

Measuring a Quantum System's Classical Information

John L. Haller Jr.

CCC Information Services, Data Science, Chicago, USA

Email: jlhaller@gmail.com

Received November 21, 2013; revised December 18, 2013; accepted January 5, 2014

Copyright © 2014 John L. Haller Jr. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. In accordance of the Creative Commons Attribution License all Copyrights © 2014 are reserved for SCIRP and the owner of the intellectual property John L. Haller Jr. All Copyright © 2014 are guarded by law and by SCIRP as a guardian.

ABSTRACT

In the governing thought, I find an equivalence between the classical information in a quantum system and the integral of that system's energy and time, specifically $I = \int 2\epsilon dt/\hbar$, in natural units. I solve this relationship in four ways: the first approach starts with the Schrödinger Equation and applies the Minkowski transformation; the second uses the Canonical commutation relation; the third through Gabor's analysis of the time-frequency plane and Heisenberg's uncertainty principle; and lastly by quantizing Brownian motion within the Bernoulli process and applying the Gaussian channel capacity. In support I give two examples of quantum systems that follow the governing thought: namely the Gaussian wave packet and the electron spin. I conclude with comments on the discretization of space and the information content of a degree of freedom.

KEYWORDS

Energy; Time; Information; Entropy; Rate; Capacity; Brownian Motion; Quantum of Action; Quantum of Information; Discrete Space; Discrete Time; Spin

1. Introduction

Anyone watching the media these days, especially the business news, knows about Big Data. The rate of data generation is growing exponentially and storage is due to multiply by 50 times between 2010 and 2020 [1]. The importance of the digital world is pronounced in almost every industry and every field of science [2]. It is not surprising that information is also important as a physical quantity in physics [3,4].

On that front, physics is challenged with many open questions including “five great problems” that would continue the march toward more knowledge [5]. I intend to contribute new insight vis-a-vis the governing thought that information equals energy times time and focus its application on two well-studied systems, the Gaussian wave packet and the electron.

While a base knowledge in information theory and physics is assumed, the arguments and derivations are intended to flow naturally from current understandings leading to new theory. I find the simplicity of the mathematics gives reason for special consideration. To show this elegance and resilient, the paper derives the

governing thought in four different proofs, then goes through two examples where the governing thought applies, and lastly ends with a couple notes in the appendix.

While the proper description of the word “information” is closer to self-information or entropy as developed by Shannon [6,7], I have chosen to use the word “information” to emphasize that a complete statistical description of a measurement of nature can be described with classical bits of information.

Shannon showed that self-information is formally defined as the negative expected log probability, $I = -E[\log(p)]$, and has meaning with both discrete and continuous probability distributions. It can even be derived from the definition of statistical entropy, $\log(\Omega)$, as given by Boltzmann [8]. The natural logarithm is used in this analysis and thus the units of information are the natural unit, or nats [9].

Energy times time, or more loosely “action”, has been a valuable concept in physics for over 2 centuries. It not only survived the migration from classical mechanics to quantum mechanics but rather it thrived with the realization that it was quantized [10]. We will start our investigation, into how information and energy times time are

equivalent, here with quantum mechanics and argue that information is one in the same as energy times time.

2. Quantum Mechanics

2.1. Schrödinger Equation

The Schrödinger equation, found during the advent of quantum mechanics, dictates how a wave function and its phase evolve through time. The Hamiltonian or energy operator, H , of a system is equal to \hbar times the imaginary derivative with respect to time; with the operator's eigenvalue, the energy, ϵ , of the system [10].

$$H\psi = i\hbar \frac{d}{dt}\psi = \epsilon\psi$$

The solution to this equation is the complex exponential,

$$\psi(t) = \psi(0)e^{-i\epsilon t/\hbar}$$

One can calculate the probability distribution associated with this wave function via its magnitude squared [10].

$$p(t) = |\psi(t)|^2 = |\psi(0)|^2 = p(0)$$

Note the phase information is lost. Calculating the information without considering the phase information one would conclude that the information is constant and a function only of its initial state, $\psi(0)$. Let me assume that $\psi(0) = 1$. However if we dig a little deeper an insight appears that when looked at in a few different ways proves resilient.

If the probability is constant, then the size of the space is equal to one over the probability, $\Omega = 1/p$. In this case S is the thermodynamic entropy [8],

$$\frac{S}{k_B} = \log(\Omega) = -\log(p)$$

By being a little more formal we can relax the condition that p is constant and instead only need to have $\psi(t)$ independent from $\psi(t+dt)$. Consider a large number N of independent steps in the particle, where $t = Ndt$. In this case, the probability of $p(t)$ is equal to $p(dt)$ raised to the N th power. I next use the weak law of large numbers through the asymptotic equipartition property (AEP) to focus on the most likely states [9]. (This is important and exemplified by the Gaussian distribution, where the Gaussian has infinite range but most likely states are limited to its standard deviation. If a large number of possible outcomes can occur each with the same probability within the range $\Delta\epsilon\Delta t$ of a system, then one can prove that information encoded into that system is equal to the negative log of the probability on average.) In this case

$$\psi(t) = \prod_{n=1}^N \psi(dt) = e^{-i\epsilon Ndt/\hbar}$$

and

$$p(t) = |\psi(t)|^2 = (|\psi(dt)|^2)^N$$

The AEP and the weak law of large numbers [9] can be used to show the negative log probability approaches the incremental entropy. Calling this the differential information, dI , I have

$$\frac{-\log(p(t))}{N} = -\log(p(dt)) \rightarrow dI \text{ as } N \rightarrow \infty$$

2.2. Dueling Information Rates

The insight comes by breaking up $p(dt) = |\psi(dt)|^2$ and looking at the differential information,

$$dI = -\log(|\psi(dt)|^2) = -\log(\psi^\dagger(dt)) - \log(\psi(dt))$$

Plugging in for $\psi(dt)$, results in the trivial answer

$$dI = \frac{(-i\epsilon + i\epsilon)dt}{\hbar} = 0$$

There are two information rates that cancel each other out, one equal to the imaginary energy divided by \hbar and the other a negative imaginary energy divided by \hbar .

If a Minkowski transformation had been performed prior to calculating the probability distribution a different answer would result. The Minkowski transformation takes imaginary time and makes it real. We see this transformation appear in relativity and analytic continuation [11,12]. After applying the Minkowski transformation $\psi(dt) = e^{-\epsilon dt/\hbar}$ such that,

$$p(dt) = |\psi(dt)|^2 = e^{-2\epsilon dt/\hbar}$$

and

$$dI = 2\epsilon dt/\hbar$$

or,

$$I = \int 2\epsilon dt/\hbar$$

This last equation is the governing thought of the paper. Assuming that the mass energy is not a function of time (which is not always the case), the following simple expression results,

$$I = 2\epsilon t/\hbar$$

2.3. Non-Commuting Operators

It is possible to remove the dependency on the Minkowski transformation and arrive at the same result by replacing the energy eigenvalue ϵ , with the energy op-

erator. When this is the case, the complex conjugate operation also requires the transpose of the operators since they do not commute. Using the power rule to expand the exponent and the logarithm, I now have, with H the energy operator,

$$-\log(\psi(t)) = \frac{-iHt}{\hbar}$$

$$-\log(\psi^\dagger(t)) = \frac{itH}{\hbar}$$

With this approach the negative log probability is now

$$-\log(p(t)) = \frac{itH - iHt}{\hbar}$$

Since the commutator $[H, t] = i\hbar$ [10], I have,

$$-\log(p(t)) = i(-i) = 1$$

I show in Section 4 that this is true of every step in the process with the appropriate step of size, $\delta t = \hbar/2\epsilon$. For a large number of K independent steps (where, $K = t/\delta t$), the negative log probability approaches the information, $-K \log(p(t)) = K \rightarrow I$, or

$$I = K = 2\epsilon t/\hbar$$

3. Signals

The initial conclusion (before I introduced imaginary time or the commutator) was that there is no information contained in the phase of a wave function; however we know from our analysis of signals that sine and cosine waves are capable of transmitting information in their phases. There are differences between the phase of the wave function and the phase of a signal [4] but it is worthwhile to pursue this approach as well.

Work by Nyquist and Hartley after the turn of the 20th century [13,14] tells us that the bandwidth of the signals is in direct proportion to the width of the signal in the frequency domain.

Gabor was even closer on track in the middle of 20th century when he tiled the time-frequency plane with quantized “logons” on information [15]—see **Figure 1**. Each logon was one degree of freedom and is represented by a shifted and modulated Gaussian wave packet. These wave packets were used as a basis to represent a signal with bandwidth f and duration t .

A more rigorous analysis of the number of degrees of freedom of a signal limited in bandwidth and time can be found by Slepian, Pollak and Landau [16-18].

They concluded the rate of information that can be encoded into a signal is linear in the bandwidth (or frequency). With Planck’s work on black body radiation and Einstein’s equation for the photo electric effect, where $\epsilon = hf$ [10], this proportion reduces to below

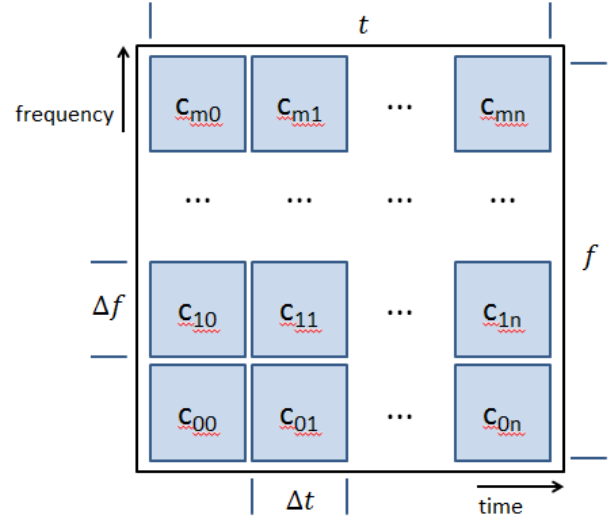


Figure 1. Time frequency plane quantized to individual degrees of freedom, each containing one natural unit of information.

where dI/dt is the information rate of the signal, f is the bandwidth of the signal, ϵ is Planck’s constant times f and t is the duration of the signal.

$$\frac{dI}{dt} \propto f \propto \epsilon$$

From here I can quickly return to the governing thought a third time by solving for the direct proportion in the equation above by dividing by the minimum width of the signal (the Heisenberg uncertainty relation).

$$dI = \frac{\epsilon dt}{\Delta \epsilon \Delta t} = 2\epsilon dt/\hbar$$

In Section 5.1, I show that the Gaussian wave packet, which obtains the minimum uncertainty, contains one natural unit of information, thus completing this proof.

An important insight to interpreting this equation is seen by again returning to **Figure 1** and looking at the time-frequency plane as a Venn diagram of entropy.

4. Brownian Motion

Before the two examples, I will re-derive the governing thought yet a 4th way by discretizing space and motion.

Building on the analysis by Kubo on the fluctuation dissipation theorem [19], I formalize the 2 time constants for a diffusing free particle; the collision time, δt and the relaxation time, τ . When the relaxation time is equal to the thermal time, $\tau = \hbar/2k_B T$, the diffusion constant becomes, $D = \hbar/2m$, [19-22] and spatial variance is $(\Delta x(t))^2 = 2Dt = \hbar t/m$.

4.1. Bernoulli Process

Introducing the Bernoulli process as reviewed by Reif

and Chandrasekhar [8,23], one can solve for the step size, δt , (or the collision time). The contribution to the spatial variance is balanced between drift and diffusion; when the probability parameter is 1/2 the variance is,

$$(\Delta x(K))^2 = \delta x^2 K + (\Delta v_K)^2 t^2$$

Here δx is the spatial step size, K is the number of steps, $t = K \cdot \delta t$ is the duration of the process and $(\Delta v_K)^2$ is the variance in velocity. From Dirac, we know that $\delta x = c \cdot \delta t$ [24] which allows us to calculate $(\Delta v_K)^2$.

When K is large, the average variance of the sum of K samples of a distribution is equal to the variance of the individual sample divided by K

$$(\Delta v_K)^2 = \frac{(\Delta v_1)^2}{K} = \frac{\frac{1}{2} \left(\frac{\delta x}{\delta t} \right)^2 + \frac{1}{2} \left(\frac{-\delta x}{\delta t} \right)^2}{K} = \frac{c^2}{K}$$

Equating $(\Delta x(t))^2$ and $(\Delta x(K))^2$, or $\hbar t/m = 2\delta x^2 K$, results in ,

$$\delta t = \frac{\hbar}{2mc^2} = \frac{\hbar}{2\epsilon}$$

Thus when the relaxation time is equal to one over twice the temperature $\tau = \hbar/2k_B T$, the collision time is one over twice the energy $\delta t = \hbar/2\epsilon$, and visa versa.

4.2. Information Content

With the details of the Bernoulli process defined, we can move onto the Gaussian channel. Combined with the Shannon-Nyquist's sampling theorem one has the channel capacity per second, C' [9],

$$C' = W \cdot \log \left(1 + \frac{P}{N_0 W} \right)$$

P is the signal power, $N_0/2$ is the noise spectral density and W is the bandwidth of the channel. (In this case, the channel is the vacuum which either has infinite bandwidth or some very large value.)

Using the assumption (aided by the insight of appendix A1) that the signal spectral density is equal to the noise spectral density, the signal power, P , is the noise spectral density times twice the bandwidth of the signal, $P = (N_0/2)(2\Delta w)$. Since the bandwidth of the signal is much smaller than the bandwidth of the channel, $\Delta w \ll W$, we can re-write the equation above as,

$$C' = \Delta w$$

The signal is the location of the particle performing the Bernoulli process with a step size of δt , thus the Shannon-Nyquist sampling theorem [25] tells us that the maximum frequency that can be represented by the discrete Bernoulli process is $\Delta w = 1/2\delta t$.

$$C' = \frac{1}{2\delta t} = \frac{\epsilon}{\hbar}$$

To finish the derivation I will take one more finding from Dirac, who showed that there is both a positive and negative solution to the energy eigenvalue [24]. Because there are two independent particles diffusing and information is generated by each particle a factor of 2 must be included; returning us to the governing thought.

$$\frac{dI}{dt} = 2C' = \frac{2\epsilon}{\hbar}$$

5. Examples

5.1. Gaussian Wave Packet

The Gaussian wave packet has many special properties, including 1) its Fourier transform is also a Gaussian [25], 2) the Gaussian obtains the minimum uncertainty relation [12], and 3) the Gaussian maximizes the differential entropy for a given variance [9]. As introduced above, Gabor [15] used the first two properties to tile the time frequency plane with shifted and modulated Gaussian wave packets. I will expand on property 3) and show that the information that can be decoded from one Gaussian wave packet is one nat.

First a result from Hirshman [26] where he proposed that to properly measure the information contained in a pair of distributions linked through the Fourier Transform (FT) one must add the differential entropy of the probability distribution in the time domain to the differential entropy of the probability distribution in the *standard* frequency domain. Given the scale property of the FT and the differential entropy, the sum of the two differential entropies is constant regardless of scale factor. Thus the information you can encode into a Gaussian wave packet is the same regardless of the relative width of the wave packet in the two domains.

Hirshman found that any FT pair contained at least $\log(e/2)$ of information and that the Gaussian has exactly $\log(e/2)$. I believe Hirshman missed an extra $\log(2)$ which nature requires. Looking at the governing thought, and applying the Heisenberg uncertainty principle when the energy is not a function of the time, shows that information in nature is greater than or equal to 1 natural unit.

$$I = 2\epsilon t/\hbar \geq 1$$

To show how this applies to the Gaussian, I again use the example of a massive particle. Dirac tells us from his work on the relativistic wave equation that there is both a positive and negative Eigen state [24]. Looking at the positive eigenvalue where the mass-energy divided by Planck's constant equals the average frequency $\bar{f} = mc^2/h$,

we have

$$\Phi_+(f) = \frac{1}{\sqrt[4]{2\pi(\Delta f)^2}} e^{-\frac{(f-\bar{f})^2}{4(\Delta f)^2}}$$

and for the negative eigenvalue,

$$\Phi_-(f) = \frac{1}{\sqrt[4]{2\pi(\Delta f)^2}} e^{-\frac{(f+\bar{f})^2}{4(\Delta f)^2}}$$

If $\hbar\Delta f = \Delta\epsilon \ll mc^2$ the two functions don't overlap, they don't interfere and thus according to Feynman [10] it's their probability distributions that add not the probability amplitudes (or wave functions). See **Figure 2**. The resulting probability distribution for the frequency domain $P_\Phi(f)$ is,

$$P_\Phi(f) = \frac{1}{2}|\Phi_+(f)|^2 + \frac{1}{2}|\Phi_-(f)|^2$$

Taking the inverse FT we have $\phi_+(t) = \int \Phi_+(f) e^{i2\pi ft} df$ and $\phi_-(t) = \int \Phi_-(f) e^{i2\pi ft} df$. The resulting probability distribution for the time domain, $p_\phi(t)$ is,

$$p_\phi(t) = \frac{1}{2}|\phi_+(t)|^2 + \frac{1}{2}|\phi_-(t)|^2$$

Given the modulation properties of the FT, $|\phi_+(t)|^2 = |\phi_-(t)|^2$ and with $4\pi\Delta f\Delta t = 1$, $p_\phi(t)$ reduces to

$$p_\phi(t) = \frac{1}{\sqrt{2\pi(\Delta t)^2}} e^{-\frac{t^2}{2(\Delta t)^2}}$$

Now using Hirshman's sum the result is, $\log(e)^1$,

$$\begin{aligned} I &= h(P_\Phi(f)) + h(p_\phi(t)) \\ &= \frac{1}{2}\log(8\pi e(\Delta f)^2) + \frac{1}{2}\log(2\pi e(\Delta t)^2) = 1 \end{aligned}$$

Coupling this back to Gabor's original analysis, I would conclude that a measurement of the Gaussian wave packet requires one natural unit of classical information to describe. This finite amount is due to the inherent noise associated with the Heisenberg uncertainty principle.

5.2. Electron Spin State

One might jump to the conclusion that a spin 1/2 particle has one natural unit of information by taking our governing thought and applying it to the spin angular momentum. However this is not correct as spin is quantized to $\hbar/2$ along of each of the three spatial dimensions. Associating one natural unit of information to each spa-

¹It is interesting to note that the Hirshman sum of the exponential wave function is also one natural unit.

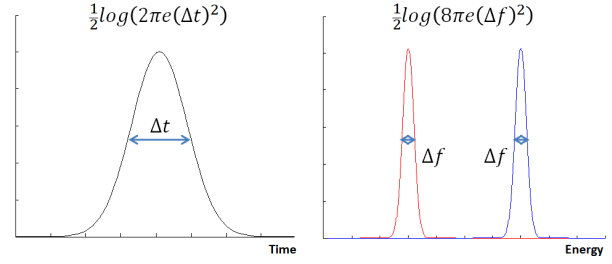


Figure 2. Probability distribution for Gaussian with non-overlapping positive and negative states. Hirshman sum is one natural unit.

tial dimension for a total of 3 natural units is also not correct since the measurements are not independent. The way to tackle this problem is by looking at the three spatial dimensions *and* the time dimension, then using our governing thought on the magnitude of the all four spin operators.

Let's start with the mathematical formulism to deal with the spin operator for the time dimension. This operator maps the wave function at the instantaneous moment in time to the discrete value $t_K = K \cdot \delta t$, or t_{K+1} when $t \in [t_K, t_{K+1})$.

Sticking with the spin operator formulism, we need a 2×2 matrix that has unity eigenvalues and returns the wave function untouched (since we are simply mapping to a discrete time value but not touching any of the spatial dimensions). You can see I have identified this Pauli matrix as the identity matrix. The spin operator S_t associated with this identity Pauli matrix is,

$$S_t = \frac{\hbar}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

S_t is now a fourth spin operator similar to S_x , S_y , and S_z .

One implication to adding the identity matrix to the formulism is that the magnitude of the spin angular momentum now takes on a more simple form with s the quantum spin number. Adding to the standard way of calculating $|S|$ [12] we have,

$$|S| = \sqrt{|S_x|^2 + |S_y|^2 + |S_z|^2 + |S_t|^2}$$

$$|S| = \sqrt{\hbar^2 s(s+1) + |S_t|^2} = \hbar(s+1/2)$$

Now applying the governing thought to angular momentum, I have the information in the spin of a particle equal to twice the magnitude divided by \hbar .

$$I_{spin} = \frac{2|S|}{\hbar} = 2s + 1$$

Applying this to the electron with $s=1/2$ the electron should have 2 natural units of information.

$$I_{spin-1/2} = 2$$

Let's see how that plays out using our current understanding of quantum information theory.

We know from Schumacher and Westmoreland [4], that the probability of error in inferring a message from a quantum measurement is at least one minus the dimension of the Hilbert space divided by the number of distinct messages, Y ,

$$P_{\text{Error}} \geq 1 - \frac{\dim(\mathcal{H})}{Y}$$

We find this channel capacity is equal to $\log(2)$ when the Hilbert space is a qubit and we choose to send spins in only the $|0\rangle$, or $|1\rangle$ state. In this case $Y = \dim(\mathcal{H}) = 2$ and P_{Error} can be zero.

However nature does not just produce electrons in only the $|0\rangle$, or $|1\rangle$ state. Sending an electron in one of only two states is a human choice and filtering or initialization is required. For nature to maintain symmetry and balance, the state must have a uniform distribution around the Bloch sphere. Thus the arbitrary state $|\xi\rangle$ is created

$$|\xi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\phi} \sin\left(\frac{\theta}{2}\right)|1\rangle$$

We also know that if we classically measure in an arbitrary direction, the wave function collapses to the state defined by that outcome [4]. This means that if we make a measurement in one direction with zero variance, any other non-commuting observable will have maximum variance.

However there is nothing stopping us from calculating the entropy we would expect if a measurement were to happen (even though don't make the measurement as that would collapse the state). Here I use the word entropy instead of information since a measurement of a state with $P_{\text{Error}} > 0$ will produce a partially random Boolean output that is not completely deterministic from knowledge of the initial state. Yet the term information is still relevant since it takes that much information to describe the outcome.

In this way we can add the entropy in each non-commuting observable in the same way Hirshman showed us.

For ξ to be uniformly distributed across the Bloch sphere we need to look at the Jacobian between spatial and spherical coordinates to seek how θ and ϕ are distributed. We find the determinant of the Jacobian is equal to $\sin(\theta)$, which means that ϕ is uniformly distributed but θ has the distribution $p(\theta) = (1/2)\sin(\theta)$ for $\theta \in [0, \pi]$.

Defining $H_2(\beta) \equiv -\beta \log(\beta) - (1-\beta) \log(1-\beta)$, where β is the probability of a positive measurement of the qubit in question, we can calculate the entropy of each spin operator by averaging $H_2\left(\beta = \cos\left(\frac{\theta}{2}\right)^2\right)$ over θ to get the average information, one will find,

$$I_{S_z} = \int_0^\pi \frac{1}{2} \sin(\theta) H_2\left(\beta = \cos\left(\frac{\theta}{2}\right)^2\right) d\theta = \frac{1}{2}$$

Since the distribution of ξ is uniform around the Bloch sphere this calculation is the same for S_x and S_y .

$$I_{S_x} = I_{S_y} = I_{S_z} = \frac{1}{2}$$

One might not be satisfied that each of S_x , S_y , and S_z can be separable, however if you go through the math and use the joint distribution on θ and ϕ and take into consideration the angle between ξ and S_x , S_y , and S_z , one gets the same answer.

The entropy of the S_t operator takes different reasoning to calculate, but the answer is the same. To start we need to review Section 4.1 on the Bernoulli process. If S_t acts to confirm a particle is occupying the time, $t_K = K \cdot \delta t$, where K is the step index and δt is the step size, the probability of a positive confirmation is equal to the relative distance the instantaneous time, t is away from t_K . A negative confirmation would mean that the particle is found in the state t_{K+1} .

Figure 3 is a picture of a particle at time t uniformly distributed between $t \in [t_K, t_{K+1})$. Finding the particle in the t_K state is equal to P_K and finding the particle in the t_{K+1} state is $1 - P_K$.

$$P_K = \int_{t_K}^t \frac{1}{\delta t} dt = \frac{t - t_K}{\delta t}$$

Without loss of generality we can set $K=0$. To complete the derivation I average $H_2(\beta = t/\delta t)$ over the uniform distribution to find I_{S_t} .

$$I_{S_t} = \int_0^{\delta t} \frac{1}{\delta t} H_2\left(\beta = \frac{t}{\delta t}\right) dt = \frac{1}{2}$$

Thus, using Hirshman's sum I find,

$$I_{\text{spin-1/2}} = I_{S_x} + I_{S_y} + I_{S_z} + I_{S_t} = 2$$

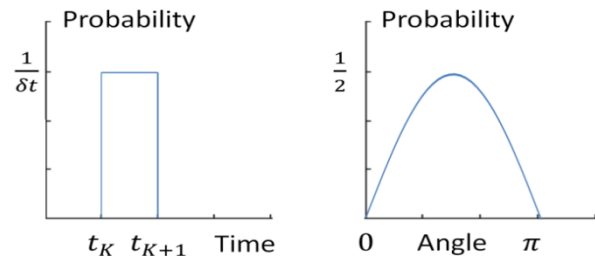


Figure 3. The action of S_t is to collapse the particle into the state, $t = t_K$, or $t = t_{K+1}$. Distribution of instantaneous time is uniform between two states. The action of S_x , S_y , and S_z is to collapse the particle into spin up or spin down. The distribution is $\sin(\theta)$.

It is interesting to note that ξ is described by 2 degrees of freedom. One real number from θ and one real number from ϕ . Thus we further support the idea that one degree of freedom is associated with one natural unit.

6. Discussion and Conclusion

The questions that I addressed are, “how much information is in an evolving system, how does one quantify it, and how much classical information is needed to describe it?”

You might have noticed that in some cases above I used the term rate, while in other cases I used capacity. We know from Shannon’s noisy-channel coding theorem that the rate must be less than the capacity for the probability of error of decoding a message to go to zero; and conversely that if the rate is greater than the capacity, an arbitrary small probability of error is not achievable [6]. However from my analysis it appears that in nature the rate of the underlying particles equals the capacity of the channel that transmits those particles through space-time; which in turn equals the energy of the particle. This insight is another reason why I am using the word information to describe $I = \int 2Rdt = \int 2C'dt = \int 2\epsilon dt/\hbar$.

I propose that the quantum state and the associated degree of freedom keep an extremely high (or possibly infinitely high) precision of its value. Yet when the wavefunction collapses and a classical measurement occurs, the capacity of that channel is only one natural unit of information per degree of freedom *and* the entropy rate of that measurement is such that one natural unit of information is needed to describe any measurement process you could ask of it per degree of freedom.

Thus, I find that precision to the infinite decimal point in the classical measurement is neither required nor possible since the classical information is finite. It thus makes sense that space is quantized. Planck showed energy is quantized [10]; Quantum Mechanics showed “action” is quantized; and from the analysis here a finite system is described by a finite amount of information; there is just not enough information in nature (or too much noise) to localize a particle to a continuous value that is not part of a finite set of values.

A broader debate is necessary to understand the physicality of classical information describing a quantum system, since its implications can be seen on both tangible qualities like the entanglement of quantum states and intangible qualities like the collapse of the wavefunction. Yet, one advantage of knowing the governing thought is that you are able to use information theoretic tools to solve physics problems and vice versa. For example, the principle of least action (which is fundamental to mechanics) can be now looked at as a principle of least information.

There is much more work to do in these areas including application to the unsolved problems of physics [5], or for that matter other unknown unsolved problems. Still having this insight into information is a good start.

Acknowledgements

JLH would like to thank many the ground breaking contributors who built the foundations of quantum mechanics and information theory, the Professors who inspire him, and specifically his family for their love, Professor Hawking for his reminder to always move “forward” and the Lord for the opportunity to contribute.

REFERENCES

- [1] J. Gantz and D. Reinsel, “The Digital Universe in 2020,” IDC, Framingham, 2012.
- [2] J. Manyika, *et al.*, “Big Data: The Next Frontier for Innovation, Competition, and Productivity,” McKinsey Global Institute, San Francisco, 2011.
- [3] L. Brillouin, “Science and Information Theory,” 2nd Edition, Academic Press Inc., Waltham, 1962.
- [4] B. Schumacher and M. Westmoreland, “Quantum Processes, Systems, and Information,” Cambridge University Press, Cambridge, 2010.
- [5] L. Smolin, “The Trouble With Physics,” First Mariner Books, New York, 2006
- [6] C. Shannon and W. Weaver, “The Mathematical Theory of Communication,” University of Illinois Press, Champaign, 1949.
- [7] M. Volkenstein, “Entropy and Information,” Birkhauser, Berlin, 2009.
<http://dx.doi.org/10.1007/978-3-0346-0078-1>
- [8] F. Reif, “Fundamentals of Statistical and Thermal Physics,” McGraw Hill, Boston, 1965.
- [9] T. Cover and J. Thomas, “Elements of Information Theory,” John Wiley & Sons Inc., New York, 1991.
<http://dx.doi.org/10.1002/0471200611>
- [10] R. Feynman, “Lectures on Physics,” Addison-Wesley Publishing, Reading, 1965.
- [11] A. Einstein, “The Meaning of Relativity,” Princeton University Press, Princeton, 1953.
- [12] R. Shankar, “Principles of Quantum Mechanics,” 2nd Edition, Plenum Press, New York, 1994.
<http://dx.doi.org/10.1007/978-1-4757-0576-8>
- [13] H. Nyquist, *Bell System Technical Journal*, Vol. 3, 1924, pp. 324-326.
<http://dx.doi.org/10.1002/j.1538-7305.1924.tb01361.x>
- [14] R. V. L. Hartley, *Bell System Technical Journal*, Vol. 7, 1928, pp. 535-563.
<http://dx.doi.org/10.1002/j.1538-7305.1928.tb01236.x>
- [15] D. Gabor, *Journal of Institution of Electrical Engineers*, Vol. 93, 1946, p. 429.
- [16] D. Slepian and H. O. Pollak, *Bell System Technical Journal*, Vol. 40, 1961, pp. 43-63.

- <http://dx.doi.org/10.1002/j.1538-7305.1961.tb03976.x>
- [17] H. J. Landau and H. O. Pollak, *Bell System Technical Journal*, Vol. 40, 1961, pp. 65-84. <http://dx.doi.org/10.1002/j.1538-7305.1961.tb03977.x>
- [18] H. J. Landau and H. O. Pollak, *Bell System Technical Journal*, Vol. 41, 1962, pp. 1295-1336. <http://dx.doi.org/10.1002/j.1538-7305.1962.tb03279.x>
- [19] R Kubo, *Reports on Progress in Physics*, Vol. 29, 1966, p. 255. <http://dx.doi.org/10.1088/0034-4885/29/1/306>
- [20] J. Haller Jr., *Journal of Modern Physics*, Vol. 4, 2013, pp. 85-95. <http://dx.doi.org/10.4236/jmp.2013.47A1010>
- [21] J. Haller Jr., *Journal of Modern Physics*, Vol. 4, 2013, pp. 1393-1399. <http://dx.doi.org/10.4236/jmp.2013.410167>
- [22] E. Nelson, *Physical Review*, Vol. 150, 1966, pp. 1079-1085. <http://dx.doi.org/10.1103/PhysRev.150.1079>
- [23] S. Chandrasekhar, *Reviews of Modern Physics*, Vol. 15, 1943, pp. 1-89.
- [24] P. A. M. Dirac, "The Principles of Quantum Mechanics," 4th Edition, Oxford University Press, Oxford, 1958, p. 262.
- [25] R. N. Bracewell, "The Fourier Transform and Its Applications," 2nd Edition, McGraw-Hill Inc., New York, 1986.
- [26] I. I. Hirshman Jr., *American Journal of Mathematics*, Vol. 79, 1957, pp. 152-156. <http://dx.doi.org/10.2307/2372390>
- [27] H. Nyquist, *Physical Review*, Vol. 32, 1928, pp. 110-113. <http://dx.doi.org/10.1103/PhysRev.32.110>

Appendix

A1. Signal Spectral Density

It is helpful to show how the signal of Brownian motion has the same spectral density as the noise. This is shown in the example of a harmonic oscillator with a thermal energy $k_B T$ (from two degrees of freedom) exactly equal to the quantum ground state energy, $\hbar\omega_0/2 = k_B T$ [20],

$$E = \frac{m}{2} \left(\frac{x}{\tau} \right)^2 + \frac{p_x^2}{2m}$$

The signal power is the time derivate of the energy, $P = \dot{E}$. As the signal is wiped-out for frequencies higher than one over twice the relaxation time, $W_\tau = 1/2\tau$ we divide the power by $2W_\tau$ to give the signal spectral density, $P\tau$,

$$P\tau = \tau \left(\frac{mx\dot{x}}{\tau^2} + \frac{p_x \dot{p}_x}{m} \right)$$

Here \dot{p}_x is the force, which is equal to $-mx/\tau^2$. Re-writing this equation and interpreting x and p_x as operators to replace with the commutator $[x, p_x] = i\hbar$ [10], we get

$$P\tau = \frac{xp_x - p_x x}{\tau} = \frac{i\hbar}{\tau} = 2k_B T$$

This is purely reactive power (as one would expect from an un-damped harmonic oscillator). But for the purposes here, one will recognize the apparent power, or magnitude of the complex power, (apart from the resistive factor) as the Johnson-Nyquist noise spectral density when spread over both positive and negative frequencies [27]. Thus we support the assumption in Section 4.2.

A2. Thermodynamic Derivation

Having in appendix A1, introduced the harmonic oscillator with quantum ground state energy equal to the temperature, it is straightforward to show that, for this example, the governing thought also applies within thermodynamics. Reviewing the power, P , we have

$$P = \frac{i\hbar}{\tau^2}$$

$$P = \frac{d(\Delta E)}{dt}$$

Since the power is reactive (imaginary), there is no work done and from the first law of thermodynamics [8] ΔE will equal the heat, Q . Our expression of for thermodynamic entropy is now ΔS

$$\Delta S = \int \frac{d(\Delta E)}{T} = \int \frac{P dt}{T}$$

Remembering $\tau = \hbar/2k_B T$ and assuming the temperature is constant gives the result,

$$\Delta S = k_B \int \frac{4k_B T i}{\hbar} dt = k_B \frac{4k_B T i t}{\hbar}$$

It is interesting that once again imaginary time is in the picture. Since $E_0/2 = \hbar\omega_0/2 = k_B T$, and upon making a Minkowski transformation [11,12],

$$\Delta S/k_B = 2E_0 t/\hbar$$

Haller derives the entropy rate of thermal diffusion for one particle in [21]; here I will use a quicker derivation for both the positive and negative energy states and use the Gaussian channel capacity. The signal in this case is the width of the diffusing particle undergoing the Bernoulli process for one step, δt ,

$$(\Delta x(\delta t))^2 = 2D\delta t = \hbar\delta t/m$$

The noise is the minimum width of the wavepacket [10],

$$(\Delta x_0)^2 = \hbar^2/(4k_B T m)$$

The capacity of one step of the Gaussian channel with this signal and noise is

$$C = \frac{1}{2} \log \left(1 + \frac{(\Delta x(\delta t))^2}{(\Delta x_0)^2} \right)$$

Taking C over the step size, δt , (and including a factor of 2 for the two particles), we have the information rate,

$$\frac{dI}{dt} = \frac{2C}{\delta t} = \frac{1}{\delta t} \log \left(1 + \frac{(\Delta x(\delta t))^2}{(\Delta x_0)^2} \right)$$

Assuming a non-relativistic particle, where $mc^2 \gg k_B T$, or equivalently $\delta t \ll \tau$, the above reduces to,

$$\frac{dI}{dt} = 4k_B T/\hbar$$

Replacing $E_0 = 2k_B T$, the governing thought is again returned,

$$I = \int 2E_0 dt/\hbar$$

As assumed before, the temperature is constant so by integrating, we finish with information being equal to the thermodynamic entropy calculated above in this section,

$$I = 2E_0 t/\hbar = \Delta S/k_B$$