Entropy Rate of Thermal Diffusion

John L. Haller Jr.

jlhaller@alumni.princeton.edu

Abstract: The thermal diffusion of a free particle is a random process and generates entropy at a rate equal to twice the particle's temperature, $R = 2k_BT/\hbar$ (in natural units of information per second). The rate is calculated using a Gaussian process with a variance of $(\Delta x_0 + \Delta p\cdot t/m)^2$. One would be keen to notice that the solution to the quantum mechanical diffusion of a free particle is $(\Delta x_0)^2 + (\Delta p\cdot t/m)^2$, however we assume that concurrent to quantum diffusion, the center of the wavepacket is also undergoing classical diffusion which adds an addition variance in the amount of $(\hbar\cdot t/m)$, making up the difference. Derivations of the variance and subsequent entropy rate are given.

I Primary Finding: When a free particle is at a non-zero temperature, it is composed of a spectrum of frequencies that interact with each other and cause the probability distribution of where one can find the particle to spread. We will show that the entropy rate, associated with the probability distribution diffusing, is equal to twice the particles temperature.

$$R = 2k_B T / \hbar \tag{1}$$

The rate, R, is calculated below using the natural logarithm and thus the units for the rate are natural units of information per second, when the temperature (T) is expressed in degrees Kelvin, Boltzmann's constant (k_B) is expressed in Joules per Kelvin and Planck's constant divided by 2π ($\hbar$) is in Joule-seconds.

This equation tells us how much information we need, each second, on average, in-order to track a diffusing free particle to the highest precision that nature requires. By quantifying this number, we are able to guarantee that a computer (with possibly large, but finite memory) can store a "perfect" replica of the particles trajectory.

II Assumptions: We prove this primary result by making the following three assumptions:

1) The diffusion of a free particle in a vacuum can be modeled as a discrete process with a small time step, $dt \ll \hbar/(2k_BT)$, where T is the temperature.

2) Knowing the particles location at time step n+1 allows one to determine the location of the particle at the previous time step n, i.e. conditional entropy is zero, $h(X_n|X_{n+1}) = 0$ where X_n is the random variable that represents where the particle can be found at time step n.

3) Concurrent to the quantum diffusion of the wavepacket, the center of the wavepacket also undergoes classical diffusion with a diffusion constant $D = \hbar/2m$, where m is the mass of the particle.

III Setup: At t=0, a free particle in vacuum is initialized into a minimum uncertainty Gaussian wavepacket with a spatial variance equal to $(\Delta x_0)^2$. As time increases so does its variance and thus its entropy.

To calculate the entropy rate of this process, it is helpful to think of time as occurring in discrete units of a small size dt (assumption one).

We can look at a Venn diagram of this process, figure (1). $X_{g,0}$ (or X0 in the figure) is a random variable, drawn from g(x,0), that describes the location of where the particle can be found at time t=0. $X_{g,1}$ (X1) is a random variable, drawn from g(x,dt), that describes the location of where the particle can be found at time t=dt. $X_{g,2}$ (X2) is drawn from g(x,2·dt) and so on up to $X_{g,n}$ which is drawn from g(x,t) when t=n·dt.

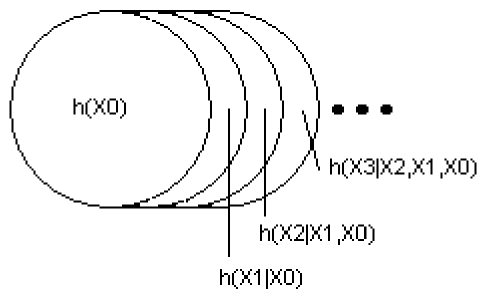

Figure (1) – Venn diagram of the conditional entropies of the diffusion process

As hinted to in the diagram (but explicitly stated here as assumption two), we will assume that $h(X_{g,n}|X_{g,(n+1)}) = 0$, where h is the differential entropy of g(x,t), $h = -\int g\cdot\log(g)\cdot dx$. This essentially means that knowledge of the location of the particle at a given time allows one to calculate where it was in the previous time step. This could hold true if there was a way to tell from which direction the particle previously came from.

In section V, we show that using assumption three, a free particle diffuses with a variance in its position (if localized) of $(\Delta x)^2$.

$$(\Delta x)^2 = \left(\Delta x_0 + \frac{\Delta p \cdot t}{m} \right)^2 \tag{2}$$

Thus $X_{g,n}$ (or simply X_n) is a Gaussian random variable with variance $(\Delta x)^2 = (\Delta x_0 + \Delta p\cdot n\cdot dt/m)^2$.

IV Entropy Rate: The entropy rate of this process is calculated from the two definitions/methods of the entropy rate. Both methods are used because the dual limit of the rate, as the number of steps goes to infinity and as the time step dt goes to zero, does not exist. However we show that in one method the rate, R, is less than or equal to twice the temperature (when Plank units are used) and in the other method R is greater than or equal to twice the temperature. Thus R is equal to twice the temperature.

In the first case, the entropy rate, R, is calculated by taking the limit as the number of steps goes to infinity of the conditional entropy of the last step given all previous steps divided by the time step.[1]

$$R \equiv \frac{\lim_{n\to\infty} h(X_{n+1} \mid X_n, X_{n-1}, ..., X_1, X_0)}{dt} \tag{3}$$

To solve for R, we first notice that since $h(X_n|X_{n+1}) = 0$ (assumption two), thus we can show by induction that

$$h(X_{n+1} \mid X_n, X_{n-1}, ..., X_1, X_0) = h(X_{n+1} \mid X_n) \tag{4}$$

Due to the symmetric nature of mutual information, we can prove the equation below.[2]

$$h(X_{n+1} \mid X_n) = h(X_{n+1}) - h(X_n) + h(X_n \mid X_{n+1}) \tag{5}$$

Brining us to the equation for R below

$$R = \frac{\lim_{n\to\infty}\left[h(X_{n+1}) - h(X_n)\right]}{dt} \tag{6}$$

Since the X_n's are Gaussian, we can easily calculate the differential entropy of each step using equation (2).[3]

$$R = \frac{\lim_{n\to\infty}\left[\frac{1}{2}\log\left(2\pi e\cdot\left(\Delta x_0 + \frac{\mathrm{dt}}{\mathrm{m}}\cdot(n+1)\cdot\Delta\mathrm{p}\right)^2\right) - \frac{1}{2}\log\left(2\pi e\cdot\left(\Delta x_0 + \frac{\mathrm{dt}}{\mathrm{m}}\cdot n\cdot\Delta\mathrm{p}\right)^2\right)\right]}{dt} \tag{7}$$

$$R = \frac{\lim_{n\to\infty}\left[\log\left(1 + \frac{\frac{dt}{m}\cdot\frac{\Delta p}{\Delta x_0}}{1 + \frac{dt}{m}\cdot n\cdot\frac{\Delta p}{\Delta x_0}}\right)\right]}{dt} \tag{8}$$

Since dt is small, we can Taylor expand the logarithm giving the first term plus the terms that are O(dt) or smaller.

$$R = \lim_{n\to\infty}\left[\frac{\frac{\Delta p}{m\Delta x_0}}{1 + \frac{\Delta p\cdot n\cdot dt}{m\Delta x_0}} + O(dt) + O(dt^2) + \ldots\right] \tag{9}$$

Ignoring the terms of O(dt) or smaller, and by observation that the denominator of the first term is greater than or equal to one, we see that for any non-negative n,

$$R \le \frac{\Delta p}{m\Delta x_0} \tag{10}$$

As solved for later in the paper, equation (31) and (32) can be use to show that,

$$R \le 2k_B T/\hbar \tag{11}$$

To calculate the other constraint on R, we start with R being equal to the limit as n goes to infinity of the entropy of all the X_n's divided by n times dt. Since we are looking at the rate of generation of the entropy (not the initial conditions), we subtract the entropy of the initial state $h(X_0)$. This also assures that R is in the correct units.

$$R \equiv \lim_{n\to\infty}\left[\frac{h(X_n, X_{n-1}, \ldots, X_1, X_0) - h(X_0)}{n\cdot dt}\right] \tag{12}$$

Since $h(X_n|X_{n+1}) = 0$ (assumption two), we know that $h(X_n, X_{n-1}, \ldots X_1, X_0) = h(X_n)$, thus

$$R = \lim_{n\to\infty}\left[\frac{h(X_n) - h(X_0)}{n \cdot dt}\right] \tag{13}$$

Plugging in for the differential entropy of the Gaussian distribution that describes X_n, and X_0, equation (2), we arrive at

$$R = \lim_{n\to\infty}\left[\frac{\frac{1}{2}\log\left(2\pi e\cdot\left(\Delta x_0 + \frac{\mathrm{dt}}{\mathrm{m}}\cdot n\cdot \Delta \mathrm{p}\right)^2\right) - \frac{1}{2}\log\left(2\pi e\cdot\left(\Delta x_0\right)^2\right)}{n\cdot dt}\right] \tag{14}$$

$$R = \lim_{n\to\infty}\left[\frac{\log\left(1 + \frac{\mathrm{dt}}{\mathrm{m}}\cdot n\cdot\frac{\Delta \mathrm{p}}{\Delta \mathrm{x}_0}\right)}{n\cdot dt}\right] \tag{15}$$

Expanding R using Taylor's theorem

$$R = \lim_{n\to\infty}\left[\frac{\Delta \mathrm{p}}{m\Delta \mathrm{x}_0} + \sum_{k=2}\frac{1}{k}\left(n\cdot dt\right)^{k-1}\cdot\left(\frac{\Delta \mathrm{p}}{m\Delta \mathrm{x}_0}\right)^k\right] \tag{16}$$

Even though the sum does not converge for dt going to zero while n goes to infinity at the same time, the sum is non-negative for all n>0 if dt≥0 (which it is). Thus

$$R \geq \frac{\Delta \mathrm{p}}{m\Delta \mathrm{x}_0} \tag{17}$$

Again plugging in equations (23) and (24)

$$R \geq 2k_B T/\hbar \tag{18}$$

Putting equation (11) and (18) together reveals our primary result.

$$R = 2k_B T/\hbar \tag{19}$$

This result is consistent with (and should be predicted by) the perspective gained from Information Mechanics. Information Mechanics states that information is equal to twice energy times time divided by ħ, $I=2\varepsilon\tau/\hbar$.[4] In this view, the temperature acts as an average energy and generates information (or entropy) at a rate equal to twice the average energy divided by ħ.

V The Variance of X_n: Given the wave particle duality, which states that a free particle is both a wave and a particle, it is sensible to assume that our free particle undergoes both quantum mechanical diffusion of the wave and classical diffusion of the particle. This is the essence of assumption three.

Introducing X_p, X_f, p(x,t), and f(x,t) makes this more clear. X_p is a random variable drawn from p(x,t), equation (40), the probability distribution associated with the quantum mechanical wavefunction which is the solution to the quantum diffusion equation, equation (33). X_f, is a random variable drawn from f(x,t), equation (43), which represents the diffusion of the center of ψ(x,t) from its original center, and is the solution to real diffusion equation, equation (42).

If we were to make an observation of where the particle is located, X_g, our answer would be the sum of a sample X_p drawn from $p(x,t) = \psi^*(x,t)\psi(x,t)$ and a sample X_f drawn from f(x,t).

$$X_g = X_p + X_f \tag{20}$$

Thus the action of f(x,t) is to translate the center of the wavefunction, ψ(x,t), by a sample of X_f.

As we know from probability theory, the resulting distribution, g(x,t) is equal to the convolution of p(x,t) and f(x,t) over the x variable (30).[5]

$$g(x,t) = p(x,t) *_x f(x,t) \tag{21}$$

Since both p(x,t) and f(x,t) are Gaussian distributions, it is easy to show that the convolution of the two is again a Gaussian distribution with an expected value being equal to the sum of the two expected values (which in this case is zero) and a variance that is equal to the sum of the variances of the individual distributions.

$$\overline{x_g} = \overline{x_p} + \overline{x_f} = 0 \tag{22}$$

$$(\Delta x_g)^2 = (\Delta x_p)^2 + (\Delta x_f)^2 \tag{23}$$

Shown in equation (41) the variance of p(x,t) is $(\Delta x_p)^2$.

$$(\Delta x_p)^2 = (\Delta x_0)^2 + \left(\Delta p \cdot \tfrac{t}{m}\right)^2 \tag{24}$$

In this equation t is the amount of time that has pasted since the particle was initialized in the minimum uncertainty state, Δx_0 is the standard deviation of the minimum uncertainty state, Δp is the standard deviation of the minimum uncertainty state in the momentum domain and m is the mass of the particle.

Shown in equation (48) the variance of f(x,t) is $(\Delta x_f)^2$.

$$(\Delta x_f)^2 = \frac{\hbar \cdot t}{m} \tag{25}$$

Thus we get $(\Delta x_g)^2$.

$$(\Delta x_g)^2 = (\Delta x_0)^2 + \left(\Delta p \cdot \tfrac{t}{m}\right)^2 + \tfrac{\hbar \cdot t}{m} \tag{26}$$

Multiplying the last term by the Heisenberg Uncertainty principle (32), $2\Delta x_0 \Delta p/\hbar = 1$, we can group.

$$(\Delta x_g)^2 = (\Delta x_0)^2 + \left(\Delta p \cdot \tfrac{t}{m}\right)^2 + 2\Delta x_0 \Delta p \cdot \tfrac{t}{m} = \left(\Delta x_0 + \Delta p \cdot \tfrac{t}{m}\right)^2 \tag{27}$$

It is helpful to the understanding of the model to look at equation (26). $(\Delta x_g)^2$ is the sum of three variances. The first is from the Heisenberg Uncertainty Principle, the second is from the thermal drift of the center of the minimum uncertainty wavepacket moving with a group momentum taken as a sample of the momentum domain, and the third is from the classical diffusion on top of the other two.

It is interesting and worthwhile to here note that the variance of g(x,t) would be the same if the classical diffusion term was not accounted for, but rather if there existed a perfect correlation between the $(\Delta x_0)^2$ term and the $(t\cdot\Delta p/m)^2$ term.

VI The Imaginary Diffusion Equation: The Kinetic Energy Hamiltonian characterizes the wave packet of a free particle in one dimension, where H is the Hamiltonian, p is the momentum along the x direction, and m is the mass of the particle.[6]

$$H = p^2 / 2m \tag{28}$$

Given that the momentum commutes with the Hamiltonian, $[p,H] = [p,p^2/2m] = 0$, each eigenvalue of the momentum is a constant of motion and thus the variance in momentum space does not grow with time. It is possible to learn the width of the variance of the momentum by looking at the equipartition of energy.[7] Using the equipartition of energy we know to equate the degree of freedom associated with the average Kinetic Energy to one half the temperature times Boltzmann's constant.

$$\overline{p^2 / 2m} = \tfrac{1}{2} k_B T \tag{29}$$

Since we will assume that the average momentum is zero, we can solve for the variance of the momentum.

$$\bar{p} = 0 \tag{30}$$

$$\Delta p^2 = \overline{p^2} - \bar{p}^2 = k_B T m \tag{31}$$

Also from the Heisenberg Uncertainty Principal, we can solve for the standard deviation of the wavefunction in the spatial domain in terms of its width in the momentum space.

$$\Delta x_0 = \hbar / 2\Delta p \tag{32}$$

With these dependencies stated, we can move onto the imaginary diffusion equation, which takes the original Hamiltonian and rewrites it in terms of operators. Interpreting the Hamiltonian as the imaginary time derivative operator and the momentum as the negative imaginary spatial derivative operator we can take equation (28) and arrive at the imaginary diffusion equation.

$$H \cdot \psi(x,t) = i \cdot \hbar \cdot d/dt\, \psi(x,t) = \tfrac{-\hbar^2}{2m} \cdot d^2/dx^2\, \psi(x,t) \tag{33}$$

Don't forget that we still have the eigenvalue equations (34,35) where H and p are the operators and ω and k are the eigenvalues.

$$H \cdot \psi(x,t) = \hbar \cdot \omega \cdot \psi(x,t) \tag{34}$$

$$p^2 / 2m \cdot \psi(x,t) = \hbar^2 \cdot k^2 / 2m \cdot \psi(x,t) \tag{35}$$

We can equate the different eigenvalues, ω and k, through equations (28,34,35) and as we should expect arrive at the equation for kinetic energy.

$$\hbar \cdot \omega = \hbar^2 \cdot k^2 / 2m \tag{36}$$

To solve equation (33), we will begin in the momentum domain $\Psi(k/2\pi)$ and take the inverse Fourier Transform to observe how $\psi(x,t)$ evolves over time.[8] We use $k/2\pi$ (the wavenumber divided by 2π) as the independent variable because we want both $\Psi(k/2\pi)$ and $\psi(x,t)$ to be normalizable to one.

$$\Psi(k/2\pi) = \left(\frac{2\pi}{(\Delta k)^2}\right)^{\frac{1}{4}} \cdot \exp\left[\frac{-k^2}{4\cdot(\Delta k)^2}\right] \tag{37}$$

Our assumption that the wavefunction of the free particle in the momentum space is a Gaussian wavepacket is quite reasonable given the nice properties of the Gaussian. Similarly this assumption is already implicit in the equipartition of energy which was used to find the width of the initial wavepacket. Because the equipartition theorem is derived from the perfect gas law (where particles are modeled using the binomial distribution, of which the Gaussian is the limit) the Gaussian is the right distribution to start with.

To properly account for the evolution of $\psi(x,t)$ governed by equation (33), $\exp[i(kx-\omega t)]$ is used as the kernel for the inverse Fourier Transform.

$$\psi(x,t) = \int_{-\infty}^{\infty} \left(\frac{2\pi}{(\Delta k)^2}\right)^{\frac{1}{4}} \cdot \exp\left[\frac{-k^2}{4\cdot(\Delta k)^2}\right] \cdot \exp\left[i\cdot(kx-\omega t)\right]\cdot \frac{dk}{2\pi} \tag{38}$$

Using equation (36) to substitute in for ω you can solve for equation (38) by completing the squares.[9]

$$\psi(x,t) = \frac{1}{\left(2\pi\left(\Delta x_0 + i\Delta p\cdot \frac{t}{m}\right)^2\right)^{\frac{1}{4}}} \cdot \exp\left[\frac{-x^2}{4\cdot\Delta x_0 \cdot \left(\Delta x_0 + i\Delta p \cdot \frac{t}{m}\right)}\right] \tag{39}$$

Where Δp was inserted in place of $\hbar\Delta k$.

To calculate the variance, we need to take the magnitude squared of the wavefunction and get the distribution of the particle.

$$p(x,t) = \psi^*(x,t)\cdot\psi(x,t) = \frac{1}{\sqrt{2\pi\left((\Delta x_0)^2 + \left(\Delta p\cdot\frac{t}{m}\right)^2\right)}} \cdot \exp\left[\frac{-x^2}{2\cdot\left((\Delta x_0)^2 + \left(\Delta p \cdot \frac{t}{m}\right)^2\right)}\right] \tag{40}$$

This is of course the know result from quantum mechanics where the variance of the particle is the sum of the initial variance from the Heisenberg Uncertainty Principal and the associated variance of the momentum domain acting like a velocity of magnitude $\Delta p/m$.[10]

$$(\Delta x_p)^2 = (\Delta x_0)^2 + \left(\Delta p \cdot \frac{t}{m}\right)^2 \tag{41}$$

VII The Real Diffusion Equation: When the diffusion constant of a diffusion process is real and does not vary with position, the resulting diffusion equation is as below.[11]

$$\frac{d}{dt} f(x,t) = D \cdot \frac{d^2}{dx^2} f(x,t) \tag{42}$$

Of course the solution to this real diffusion equation is the Gaussian with zero mean and variance equal to 2Dt.[12]

$$f(x,t) = \sqrt{\frac{1}{4\pi Dt}} \cdot \exp\left(-\frac{x^2}{4Dt}\right) \tag{43}$$

$$\left(\Delta x_f\right)^2 = 2Dt \tag{44}$$

To find D, we will start with the imaginary diffusion operator and perform a Minkowski transformation.[13] The imaginary diffusion operator (33) is

$$\frac{d}{dt} = \frac{i \cdot \hbar}{2m} \cdot \frac{d^2}{dx^2} \tag{45}$$

Upon applying the Minkowski transformation, imaginary time is replaced with real time, $i \cdot t \rightarrow \tau$.[14] Applied on the imaginary diffusion operator, the Minkowski transformation brings out the real diffusion constant we are looking for.[15]

$$\frac{d}{dt} = \frac{i \cdot \hbar}{2m} \cdot \frac{d^2}{dx^2} \rightarrow \frac{d}{d\tau} = \frac{\hbar}{2m} \cdot \frac{d^2}{dx^2} \tag{46}$$

By observation we see that

$$D = \frac{\hbar}{2m} \tag{47}$$

Thus we can calculate the variance of f(x,t).

$$\left(\Delta x_f\right)^2 = \frac{\hbar \cdot t}{m} \tag{48}$$

The resulting diffusion constant, D, and the variance, $\left(\Delta x_f\right)^2$, can also be derived by assuming the free particle undergoes a Bernoulli random walk as the source of the diffusion.[16]

VIII Conclusion: We have seen that by making three assumptions about the thermal diffusion of a free particle, we are able to show that entropy is generated at a rate equal to twice the particles temperature (when expressed in the correct units).

This result will be applicable to all studies on free particles and other environments that are governed by similar equations. Also, as hinted earlier, a myriad of applications exist in computer modeling, including but not limited to: Finite Difference Time Domain methods, Block's equations for Nuclear Magnetic Resonance Imaging, and plasma and semiconductor physics.

To check the third assumption, one would measure the location of free particle, with time since initialization, as a parameter, and show that the variance of an ensemble of free particles includes the term linear in t.

To check the primary result, one would perform a quantum non-demolition measurement on the quantum state of an ensemble of free particles. The minimum bit rate needed to describe the resulting string of numbers that describe the trajectory would be the entropy rate.

In the experiment, the time step should be as small as possible but needs not go to zero because a sensitivity analysis on how the time step affects the entropy rate is possible and thus a verification of this result would not require the time step to actually go to zero.

However even before an experiment needs to be conducted, this result is useful by suggesting the use of different information theoretical techniques to examine problems with de-coherence and might give a different perspective on the meaning of temperature.

---

[1] Cover, *Elements of Information Theory*, John Wiley & Sons, New York 1991, Ch 4

[2] Ibid, Ch 3

[3] Ibid, Ch 16

[4] Haller, "Information Mechanics", http://vixra.org/abs/0908.0097

[5] Bracewell, *The Fourier Transform and Its Applications*, 2nd ed. Mc Graw Hill, New York 1986, Ch 15

[6] Shankar, *Principles of Quantum* Mechanics, Plenum Press, New York 1994, Ch 5

[7] Halliday and Resnick, *Fundamentals of Physics*, 3rd ed. John Wiley & Sons, New York 1988, Ch 21

[8] Bohm, *Quantum Theory*¸ Dover Publications, Mineola, N.Y. 1989, Ch 3

[9] Ibid.

[10] Shankar, *Principles of Quantum Mechanics*, Plenum Press, New York 1994, Ch 5

[11] Bittencourt, *Fundamentals of Plasma Physics*, 2nd ed. Sao Jose dos Campos, SP 1995 Ch 10

[12] Einstein, "Investigation on the Theory of, The Brownian Movement", Translated by Cowper, Dover 1956

[13] R. Shankar, Principles of Quantum Mechanics 2nd ed, Plenum Press, New York, NY 1994 Ch 21

[14] Einstein, *The Meaning of Relativity*, 5th Edition, Princeton University Press, Princeton NJ 1956, Ch 1

[15] Stephen W. Hawking, *The Theory of Everything: The origin and fare of the universe, new edition*, Phoenix Books, Beverly Hills, CA 2007, Ch 5

[16] Haller, "Advances in Black Body Radiation", http://vixra.org/abs/0909.0024