Nature of Entropy

Francis T. S. Yu

Emeritus Evan Pugh (University) Professor of Electrical Engineering, Penn State University, University Park, PA 16802, USA

Email: fty1@psu.edu

Abstract

One of the most intriguing laws in modern physics must be the Boltzmann's entropy principle. As it stated, entropy within an isolated subspace increases with time, but it does not tell us why it has to be changing naturally with time. Since any evolution of our universe has a profound connection on her subspaces, from which we see that; there is a connection of entropy changes within an isolated (passive) subspace. Based on the big bang creation of our universe we see that; entropy within an isolated subspace increases naturally with time, since every subspace is a temporal (t > 0) subspace. In other words every subspace has a price tag; an amount of energy ΔE and a section of time Δt to create. Yet without an amount of information ΔI it is impossible to created. Although entropy can be traded with information, but it is the amount (e.g., in bits) but not the actual information that most physicists assumed. I have also shown that Shannon type communication is for reliable information transmission and it is not purposely to make the transmitted signal more ambiguous as quantum entanglement communication has proposed.

1 Boltzmann's Entropy

As accepted by most cosmologists our universe was created with a huge energy explosion (i.e., big bang) with time within a "non-empty" space. But empty space and temporal space are mutually exclusive, we see that every subspace no matter how small it is created by an amount of energy ΔE and a section of time Δt (i.e., ΔE, Δt) within our universe [1, 2]. From which we see that; every subspace is temporal (t > 0) (i.e., existed with time). From which we see that; every subspace or substance changes "naturally" with time.

In view of entropy universe, there are two simple and elegant laws which must be the Boltzmann's entropy principle [3] and Shannon's information theory [4]. As from their mathematical similarity it exists a profoundly relationship between them as given by.

$$S = - k \ln p \qquad (1)$$

$$I = - \log_2 p \qquad (2)$$

where S is the Boltzmann's entropy in joules per degree kelvin and I is the Shannon's amount information in bits, p represents a probabilistic distribution of particles within an isolated system or subspace, k is the Boltzmann constant, ln and $\log_2$ are the natural log and log of 2. From which we see that an amount of information can be simply traded with an amount of entropy as given by.

$$S \leftrightarrow I \qquad (3)$$

In other words, an amount of information in bits is equivalently equaled to an amount of entropy as given by.

$$S = k I \ln 2 \qquad (4)$$

Which is precisely the relationship what we were looking for the application of information in science. But we have frequently misused interpreted an amount of entropy as the actual information, since a quantity of information is "not" representing the actual information since it represents as a "cost" (e.g., in bits). For example, there are thousands of articles having "the same bits" of information but different articles. In fact, there are numerous physical substances

other than articles may have the same amount of information in bits. For instance copies of books may have the same bits but different books.

Nevertheless, within our temporal universe, every subspace is changing with time [3, 4], from which we see that; Boltzmann's entropy as well Shannon's information cannot be the exception. As Boltzmann's entropy stated, within an isolated subspace entropy increases with time or remains constant, but Eq. (1) does not show entropy changes with time as the principle did? The reason is that entropy increases naturally with time is due to time dependent universe by which her boundary expands constantly at the speed of light [3, 4].

Similarly to Eq. (2) we see that; the higher the amount of information is given by a source, the more uncertain its information is provided. Again we see that p is independent of time. However in principle every physical law is changing naturally with time (i.e., t > 0), strictly Eq. (1) and Eq. (2) should be written as time variable equations as given by,

$$S(t) = - k \ln p(t), \ t > 0 \tag{5}$$

$$I(t) = - \log_2 p(t), \ t > 0 \tag{6}$$

where t> 0 denotes equation exists if an only if within the positive time domain and k represents the Boltzmann's constant. We see that information or entropy "increases" (i.e., changes) with time. And an amount of entropy is equivalently equaled to an amount of I(t) in bits.

$$S(t) = k I(t) \ln 2, \ t > 0 \tag{7}$$

With reference to preceding equation we see that entropy increases as amount information increases. In other words "every bit" of information ΔI takes an amount of energy ΔE and a section of time Δt to "transmit" or to create such as given by,

$$\Delta I \sim \Delta E \, \Delta t = h \tag{8}$$

per bit of information. In view of Eq. (7) an amount of entropy ΔS per bit of information can be written as,

$$\Delta S = k \ln 2 \tag{9}$$

per bit of information. Since entropy ΔS is equaled to an amount of energy ΔQ divided by Kelvin temperature T as given by [5];

$$\Delta S = \Delta Q/T \tag{10}$$

every quantum state energy is approximately equaled to an amount of entropy,

$$\Delta S \sim \Delta E \, \Delta t/T = h/T, \text{ per bit of information} \tag{11}$$

where every ΔE is "coexisted" with Δt, $\Delta Q = \Delta E \, \Delta t$, T is the absolute thermal noise temperature in Kelvin (i.e., T = C + 273), and C is the temperature in degree Celsius.

In which we see that every bit of information is profoundly connected with the law of Heisenberg's uncertainty[6]. Since every subspace is created by ΔE and Δt, Boltzmann's entropy as well Shannon's information is also profoundly connected with uncertainty principle as given by,

$$\Delta E \, \Delta t \geq h \tag{12}$$

Or equivalently,

$$\Delta v \, \Delta t \geq 1 \tag{13}$$

where Δv is the bandwidth. Thus we see that every law, principle, theory, and paradox has a necessary cost of ΔE and Δt, otherwise law, principle, theory, and paradox cannot guarantee to be existed within our universe [7, 8].

Since increasing entropy is regarded as a "degradation" of energy by Kelvin, but entropy was actually originated by Clausius [9]. Yet Kelvin might have intended it to be used as a "negative" of entropy (i.e., neg-entropy). From which we see that; as entropy (or amount of information) increases meant that; there is an amount of "energy degradation" within an isolated subspace. Which is also meant that; an amount of entropy or amount of information "degrades with time" within an isolated (i.e., passive) subspace. Let me stress again; "energy degradation" naturally with time is due to boundary expansion of our universe at the speed of light [1, 2]. From which as I see it; entropy increases naturally with time is "no longer" a myth, as some scientists believed it is.

In view all the laws and principles are attached with a cost (i.e., ΔE, Δt), however it is the section of time that will be used can be squeezed to zero (i.e., $\Delta t \rightarrow 0$). But we can never be able to achieve it, even though we assume that we have all the energy of ΔE allows us to pay for it! That is the physical limit of temporal (t > 0) subspace. In other words the instantaneous moment at t = 0 can only be approached but can "never" be able to attend. This is precisely the law of entropy within an isolated (passive) subspace increases "naturally" with time.

2. Quantum limited Information

Every substance or subspace (no matter how small it is) has a piece of information. For example, the universe is filled with information (i.e., spatial, and temporal), or information is compacting the entire universe. Strictly speaking, if one is dealing with the origin of the universe, the aspect of information has never been absence. Then one would ask: What would be the amount of information, aside the necessary cost of ΔE and Δt, is required to create a specific subspace or substance? Equivalently what would be the equivalent amount of entropy to create it? This is a very viable question to be answer, let me start with the law of uncertainty as given by,

$$\Delta v \, \Delta t \geq 1 \qquad (14)$$

where Δv is the bandwidth. Preceding equation shows a profound relationship of an "information cell" is limited by $\Delta v \, \Delta t = 1$ [9], as illustrated in Figure 1. As we know that every bit of information can be effectively transmitted, if and only if it is limited with the limit of the uncertainty principle (i.e., $\Delta v \cdot \Delta t \geq 1$). This relationship implies that bandwidth of the signal should be either equal or smaller than bandwidth of the system (i.e., $1/\Delta t \leq \Delta v$). From which Δt and Δv can be "traded".
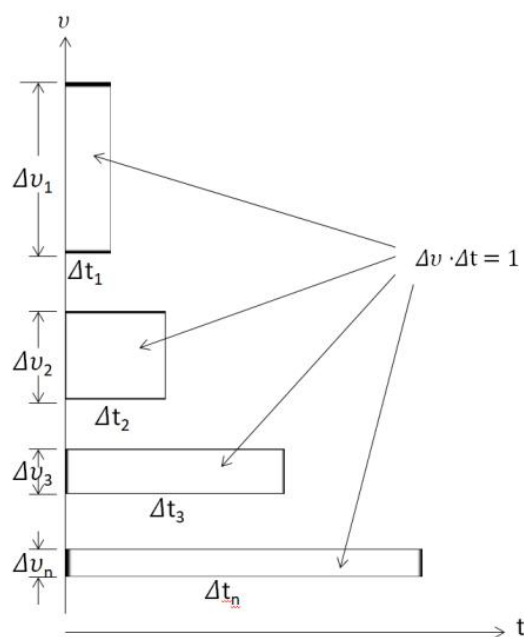

Figure 1. Shows various (Δv, Δt) sizes of information cells; where Δv_n and Δt_n are the bandwidths and time-limited sections, and $v_1 > v_2 > v_3 > ... > v_n$ are the center frequencies.

We stress that this is the unit region but not the shape of the information. From which we see that within each region, that is (Δv, Δt) [or equivalently (ΔE, Δt)] can be mutually traded. But it is from Δv to Δt or from ΔE to Δt, since Δv and ΔE are physical quantities. Therefore we see that; once a section of Δt is expensed, we cannot get back the same moment of Δt, although we can create the same section of Δt since time is a forward dependent variable.

From which we see that there are two types of information transmission orientations; one is limited by uncertainty Principle and the other is constrained within the certainty subspace. And the boundary between these two regimes is given by $\Delta v \cdot \Delta t = 1$ (or $\Delta E \cdot \Delta t = h$) which is known as Quantum Unit [10]. In which we see that under uncertainty regime, information is carried by means of intensity (i.e., amplitude square) variation. Yet, information can also be transmitted within the certainty regime, for instance such as applied to complex-amplitude communication [11, 10]. Since it is limited by the law of uncertainty, a quantum unit subspace (QLS) is depicted in Figure 2.
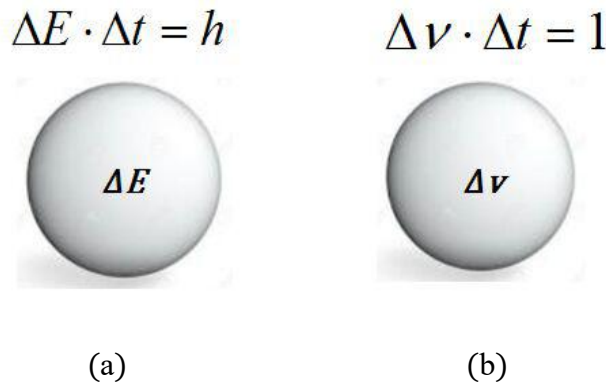
$$\Delta E \cdot \Delta t = h \qquad \Delta v \cdot \Delta t = 1$$

$$\Delta E \qquad \qquad \Delta v$$

(a) \qquad\qquad\qquad (b)

Figure 2. A set of Quantum Limited Subspaces (QLS)

As we know that every subspace within our universe is a temporal ($t > 0$) subspace, the radius of any temporal ($t > 0$) subspace can be described by a time dependent variable as given by.

$$r = c \cdot \Delta t \tag{15}$$

where c is the speed of light and Δt represents a section of time. In which we see that the size of the subspace enlarges rapidly as Δt increases as given by.

$$V = (\tfrac{3}{4}) \pi (c \Delta t)^3 \tag{16}$$

Since carrier bandwidth Δv and time resolution Δt are exchangeable, the size of the QLS enlarges as the carrier bandwidth Δv decreases. In which we see that narrower the carrier

bandwidth Δv has the advantage of having a larger quantum limited subspace for complex-amplitude communication as shown in Figure 3.
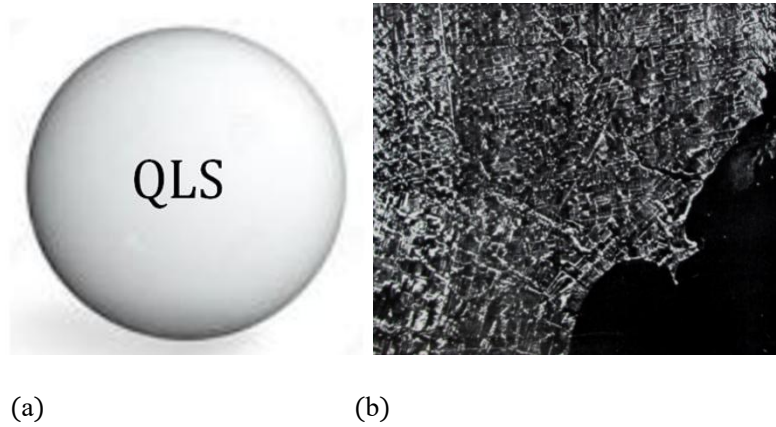

(a)   (b)

Figure 3. A "very large" quantum limited subspace as depicted in Figure 3(a) can be realized in practice within our temporal (t > 0) space: For example such as applied to synthetic aperture radar imaging shown in Figure 3(b).

From which we see that it is possible to create a temporal (t > 0) coherent subspace within a temporal (t > 0) space (i.e., our universe) for complex-amplitude communication. But it is "not" possible to create any time independent or timeless (t = 0) subspace within our temporal universe, since timeless (t = 0) subspace or time independent subspace cannot exist within our temporal (t > 0) universe. We note that; timeless (t = 0) space is a virtual subspace without time, it is an instantaneous limit (i.e., $\Delta t = 0$) of our temporal (t > 0) universe which is the well-known causality principle of our time-space [13]. From which we see that; response can only occur after the excitation within our temporal (t > 0) subspace. As we have shown in the preceding the limit of $\Delta t \rightarrow 0$ within a temporal subspace can only be approaching at the expense of a huge amount of energy ΔE, but we will never be able to reach it!

Since timeless (t = 0) space cannot exist within temporal space, which is known as "temporal space exclusive principle", I note that timeless (t = 0) or time independent subspace is not an inaccessible subspace as some scientists claimed, since inaccessibility implies it is existed within our temporal (t > 0) universe. One of the apparent reasons for the using a large quantum limited subspace for complex information transmission must be for the application to complex wave front transmission (e.g., holography) [12], complex-match filter synthesis [14] and as well apply to synthetic aperture radar imaging [11]. But there is a price to pay, since it requires a longer section of Δt .

3. Reliable Information Transmission

One of the aspects in communication is that; reliable information can be transmitted at high degree of certainty such that information can be properly receive. In which there are two key equations from mutual communication perspective are: mutual information transmission through a passive additive noise channel as given by [5],

$$I(A;B) = H(A) - H(A/B) \tag{17}$$

and

$$I(A;B) = H(B) - H(B/A) \tag{18}$$

where $H(A)$ represents the entropy or information provided by the sender, $H(A/B)$ is the information loss (or equivocation) through communication channel due to noise, $H(B)$ represents the amount of information received at the receiving end and $H(B/A)$ is noise entropy of the transmitting channel.

Yet, there is a distinction between these two equations; one is for "reliable" information transmission and the other is for "retrieving" the information that has been send. In other words an information will be reliably convey to the receiver and the other is to retrieve an information that has been already received by the receiver. Although both equations represent the mutual communication between sender and receiver; but their objectives are rather different. Example: Eq. (17) is developed for reliable information transmission for which transmitted information has a high degree of certainty to reach the receiver. While Eq. (18) is purposely used to retrieve information from a noisy received information. For which we see that; for reliable information transmission, one can simply increase the signal energy at the transmitting end such that is has a higher signal to noise ratio at the receiving end. While for unreliable information transmission is to extract information from an ambiguous corrupted received information. In other words, one is to be sure that information will be receiving by the receiver before transmission is started and the other is to retrieve the information after information has been received by the receiver.

In principle we have two kinds of information transmission orientations; one is information retrieval orientation developed by Norbert Wiener's [15,16] and the other is reliable transmission orientation developed by Claude Shannon [4]. However there is a major distinction between them; Wiener's communication strategy is that; if the information is corrupted through transmission, it may be recovered at the receiving end, but with a "heavy

price" to pay by post processing. Yet Shannon's information-transmission carries a step further by encoding the information before it is transmitted such that, information can be "reliably" received, but also with a "price" which is mostly at the transmitting end. In view of the Wiener and Shannon orientations; mutual information transfer of Eq. (17) is kind of Shannon type, while Eq. (18) is kind of a Wiener type. From which we see that; reliable information transmission is basically manipulated by the sender to minimize the noise entropy H(A/B) (or equivocation) of the channel, as shown by.

$$I(A;B) \approx H(A) \tag{19}$$

Nevertheless a simple way to achieve is by increasing the signal energy (i.e., ΔE) to noise ratio. And other is by redundancy coding.

But as for Wiener type information retrieval it is to recover the transmitted information after it was corrupted. Which is to maximize H(B/A) (i.e., the channel noise). Since the entropy H(B) at the receiving end is larger than the entropy at the sending end [i.e., H(B) > H(A)] we have,

$$I(A; B) = H(B) - H(B/A) \approx H(A) \tag{20}$$

From which we see that; information can be recovered after being received, again has a price; ΔE and Δt. In view of these strategies, we see that the price paid for using Weiner type for communication would be much higher than the Shannon type. For example aside the price paid for higher amount of energy ΔE it required a much longer section of time Δt to pay for post processing. Thus we see that Wiener communication strategy has an advantage for a none-cooperative sender; for examples such as applied to radar detection, and others. One the other hand, Shannon type provides a more reliable information transmission strategy, by simply increasing the signal to noise such that every bit of information can be "reliably transmitted" to the receiver.

In view of quantum entanglement communication [17], we see that it is basically using Wiener communication strategy. We see that the price pay by the receiver will be much higher and also much more costly. For instance post processing takes more energy (i.e., $\Delta E0$ and it requires much longer time (i.e., Δt) to post processing that deviates further away from real time to receiving the message. From which we see that; quantum entanglement information transmission requires that received signal [i.e., H(B)] be "more equivocal" (i.e., uncertain). In other words quantum entanglement communication has treated H(B) as an information source instead of corrupted information entropy at the receiving end. Since larger entropy H(B) at

the receiving end as compared with the information source H(A) [i.e., H(B) ≥ H(A)], it does not mean that increases H(B) entropy at the receiving end improves the information source entropy H(A) at the transmitting end. For example entropy for binary digital source is H(A) = 1 bit per unit Δt, it is wrong to treat the received end entropy H(B) as an information source since H(B) > H(A). However if the binary digital source (e.g., 0 and 1) can simultaneously and instantaneously be transmitted (i.e., a qubit or quantum bit) then H(B) can be interpreted as the information source, but H(A) would be a qubit source. But as I had have shown that the simultaneously and instantaneously superposition principle only existed within an empty timeless (t = 0) space but not exist within our temporal (t > 0) universe [8, 18]. Nevertheless this is precisely a wrong strategy for reliable information transmission as Shannon. In other words by treating H(B) as an information source at the receiver it does not represent as the information source of H(A). From which as I see it, quantum entanglement communication is basically designed for extracting information as Weiner type communication. Anyway it is not the purpose of reliable information-transmission of Shannon. Nevertheless quantum entanglement information-transmission dependents on the principle of superposition principle, but superposition principle does "not" exist within our universe since empty space. Which is a mathematical virtual subspace that is not a physically realizable subspace since empty and temporal spaces are mutually excluded.

4. Conclusion

It is reasonable to assume that our universe was created from a big bang explosion, since our universe has been observed as an expanding universe. This has been an enticing reason for astrophysicists to assume that our universe was created within an "absolute empty" space about 14 billion light years ago. But from physical realizable standpoint, it is not possible that for our universe was created within an empty space since temporal (t >0) and timeless (t = 0) are mutually exclusive. Of which we see that; every subspace within our universe, that is no matter how small it is, is created by an amount of energy ΔE and a section of time Δt (i.e., ΔE, Δt). From which we see that every subspace within our universe is a temporal (t > 0) subspace, otherwise it cannot exist within our time-space. From which we see that entropy of an isolated (passive) subspace is constantly changing naturally with time. Since both Shannon information measure and Boltzmann entropy expression were described in logarithmic forms, entropy and information can be simply traded. By which we see that trading an amount of

information equals to an amount of entropy is possible. Nevertheless the cost of information is not representing the actual information. For instance, a 100,000-bit book has 2^{100000} possible volumes of books which is also not included all other books in different languages. Nevertheless, the essence of this article is to show that entropy within an isolated subspace is changing naturally with time, which is primarily due to the constantly dynamic expansion of our universe. Finally, I have also shown that Shannon type communication is primarily for "reliable" information transmission so that higher degree of certainty information can be received at the transmitted end. It is "not" by means of making the received signal more ambiguous as quantum entanglement communication has proposed.

References

[1] F.T.S. Yu, "Time: The Enigma of Space", Asian Journal of Physics, Vol. 26, No.3, 143-158, 2017.

[2] F.T.S. Yu, "From Relativity to Discovery of Temporal (t > 0) Universe", Origin of Temporal (t > 0) Universe: Correcting with Relativity, Entropy,

[3] Ludwig Boltzmann, "Über die Mechanische Bedeutung des Zweiten Hauptsatzes der Wärmetheorie". Wiener Berichte. 53: 195–220, (1866).

[4] C. E. Shannon and W. Weaver, The Mathematical Theory of Communication, University of Illinois Press, Urbana, IL, 1949.5. F. T. S. Yu, Optics and Information Theory, Wiley-Interscience, New York, 1976. Information Theory, Wiley

[5] F. T. S. Yu, Optics and Information Theory, Wiley-Interscience, New York, 1976. Information Theory, Wiley

[6] W. Heisenberg, "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik," Zeitschrift Für Physik, vol. 43, no. 3–4, 172 (1927).

[7] F. T. S. Yu, "Nature of Temporal (t > 0) Quantum Theory: I" Quantum Mechanics, Edited by P. Bracken, Chapter 8, p 139-160, IntechOpen, London (2020).

[8] F. T. S. Yu, "Nature of Temporal (t > 0) Quantum Theory: II" Quantum Mechanics, Edited by P. Bracken, Chapter 9, p 161-188, IntechOpen, London (2020).

[9] D. Gabor, "Communication Theory and Physics," Phil. Mag, vol. 41, no. 7, 1161, 1950.

[10] F. T. S. Yu, "Information Transmission with Quantum Limited Subspace", Asian Journal of Physics, Vol. 27 , no 1, 1-12 (2018).

[11] L. J. Cultrona, E. N. Leith, L. J. Porcello and W. E. Vivian, "On the application of Coherent Optical Processing Techniques to Synthetic-Aperture Radar," Proc. IEEE, 54, 1026 (1966).

[12] E. N. Leith and J. Upatniecks, "Reconstructed Wavefront and Communication Theory", J. Opt. Soc. Am., 52 (1962)1123.

[13] M. Bunge, Causality: the place of the causal principle in modern science. Harvard University

[14] F.T.S. Yu, Introduction to Diffraction, Information Processing and Holography, Chapter 10, MIT Press, Cambridge, Mass., 91-98 (1973).

[15] N. Wiener, Cybernetics, MIT Press, Cambridge, MA, 1948.

[16] N. Wiener, Extrapolation, Interpolation, and Smoothing of Stationary Time Series, MIT Press, Cambridge, MA, 1949.

[17] K. Życzkowski,, P. Horodecki, M. Horodecki, and R. Horodecki, "Dynamics of quantum entanglement". Phys. Rev. A 65, 1-10, 2001.

[18] F. T. S. Yu, The Fate of Schrodinger's Cat, Asian J. Phys., vol. 28, no 1, 63-70 (2019).