Lenny, Me, and Entropy

Ken Matusow

A simple general system is defined that has exactly three properties: the system is non-continuous, non-deterministic, and has a well-defined, consistent metric. It is asserted that many of the basic concepts, principles, and equations of quantum mechanics and classical mechanics, including those of special and general relativity, are derivable, and indeed are unavoidably emergent from first principles given these three properties, and these three properties only.

For the past four years I've been attending a series of lectures given by Profession Leonard Susskind, the Felix Bloch professor of physics at Stanford University. During the course of these lectures, lectures that explore the mathematical foundations of modern physics, I have been struck by common lines of reason that parallel the mathematics of physics with the mathematics of systems theory and information theory. This paper is the direct result of a series of brief conversations I had and questions I posed to Susskind regarding the relationship between physics and systems theory.

In the spring of 2009 I met with Professor Susskind to discuss some ideas I had regarding the modelling of physical systems as stochastic processes. After some discussion Susskind challenged me to describe a stochastic system as a function of a complex wavefunction. During two years I spent to trying to resolve this still outstanding problem it became apparent that the solution could be easily expanded to describe a wide variety general physical processes. This paper represents the culmination of my efforts to complete this challenge.

This paper treats the description of physical systems as a systems theoretic construct rather than an extension of long accepted physical principles. In essence, the purpose of this exercise is to look at physical systems with 'fresh eyes', to view physical systems as general systems which are imbued with invariant modes of emergent behavior. Rather than building on the long and storied legacy of theoretical physics, the resulting general system is axiomatic. It stands alone as a self-consistent mathematical object independent from other physical theories.

The reason for this approach is based on practical considerations. Exceptional claims must be accompanied by exceptional proof and evidence. As an axiomatic system, the model is subject to formal proof, or disproof. If mathematical consistency can be claimed, then the following question can be posed:

Do the behaviors of this formal model accurately mimic those of real physical systems?

This question can only be answered by experimentation and validated by empirical evidence. If the evidence points to an answer of 'yes, the model does appear to behave in a manner similar to physical systems', then the general system can then be elevated to the level of 'candidate physical theory'.

This paper outlines a formal model of a one-dimensional stochastic system. The system, called S, is modelled as an abstract random walk. That is, the system S is associated with the notion of state and the state of S changes in a non-deterministic manner. Each change of state, or state transition, is labelled an **event**. Furthermore, the system S is subject to the constraint that it satisfies a metric with the property that the distance to the n-th event is proportional to n. The system described by S is axiomatic and is an example a general system. The term general system implies that the model can used to describe any real life system that satisfies the defined criteria, the axioms that define the model. Examples of such systems may be found is a disparate set of disciplines ranging from biological systems, chemical systems, economic systems, and others. A specific example of this kind of system, and one that is familiar to the physics community is a one-dimensional random walk where the system state represents the position of a particle.

In summary, let the system S be is an axiomatic system that satisfies the two postulates:

- Let S be a random walk
- Let the distance to the n-th event in S be proportional to n

Once the behaviour of and characteristics of the system S are thoroughly explored, the model can shown to be equivalent to the following three, more general axioms:

- Let the system S be discrete, not continuous
- Let the system S be non-deterministic; the system is stochastic
- Let the system S be associated with a well-defined, consistent metric

Restated in terms of a physical system, this paper claims that defining a general system to be both discrete and non-deterministic, and also associated with a consistent metric is sufficient to construct a model of a physical system that is consistent with many of the generally accepted laws and theories of physics.

In order to meet Lenny Susskind's challenge to model a stochastic system through the use of complex wavefunctions the model must satisfy the following criteria:

- *The model is axiomatic*
- *Each step is subject to formal proof*
- *The stochastic model may be completely described by the equation* $1 = \sum_{d=-\infty}^{\infty} |\bar{\psi}(d)|^2$,

  where d represents the discrete position of a particle in a single spatial dimension, and $\bar{\psi}$ *is a discrete complex wavefunction such that*
  $P(d) = \bar{\psi}(d)\bar{\psi}^*(d) = |\bar{\psi}(d)|^2$

1 Introduction

This paper describes a stochastic system from the perspective of a systems theorist. An abstract stochastic model is first presented. The properties of the model are examined, and then compared to the properties of a physical system. If a close match between the abstract system and the physical system is established, then the abstract model is elevated to candidate status as a formal description of the physical model.

The abstract system is axiomatic. That is, a set of two axioms is presented and examined. The properties and behaviours of the axiomatic system are elucidated through the use of formal proofs based solely on the two axioms.

This paper is a synopsis of a larger, more complete, and more formal model that is available as a draft.

2 The Two Axioms

Let a system called S be stochastic. Associated with S is the notion of state. An *event*, or state transition, is said to occur when the state of S changes. The state of the system S after a series of n events is defined to be $\beta(n)$. Because the system is stochastic, associated with each value of $\beta(n)$ is the probability that the system will end up in that state, $P(\beta(n))$. The probability $P(\beta(n))$ represents the probability that the system S is in state $\beta(n)$ after n events. This type of system is typically called a 'random walk'.

The properties and behaviors of random walk systems are well known.

Now, let a metric be assigned to S such that the distance to the n-th event is proportional to n. Let the proportionality factor be called $\bar{c}$.

The two axioms may now be made explicit:

- **Let S be a random walk**
- **Let the distance to the n-th event in S be proportional to n**

The system S will be shown to exhibit a remarkably rich variety of emergent properties and behaviors. Furthermore, these properties and behaviors closely mimic those of traditional physical systems.

3 The Structure of the Model

The formal description of the system S will be broken down into the number of discrete sections. Each section iteratively builds on the results of the previous section until the criteria of the model are satisfied. The sections, or steps, used to prove the Susskind challenge are:

- Determination of the coordinate system demanded by the axioms
- Analysis of the probabilities associated with each of the coordinates
- Introduction of complex wavefunctions as a function of state
- Derivation of the displacements associated with a state
- Assignation of units of measurement to each of the two axes
- Renaming of constants, relationships, and behaviors of the system with labels familiar to the physics community
- Final proof of the prime assertion in terms familiar to familiar to the physics community

If each of the steps can be formally proven then it will be claimed that the system S is mathematically consistent.

The final step is to examine and compare the emergent behaviors of the general system S with the behaviors physical phenomena. In other words, can the system S be used to explain and predict the behaviors of physical phenomena, and furthermore does the model lead to results and predictions and can be experimentally validated that are not explained by current, more standard physical models.

3.1 The Determination of the Resulting Coordinate System

Examine a system representing a single event, where $n = 1$, or $S(1)$. Assuming an initial state of 0, the final state system of a single event can be represented by the two coordinates {1,-1}. The system $S(1)$ satisfies both axioms. The system is random walk and the distance from the coordinate 0 to both 1 and -1 is proportional to the number of events. As $\bar{c}$ is the proportionality constant, the distance to both 1 and -1 is equal to $1 * \bar{c} = \bar{c}$.

Now examine $S(2)$. The system $S(2)$ must be in one of the states {2,1,0,-1,-2). The probability that the system S is in state 1 or state -1 is 0. The probabilities oscillate between positive and 0 values. The probability that the system S is in state 2, state 0, or state -2 is positive. The system is a random walk and therefore satisfies the first axiom. However the system, as currently described, does not satisfy the second axiom. This may be shown as follows.

If the state of the system $\beta(n)$ is either 2 or -2, the second axiom is satisfied. The distance to the second event is $2 * \bar{c}$. However, if the resulting state of the system $S(2)$ is 0, then the second axiom is not satisfied. The distance from the initial state 0 to the final state in $S(2)$, also 0, by definition, is 0 not 2. Since the system is axiomatic, the model must be modified in order to satisfy the second axiom: **Let the distance to the n-th event in S be proportional to n.**

The only way to satisfy the second axiom is to assign the state of $S(2)$ to the coordinate $(0,2)$, where the X-coordinate represents the state of the system, $\beta(n)$, and the Y-coordinate represents an *imaginary state*, $i\beta(n)$. Thus for $S(2)$ to satisfy the second axiom the complete state the system must be represent by the complex coordinate $(\beta(n), i\beta(n))$.

If an analysis of all possible states of $S(n)$ is undertaken a list of complex coordinates results. An example for small values of n is seen below:

$S(0) = \{0\}$
$S(1) = \{1,-1\}$
$S(2) = \{(2,0),(1,1),(0,2),(-1,1),(-2,0)\}$
$S(3) = \{(3,0),(2,1),(1,2),(0,3),(-1,2),(-2,1),(-3,0)\}$

The union of all values of $S(n)$, where $0 \leq n < \infty$ defines a coordinate system, namely the set of all ordered pairs on or above the X-axis, the northern semi-plane. Actually, this is not exactly correct. The ordered pair (0,1) is not defined for the coordinate system. The reason is that the ordered pair (0,1) it not consistent with the first axiom, since there is no change of state. The system starts out in state 0 and after a single event, the system remains in state 0. Therefore there is no state transition. The coordinate (0,1) is an example of a mathematical singularity. It turns out that the implications of this singularity are profound.

Associated with each integer coordinate (X,Y) defined by the system S are the following.

- X : The value of the state of the system: $\beta(n)$
- Y : The value of the imaginary state of the system: $i\beta(n)$
- n : The unique number of events associated with the coordinate: $n = X + Y$
- x : The value of the displacement in the X direction
- y : The value of the displacement in the Y direction
- P : The probability that a system of n events will end up in the state (X,Y): $P(X,Y)$

A unique 6-tuple (X,Y,n,x,y,P) is associated with every defined point in the coordinate system.

With the sole exception of the probability P, the other 5 values of every 6-tuple must be identical for all systems defined by the two axioms. Let S_1 and S_2 represent any two instances of S. The coordinate systems of the two systems must be identical.

3.2 Analysis of the Probabilities Associated with Each of the Coordinates

To this point in the discussion the system S is static. There is no dynamic component, no mechanism that drives the system to new states. This is like trying to build a model of physics without acknowledging the notion of energy. For the axiomatic system S the dynamic component is driven by the probabilities associated with each coordinate. The idea of least action is replaced by the concept of entropy. The informational, or Shannon entropy, associated with the set of probabilities is the dynamic that drives the system.

Basic probability theory states that the sum of the probabilities associated with all coordinates associated with $S(n)$ must be equal to 1.

(1) $$1 = \sum_{k=-n}^{n} P\big(k,(n-|k|)\big)$$

Furthermore, the discipline of information theory demands that the probability density function (PDF) associated with $S(n)$ must be the one that is described where the entropy functional is maximized, subject to possible additional informational constraints[i]:

(2) $$H(n) = -\sum_{k=-n}^{n} P\left(k,(n-|k|)\right)\log_2 P\left(k,(n-|k|)\right)$$

Given no further constraints the PDF associated with $S(n)$ is binomial. However, it was pointed out in section 3.1 that the probabilities associated with state of the system oscillate between a positive value and 0. The probabilities exhibit a wavelike characteristic.

3.3 Introduction of Complex Wavefunctions as a Function of State

As the collective probabilities or PDF of $S(n)$ exhibits wavelike characteristics, let the probabilities of $S(n)$ be described by the discrete wavefunction $\overline{\psi}$. As the wavefunction $\overline{\psi}$ operates on a set of complex coordinates, $\overline{\psi}$ is by definition complex, $\overline{\psi}\left(\beta(n),i\beta(n)\right)$.

Therefore the specific PDF associated with system S, as defined by the entropy functional $H(n)$, is described by the wavefunction $\overline{\psi}$.

The probabilities $P\left(\beta(n),i\beta(n)\right)$ may now be re-written in terms of the wavefunction:

(3) $$P\left(\beta(n),i\beta(n)\right) = \overline{\psi}(\beta(n),i\beta(n))$$

The normalized equation $1 = \sum_{k=-n}^{n} P\left(k,(n-|k|)\right)$ may be re-written in terms of the complex wavefunction $\overline{\psi}$.

(4) $$1 = \sum_{k=-n}^{n} \overline{\psi}\left(k,(n-|k|\right)$$

Now define the complex wavefunction in terms of displacement rather than state.

3.4 Derivation of the Displacements Associated with a State

The displacement x along the X-axis is an independent variable and can be written as:

(7) $$x = \beta(n)\overline{c}$$

The same cannot be said for displacement in the Y dimension since the displacement in the Y dimension is always a function of the state of S, $\beta(n)$, and not a function of the imaginary state of S, $i\beta(n)$. The displacement in the Y dimension is a dependent variable.

Assuming the underlying coordinate system is geometrically flat the second axiom demands that $n\bar{c} = \sqrt{x^2 + y^2}$ for any system of n events, where x and y are the displacements on the X and Y-axes, respectively. As both the distance along the diagonal and the distance along the X-axis are well defined, it is possible to determine the distance in the Y-dimension.

Solving for y:

(8) $$y = n\bar{c}\sqrt{1 - \left(\frac{\beta(n)}{n}\right)^2}$$

The displacements for both the X and Y coordinates may be explicitly defined and written in terms of the proportionality constant $\bar{c}$.

3.5 Assignation of Units of Measurement to Each of the Two Axes

Utilizing the principles of dimensional analysis units of measurement will be assigned to both axes.

Assign a unit of measurement on the X-axis. Call the unit of measurement m. Let the relationship between $\bar{c}$ and m be defined as $\bar{c} = k_m\langle m\rangle$, where k_m is a constant that is measured in units of $\langle m\rangle$.

The displacement on the X-axis, $\beta(n)\bar{c}$, may now be written in unit of m, namely:

(9) $$x = \beta(n)k_b\langle m\rangle$$

The Y-axis can similarly be written in units of m. Care must be taken to guarantee units of measurement are handled properly.

(10) $$y\langle m\rangle = nk_m\langle m\rangle * \sqrt{1 - \left(\frac{\beta(n)}{n}\right)^2}$$

However, since $\bar{c} = k_m\langle m\rangle = k_s\langle s\rangle$ the original equation may be rewritten as:

(11) $$y\langle m\rangle * \frac{k_s\langle s\rangle}{k_s\langle s\rangle} = \frac{k_s\langle s\rangle}{k_s\langle s\rangle} * \sqrt{\left(nk_m\langle m\rangle\right)^2 - \left(\beta(n)k_m\langle m\rangle\right)^2} = nk_s\langle s\rangle * \sqrt{\frac{\left(nk_m\langle m\rangle\right)^2}{\left(nk_s\langle s\rangle\right)^2} - \frac{\left(\beta(n)k_m\langle m\rangle\right)^2}{\left(nk_s\langle s\rangle\right)^2}}$$

Let the normalized ratio between $k_m\langle m\rangle$ and $k_s\langle s\rangle$, namely $\frac{\beta(n)}{n} * k_{ms}\langle m/s\rangle$, be called the *Displacement Velocity* v_D. The only thing that has changed is the unit of measurement. Since

v_D is by definition normalized, v_D has the property $1 \le v_D \le 0$. For convenience purposes let the maximum value of v_D be labelled as c.

Substituting in the new terms the Y displacement y is now defined in units of $\langle s \rangle$ and $\langle m/s \rangle$. The equation may now be written as:

$$(12) \qquad y\langle s \rangle = nk_s\langle s \rangle\sqrt{1 - \frac{v_D^2}{c^2}}$$

Where $nk_s\langle s \rangle$ is the maximum displacement on the Y-axis for a system of n events measured in units $\langle s \rangle$.

The displacements $x\langle m \rangle$ and $y\langle s \rangle$ are the two displacements x and y of the coordinate (X,Y).

3.6 Renaming of Constants, Relationships, and Behaviors of the System with Labels Familiar to the Physics Community

The X and Y coordinates will now be assigned names. As with the units of measurement, the names (to this point) are syntactic.

Let the X-axis be labelled the ***<distance axis>, <d>***.
Let the Y-axis be labelled the ***<time axis>, <t>***.

As such, the unit of measurement $\langle m \rangle$ quantifies the amount of displacement on the distance axis, while the unit of measurement $\langle s \rangle$ quantifies the amount of displacement on the time axis. It must be re-emphasized that, although both terms are loaded, to this point the assignations **<time>** and **<distance>** to the X and Y-axes are symbolic, pseudo-random, and entirely syntactic. No semantic content is demanded, although the implications are obvious.

Furthermore, additional labels will be syntactically assigned to key relationships associated with time axis and distance axis.

- Let the unit value of distance be called ***Planck length***, $1 * k_m\langle m \rangle = \ell_P$
- Let the unit value of time be called ***Planck time***, $1 * k_s\langle s \rangle = t_P$
- Let the maximum displacement on the <distance axis>, that is let $\beta(n) = n$, be called ***proper distance***, or $nk_m\langle m \rangle = n\ell_P$
- Let the maximum displacement on the <time axis>, that is let $\beta(n) = 0$, be called ***proper time***, τ, or $nk_s\langle s \rangle = nt_P$
- As described previously, let the maximum value of the displacement velocity, $v_D = nk_{ms}\langle m/s \rangle = \dfrac{nk_M\langle m \rangle}{nk_s\langle s \rangle} = \dfrac{n\ell_P}{nt_P} = \dfrac{\ell_P}{t_P}$, be called *the luminal velocity*, or c
- Let the system S itself be labelled a ***frame of reference***

3.7 Final Proof of the Prime Assertion in Terms Familiar to the Physics Community

The basic probabilistic relationships may now be re-written in terms of the *distance* and *time* displacements of the discrete wavefunction $\overline{\psi}$. It must again be re-emphasized that the terms the *distance* and *time* are syntactic assignations and to this point have nothing to do with commonly accepted notions of time and distance.

(13) $$P\big(\beta(n),i\beta(n)\big) = \overline{\psi}\,(\beta(n),i\beta(n)) = \overline{\psi}\,(d,t)$$

(14) $$1 = \sum_{k=-n}^{n} P\big(k,(n-|k|)\big) = \sum_{k=-n}^{n} \overline{\psi}\big(k,(n-|k|)\big) = \sum_{d=-n}^{n} \overline{\psi}\big(d = n\ell_P, t = f(d,n)\big)$$

The next step is to write the last equation is terms of distance, rather than in terms of distance and time. Multiplying the wavefunction by its complex conjugate is the final step in this process. The unitary property of probability theory leads to:

(15) $$1 = \sum_{d=-n\ell_P}^{n\ell_P} \overline{\psi}\big(d,t\big) = \sum_{d=-n\ell_P}^{n\ell_P} \overline{\psi}(d)\overline{\psi}^*(d) = \sum_{d=-n\ell_P}^{n\ell_P} \left|\overline{\psi}(d)\right|^2$$

Finally, as $n \rightarrow \infty$ the system S can be completely described by the equation:

(16) $$1 = \sum_{d=-\infty}^{\infty} \left|\overline{\psi}(d)\right|^2$$

Where the displacement d represents the discrete position of a particle on a one-dimensional line and $\overline{\psi}$ is a discrete complex wavefunction.

4 Conclusion

It has been formally shown that a stochastic system can be modelled using complex wavefunctions. Furthermore, the general model can be reconstituted into a form consistent with standard physics terminology. The reader will be hard pressed to find an inconsistency between the properties of time and distance and the properties defined by the axiomatic constructs called *time* and *distance*. As such, it is suggested that the underlying structure of physics is both discrete and non-deterministic. Lenny Susskind's challenge has been met.

References

[i] Claude E. Shannon, Warren Weaver. *The Mathematical Theory of Communication*. Univ of Illinois Press, 1949