Prepared for submission to the FQXi essay contest

Does God play dice with time itself?

by Del Rajan

Contents

1 Introduction

"*You believe in the God who plays dice and I in complete law and order in a world which objectively exists and which I, in a wildly speculative way, am trying to capture*," wrote Albert Einstein in one of his letters to Max Born [1]. It conveyed Einstein's dissatisfaction with the intrinsic randomness in the physical world that had been discovered through quantum physics. In fact it was Born himself who first enunciated the probability interpretation of the wave function. Einstein's words also articulated the biases a physicist can bring upon the work they do. Such a psychological state is not constrained to merely physics alone. David Hilbert wanted to acquire an analogous 'complete law and order' in the mathematical world through formal axiomatic systems. In rather a shocking manner, Kurt Gödel showed this was impossible through his incompleteness theorems.

The pertinent question is that given that the laws of physics are codified as mathematical statements, would the incompleteness of formal axiomatic systems be a further hindrance to the pursuit of 'capturing' the physical world? In this essay we argue to the contrary. Our thesis is that aspects on the work on incompleteness could be harnessed as primary mathematical tools for deeper discovery in fundamental physics.

2 Nobody 'understands' quantum physics.

"*I think I can safely say that nobody understands quantum mechanics*," said Richard Feynman in one of his captivating Messenger lectures. It is worth emphasizing that the computations of quantum-theoretic quantities for various systems is certainly well understood. What he meant by understanding is a physical explanation of the quantum effects and the equations of quantum theory in a manner that was employed in other fields of physics. Take the case of special relativity where the time dilation effect

$$\Delta t = \gamma \Delta \tau, \tag{2.1}$$

(where γ is the gamma factor) can be derived from the temporal parts of the Lorentz transformations. Initially these equations were not understood. Hendrik Lorentz, Joseph Larmor and others interpreted these various formulae due to effects involving the ether. It was only through Einstein that we understood the phenomena as a manifestation of an underlying physical principle of the Universe. He postulated that the speed of light has the same value in all inertial frames of reference. The unavoidable consequence of this understanding was a correct (or at least deep) physical interpretation of the Lorentz transformations, and the discovery that time is not absolute.

Reverting to the confusion around quantum theory, the postulates say that to describe the state of a physical system one needs to assign a complex number, known as the amplitude, for each possible configuration that one would find the system in upon measurement. As an example, the system can be in state

$$|\psi\rangle = \sum_{i=1}^{n} c_i |u_i\rangle = c_1 |u_1\rangle + c_2 |u_2\rangle + \cdots + c_n |u_n\rangle, \tag{2.2}$$

where c_i represents the complex amplitudes. Furthermore the theory states that the probability a measurement finds the system in a particular configuration $|u_i\rangle$ is given by

$$|c_i|^2, \tag{2.3}$$

known as the Born rule. This lack of a direct correspondence between the physical system and the equations gives rise to a pseudo-explanation that the system prior to observation is in multiple configurations at the same time. There is no genuine evidence to support such a view. A more sober approach is to directly face the inceptions of at least two separate problems:

- Physical problem: What is the configuration of the system prior to measurement?
- Theoretical problem: What do the complex amplitudes represent?

Stating the physical problem more concretely, before the measurement is the mass of a particle localized or distributed across space? The significance of the theoretical problem becomes apparent in the regime where quantum algorithms [2] exponentially outperform their classical counterparts, through the interference of such complex numbers.

Historically, questions such as the above were simply denied any attention from most physicists as prescribed to them by the Copenhagen interpretation. Despite having an aura of a submissive doctrine, there is in fact no consensus on what the interpretation actually states. The overarching theme of it is that a deeper description beyond quantum theory is not needed. The technological aims associated with the Second World War brought upon a refined version of this attitude known as the 'radar philosophy' [3] or more famously phrased as 'shut up and calculate'. It involved adopting a pragmatic approach of matching the outputs of quantum theory with experimental outputs, without a physical understanding. In the latter half of the 20th century this resulted in an extraordinary amount of research productivity while paradoxically providing little aid in answering the most basic of questions: What a superposition physically represents? Crudely stated, progress was measured in every parameter except understanding.

It is our view that an understanding of quantum physics is necessary for genuine progress. There are two reasons. The first is that the depth to which we comprehend the Universe we inhabit is a reflection of our species' intelligence. To support such a claim, let us direct our attention to natural language processing where computers can perform machine translation (translate text from one language to another). Modern machines can provide a striking degree of accuracy in this task using mathematical models. However they are not regarded as "intelligent" given they do not understand the meaning in a body of text. This task of reading comprehension is part of the artificial intelligence aim known as natural language understanding. Given we have such stringent standards for machines to be regarded as intelligent, it would only be fair that humans adopt an analogous measure. Mathematically modelling the physical world without deep understanding can be compared to machine translation without comprehension. Only with the addition of understanding in our physics can we regard ourselves as an advancing intelligent species.

The second and far more pragmatic reason is that a deeper understanding leads to finer mathematical models to calculate with. In the case of relativity, Einstein's physical principle paved the way for Hermann Minkowski's conceptualization that space and time are one unified object. Within that picture held the seed to Einstein's greatest calculational model; he advanced the spacetime picture by suggesting that spacetime is curved. Utilizing that understanding, he constructed the appropriate differential geometric structure that computed the perihelion effect of Mercury and ultimately quantified the large-scale Universe. Stated in another way, the "search space" of all possible mathematical objects is far too large, and only by a physical understanding can one find their way from the "local neighborhood" of the mathematics of Newton's gravity to theory of manifolds. Confining this reasoning to the task of deriving quantum gravity means we must understand quantum physics. Stated more dramatically, countless minds of the 20th century were lost through 'shut up and calculate' and the price we pay for that today is a lack of a sufficient theory of quantum gravity.

3 Forget mathematical beauty. Think mathematical randomness.

To make progress in genuinely understanding quantum physics, it is wise to first heed the words of the person who arguably understood it the most. Paul Dirac distilled his research guidance through a lecture on the relationship between mathematics and physics [4]. He conveyed how it was advantageous for the 19th century physicist to assume that the Universe had a characteristic that its equations of motion must be of simple form. Such a bias was to be viewed as an instrument of research which Dirac called the 'principle of simplicity.' Observing the physics of the early 20th century gave him the conviction that the Universe had an inbuilt mathematical beauty. He proceeded to describe the tenet of the lecture in that the researcher should and we quote "*...change the principle of simplicity into a principle of mathematical beauty.*" He assisted by providing a practical step and we quote, "*The method is to begin by choosing that branch of mathematics which one thinks will form the basis of the new theory. One should be influenced very much in this choice by considerations of mathematical beauty.*" He alluded to a direction and we quote "*I would suggest,* *as a more hopeful-looking idea for getting an improved quantum theory, that one take as basis the theory of functions of a complex variable. This branch of mathematics is of exceptional beauty...*"

The wide adoption of the 'principle of mathematical beauty' can be witnessed in the geometrization of physics in the latter half of the 20th century. Complex variables did indeed play a central role. The subject of complex manifolds, which represents the synthesis of complex variables with differential geometry, infused various parts of fundamental physics that aimed towards quantum gravity. We highlight two of various examples at a coarse level.

The first direction built on the theory of relativity. Geometric reformulations of general relativity were introduced through techniques such as null tetrads and two-component spinors [5]. Based on this, Roger Penrose and his Oxford group constructed twistor theory [6] where complex manifolds such as a (projective) twistor space became the abstract arena of interest. Despite the immense beauty of the utilization of geometry, quantum gravity was not achieved. Meanwhile another direction stemmed from the quantum theory of fields or more precisely non-Abelian gauge theories. These form the frameworks for modern particle physics. A geometrization of this [7] involved vector bundles, connections and other related techniques. Developing on this style, trained particle physicists most notably Edward Witten led the development of string theory with the intention to describe the regime of quantum gravity. Geometrically the theory also harnessed a subset of complex manifolds known as the Calabi-Yau manifolds [8]. As mesmerizing as this path was, over time it did not lead to a sufficient theory of quantum gravity.

Perhaps the time is ripe for a new principle. Through the innovator's spirit that thrives on creative destruction, we make the assertion to abandon the 'principle of mathematical beauty.' Analogous to Dirac, we make observations from relevant physics to suggest an alternative instrument of research. One with an intention for a physical understanding of quantum phenomena.

The first observation is that despite various efforts it has not been possible to derive the probabilistic Born rule (2.3) from the non-probabilistic postulates of quantum theory. It is an erroneous to say that Gleason's theorem provides such a derivation [9]. Rather the correct view is that if one is given the non-probabilistic structures of quantum theory and one also assumes that the theory has a probabilistic character, then Gleason's theorem states that such a character must be expressed in no other way than the Born rule. Note the assumption of a probability. The lack of a genuine derivation can be a hint that the randomness is fundamental.

The second observation stems from the work of John Bell [10]. By invoking the assumptions of locality, realism and free choice he derived his famous inequality. Certain entangled systems violate this inequality. This implies a deep understanding of our Universe: One or more of the assumptions must be false! The consensus is that the assumption of locality is incorrect. However, these is no mathematical or experimental evidence for the decision to forgo that particular assumption. It may very well be the case that one should abolish free will, or that our concept of physical realism needs to be radically altered. Perhaps even more remarkable is that this theorem can be derived with elementary use of classical probability theory, and without reference to quantum theory. Bell's theorem suggests that to understand the Universe, with its quantum physics, it may be favorable to use mathematical tools that are designed to study randomness as opposed to utilizing the quantum theory.

Using these observations, we make the hypothesis that it is advantageous for the 21st century physicist to assume that at a deeper level the Universe is not elegantly beautiful, but rather that it is physically random in a shockingly monstrous way. We have only seen hints of such a structure through quantum phenomena and have far more to uncover. Rather than view intrinsic randomness as a sin like Einstein, or worse deny the physical world altogether like the disciples of the 'shut up and calculate' method, we want to embrace the increasing physical depths of sophisticated uncertainty. This bias provides an instrument of research which we call the 'principle of mathematical randomness.' A practical step, analogous to Dirac's step, is to look at an area of mathematics where randomness is captured in a rigorous manner. Rather elegantly, this is found in the unpalatable work on the incompleteness of formal axiomatic systems.

4 Incompleteness: Undecidability, Uncomputability, and Incompressibility.

"*I have shown that God not only plays dice in physics, but even in pure mathematics!*" describes Gregory Chaitin [11] with respect to his contribution towards a modern understanding of incompleteness. However, the historical inception of the work on incompleteness can be attributed to issues regarding set theory. The theory allowed for the possibility of a contradictory set, namely the set of all sets that are not members of themselves. Hilbert assumed that such an anomalous construction reflected the lack of a single formal axiomatic system for all of mathematics. This meant that all valid mathematical proofs should in principle be stated in a single agreed system that had some alphabet, grammar, axioms, rules of interference, and a proof-checking algorithm that was mechanical. The aim was to preclude all contradictions from arising, and for achieving absolute certainty in the domain of pure mathematics.

Such a vision was violently shattered by Gödel with his undecidability results [12, 13]. He showed that there are undecidable statements in pure mathematics, which are statements that are true but cannot proved to be true. Moreover any attempt to mend the system by adding extra axioms only leads to further assertions that are true but unprovable. Stated more precisely, Gödel's incompleteness theorem showed that every formal axiomatic system which is finitely specified, strong enough to include arithmetic, and consistent, is incomplete. One such example is Zermelo-Fraenkel set theory with the Axiom of Choice. The condition relating to arithmetic is crucial since the theories that do not have this property can be complete. One notable example is plane Euclidean geometry which is complete.

The situation can be seen in a concrete manner as we move our gaze from the mathematical world to the computational world. One can deduce incompleteness from uncomputability which was discovered by Alan Turing. He invented a theoretical model of a machine that in principle could compute anything that was computable.

For incompleteness, the most relevant aspect of Turing's work is the halting problem. This is an instantiation of the broad question regarding the computational world on what things can or cannot be computed. The problem involved computing whether or not a computer program would eventually halt, without actually running it. It is surprising that such a simply stated task cannot be solved by a Turing machine. The problem is uncomputable. A rough derivation of incompleteness from the halting problem is as follows: Assume a formal axiomatic system that can always prove whether or not individual programs halt. Then by systematically going through all possible proofs in size order one can find a proof that the particular program one is examining will halt or one will find a proof that it will never halt. However Turing's work on the halting problem showed that this was not possible. Hence the formal system must be incomplete.

Despite being arguably the greatest results in logic, the relevance of Godel's incompleteness theorems on formalism was later largely ignored by the mathematical community. Turing's work on uncomputability did not change this view to any significant degree. From such observations on historical matters, it is not unfair to say that a psychological state resembling 'shut up and calculate' hovered over pure mathematics. Hence it is ironic that it is randomness, that so distinctly characterizes quantum physics, is also the concept that makes incompleteness emerge as a pervasive phenomenon throughout pure mathematics.

To see this we look into the field of algorithmic information theory [14] which was independently developed by Chaitin as well as Ray Solomonoff and Andrey Kolmogorov. It is based on a notion of the information content of an individual object using concepts from computability theory. This is in contrast to classical information theory which is built on probability distributions. To elaborate, the information content of an object such as a finite binary string x can be captured by the algorithmic entropy $K(x)$ (also known as the Kolmogorov complexity) which is defined as the length of the shortest binary computer program that can produce x as the output. In other words, the algorithmic entropy of a string provides the length of its shortest possible compression. As an example, the string

$$01010101010101010101010101010101010101010101010101010101010101 \tag{4.1}$$

has the compressed program description "31 repetitions of 01". This is opposed to the string

$$11001000011000011101111011101100111110100100001001010111100101 \tag{4.2}$$

which has no shorter description than writing down the string itself.

Using the algorithmic entropy, one can arrive at a definition of randomness of a single individual sequence which is not possible using probabilistic theories. The fundamental idea is that randomness equates to incompressibility. The inability to compress highlights that there is no structure of pattern to provide a concise description (as witnessed in string (4.2)). Stated more formally, a finite string x is random if and only if the algorithmic entropy of the string $K(x)$ is no shorter than the length of the string. The extension of this definition to infinite strings, and other equivalent formulations has been widely explored [15].

By harnessing this definition of randomness, Chaitin derived incompleteness in a variety of ways. One path involved utilizing a quantity known as the Omega number. This number represents the probability that a randomly chosen computer program will eventually halt. It turns out to be an infinitely incompressible number, and hence truly random. In this context, the incompleteness theorem is based on the inability of a formal system to compute the bits of the Omega number. The main message of these information-theoretic derivations of incompleteness is that the mathematical world holds an infinite amount of information while any given set of axioms only captures a finite subset of this information. Hence incompleteness coveys itself as a natural and pervasive phenomenon in mathematics.

5 Forget information. Think compression.

We have seen that randomness is quantified through compression. Chaitin has mooted [16] that compression can be regarded as a universal concept. He provides various examples outside of mathematics, including that a scientific theory itself can be thought of as a compression of experimental data. This universality of compression aligns with our 'principle of mathematical randomness.' Specifically, we want to harness compression as a primary mathematical technique for the aim to make deeper discoveries of the physical world.

One can also arrive at the previous statement through an alternative direction, namely the dogma that information is fundamental. Besides algorithmic information theory, compression exists at a foundational level in the information theories associated to physical systems.

In classical information theory, the fundamental result is the noiseless coding theorem. It highlights that the Shannon entropy $H(X)$ can be operationally defined in terms of optimal compression. More precisely, for a sequence of n two-outcome random variables, uncompressed n bits can be optimally compressed to $H(X)n$ bits:

$$H(X)n_{uncompressed} = n_{compressed}. \tag{5.1}$$

Quantum information theory provides a reconceptualization of quantum physics in terms of information. The fundamental result in this theory is the Schumacher's noiseless coding theorem. This articulates that the von Neumann entropy $S(\rho)$ can be operationally defined in terms of optimal compression. More specifically, uncompressed n qubits can be optimally compressed to $S(\rho)n$ qubits:

$$S(\rho)n_{uncompressed} = n_{compressed}. \tag{5.2}$$

The latter theory has added support to John Wheeler's phrase 'It from Bit' [17]. This puts forth the notion that the physical world emerges from information. A 'chicken or the egg' paradox arises when ones combines this with Rolf Landauer's assertion [18] that 'Information is Physical.' Stated more directly, what is the physical substrate that asks the yes-no question which forms the Bit that creates the It? The answer is unclear.

These ideas are muddled further when one probes what the information in each of the information theories is supposed to represent; classical information is fit for engineering purposes with no regard to meaning [19]; nobody knows what quantum information is [20]; a similar ambiguity exists on what algorithmic information is about [14]. Given each of the information theories has an associated entropy, perhaps that may add some clarification. Such a hope is quickly diminished when one reads that Shannon called his term the entropy because von Neumann suggested one reason as "*...nobody knows what entropy really is, so in a debate you will always have the advantage.*"

Our view to resolve the confusion is remove the focus on information in information theories. Rather it is compression in these theories that is fundamental, and whose character we imagine will be found in the deeper laws of the Universe. Building on these established theories, we state that any physical structure that is compressible is what should constitute information, and the quantity involved in optimal compression is what should be termed the entropy.

6 Spacetime compression and 'understanding' quantum physics.

"*Not only does God definitely play dice, but He sometimes confuses us by throwing them where they can't be seen*," is a quote assigned to Stephen Hawking regarding the added unpredictability when one introduces spacetime structures to quantum physics. This alludes us to the question of whether spacetime is intrinsically random? Hence could one employ the tool of compression and develop a 'spacetime-information' theory?

To explore this, we provide a speculative path. We emphasize that these ideas are under-developed but proceed with the intention of conveying possibilities. Let us go back to the time dilation formula (2.1). There has been work [21] arguing that time dilation (as well as length contraction) is more fundamental than the Minkowski geometry. Utilizing these various influences and given our aim for a 'spacetime-information' theory, we postulate that time dilation (2.1) is a de-compression. Stated mathematically, we rewrite the time dilation equation (2.1) as an information-theoretic compression formula

$$\alpha \Delta t_{uncompressed} = \Delta t_{compressed} \tag{6.1}$$

where $\alpha \equiv 1/\gamma$. This is analogous to compression formulas (5.1) and (5.2). Using our compression tool, this implies that a time interval can be treated as information. Given that the reciprocal of the Lorentz factor optimally compresses a time interval, it can be viewed as an entropy. In fact, this α-entropy shares many similarities to the Shannon and the von Neumann entropy.

To draw out physical implications, we assume that this information-theoretic compression has the probabilistic backbone of classical information theory. (Quantum information theory with its compression formula (5.2) was derived in a similar manner.) Such a mathematical framework assumes an information source that is intrinsically random, and outputs two types of sequences (or states in the quantum case). These are known as typical and atypical. Typical sequences are those that can be optimally compressed and obey (5.1). Atypical sequences cannot be compressed. The crucial result is that typical sequences are overwhelmingly likely to occur in the asymptotic limit and therefore compression succeeds with probability approaching one.

Harnessing this framework for our time compression formula (6.1) suggests that a time interval emerges in an intrinsically random manner. It can be classified as a typical or atypical time interval. Typical time intervals are those that can be compressed. This means they obey relativistic time dilation (6.1) and are overwhelmingly likely to occur for sufficiently large time intervals. Far more profound are atypical time intervals. These are ones that exhibit Lorentz violations as they cannot be compressed. They do occur but very rarely. This means that a Lorentz violation is not about scale, but about probability.

Can a generalization of spacetime compression help us understand quantum physics? The physical problem of quantum physics, as described in a previous section, becomes mysterious through various scenarios. One example being quantum jumps and another example being the case where one can measure a particle and after some time, it is measured again but found in another location that could not classically occur given the duration.

Compression techniques can involve a reduction of data or even data reordering. Perhaps at the quantum scale, a subset of events are deleted or unordered due to a compression of time. When one measures the particle at various times, it gives the impression that the particle is behaving in a shocking manner. What is actually occurring is that the observer is seeing the output of a temporally unordered classical trajectory. A measurement can be defined as the moment at which the temporal ordered classical world meets the temporal unordered quantum world. The concept of unordered time has been rigorously analyzed in other settings [22].

The theoretical problem of quantum physics involves complex numbers. Perhaps the amplitudes relate to the unordered time given that complex numbers are mathematically unordered. Certainly the complexification of time is not well understood in established subjects such as the Wick rotation and the Newman-Janis trick [5].

The psychological state that thirsts for physical understanding wonders whether these speculations could placed on firm footing using a formal and more sophisticated theory of spacetime compression. Would the expansion of the Universe turn out to be an information-theoretic decompression? Would the Bekenstein-Hawking entropy correspond to an optimal compression? Would we ever reach the promised land of quantum gravity?

7 Conclusion

Newton thought that time was absolute. Einstein showed that it was relative. In this essay we speculate that it comes in two forms: typical and atypical. The emergence of time in this intrinsic random manner suggests that God not only plays dice but plays dice all the time and with time itself. This activity may occur in the same ontological place as the answer to the question 'What was God doing before He made heaven and earth?' St Augustine in his Confessions recounts an answer someone gave as 'He was preparing hell for those who pry into mysteries' which can be interpreted as 'shut up and calculate!'

References

[1] Einstein, A., Born, M. *"Born-Einstein Letters, 1916-1955: Friendship, Politics and Physics in Uncertain Times,"* Palgrave Macmillan (2004).

[2] Chuang, I., Nielsen, M. *"Quantum Computation and Quantum Information,"* Cambridge University Press (2000).

[3] Kaiser, D. *"History: Shut up and calculate!"* Nature 505, 153–155 (2014).

[4] Dirac, P.A.M. *"The Relation between Mathematics and Physics,"* Proceedings of the Royal Society of Edinburgh 59, 122–129 (1940).

[5] Flaherty, E. J. *"Hermitian and Kählerian geometry in Relativity,"* Springer (1976).

[6] Huggett, S. A., Tod, K. P. *"An Introduction to Twistor Theory,"* Cambridge University Press (1994).

[7] Atiyah, M., Dijkgraaf, R., Hitchin, N., *"Geometry and Physics,"* Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 368, 913–926 (2010).

[8] Green, M. B., Schwarz, J. H., Witten, E. *"Superstring theory: Volume 2, Loop amplitudes, Anomalies and Phenomenology,"* Cambridge University Press (2012).

[9] Gleason, A. M. *"Measures on the closed subspaces of a Hilbert space,"* Journal of Mathematics and Mechanics 6, 885–893 (1957).

[10] Brunner, N., Cavalcanti, D., Pironio, S., Scarani, V., Wehner, S. *"Bell Nonlocality,"* Reviews of Modern Physics 86, 419 (2014).

[11] Calude, C. S., Chaitin, G. J. *"Randomness and Complexity: from Leibniz to Chaitin,"* World Scientific (2007).

[12] Calude, C. S., Stay, M. A. *"From Heisenberg to Gödel via Chaitin,"* International Journal of Theoretical Physics 46, 2013–2025 (2007).

[13] Calude, C. S. *"Algorithmic Randomness, Quantum Physics, and Incompleteness,"* International Conference on Machines, Computations, and Universality, 1–17 (2004).

[14] Grünwald, P. D., Vitányi, P.M.B. *"Algorithmic Information Theory,"* Handbook of the Philosophy of Information, 281–320 (2008).

[15] Downey, R. G., Hirschfeldt, D. R. *"Algorithmic Randomness and Complexity,"* Springer (2010).

[16] Chaitin, G. J. *"Meta Math!: The Quest for Omega,"* Vintage (2008).

[17] Wheeler, J. A. *"Information, Physics, Quantum: The search for links,"* Complexity, entropy, and the physics of information 8, (1990).

[18] Landauer, R. *"Information Is Physical,"* Physics Today 44, 23–29 (1991).

[19] Shannon, C. E. *"A Mathematical Theory of Communication,"* Bell System Technical Journal 27, 379–423 (1948).

[20] Leifer, M.S. *"Is the Quantum State Real? An Extended Review of Ψ-Ontology Theorems,"* Quanta 3, 67-155 (2014).

[21] Brown, H. *"Physical relativity: Space-time structure from a Dynamical perspective,"* Clarendon Press (2006).

[22] Zych, M., Costa, F., Pikovski, I., Brukner, Č., *"Bell's theorem for temporal order,"* Nature Communications 10, 1–10 (2019).

A Technical endnotes

In this section, we provide a mathematical sketch behind the development of this time compression (6.1). It follows the structure of classical and quantum information theory. We emphasize that these speculations are underdeveloped.

Classical compression

In classical information theory, the Shannon entropy of a discrete random variable X is:

$$H(X) \equiv H(p_1, \dots, p_m) \equiv -\sum_x p_x \log_2 p_x. \tag{A.1}$$

It is conventionally agreed that $0 \log 0 \equiv 0$ in the use of the above definition since $\lim_{x\to 0} x \log x = 0$. The entropy has bounds, $0 \leq H(X) \leq \log_2 m$. For the simple case of a two-outcome random variable, this results in bounds, $0 \leq H(X) \leq 1$. Shannon's noiseless channel coding theorem states: Suppose $\{X_i\}$ is an independent and identically distributed information source with entropy rate $H(X)$. Suppose $R > H(X)$. Then there exists a reliable compression scheme of rate R for the source. Conversely, if $R < H(X)$ then any compression will not be reliable. Hence for a sequence of n random variables, a trivially uncompressed representation uses $n \log_2 m$ bits. According to the theorem, the optimal compression without loss of information uses $nH(X)$ bits. For the case of two-outcome random variables, uncompressed n bits can be optimally compressed to $nH(X)$ bits.

Quantum compression

The von Neumann entropy is defined as $S(\rho) \equiv -\mathrm{tr}(\rho \log_2 \rho)$, where ρ is a density operator in a d-dimensional Hilbert space, $\mathcal{H}$. If the eigenvalues of ρ are λ_x, then $S(\rho)$ can be rewritten as

$$S(\rho) \equiv -\sum_x \lambda_x \log_2 \lambda_x, \tag{A.2}$$

with $0 \log 0 \equiv 0$. The bounds are $0 \leq S(\rho) \leq \log_2 d$. For the simple case of a qubit, this amounts to $0 \leq S(\rho) \leq 1$. The von Neumann entropy is just the Shannon entropy with the probabilities p_i replaced by eigenvalues λ_x, ie $S(\rho) = H(\lambda_x)$. Schumacher's noiseless channel coding theorem states: Let $\{\mathcal{H}, \rho\}$ be an independent and identically distributed quantum source. If $R > S(\rho)$ then there exists a reliable compression scheme of rate R for the source. If $R < S(\rho)$, then any compression scheme of rate R will not be reliable. For a qubit source, this means a trivially uncompressed n qubits can be optimally compressed to $nS(\rho)$ qubits.

Time compression

We postulate that

$$\alpha = H(X), \tag{A.3}$$

where $\alpha \equiv 1/\gamma = \sqrt{1 - (v^2/c^2)}$. (This is analogous to $S(\rho) = H(\lambda_x)$ in quantum information theory). For an uncompressed time interval, optimal compression is achieved using the time dilation formula (6.1). This is analogous to classical (5.1) and quantum (5.2) case. The bounds for this α-entropy are $0 \leq \alpha \leq 1$, with $\alpha = 0$ when $v = c$. The uncompressed interval can be thought of as the case when maximum entropy occurs. A particular physical realization of this is when a given coordinate time Δt is contracted to various proper times $\Delta\tau$ depending on the velocity of different observers, ie $\Delta t = \gamma_1 \Delta\tau_1$, $\Delta t = \gamma_2 \Delta\tau_2$, etc.

Properties of α-entropy

We have seen that different velocities compress a time interval differently hence we can identify this entropy with velocity. In classical information theory, the mutual information, $H(X{:}Y)$, of X and Y measure how much information X and Y have in common; the quantum mutual information for systems A and B is denoted by $S(A{:}B)$. This notion of commonality can be captured using the physical scenario of the relativistic velocity, $v_r = (v_1 - v_2)/(1 - (v_1 v_2/c^2))$, of two observers, v_1 and v_2 with respect to a fixed coordinate time interval. Each observer can be identified with a compression entropy, $\alpha_1 = 1/\gamma_1$ and $\alpha_2 = 1/\gamma_2$. Their relative Lorentz factor, $\gamma_r = \gamma_1\gamma_2(1 - (v_1 v_2/c^2))$, provides the inspiration to define the α-mutual information between α_1 and α_2:

$$\alpha_{1:2} \equiv \frac{1}{\gamma_r} = \frac{1}{\gamma_1\gamma_2(1 - \frac{v_1 v_2}{c^2})} = \frac{\alpha_1\alpha_2}{(1 - \frac{v_1 v_2}{c^2})} \tag{A.4}$$

The classical joint entropy, $H(A,B)$, and the quantum joint entropy, $S(A,B)$, help us define the α-joint entropy:

$$H(X,Y) = H(X) + H(Y) - H(X{:}Y), \tag{A.5}$$

$$S(A,B) = S(A) + S(B) - S(A{:}B), \tag{A.6}$$

$$\alpha_{1,2} \equiv \alpha_1 + \alpha_2 - \alpha_{1:2}. \tag{A.7}$$

Similarly, the α-conditional entropy can be developed using the classical and quantum analogue:

$$H(X|Y) \equiv H(X,Y) - H(Y), \tag{A.8}$$

$$S(A|B) \equiv S(A,B) - S(B), \tag{A.9}$$

$$\alpha_{1|2} \equiv \alpha_{1,2} - \alpha_2. \tag{A.10}$$

We'll derive entropic properties concerning two systems, and compare them with the classical theory. For example, it can easily be shown that $\alpha_{1:2} = \alpha_{2:1}$ and $\alpha_{1,2} = \alpha_{2,1}$ which is like the classical case of $H(X{:}Y) = H(Y{:}X)$ and $H(X,Y) = H(Y,X)$. Furthermore, $H(Y|X) \geq 0$ and $H(X) \leq H(X,Y)$ with equality satisfied for each inequality, if and only if Y is a function of X; the last two properties fail for the quantum case, in particular for systems involving entanglement; for our α case, $\alpha_{2|1} \geq 0$ and $\alpha_1 \leq \alpha_{1,2}$ is satisfied as long as $\gamma_r \geq \gamma_2$, ie physically the relative Lorentz factor has to be greater than or equal to the Lorentz factor of the second observer. The property of subadditivity holds for all three cases: $H(X,Y) \leq H(X) + H(Y)$, $S(A,B) \leq S(A) + S(B)$, and $\alpha_{1,2} \leq \alpha_1 + \alpha_2$ (which can be reduced to $\alpha_{1:2} \geq 0$). Classically, $H(Y|X) \leq H(Y)$ and it can be shown that $\alpha_{1|2} \leq \alpha_1$ (which can be reduced to $\alpha_{1:2} \geq 0$). In the last classical equation as well as second to last equation of classical subadditivity, equality is expressed if X and Y are independent variables. In the last two analogous relativistic equations, equality is expressed if $\alpha_{1:2} = 0$, which is when one of the observers has velocity c. Thus, the notion of 'independence' enters when one of the observers is moving at the speed of light.

Typical and atypical time intervals

The crucial idea underlying classical data compression is to divide the sequences into two types, namely typical sequences and atypical sequences. Given $\epsilon > 0$, a sequence $x_1, \ldots, x_n$ is ϵ-typical if it satisfies $2^{-n(H(X)+\epsilon)} \leq p(x_1, \ldots, x_n) \leq 2^{-n(H(X)-\epsilon)}$, and the set of all ϵ-typical sequences of length n is denoted by $A_\epsilon^{(n)}$. Atypical sequences are sequences that are not typical. The theorem of typical sequences states that: (a) The probability that a sequence is ϵ-typical is: $\Pr\{A_\epsilon^{(n)}\} > 1 - \epsilon$, for sufficiently large n. (b) The number of ϵ-typical sequences satisfies $(1-\epsilon)2^{n(H(X)-\epsilon)} \leq |A_\epsilon^{(n)}| \leq 2^{n(H(X)+\epsilon)}$, for sufficiently large n.

With respect to classical data compression, since there are at most $2^{nH(X)}$ typical sequences, to uniquely identify them requires only $nH(X)$ bits. If a sequence is atypical, compression fails. Hence, typical sequences are those that can be optimally compressed, whereas atypical sequences are those that can't. But from the theorem, we can see that typical sequences are overwhelmingly likely to occur in the asymptotic limit, and compression succeeds with probability approaching one; atypical sequences do occur for large n, albeit very rarely. To build quantum information theory, one develop a quantum form of the typical sequence theorem and derives the quantum compression.

In our time compression case involving the α-entropy, one would ideally need to define an time interval Δt that is ϵ-typical. If this relativistic case is similar to the classical and quantum case, then one can easily see that typical time intervals are those that can be compressed from Δt to $\alpha\Delta t$. Physically, this means typical time intervals obey the relativistic time dilation formula (5.2). Therefore, atypical time intervals are those that exhibit Lorentz violations. Furthermore, if a similar information theory result occur in this case, then the probability that a time interval is ϵ-typical is: $\Pr\{T_\epsilon^{(\Delta t)}\} > 1 - \epsilon$, for sufficiently large duration Δt. Hence Lorentz violations do occur for large time intervals but very rarely. In this sense, Lorentz violations is fundamentally not a matter of scale, but rather of probability; hence this violation may be experimentally detectable at large scales given a very large sample size.

It's important to note that this idea also applies to space intervals. For the x-direction, $\alpha\Delta x_{uncompressed} = \Delta x_{compressed}$, where the compressed space interval is due to length contraction. In the data compression subsection, it can be seen that optimal compression is achieved using $nH(X)$ bits (or $nS(\rho)$ qubits). For the case $nR > nH(X)$, compression without loss of information is achieved, but is not optimal. For $nR < nH(X)$, information is lost and compression is not reliable. These results seem to align to Lorentz transformations, which can be re-written as $\alpha\Delta t_{uncompressed} = \Delta t \pm \frac{v}{c^2}\Delta x$ where $\alpha\Delta t_{uncompressed} = \Delta t_{compressed}$ is the optimally compressed interval. The Lorentz equation for the the minus case, leads to $\Delta t > \Delta t_{compressed}$; this can be interpreted as compression is achieved but not optimally. For the plus case, $\Delta t < \Delta t_{compressed}$, which means some information is lost and gone to Δx.