropic Dynamics and the Quantum Measurement Problem," in *American Institute of Physics Conference Series*, P. Goyal, A. Giffin, K. H. Knuth, and E. Vrscay, eds., vol. 1443 of *American*

Institute of Physics Conference Series, p. 104. 2012. arXiv:1108.2550 [quant-ph].

[83] D. T. Johnson, *Generalized Galilean Transformations and the Measurement Problem in the Entropic dynamics approach to quantum theory*. PhD thesis, State University of New York at Albany, 2011.

[84] A. Caticha, "The Entropic Dynamics of Relativistic Quantum Fields," *ArXiv e-prints* (2012) , arXiv:1212.6946 [quant-ph].

[85] S. Ipek and A. Caticha, "Entropic Quantization of Scalar Fields," in *34th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering (21 – 26 September 2014), Château Clos Lucé, Parc Leonardo Da Vinci, Amboise, France.*