A Computational Model of Time-Dilation

Charles Davi

June 26, 2019

Abstract

We propose a model of time-dilation that follows from the application of concepts from information theory and computer theory to physical systems. Our model predicts equations for time-dilation that are identical in form to those predicted by the special theory of relativity. In short, our model can be viewed as an alternative set of postulates rooted in information theory and computer theory that imply that time-dilation will occur. We also show that our model is consistent with the general theory of relativity's predictions for time-dilation due to gravity.

1 Introduction

Prior to the twentieth century, physicists appear to have approached nature with a general presumption that fundamental physical properites such as energy and charge are continuous. This was likely the result of a combination of prevailing philosphocial views, and mathematical convenience, given that continuous functions are generally easier to calculate than discrete functions. However, this view of nature began to unravel in the early twentieth century as a result of the experiments of Robert A. Millikan, and others, who demonstrated that charge appeared to be an integer multiple of a single value, the elementary charge e [17], and confirmed Einstein's predictions for the energy of electrons ejected due to the photoelectric effect, which suggested a quantized, particle theory of light [18]. In the century that followed these experiments, the remarkable success of quantum mechanics as a general matter demonstrated that whether or not the ultimate, underlying properties of nature are in fact discrete, the behavior of nature can nonetheless be predicted to a high degree of precision using models that make use of discrete values. This historical progression from a presumption of continuous values, towards the realization that fundamental properties of nature such as charge are quantized, was facilitated in part by the development of experimental techniques that were able to make measurements at increasingly smaller scales, and the computational power of the computer itself, which facilitated the use of discrete calculations that would be impossible to accomplish by hand. Though admittedly anecdotal, this progression suggests the possibility that at a sufficiently small scale of investigation, perhaps we would find that all properties of nature are in fact discrete, and thus, all apparently

continuous phenomena are simply the result of scale. Below we show that if we assume that all natural phenomena are both discrete, and capable of being described by computable functions, then we can achieve a computational model of time-dilation that predicts equations that are generally identical in form to those predicted by the special theory of relativity.

1.1 The Information Entropy

Assume that the distribution of characters in a string x is $\{p_1, \ldots, p_n\}$, where p_i is the number of instances of the i-th character in some alphabet $\Sigma = \{a_1, \ldots, a_n\}$, divided by the length of x. For example, if $x = (ab)$, then our alphabet is $\Sigma = \{a, b\}$, and $p_1 = p_2 = \frac{1}{2}$, whereas if $x = (aaab)$, then $p_1 = \frac{3}{4}$ and $p_2 = \frac{1}{4}$. The minimum average number of bits per character required to encode x as a binary string without loss of information, taking into account only the distribution of characters within x, is given by,

$$H(x) = -\sum_{i=1}^{n} p_i \log(p_i).^{1}$$

(1)

We call $H(x)$ the **information entropy** of x. The intuition underlying the information entropy is straightforward, though the derivation of equation (1) is far from obvious, and is in fact considered the seminal result of information theory, first published by Claude Shannon in 1948 [20]. To establish an intuition, consider the second string $x = (aaab)$, and assume that we want to encode x as a binary string. We would therefore need to assign a binary code to each of a and b. Since a appears more often than b, if we want to minimize the length of our encoding of x, then we should assign a shorter code to a than we do to b. For example, if we signify the end of a binary code with a 1, we could assign the code 1 to a, and 01 to b.[2] As such, our encoding of x would be 11101, and since x contains 4 characters, the average number of bits per character in our encoding of x is $\frac{5}{4}$. Now consider the first string $x = (ab)$. In this case, there are no opportunities for this type of compression because all characters appear an equal number of times. The same would be true of $x = (abcbca)$, or $x = (qs441z1zsq)$, each of which has a uniform distribution of characters. In short, we can take advantage of the statistical structure of a string, assigning longer codes to characters that appear less often, and shorter codes to characters that appear more often. If all characters appear an equal number of times, then there are no opportunities for this type of compression.

---

[1]Unless stated otherwise, all logarithms referenced in this paper are base 2.

[2]Rather than make use of a special delimiting character to signify the end of a binary string, we could instead make use of a"prefix-code". A prefix code is an encoding with the property that no code is a prefix of any other code within the encoding. For example, if we use the code 01 in a given prefix code, then we cannot use the code 010, since 01 is a prefix of 010. By limiting our encoding in this manner, upon reading the code 01, we would know that we have read a complete code that corresponds to some number or character. In contrast, if we include both 01 and 010 in our encoding, then upon reading an 01, it would not be clear whether we have read a complete code, or the first 2 bits of 010.

In general, if a string x is drawn from an alphabet with n characters, and the distribution of these characters within x is uniform, then $H(x) = \log(n)$, which is the maximum value of $H(x)$ for a string of any length drawn from an alphabet with n characters.

Shannon showed in [20] that a minimum encoding of x would assign a code of length $l_i = \log(\frac{1}{p_i})$ to each $a_i \in \Sigma$. If the length of x is N, then each a_i will appear Np_i times within x. Thus, the minimum total number of bits required to encode x using this type of statistical compression is $\sum_{i=1}^{n} Np_i l_i = NH(x)$. Therefore, the minimum average number of bits per character required to encode a string of length N is $H(x)$. Note that $H(x)$ is not merely a theoretical measure of information content, since there is always an actual binary encoding of x for which the average number of bits per character is approximately $H(x)$ [14]. Thus, $H(x)$ is a measure of the average number of bits required to actually store or transmit a single character in x. However, the value of $H(x)$ is a function of only the distribution of characters within x, and therefore, does not take into account other opportunities for compression. For example, a string of the form $x = a^N b^N c^N$ has an obvious structure, yet $H(x) = \log(3)$ is maximized, since x has a uniform distribution of characters.[3] Thus, even if a string has a high information entropy, the string could nonetheless have a simple structure.

1.2 The Information Content of a System

Despite the limitations of $H(x)$, we can still use $H(x)$ to measure the information content of representations of physical systems, understanding that we are able to account for only the statistical structure of the representation. We begin with a very simple example: consider a system comprised of N particles that initially all travel in the same direction, but that over time have increasingly random, divergent motions. We could represent the direction of motion of each particle relative to some fixed axis using an angle θ. If we fix the level of detail of our representation of this system by breaking θ into groups of $A = [0, \frac{\pi}{2})$, $B = [\frac{\pi}{2}, \pi)$, $C = [\pi, \frac{3\pi}{2})$, and $D = [\frac{3\pi}{2}, 2\pi)$, then we could represent the direction of motion of each particle in the system at a given moment in time as a character from $\Sigma = \{A, B, C, D\}$ (see Figure 1 below). Note that this is clearly not a complete and accurate representation of the particles, since we have, for example, ignored the magnitude of the velocity of each particle. Nonetheless, we can represent the direction of motion of all of the particles at a given moment in time as a string of characters drawn from Σ of length N. For example, if at time t the direction of motion of each particle is $\theta = 0$, then we could represent the motions of the particles at t as the string $x = (A \cdots A)$, where the length of x, denoted $|x|$, is equal to N. As such, the distribution of motion is initially entirely concentrated in group A, and the resultant distribution of characters within x is $\{1, 0, 0, 0\}$. The information entropy of $\{1, 0, 0, 0\}$ is $-\sum_{i=1}^{4} p_i \log(p_i) = 0$ bits,[4] and therefore, the minimum average number of bits per character required

---

[3]For example, if $N = 2$, then $x = aabbcc$.
[4]As is typical when calculating $H(x)$, we assume that $0 \log(0) = 0$.

to encode this representation of the particles at t is 0 bits. Over time, the particles will have increasingly divergent motions, and as such, the distribution of characters within x will approach the uniform distribution, which is in this case $\{\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\}$, which has an information entropy of $\log(4) = 2$ bits. Thus, the information entropy of this representation of the particles will increase over time.
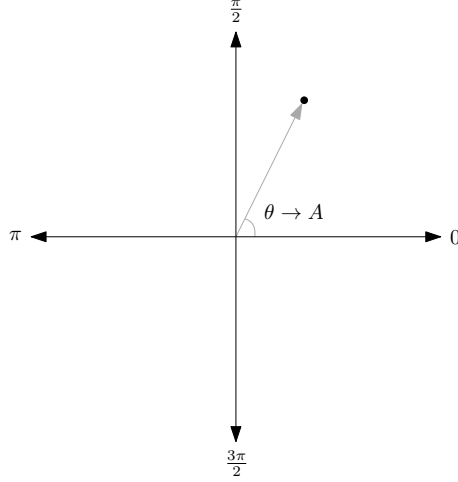

Figure 1: A mapping of angles to $\Sigma = \{A, B, C, D\}$.

We could argue that, as a result, the information content of the system itself increases over time, but this argument is imprecise, since this particular measure of information content is a function of the chosen representation, even though the behavior of the system can impact the information content of the representation. For example, if we made use of a finer gradation of the angle θ above, increasing the number of groups, we would increase the number of characters in our alphabet, thereby increasing the maximum information content of the representation, without changing the system in any way. However, this does not imply that representations are always arbitrary. For example, if some property of a system can take on only n discrete values, then a representation of the system that restricts the value of this property to one of these n values is not arbitrary. The point is that as a practical matter, our selection of certain properties will almost certainly be incomplete, and measured at some arbitrary level of precision, which will result in an arbitrary amount of information. Thus, as a practical matter, we probably cannot answer the question of how much information is required to completely and accurately represent a physical system. We can, however, make certain assumptions that would allow us to construct a complete and accurate representation of a system, and then measure the information content of that representation.

Assumption 1.1. *There is a finite set of n measurable properties Γ such that*

4

(1) for any measurable property $P \notin \Gamma$, the value of P can be derived from the values of the properties $P_i \in \Gamma$, and (2) there is no $P_i \in \Gamma$ such that the value of P_i can be derived from the other $n - 1$ properties in $\Gamma - P_i$.

We call each of the properties $P_i \in \Gamma$ a **basis property**. Note that we are not suggesting that all properties are a linear combination of the basis properties within Γ. Rather, as discussed in Sections 1.3 and 2 below, we assume that all other measurable properties can be derived from Γ using computable functions. For example, if mass and velocity are included in Γ, then Assumption 1.1 implies that momentum would not be included in Γ, since momentum can be derived from mass and velocity using a computable function.[5]

Assumption 1.2. *For any closed system,[6] each $P_i \in \Gamma$ can take on only a finite number of possible values.*

For example, assume that a closed system S contains a finite number of N elements.[7] Assumption 1.1 implies that there is a single set of basis properties Γ from which all other measurable properties of any given element can be derived. As such, in this case, S consists of a finite number of elements, each with a finite number of measurable basis properties, and Assumption 1.2 implies that each such basis property can take on only a finite number of possible values.

Assumption 1.3. *For any system, all measurable properties of the system, as of a given moment in time, can be derived from the values of the basis properties of the elements of the system, as of that moment in time.*

Together, Assumptions 1.1 and 1.3 allow us to construct a complete and accurate representation of a system at a given moment in time. Specifically, if we were able to measure the basis properties of every element of S at time t, then we could construct a representation of the state of S as a set of N strings $S(t) = \{s_1, \ldots, s_N\}$, with each string representing an element of S, where each string $s_i = (v_1, \ldots, v_n)$ consists of n values, with v_j representing the value of the j-th basis property of the i-th element of S at time t. Because $S(t)$ contains the values of the basis properties of every element of S at time t, Assumption 1.3 implies that we can derive the value of any property of S at time t from the representation $S(t)$ itself. For example, Assumption 1.3 implies that there is some computable function f that can calculate the momentum ρ of S at time t when given $S(t)$ as input. Expressed symbolically, $\rho = f(S(t))$.[8] Thus, $S(t)$ contains all of the information necessary to calculate any measurable property of S at time t, and therefore, we can take the view that $S(t)$ constitutes a complete

---

[5]We will discuss computable functions in greater detail in Sections 1.3 and 2 below, but for now, a computable function can be defined informally as any function that can be implemented using an algorithm.

[6]We view a system as closed if it does not interact with any other systems or exogenous forces, and is bounded within some definite volume.

[7]We deliberately use the generic term "element", which we will clarify in Section 2 below.

[8]We assume that all particles and systems have definite locations, and definite values for all of their measurable properties. As such, our model, as presented herein, is necessarily incomplete, since it does not address systems governed by the laws of quantum mechanics.

and accurate representation of the state of S at time t. Note that Assumption 1.3 does not imply that we can determine all future properties of S given the values of the basis properties of its elements at time t, but rather, that the value of any property of S that exists at time t can be derived from the values of the basis properties of its elements as of time t.[9] Of course, we are not suggesting that we can construct such a representation as a practical matter, but rather, we will use the concept of $S(t)$ as a theoretical tool to analyze the information content of systems generally, and ultimately, construct a model of time-dilation.

Recall that Assumption 1.2 implies that each basis property of S can take on only a finite number of possible values. If we assume that the n basis properties are independent, and that basis property P_i can take on k_i possible values, then the basis properties of each element of S can have any one of $K = k_1 \cdots k_n$ possible combinations of values. If we distinguish between the N elements of S, and assume that the basis properties of each element are independent from those of the other elements, then there are K^N possible combinations of values for the basis properties of every element of S. Since the values of the basis properties of the elements of S determine all measurable properties of S, it follows that any definition of the overall state of S will reference either the basis properties of the elements of S, or measurements derived from the basis properties of the elements of S. As such, any definition of the overall state of S will ultimately reference a particular combination of values for the basis properties of the elements of S. For example, Assumption 1.3 implies that the temperature of a system is determined by the values of the basis properties of its elements. Therefore, the maximum number of states of S is equal to the number of unique combinations of values for the basis properties of the elements of S, regardless of our choice of the definition of the overall state of S.[10] We can assign each such state a number from 1 to K^N, and by S_i we denote a representation of the i-th state of S. That is, $S(t)$ denotes a representation of the state of S at time t, whereas S_i denotes a representation of the i-th possible state of S in some arbitrary ordering of the states of S. Thus, for any given moment in time t, there exists an S_i such that $S(t) = S_i$. By $|S| = K^N$ we denote the number of possible states of S.

Now imagine that we measure the value of every basis property of every element of S over some long period of time, generating M samples of the state of S, and that for each sample we store the number assigned to the particular state of S we observe. For example, if we observe S_j, then we would add the number j to our string. Thus, in this case, $\Sigma = \{1, \ldots, |S|\}$ is the alphabet, and the resultant string is a string of numbers $x = (n_1 \cdots n_M)$, representing the M states of S observed over time. Further, assume that we find that the distribution of the states of S over that interval of time is $\varphi = \{p_1, \ldots, p_{|S|}\}$, where p_i is the number of times S_i is observed divided by the number of samples M. We could then encode x as a binary string, and the minimum average number of

---

[9]We will discuss calculating future states of systems in Section 1.4 below.

[10]We will revisit this topic in the context of computable functions on the basis properties of a system in Section 2 below.

bits required to identify a single state of S would be $H(x) = -\sum_{i=1}^{|S|} p_i \log(p_i)$. Note that we are not encoding the values of the basis properties of the elements within $S(t)$, but we are instead representing each observed state of S with a number, and encoding the resultant string of numbers. That is, each possible combination of values for the basis properties of the elements of S corresponds to a particular unique overall state of S, which, when observed, we represent with a number. In contrast, in Section 1.4 below, we will use the Kolmogorov complexity to measure the information contained in an encoding of $S(t)$ itself. If the distribution φ is stable over time, then we write $H(S)$ to denote the information entropy of x, which we call the **representational entropy** of S. Thus, $H(S)$ is the average number of bits per state necessary to identify the particular states of S that are observed over time. If φ is the uniform distribution, then we have,

$$H(S) = \log(|S|). \tag{2}$$

Thus, the representational entropy of a system that is equally likely to be in any one of its possible states is equal to the logarithm of the number of possible states. We note that equation (2) is similar in form to the thermodynamic entropy of a system $k_B \ln(\Omega)$, where k_B is the Boltzmann constant, and Ω is the number of microstates the system can occupy given its macrostate. This is certainly not a novel observation, and the literature on the connections between information theory and thermodynamic entropy is extensive. (See [3] and [16]). In fact, the similarity between the two equations was noted by Shannon himself in [20]. However, the goal of our model is to achieve time-dilation, and thus, a review of this topic is beyond the scope of this paper. Finally, note that if φ is stable, then we can interpret p_i as the probability that $S(t) = S_i$ for any given t, and therefore, we can view $H(S)$ as the expected number of bits necessary to identify a single state of S.

We can also view S as a medium in which we can store information. The number of bits that can be stored in S is also equal to $\log(|S|)$, regardless of the distribution φ, which we call the **information capacity** of S. Under this view, we do not observe S and record its state, but rather, we "write" the current state of S by fixing the value of every basis property of every element of S, and use that state to represent a number or character. As such, when we "read" the current state of S, measuring the value of every basis property of every element of S, we can view each possible current state S_i as representing some number or character, including, for example, the number i. As such, S can represent any number from 1 to $|S|$, and thus, the information contained in the current state of S is equivalent to the information contained in a binary string of length $\log(|S|)$.[11] For example, whether the system is a single switch that can be in any one of 16 states, or a set of 4 switches that can be in any one of 2 states, in either case, measuring the current state of the system can be viewed as equivalent to reading $\log(16) = 4$ bits of information. Thus, each state of S can be viewed as

---

[11] Note that a binary string of length $\log(|S|)$ has $|S|$ states, and as such, can code for all numbers from 1 to $|S|$.

containing $\log(|S|)$ bits of information, which we call the **information content** of S.

Note that the representational entropy of a system is a measure of how much information is necessary to identify the states of a system that are observed over time, which, although driven by the behavior of the system, is ultimately a measure of an amount of information that will be stored outside of the system itself. In contrast, the information capacity and information content of a system are measures of the amount of information physically contained within the system. Though the information capacity and the information content are always equal, conceptually, it is worthwhile to distinguish between the two, since the information capacity tells us how much information a system can store as a general matter, whereas the information content tells us how much information is observed when we measure the basis properties of every element of a given state of the system.

Finally, note that if a system is closed, then no exogenous information has been "written" into the system. Nonetheless, if we were to "read" the current state of a closed system, we would read $\log(|S|)$ bits of information. The information read in that case does not represent some exogenous symbol, but is instead the information that describes the basis properties of the system. Thus, the amount of information observed when we measure the basis properties of every element of a given state of the system is $\log(|S|)$ bits.

1.3 The Kolmogorov Complexity

Consider again a string of the form $x = a^N b^N c^N$. As noted above, x has an obvious structure, yet $H(x) = \log(3)$, which is the maximum information entropy for a string drawn from an alphabet with 3 characters. Thus, the information entropy is not a measure of randomness, since it can be maximized given strings that are clearly not random in any sense of the word. Assume that $N = 10^8$, and that as such, at least $|x|H(x) = 3 \times 10^8 \log(3)$ bits are required to encode x using a statistical encoding. Because x has such an obvious structure, we can write and store a short program that generates x, which will probably require fewer bits than encoding and storing each character of x. Note that for any given programming language, there will be some shortest program that generates x, even if we can't prove as a practical matter that a given program is the shortest such program.

This is the intuition underlying the **Kolmogorov complexity** of a binary string x, denoted $K(x)$, which is, informally, the length of the shortest program, measured in bits, that generates x as output. More formally, given a Universal Turing Machine U (a "UTM") and a binary string x, $K(x)$ is the length of the shortest binary string y for which $U(y) = x$ [12].[12] Note that $K(x)$ does not consider the number of operations necessary to compute $U(y)$, but only the

---

[12]Note that some applications of $K(x)$ depend upon whether the UTM is a "prefix machine", which is a UTM whose inputs form a prefix-code, and thus, do not require special delimiters to indicate the end of a string. For simplicity, all UTMs referenced in this paper are not prefix machines, and thus, an integer n can be specified as the input to a UTM using $\log(n)$ bits.

length of y, the binary string that generates x. Thus, $K(x)$ is not a measure of overall efficiency, since a program could be short, but nonetheless require an unnecessary number of operations. Instead, $K(x)$ is a measure of the information content of x, since at most $K(x)$ bits are necessary to generate x on a UTM.

We will not discuss the theory of computability in any depth, but it is necessary that we briefly mention the Church-Turing Thesis, which, stated informally, asserts that any computation that can be performed by a device, or human being, using some mechanical process, can also be performed by a UTM [22] [21].[13] In short, Turing's formulation of the thesis asserts that every mechanical method of computation can be simulated by a UTM. Historically, every method of computation that has ever been proposed has been proven to be either equivalent to a UTM, or a more limited method that can be simulated by a UTM. As such, the Church-Turing Thesis is not a mathematical theorem, but is instead a hypothesis that has turned out to be true as an empirical matter. The most important consequence of the thesis for purposes of this paper, is that any mathematical function that can be expressed as an algorithm is assumed to be a **computable function**, which is a function that can be caculated by a UTM. However, it can be shown that there are **non-computable functions**, which are functions that cannot be caculated by a UTM, arguably the most famous of which was defined by Turing himself, in what is known as the "Halting Problem" [21].[14] Unfortunately, $K(x)$ is a non-computable function [23], which means that there is no program that can, as a general matter, take a binary string x as input, and calculate $K(x)$. However, $K(x)$ is nonetheless a powerful theoretical measure of information content.

For example, consider the string $x = (aaabbaaaabbbb)^N$. The distribution of characters in this string is uniform, and as such, $H(x) = \log(2)$ is maximized. However, we could of course write a short program that generates this string for a given N. Because such a program can be written in some programming language, it is therefore computable, and can be simulated by $U(y) = x$, for some y. Therefore, $K(x) \leq |y|$. While this statement may initially seem trivial, it implies that $K(x)$ takes into account any opportunities for compression that can be exploited by computation, unlike $H(x)$, which is a function of only the statistical structure of x. Further, we could write a generalized program that takes N as input, and generates the appropriate x as output. Because such a program is computable, we can simulate it on U using some procedure f, and we write $U(f, N)$ to denote U running f with N as input. As such, $K(x) \leq |f| + \log(N)$. Similarly, we can write a program that takes a string x as input, and generates that same string x as output, and as such, we can do

---

[13]In [22], Turing stated that, "A function is said to be effectively calculable if its values can be found by some purely mechanical process." He then went on to specificy in mathematical terms, exactly what constitutes a purely mechanical process, which lead to the definition of a UTM. Thus, he asserted an equivalence between the set of functions that can be calculated by some purely mechanical process, and the set of functions that can be calculated by a UTM.

[14]Stated informally, in [21], Turing asked whether or not there was a program that could decide, as a general matter, whether another program would, when given a particular input, halt or run forever. He showed that assuming that such a program exists leads to a contradiction, implying that no such program exists.

the same on U with some procedure I, such that $U(I, x) = x$. It follows that $K(x) \leq |x| + C$, where $C = |I|$. Note that C will not depend upon our choice of x, and thus, $K(x)$ is always bounded by the length of x plus a constant. If $K(x) = |x| + C$, then there is no program that can compress x into some shorter string, and therefore, x is arguably devoid of any patterns. As such, we can take the view that if $K(x) = |x| + C$, then x is a **random** string. Thus, $K(x)$ could be used to distinguish between a string that has structure, but also happens to have a uniform distribution of characters, and a truly random string with no structure. As such, $K(x)$ is a measure of both information content, and randomness. Finally, it can be shown that for any two UTMs, $U1$ and $U2$, $|K(x)_{U1} - K(x)_{U2}| \leq C$, for some constant C that does not depend upon x [23]. Because all forms of computation are presumed to be equivalent to a UTM, and $K(x)$ does not depend upon our choice of UTM beyond a constant, $K(x)$ can be viewed as a universal, objective measure of both information content and randomness.

1.4 The Representational Complexity of a System

Assumption 1.4. *There is a computable function R, such that when given the current state of the basis properties of a closed system, R generates the next state of the basis properties of that system.*

As such, R is a computable function that is applied to a representation of the current state of a system, thereby generating a representation of the next state of the system. Thus, we assume that time is effectively discrete, and by t_0 we denote the minimum increment of time, which we call a **click**.[15] Specifically, we assume that the application of R to a representation of the current state of a system $S(t)$ will generate a set of representations of states $\{S_{j_1}, \ldots, S_{j_q}\}$ corresponding to all possible next states of S. Note that in this case, we are performing mathematical operations on a representation of the current state of a system $S(t)$, and asserting that the result $R(S(t))$ will include an objective representation of the actual, physical next state of S. That is, first we measure the value of every basis property of every element of S at time t, and then represent that current state as $S(t)$. Then we apply R, which is a computable mathematical function, to $S(t)$, which will generate a set of representations $\{S_{j_1}, \ldots, S_{j_q}\}$. Assumption 1.4 implies that if we physically measure the value of the basis properties of every element of the next state of S, we would find that this next state is represented by some $S_{j_a} \in R(S(t))$. Thus, $S(t)$ is a representation of the state of S at time t that is the result of actually physically measuring the value of every basis property of every element of S, whereas $R(S(t))$ is a set of representations that is the result of mathematical operations performed on $S(t)$. In short, $S(t)$ is the product of observation, and $R(S(t))$ is

---

[15]As such, we assume that the time it takes for a system to progress from one state to the next is necessarily discrete, whereas time itself can still be viewed as continuous. Thus, we assume that systems change states upon clicks, but we can nonetheless conceive of time between clicks.

the product of computation, and Assumption 1.4 implies that we always find that $S(t + t_0) \in R(S(t))$. That is, the set of states generated by $R(S(t))$ always includes a correct representation of the next state of S. Note that in all cases, $|R(S(t))| \leq |S|$, which we assume to be finite. This does not imply that all states generated by $R(S(t))$ are equally likely, but rather, that R will generate all of them given $S(t)$.[16]

If S is deterministic, then $R(S(t)) = \{S_j\}$ contains exactly one next state of S. Assume that S is deterministic. It follows that, given any initial state $S(t)$, we can generate all future states of S through successive applications of R. For example, $R(S(t)) = S(t + t_0)$ is the next state of S, and $R(R(S(t))) = S(t + 2t_0)$ is the state after that, and so on. Let $R(S(t))^m$ denote m successive applications of R to $S(t)$. For example, $R(S(t))^2 = R(R(S(t))) = S(t + 2t_0)$. In general, if $\Delta t = t_f - t$, then $R(S(t))^{\frac{\Delta t}{t_0}} = S(t_f)$.[17] That is, we can view R as being applied every click, and thus, $S(t_f)$ is the result of $\frac{\Delta t}{t_0}$ successive applications of R to $S(t)$.

Since $S(t)$ is a finite set of strings $\{s_1, \ldots, s_N\}$, we can encode $S(t)$ as a binary string x. Since x is a binary string, at most $K(x)$ bits are necessary to generate x on a UTM. In short, we can encode $S(t)$ as a binary string x, and then evaluate $K(x)$, and thus, we can adapt $K(x)$, which is typically defined using binary strings, to $S(t)$. We will apply $K(x)$ to $S(t)$ only to establish upper bounds on $K(x)$ that hold regardless of our choice of encoding, and as such, our choice of encoding is arbitrary for this purpose. We write $S(t)_x$ to denote the binary encoding of $S(t)$, and S_{i_x} to denote the binary encoding of S_i. We call $K(S(t)_x)$ the **representational complexity** of $S(t)$, which, as noted above, will depend upon our choice of encoding, since our choice of encoding will determine $S(t)_x$.

Since we assume that R is a computable function, there is some f_x that can simulate R on U, such that if $R(S_i) = S_j$, then $U(f_x, S_{i_x}) = S_{j_x}$. Further, we assume that f_x can be applied some specified number of times by including an encoding of an integer as a supplemental input that indicates the number of times f_x is to be applied. Expressed symbolically, $U(f_x, S_{i_x}, m_x)$ will generate an encoding of $R(S_i)^m$. Therefore, if $t_f = t + \Delta t$, then $S(t_f)_x = U(f_x, S(t)_x, (\frac{\Delta t}{t_0})_x)$. Note that f_x will depend upon on our choice of encoding. As such, we assume that our choice of encoding is fixed, and therefore, $|f_x|$ is a constant. Similarly, given any initial state $S(t)$, $K(S(t)_x)$ can be viewed as a constant as well. As such, let $C = |f_x| + K(S(t)_x)$. It follows that if S is deterministic, then,

$$K(S(t_f)_x) \leq C + \log(\frac{\Delta t}{t_0}). \tag{3}$$

Note that unlike the information content of a system, the representational complexity of a system is not a measure of how much information can be physically stored within the system. It is instead a measure of how much information

---

[16] These probabilities are irrelevant to our analysis in this section, and thus, we ignore them.

[17] Note that because we have assumed that the time between states is always t_0, it follows that $\frac{\Delta t}{t_0}$ is always an integer.

is necessary to generate an encoding of a particular state of the system on a UTM. In more practical terms, we can think about the representational complexity of $S(t)$ as a measure of the amount of information necessary to store a compressed representation of the state of S at time t, which would consider all possible methods of compressing that representation. We have simply expressed an upper bound that makes use of Assumption 1.4, which implies that we can generate a representation of any state of any system by applying a single computable function R to a representation of an initial state of that system. Further, note that our choice of encoding will impact only C, which is a constant. Also note that equation (3) does not imply that the representational complexity of a system will necessarily increase over time, but rather, equation (3) is an upper bound on the representational complexity of a system that happens to increase as a function of time. That is, there could be some other upper bound that has eluded us that does not increase as a function time. Whether or not $K(S(t)_x)$ actually increases as a function of time depends upon R and S. That is, if the repeated application of R somehow organizes the elements of S over time, then the representational complexity of S would actually decrease over time. For example, a system of objects subject to a gravitational field could initially have random motions that, over time, become dominated by the force of the gravitational field, and therefore grow more structured over time, perhaps forming similar, stable orbits that would lend themselves to compression. If however, the repeated application of R causes a system to become increasingly randomized over time, then equation (3) places a limit on the amount of information necessary to generate an encoding of any given state of that system. Thus, even if a system appears to be complex, if the repeated application of a rule generated the system, then the system can nonetheless be generated by a UTM using an encoding of some initial state, and an encoding of the rule.[18] However, as noted above, $K(x)$ does not consider the number of computations necessary to generate x, and as such, a system with simple initial conditions and simple rules could nonetheless require a large number of computations to generate.

Since R can generate any future state $S(t_f)$ given $S(t)$, there is some procedure g_x that can generate not only $S(t_f)_x$, but a string of encoded states beginning with $S(t)_x$ and ending with $S(t_f)_x$, which we denote by $S(t, t_f)_x = (S(t)_x \cdots S(t_f)_x)$. The number of encoded states in that string is $M = \frac{t_f - t}{t_0} + 1$, and as such, if we let $C = |g_x| + K(S(t)_x)$, then the average number of bits per state necessary to generate the encoded string of states $S(t, t_f)_x$ is given by,

$$\frac{1}{M} K(S(t, t_f)_x) \leq \frac{1}{M}(C + \log(\frac{\Delta t}{t_0})). \tag{4}$$

Note that equation (4) approaches 0 bits as M approaches infinity. Thus, if a system is deterministic, then the average number of bits per state necessary to generate a string of M encoded states approaches 0 bits as M approaches

---

[18]For example, cellular automata can generate what appear to be very complex patterns, though they are in fact generated by the repeated application of simple rules to simple initial conditions.

infinity. In contrast, the representational entropy of S is fixed, and is always equal to $H(S)$, regardless of the number of states sampled.[19] This is consistent with the fact that $H(S)$ takes into account only the distribution of states of S over time, whereas if S is deterministic, then all future states of S can be generated by the repeated application of a rule to an initial state. Since $S(t, t_f)$ represents the states of S over some interval of time, we can view $K(S(t, t_f)_x)$ as a measure of the amount of information necessary to generate a model of the states of S over that interval of time.

Now assume that S is non-deterministic, and thus, $R(S(t))$ could contain more than one state for any given t. We can still generate all possible future states of $S(t)$ through successive applications of R, but this will require applying R to multiple intermediate states, which will generate multiple final states. If we want to evaluate $K(S(t_f)_x)$ for a particular final state, we will need to distinguish between each such final state. We can do so by supplementing f_x in a away that causes f_x to not only generate all possible final states, but also assign each such final state a number, and output only the particular final state we identify by number as an input to this supplemented version of f_x. Call this supplemented procedure h_x. Note that in this case, we are not using R to predict the final state $S(t_f)$, since by definition we are identifying the particular final state we are interested in as an input to h_x. We are instead using R to compress a representation of that particular final state. Therefore, if we provide U with h_x, $S(t)_x$, an encoding of $\frac{\Delta t}{t_0}$, and an encoding of the number assigned to the particular final state $S(t_f)$ we are interested in, then U could generate just that particular final state as output. Expressed symbolically, $S(t_f)_x = U(h_x, S(t)_x, (\frac{\Delta t}{t_0})_x, m_x)$, where m_x is an encoding of the number assigned to the particular final state we are interested in. Because $m \leq |S|$, we can treat the length of this input as a constant bounded by $\log(|S|)$. As such, let $C = |h_x| + K(S(t)_x) + \log(|S|)$. It follows that if S is non-deterministic, then,

$$K(S(t_f)) \leq C + \log(\frac{\Delta t}{t_0}). \tag{5}$$

Thus, the upper bound on equation (5) also grows as a function of the logarithm of the amount of time that has elapsed since the initial state. However, the constant C will depend upon $|S|$, which is not the case for equation (3).

Finally, we can also consider the average number of bits per state necessary to generate a string of encoded states that begin with some initial state $S(t)_x$ and end with some final state $S(t_f)_x$ for a non-deterministic system. Because $|R(S(t))| \leq |S|$ for all t, it follows that the number of strings of length M that begin with $S(t)$ is less than or equal to $|S|^{M-1}$. That is, the number of possible strings is maximized if each of the $M-1$ applications of R results in $|S|$ possible states. Thus, if we wish to generate a particular string of length M beginning with $S(t)_x$, and not all such strings, then we can assign each such string a number m from 1 to $|S|^{M-1}$. We denote any such string as $S(t, t_f, m)_x$, where

---

[19]We are assuming that the distribution of states is stable, and that a sufficient number of observations have been made in order to determine the distribution.

$m \leq |S|^{M-1}$ is the number assigned to the particular string we are interested in. Because R can generate any future state of S, there is some procedure d_x that generates not only a particular final state state $S(t_f)_x$, but a particular string of states beginning with $S(t)_x$ and ending with the particular final state $S(t_f)_x$ we are interested in. Thus, if we provide U with d_x, $S(t)_x$, an encoding of $\frac{\Delta t}{t_0}$, and an encoding of m, then U could generate just that particular string as output. Expressed symbolically, $S(t, t_f, m)_x = U(d_x, S(t)_x, (\frac{\Delta t}{t_0})_x, m_x)$, where m_x is an encoding of the number assigned to the particular string we are interested in. In this case, the value of m is unbounded, and will depend upon M, and thus, we cannot treat the length of the encoding of m as a constant. As such, let $C = |d_x| + K(S(t)_x)$. Therefore, if S is non-deterministic, then the average number of bits per state necessary to generate the encoded string of states $S(t, t_f, m)_x$ is given by,

$$\frac{1}{M} K(S(t, t_f, m)_x) \leq \frac{1}{M}(C + \log(\frac{\Delta t}{t_0}) + (M-1)\log(|S|)). \qquad (6)$$

Note that equation (6) approaches $\log(|S|)$ as M approaches infinity, which is also the information content of S. As a result, in practical terms, since t_0 is presumably extremely small, M will be extremely large,[20] and thus, the average amount of information per state necessary to model the behavior of a non-deterministic system is bounded by $\log(|S|)$ from above.

2   Energy and Information

We assume that energy is always quantized, and by E_0 we denote the minimum possible energy for any particle. We will discuss the value of E_0 in Section 3.5 below. Further, we assume that each such quantum of energy is an actual physical quantity that can change states over time. This is not to suggest that energy has mass, but rather, that energy is a physical substance, with quantity that can be measured, which we assume to be an integer multiple of E_0. For example, photon pair production demonstrates that the kinetic energy of a photon can be converted into the mass energy of an electron and positron [5]. Similarly, electron-positron annihilation demonstrates that the mass energy of an electron and positron can be converted into the kinetic energy of a photon [7]. Moreover, that total energy is conserved in these interactions [5] [4], and not mass or velocity, suggests that energy is a physical substance, as opposed to a secondary property due to mass or velocity. Finally, the photon itself is further evidence for this view of energy as a physical substance. That is, since the velocity of a photon is constant regardless of its energy, the energy of a photon cannot be a consequence of its velocity. Since the photon appears to be massless, the energy of a photon cannot be a consequence of its mass. Yet a photon is nonetheless capable of collisions, as is demonstrated by Compton scattering [8]. As such, the photon must have some physical substance, even

---

[20]For example, if we assume that t_0 is on the order of the Planck time, then M would increase by approximately 10^{44} states per second.

though it appears to be massless. To reconcile all of these observations, we view the energy of a photon as the primary substance of the photon, and not a consequence of its other properties. Thus, we view energy itself as a primary physical substance that can give rise to either mass or velocity, depending upon its state. In short, we view energy as a quantized physical substance that can change states over time.

Assumption 2.1. *Each quantum of energy is always in one of some finite number of K states, where K is constant for all quanta of energy.*[21]

Additionally, we assume that each quantum of energy is always in one of two categories of states: a **mass state**, or a **kinetic state**. This is not to say that $K = 2$, but rather, that each of the K possible states can be categorized as either a mass state or a kinetic state. Further, we assume that a quantum of energy in a mass state contributes to the mass energy of a system, and a quantum of energy in a kinetic state contributes to the kinetic energy of a system. For example, we assume that the photon consists entirely of quanta in a kinetic state, which we will discuss in greater detail in Sections 3.3, 3.5, and 5.2 below.

Further, we assume that the basis properties of a system are coded for in the energy of the system. Thus, we assume that the quanta of energy within a system constitute the elements of the system, and that as such, all measurable properties of the system follow from the information contained within those quanta of energy. Specifically, we assume that the basis properties of a quantum of energy consist solely of its state and its position. It follows that if we could measure the state and position of each quantum of energy within a system, then we could derive all other measurable properties of the system. Note that we are not suggesting that we can, as a practical matter, actually determine the state and position of each quantum of energy within a system. Nor are we suggesting that we can, as a practical matter, calculate all of the measurable properties of a system given the state and position of each quantum of energy within the system. Rather, we are assuming that the energy of a system contains all of the information necessary to determine all of the measurable properties of the system, and thus, if the means existed to obtain that information, and we knew which functions to apply to that information, then we could determine any measurable property of the system.

Thus, we assume that the collective states and positions of the quanta of energy within a system completely characterize the measurable properties of the system. For example, assume that a system S with a total energy of E_T consists of $N = \frac{E_T}{E_0}$ quanta of energy, and let X_i be a vector representing the position of the i-th quantum. Using the notation developed in the sections above, $S(t)$ is therefore a set of N vectors $\{(n_1, X_1), \ldots, (n_N, X_N)\}$, with each $n_i \leq K$ representing the state of the i-th quantum of energy within the system at time t, and X_i representing its position at time t. Thus, we view Γ as consisting of two properties: the state of a quantum, and its position. As such, given the state

---

[21]Note that Assumption 2.1 accounts for only the K intrinsic states that a quantum can occupy, and not its position.

and position of every quanta of energy within a system, we can derive all other measurable properties of the system. While this characterization of Γ might seem minimalist, recall that we defined Γ as a set of basis properties from which all other measurable properties of a system can be derived using computable functions. It follows that we can view $S(t)_x$ as the input to a UTM running some procedure f that calculates the particular property we are interested in. For example, there will be some function f_ρ that can calculate the momentum of a system given the basis properties of its elements. Expressed symbolically, $U(f_\rho, S(t)_x)$ will calculate the momentum of S at time t.[22]

As a result, there is a tacit equivalence between the act of measuring a particular property of a system, and executing the computable function that calculates that property. If we view the act of measuring a particular property of a system as a purely mechanical process, then the Church-Turing Thesis suggests that the actions taken in order to effectuate that measurement could be viewed as an algorithm. If we consider the nature of measurement, which requires following certain specified procedures for interacting with a physical object that ultimately result in a numerical output, then the characterization of measurement as a type of algorithm seems reasonable. Further, if we consider the nature of computation, which requires following certain specified procedures for manipulating a physical medium that ultimately result in a numeral output, then the characterization of computation as a type of measurement also seems reasonable. Thus, the Church-Turing Thesis offers a reasonable explanation as to why measurement can be simulated by computation, since both computation and measurement can be viewed as mechanical processes. Moreover, since the set of computable functions is infinite, a system can have an infinite number of measurable properties, even though the information content of the system is finite.[23]

Thus, any measurement of the overall state of S will be given by some function f, for which $U(f, S_{i_x})$ gives the state of S as defined by the measure f for the particular values of the basis properties given by S_i. For any f, each S_i will correspond to some value $U(f, S_{i_x})$, which might not be unique to S_i. That is, it could be the case that $U(f, S_{i_x}) = U(f, S_{j_x})$ for some S_i, S_j.[24] As such, the number of states of S will be maximized if $U(f, S_{i_x})$ is unique for each S_i. Thus, the number of states of S, under any definition of the state of S, will be limited to the number of unique combinations of states and positions of

---

[22]Technically, $U(f_\rho, S(t)_x)$ will generate an encoding of the numerical value of the momentum.

[23]We are not suggesting that there are in fact an infinite number of useful measurable properties, but rather, illustrating the point that even using a finite representation of a system, an infinite number of measurable properties can nonetheless be generated. We are, however, suggesting an equivalence between the set of computable functions on the basis properties of a system, and the set of measurable properties of the system. Thus, we are suggesting the possibility that for every computable function on the basis properties of a system, there is some physical measurement of the system itself that will always yield the same values as that computable function.

[24]For example, multiple microstates of a system can produce the same overall temperature for the system.

the quanta within S. It follows that the number of states a system can occupy will be maximized if the state and position of a quantum are independent, and the state and position of each quantum are independent from those of all other quanta.[25] Let X denote the total number of positions that a quantum of energy can occupy within S.

Assumption 2.2. *Given a system S that consists of $N = \frac{E_T}{E_0}$ quanta of energy, the maximum number of states of S is given by,*

$$|S| \leq K^N X^N. \tag{7}$$

Thus, the number of states a system can occupy will increase as a function of its total energy, and X, which is presumably a function of its volume. The greater the number of states a system can occupy, the greater the information capacity and information content of the system. Therefore, the greater the total energy of a system, the greater the information capacity and information content of the system. Moreover, unlike the representational complexity of a system, and the representational entropy of a system, the information capacity and information content of a system do not change as a result of compressible patterns in the system, or its statistical structure, but are instead simply a function of the number of states of the system. Similarly, the physical properties of a system cannot be "compressed" even if there are representational patterns to the values of the properties of the system. For example, if a system consists of identical elements, then we could represent and encode a single element, and store a program that iteratively generates a specified number of copies of that encoding, as opposed to representing and encoding each element. However, as a physical matter, we cannot "compress" the elements of that system simply because they are identical, or generate an arbitrary number of identical elements given all of the properties of a single element. Similarly, if a system has a particularly high probability state, we cannot reduce the physical substance of the system in that state, the way we can reduce the length of a statistical encoding of that state. Thus, as a general matter, the notion of compression seems inapplicable to any measure of the amount of information physically contained within a system. Therefore, if we want to establish a physical notion of information, then the information content and information capacity of a system seem to be the appropriate measures. Because these two measures are equal, for simplicity, we reference only the information content of a system as the measure of the amount of information physically contained within the system.

Assumption 2.3. *The information content of a closed system S with a total energy of E_T is given by,*

$$\log(|S|) = \frac{E_T}{E_0}(\log(K) + \log(X)). \tag{8}$$

---

[25]We present no theory as to the rules governing the interactions between quanta, other than the very limited assumptions set forth in Section 3 below. As such, for simplicity, we assume that any number of quanta can occupy the same position.

Note that we can view equation (8) as the actual, physical information content of a closed system, since it is the maximum amount of information that can be physically stored (as a theoretical matter) within the system.[26] Further, note that for a closed system, X is presumably constant, and thus, the maximum number of states the system can occupy will be constant. Therefore, the information content of a closed system is a conserved quantity. Moreover, because each quantum of energy can be in any one of K states, each quantum of energy itself store can $\log(K)$ bits of information, independent of its position. Thus, we can calculate the amount of energy required to store a given number of bits I using the following:

$$E = E_0 \frac{I}{\log(K)}. \tag{9}$$

Because E_0 and $\log(K)$ are both constants, it follows that energy can be viewed as proportional to information. Similarly, we can calculate the number of bits that can be stored in a given amount of energy using the following:

$$I = \log(K) \frac{E}{E_0}. \tag{10}$$

Finally, if we observe a given number of bits I by repeatedly measuring the basis properties of a system S upon each click over time, then we can express the amount of time that has elapsed as $t = t_0 \frac{I}{\log(|S|)}$.

3 Time-Dilation

In this section, we present a model of matter as a combinatorial object. We begin with a model for how quanta of energy change states over time, thereby causing systems to change states over time, and show that this process can serve as an explanation for elementary particle decay, the constant velocity of light, and time-dilation itself. This requires making certain assumptions about the structure of elementary particles, that, while presented in a technical manner, amount to little more than assuming that, in the absence of an interaction with another particle, the mass energy and kinetic energy of an elementary particle are constant, and that the kinetic energy of an elementary particle is distributed evenly within its mass.

3.1 Elementary Particles

We assume that the collective states of the mass quanta within an elementary particle generate the particle itself, and can be thought of as collectively "coding" for the particular particle that is generated. For example, assuming m_e

---

[26]Since X will presumably be a function of the volume of the system, and $\log(K)$ is a constant, we can also express equation (8) as $\log(|S|) \approx \frac{E_T}{E_0}(r + \log(sV))$. Thus, if the amount of information observed per measurement of S is fixed, then the number of measurements necessary to observe all of the information contained within S will increase as a function of the energy of S, and its volume.

is the mass of an electron, we would view an electron as being generated by $N = \frac{m_e c^2}{E_0}$ mass quanta, in some particular set of collective states, that together code for the properties of an electron. As a result, we represent the mass quanta within an elementary particle as a complete graph, where each mass quantum is represented as a vertex, with each mass quantum vertex connected to all other mass quanta vertices. Note that we are not suggesting that the quanta are necessarily "bonded" together in some fashion, but rather, we use discrete graphs because they are a convenient representational schema for the concepts we will make use of throughout the remainder of this paper. Further, we assume that the state of each quantum of kinetic energy within an elementary particle is fixed, and codes for a particular amount of displacement in a particular direction of motion, that, in the absence of an interaction with some other force or particle, never changes. As a result, we assume that direction of motion is quantized. Since direction of motion appears to be continuous, K must be extremely large.[27] Finally, we assume that each kinetic quantum can cause only one mass quantum to change position upon any given click, and thus, we represent each kinetic quantum as an "additional" vertex that is adjacent to exactly one mass quantum vertex. Thus, we assume that the mass quanta within an elementary particle can change positions independently of each other. However, we show in Section 3.3 below that the relative positions of the mass quanta within an elementary particle are approximately constant over time.
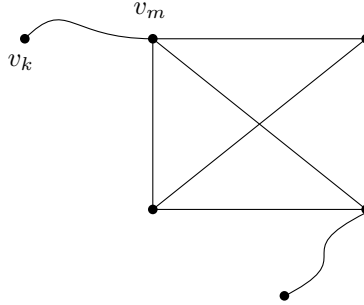

Figure 2: A network of quanta with four clusters.

We call each connected graph of quanta a **network**, and the set of all kinetic quanta adjacent to a given mass quantum, together with that mass quantum, a **cluster**. In Figure 2, v_k is a kinetic quantum, and v_m is a mass quantum. Together, v_k and v_m constitute a cluster. Thus, each network represents a single **elementary particle**,[28] and each cluster within a network will contain exactly one mass quantum vertex, and any number (including zero) of kinetic quanta vertices. The particular particle that is generated by the network will be

---

[27]This also implies an absolute frame of reference from which we can determine direction of motion, which we will address in Section 5 below.

[28]We are not suggesting the existence of particles outside of the Standard Model of physics, but there is nothing within this paper that would prohibit the existence of such particles.

determined by the collective states of the mass quanta within the network, and the kinetic energy of the resultant particle will be determined by the number of kinetic quanta within the network. Further, we assume that upon each click, exactly one quantum of energy within each cluster will become **active**, which means that either the active quantum will change states, or its cluster will change position, depending upon whether the active quantum is in a mass state or a kinetic state, respectively.

Assumption 3.1. *Upon each click, exactly one quantum of energy contained within each cluster will become active, and the probability that a given quantum of energy within a cluster will become active is equal for all quanta within the cluster.*

Thus, upon each click, we assume that exactly one quantum of energy within each cluster will become active, and the action coded for by the active quantum will determine the next state of the cluster. As such, we can think of R as being applied to the particle upon each click, generating the next state of each cluster within the particle, which will cause a given cluster to either change position, or cause the mass quantum within the cluster to change states. Stated informally, each mass quantum within an elementary particle can be thought of as having a single "vote" on what actions take place within the particle upon each click, but these actions generally affect each cluster individually, and do not impact the particle as a whole. However, in Section 3.2 below, we discuss elementary particle decay, which we assume to be the result of mass quanta changing states over time, and thereby, eventually coding for another particle, or set of particles. Because there is only one quantum of mass energy within each cluster, the greater the kinetic energy within a cluster, the less likely it is that the mass quantum within the cluster will become active upon any given click.

Assumption 3.2. *If the mass quantum contained within a cluster becomes active upon a given click, then that mass quantum will change from its current mass state to its next mass state, and no other quanta within the cluster will change state or position.*

Note that Assumption 3.2 assumes that the mass energy within an elementary particle will spontaneously change states. Thus, even in the absence of an interaction with another particle, a quantum of energy that is in a mass state will transition through mass states over time, but will not transition to a kinetic state.

Assumption 3.3. *If a quantum of kinetic energy becomes active upon a given click, then the entire cluster of which the quantum is a part will change position, traveling the distance coded for by the active quantum, in the direction coded for by the active quantum, and no other quanta of energy within the cluster will change states.*

Thus, an active quantum of kinetic energy will cause its entire cluster to travel the distance coded for by the quantum, in the direction coded for by the

quantum, but will not cause any quanta within the cluster to change states. We define the position of a cluster as the position of its mass quantum, and assume that the relative positions of all quanta within a cluster are constant over time. Since mass quanta always remain in a mass state, and kinetic quanta never change states, in the absence of an interaction with another particle, both the mass energy and kinetic energy of an elementary particle will remain constant.

Assumption 3.4. *Given an elementary particle with a kinetic energy of E_K and mass energy of E_M, the time-average of the number of kinetic quanta attached to each cluster within the particle is equal for all clusters, and is given by,*

$$q_\mu = \frac{\frac{E_K}{E_0}}{\frac{E_M}{E_0}} = \frac{E_K}{E_M}. \tag{11}$$

Thus, we assume that the kinetic quanta within an elementary particle will move from one cluster to another over time, and that as a result, the time-average of the kinetic energy of each cluster within the particle will be equal, and given by $q_\mu E_0$. In short, we assume that the kinetic quanta within an elementary particle are equally distributed among its mass quanta over time. Note that the distribution of kinetic quanta among the clusters over time has no effect on the total energy of the particle, which remains constant.

3.2   Elementary Particle Decay

In Section 3.1 above, we noted that since the probability that any quantum within a given cluster becomes active is equal for all quanta within the cluster, it follows that the greater the kinetic energy within a cluster, the lower the probability that the single mass quantum within the cluster will become active and thereby change states. Specifically, Assumption 3.4 implies that the expected probability that the mass quantum within a given cluster will become active and thereby change states upon any given click is $p = \frac{1}{q_\mu + 1} = \frac{E_M}{E_T}$. We can find the probability that k mass quanta will become active upon a given click by assuming that the clusters become active independently of each other, which results in a binomial distribution where the number of trials equals the number of clusters, which is also the number of mass quanta, and each active mass quantum is treated as a "success" event.

Assumption 3.5. *The probability that k mass quanta change states upon any given click within an elementary particle with a total of N_M mass quanta is given by,*

$$p_k = \binom{N_M}{k} p^k (1 - p)^{N_M - k}. \tag{12}$$

Thus, the expected number of active mass quanta upon any given click is $p N_M$. If we assume that the particle is stationary, with no kinetic energy, then

$p = 1$, and as such, every quantum of mass energy within the particle will change states upon each click. Assume that the expected lifetime of the particle is t_μ seconds when the particle is stationary. We can express this expected lifetime as a number of clicks given by $k_\mu = \frac{t_\mu}{t_0}$. Since the particle is stationary, upon each click, every mass quantum within the particle will change states, and thus, the total number of state changes after k_μ clicks is simply $k_\mu N_M$. We view the particle's decay as the result of the mass quanta within the particle changing states over time, thereby eventually coding for a different particle, or set of particles. As such, we view $k_\mu N_M$ as the expected total number of state changes that will cause the particle to spontaneously decay.

We view the conservation of total energy during this process as the result of the nature of energy itself, which we view as a physical substance that cannot be destroyed, but instead, changes states over time. In contrast, we view the conservation of other properties, such as mass, momentum, and charge, as the result of the physical transitions codified by R. That is, given the present state of some set of quanta of energy, only certain future states are accessible from that present state, namely those that can be generated by successive application of R to the present state. Thus, given any elementary particle, the set of particles that it can decay into will be a function of the total energy of the particle, and R. Similarly, the expected lifetime of the particle will also be a function of R. For example, if the next state of each mass quantum within an elementary particle is always the current state, then the particle will be perfectly stable. Similarly, if the mass quanta are stuck in a "loop" of states that all code for the same particle, then the particle will be perfectly stable in that case as well.
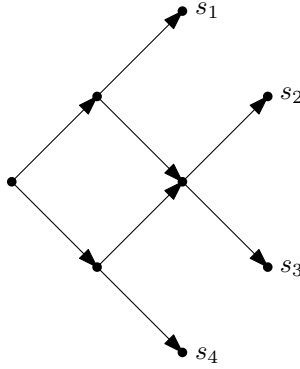

Figure 3: Sequences of collective states.

As a general matter, we can create any expected lifetime by simply assuming that at some point, the collective states of the mass quanta within the particle will eventually code for another particle, and adjust the length of that sequence to the desired lifetime of the particle. More realistically, if we allow for n possible sequences of collective states, each of which leads to a collective state that codes for another particle, or set of particles, then the expected lifetime of the particle

is the expected value of the amount of time until decay across all such sequences of collective states. For example, in Figure 3, each non-terminal node represents a collective set of states that all code for the same particle, and each terminal node represents a collective set of states that causes the particle to decay into some other particle, or set of particles. As such, the number of clicks until decay along each sequence of states terminating at s_2 is 3 clicks. Assuming the branching probability is $\frac{1}{2}$ at each node, the total probability of arriving at any s_i is $\frac{1}{4}$,[29] and thus, the expected lifetime of the particle is $\frac{5}{2}$ clicks. As a general matter, if k_i is the number of clicks until decay along sequence i, and p_i is the probability of sequence i, then the expected lifetime of the particle is $t_\mu = t_0 \sum_{i=1}^{n} p_i k_i$.

Now assume that a particle with an expected lifetime of t_μ seconds when stationary has some velocity, and thus, some kinetic energy. In that case, $p < 1$, and as such, the total expected number of state changes after k clicks is $kpN_M < kN_M$. Thus, the number of clicks required to achieve any given expected number of state changes will increase as a function of the kinetic energy of the particle. Specifically, the number of clicks required to achieve an expected number of state changes equal to $k_\mu N_M$ will increase. Assuming $kpN_M = k_\mu N_M$, it follows that $k = k_\mu \frac{E_T}{E_M}$, which is the number of clicks required to achieve an expected number of state changes equal to $k_\mu N_M$. Since $k_\mu N_M$ is the expected number of state changes that will cause the particle to spontaneously decay, it follows that k is in this case the number of clicks that we expect the particle to survive. Let $\gamma = \frac{E_T}{E_M}$. Multiplying by t_0, we can find the expected lifetime of the particle generally, which is given by,

$$t = t_\mu \gamma. \tag{13}$$

If we assume that $\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$, then equation (13) is consistent in form with the special theory of relativity. However, the reason for this result is due to the presence of kinetic energy within the particle. Further, t_μ is a measure of objective time, which we will discuss in greater detail in Sections 3.4 and 5.1 below. Thus, over time, the particle will mature through its states at an objectively slower rate, simply because it is less likely that its mass quanta will become active upon any given click, resulting in a longer expected lifetime.[30]

3.3   Kinetic Energy and Velocity

Because each quantum of kinetic energy within an elementary particle causes the cluster of which it is a part to travel a particular distance in a particular direction when active, we can think of each kinetic quantum within an elementary particle

---

[29]Note that there are two sequences of collective states that terminate at each of s_2 and s_3, each of which has a probability of $\frac{1}{8}$.

[30]Due to the complexity of composite particles, we do not present an analogous model that is specific to the spontaneous decay of composite particles. However, we do present a general model of time-dilation in Section 3.4 below, which would apply to all systems, including composite particles.

as a velocity vector v_i, such that $\|v_i\|$, which denotes the norm of the vector, gives the distance traveled by the cluster when that quantum is active. However, we do not assume that all kinetic quanta within a particle code for the same amount of displacement or direction of motion. Thus, both the displacement per click and direction of motion of a cluster can change upon any given click. If p_0 is the initial position of a cluster, and the cluster changes position m times, then we can express the final position of the cluster as,

$$p_f = p_0 + v_{i_1} + \ldots + v_{i_m}.$$

That is, because each kinetic quantum corresponds to a velocity vector, to find the final position of the cluster, we can simply add the velocity vectors that correspond to the kinetic quanta that were active over the interval of time under consideration to the initial position of the cluster. Note that the expected probability that a cluster will change position upon a given click is $1 - p = \frac{E_K}{E_T}$. As such, we can view m as an expectation value, where if k is the number of clicks that have elapsed, then $m = (1 - p)k = \frac{E_K}{E_T}k$. Let $N_K = \frac{E_K}{E_0}$ be the number of kinetic quanta within an elementary particle, and let $v_T = \sum_{i=1}^{N_K} v_i$, which is the sum over all of the velocity vectors corresponding to the kinetic quanta within the entire particle.

Assumption 3.6. *For any given cluster within an elementary particle, the frequency of each v_i in the sum $p_0 + v_{i_1} + \ldots + v_{i_m}$ approaches $\frac{1}{N_K}$ as k increases.*

That is, because we assume that kinetic quanta move from one cluster to another over time, we assume that, as a result, the frequency of each v_i in the sum $p_0 + v_{i_1} + \ldots + v_{i_m}$ approaches $\frac{1}{N_K}$ as k increases, and thus, the number of instances of each v_i approaches $\frac{m}{N_K}$.[31] Therefore, over any interval of time, for each cluster, $p_f \approx p_0 + \frac{m}{N_K}v_T$. Further, let p_{0_i} and p_{0_j}, and p_{f_i} and p_{f_j}, be the initial and final positions of two clusters, respectively. It follows that, $\|p_{f_i} - p_{f_j}\| \approx \|p_{0_i} - p_{0_j}\|$, and thus, the relative positions of the clusters are approximately constant over time. In short, we assume that all of the kinetic quanta within an elementary particle circulate among the clusters, causing the overall motion of the individual clusters to be nearly identical over time.

Because each quantum of kinetic energy can code for different directions of motion, the actual total displacement of a given cluster could exceed the straight line distance between its initial and final positions. However, we assume that the displacement per click, and the distances between the clusters, are both so small, that the overall motion is perceived as a single particle traversing

---

[31] For example, if all of the kinetic quanta within an elementary particle circulate among the mass quanta within the particle as a single group, in a "loop" around the mass quanta, then each kinetic quantum will be attached to each mass quantum exactly once over each pass around the loop. That is, the kinetic quanta repeatedly traverse a Hamiltonian circuit around the complete graph formed by the mass quanta. In that case, each kinetic quantum will be equally likely to become active within any given cluster, and thus, the frequency with which each kinetic quantum is active within a given cluster will be equal for all kinetic quanta over time, and equal within all clusters over time. Because t_0 is presumably extremely small, we assume that this frequency is almost exactly $\frac{1}{N_K}$ for all intervals of time.

a straight line. Specifically, we assume that the observed displacement of a particle after k clicks is $\frac{m}{N_K}\|v_T\|$, and thus, the observed velocity of the particle is given by,

$$v = \frac{\frac{m}{N_K}\|v_T\|}{kt_0} = \frac{\|v_T\|E_0}{E_T t_0}. \tag{14}$$

Let $\mu = \frac{\|v_T\|}{N_K}$, which can be thought of as the effective displacement per kinetic quantum.[32] As such, we can express equation (14) as,

$$v = \frac{E_K \mu}{E_T t_0}. \tag{15}$$

Note that none of the analysis above places any restrictions on the value of μ, and thus, we can always assume that $\frac{E_K \mu}{E_T t_0} = c\sqrt{1 - (\frac{E_M}{E_T})^2}$, which implies that $\mu = ct_0 \sqrt{\frac{E_T^2 - E_M^2}{E_K^2}}$.[33] In that sense, equation (15) is consistent with the special theory of relativity. However, because our analysis places no restrictions on the value of μ, equation (15) does not imply a maximum velocity, which is not consistent with the special theory of relativity. Specifically, equation (15) does not imply that velocity is always bounded by c. This would allow, for example, a particle with mass to travel at a velocity of c, which is consistent with experimental evidence that suggests that neutrinos have a non-zero mass [9], and a velocity of c [2]. Thus, while we assume that $\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$ throughout this paper in order to demonstrate consistency between our model and the special theory of relativity, equation (15), and the neutrino itself, suggest the possibility that there is a more general formulation of the value given by $\gamma = \frac{E_T}{E_M}$ that does not restrict the velocity of a particle to a value that is bounded by c.

Further, note that for any given μ, Equation (15) is maximized if $E_T = E_K$. That is, equation (15) is maximized when the particle consists solely of kinetic quanta. Further, if a given particle consists solely of kinetic quanta, and all such quanta are identical, then each quantum will code for the same amount of displacement in the same direction. As such, in that case, $\|v_T\| = N_K \|v_i\|$, and thus, $\mu = \|v_i\|$ will be constant, regardless of the total energy of the particle. Thus, if a particle consists solely of identical kinetic quanta, then the particle will have a constant velocity of $\frac{\mu}{t_0}$, regardless of its total energy.

As such, we assume that every quantum of energy within a photon is in a kinetic state, and that for any given photon, all quanta within the photon are identical. Further, we assume that all quanta within all photons code for the same amount of displacement. However, unlike elementary particles with mass,

---

[32]Note that if $E_K = 0$, then μ is undefined. However, in that case, the particle is by definition stationary, and thus, we assume that μ is 0.

[33]That said, because we have not proposed a value for t_0, we cannot, as a practical matter solve for μ. Nonetheless, the point remains, that as a theoretical matter, there are no restrictions on the value of μ that would prevent us from solving for its value. Further, note that equation (15) is an example of a computable function on the basis properties of an elementary particle.

we assume that the photon consists of a single cluster, where the "central" quantum of that cluster is in a kinetic state, not a mass state.[34] Note that because we assume that a photon is comprised of identical kinetic quanta, it follows that $\frac{E_K}{E_T} = 1$, and that μ is constant regardless of the photon's energy. Further, because all quanta within all photons code for a single displacement, the value of μ is uniform for all photons. As such, all photons will have a velocity of $\frac{\mu}{t_0}$, which we assume to be c, and a mass of 0, regardless of whether the source that ejected the photon is moving or stationary, and regardless of the energy of the photon.

Finally, note that because we allow for different kinetic quanta to code for different displacements per click, two particles could have the same kinetic energy and mass energy, yet have different values for μ, and therefore, different velocities, even if all kinetic quanta within both particles code for the same direction of motion. Thus, equation (15) implies that we cannot determine the kinetic energy of a particle by measuring its mass and velocity alone, without making additional assumptions about the value of μ. Similarly, we cannot determine the velocity of a particle based upon its kinetic energy and mass energy alone. As such, equation (15) would allow for particles that, like the photon, have constant velocities regardless of their energy.

3.4 Moving Clocks and Absolute Time

Consider a stationary clock. Even though the clock is stationary, because it is a clock, it will have some internal machinations that allow it to keep time, and thus, some small amount of kinetic energy. When the clock is stationary, all of the energy contained within the clock codes for the clock itself, which we call "clock quanta". That is, the mass energy of the clock generates the substance of the clock, and the small amount of kinetic energy within the clock generates the machinations of the clock that cause it to keep time. However, if the clock has some overall velocity, then the kinetic energy associated with that velocity will code for the overall velocity of the clock, and not its mass or its internal machinations. In order to analyze the energy of a moving clock, we will first need to expand Assumption 3.4, which is limited to elementary particles, to all closed systems generally.

Assumption 3.7. *Given a closed system with a total mass energy of E_M, the time-average of the number of kinetic quanta that code for the overall velocity of the system that are attached to each cluster within the system is equal for all clusters, and is given by $Q_\mu = \frac{E_K}{E_M}$, where E_K is the total kinetic energy of the system that codes for its overall velocity.*

Assumption 3.7 is a generalization of Assumption 3.4, and implies that, over time, the kinetic quanta that code for the overall velocity of a system are evenly distributed among the mass quanta within the system. As such, Assumption 3.7

---

[34] We will discuss the structure of the photon, and the properties of light in greater detail in Sections 3.5 and 5.2 below.

would not apply, for example, to the kinetic energy that codes for the internal machinations of a stationary clock, since that energy does not code for some overall velocity, and therefore, would not be evenly distributed among the mass of the clock, but would instead, generally be concentrated in the parts of the clock that cause it to keep time. In contrast, Assumption 3.7 would apply to the kinetic energy that codes for the overall velocity of a moving clock. Further, the amount of time required to achieve this uniform distribution will of course depend upon the system. For example, the distribution of kinetic energy that codes for the overall velocity of a solid object should generally be uniform among the mass of the object, since the rate of collisions between the particles within the object is likely going to be high, causing any uneven distributions of that kinetic energy to be transitory. In contrast, a gas trapped in a volume that is accelerated in a particular direction could require a significant amount of time before the kinetic energy that codes for the resultant overall velocity of the volume is evenly distributed among the particles within the gas. Note that Assumption 3.7 does not imply that the total kinetic energy of each particle within the gas will be equal, or that the motions of the particles within the gas will be identical, but rather, that the average number of kinetic quanta that code for the overall velocity of the volume attached to a given cluster will eventually be equal for all clusters within the volume. Thus, the particles within the gas will still have idiosyncratic motions, just like a moving clock will still have internal machinations, but Assumption 3.7 implies that the particles within the gas will eventually all share in the kinetic energy that codes for the overall velocity of the volume as a whole, in an amount that is, on average over time, in proportion to their respective masses.

In the case of a moving clock, we want to know, upon any given click, how many active quanta actually code for the clock itself, as opposed to its overall velocity. Thus, we are not interested in the expected number of active mass quanta per click, but rather, the expected number of active clock quanta per click. We can calculate this by first calculating the probability that any given clock quantum is active within a given cluster, and from that probability, calculate the probability that any given number of clock quanta are active within the entire clock, and finally, find the expected value of that distribution. As such, we view any active clock quantum within a cluster as a "success" event for purposes of our probability distribution. Note that the number of trials is equal to the number of clusters $N_M = \frac{E_M}{E_0}$ within the system, just as it is in the case of an elementary particle. Further, we assume that the clusters are independent of each other, in that the particular quantum that is active within a given cluster is not in any way a function of which quanta are active in any other clusters, and thus, the resultant distribution is again a binomial distribution.

Because the clock has some overall velocity, the expected energy of each cluster within the clock will consist of one quantum of mass energy, some very small fractional number of kinetic quanta that code for the machinations of the clock, which we denote by e, and the cluster's share of kinetic quanta that code for the overall velocity of the clock, which we denote by Q_μ. Since an active mass quantum or an active kinetic clock quantum within a cluster would constitute a

success event, the exact expected probability of success is $\frac{1+e}{1+e+Q_\mu}$. The expected value of e is the number of kinetic clock quanta divided by the number of mass quanta, which is also the ratio of the kinetic energy divided by the mass energy of the clock when stationary. It follows that $\frac{1+e}{1+e+Q_\mu} = \frac{E_M+E_C}{E_T}$, where E_C is the kinetic energy of the clock when stationary. However, we assume that E_C is so small when compared to E_M, that the actual distribution is approximated by a distribution with a probability of success given by $p = \frac{E_M}{E_T}$, and that as a result, the expected number of active clock quanta per click is approximately pN_M.[35]

When the clock is stationary, upon each click, all active quanta are clock quanta, and thus, the total number of active clock quanta after k clicks is simply kN_M. We assume that for sufficiently large k, for every kN_M of active clock quanta, there is a single value τ such that the time reported by the clock is incremented by approximately $kN_M\tau$. That is, each active clock quantum can be thought of as causing the time reported by the clock to increase by τ.[36] If the clock is accurate, then when it's stationary, after k clicks, the time reported by the clock should increase by kt_0. Thus, $kt_0 = kN_M\tau$, which implies that $N_M\tau = t_0$. When the clock is moving, the expected number of active clock quanta after k clicks is kpN_M. Therefore, the expected total change in time reported by the clock after k clicks is $t = k\frac{E_M}{E_T}N_M\tau = \frac{kt_0}{\gamma}$.

We assume that $\bar{t} = kt_0$ is measured by a **zero energy clock**, which we define as a clock that consists of only clock quanta, and thus, no kinetic energy beyond that which is necessary to allow the clock to keep time.[37] As a theoretical matter, a single mass quantum in isolation would constitute a zero energy clock, since its state would change once per click. Thus, we assume that all active quanta become active simultaneously everywhere, and therefore, we treat the number of clicks that have elapsed as an objective measure of time. It follows that,

$$t = \frac{\bar{t}}{\gamma}. \tag{16}$$

Note that equation (16) calculates time-dilation relative to $\bar{t}$, which we as-

---

[35]Disregarding the value of E_C would be reasonable for any system with internal machinations that have a kinetic energy that is several orders of magnitude smaller than the system's mass energy, which is likely to be the case for any reasonable construction of a clock. For example, even in the case of a clock driven by a pendulum, the value of $\frac{E_C}{E_M}$ will likely be proportional to $\frac{v^2}{c^2}$, where v is the velocity of the pendulum.

[36]As a practical matter, each clock quantum will not contribute equally to the clock's time keeping processes. For example, the mass quanta of an ordinary clock presumably code for stable matter, and thus, their changing states presumably do not contribute to the clock's time keeping processes. Thus, we are tacitly assuming that after a sufficiently large number of k clicks have elapsed, the average increment in time per active clock quantum always converges to a single value $\tau \approx \frac{\Delta_1+\Delta_2+\cdots+\Delta_k}{kN_M}$, where Δ_i is the increment in time reported by the clock upon the i-th click. Because t_0 is presumably extremely small, we assume that the total change in time reported by the clock can always be expressed as $kN_M\tau = \Delta_1+\Delta_2+\cdots+\Delta_k$.

[37]We will revisit this topic in greater detail in the context of inertial frames in Section 5.1 below.

sume to be a measure of absolute time, as measured by a zero energy clock, and thus, while equivalent in form, equation (16) is not entirely consistent with the special theory of relativity. Further, note that equation (16) is the mechanical result of the presence of additional kinetic energy that codes for the overall velocity of the clock. Moreover, while we have framed this discussion in terms of a moving clock, the same reasoning applies to any system with some overall velocity. That is, a system with some overall velocity will take more objective time to progress through the internal changes that would have occurred in an identical stationary system. Thus, all moving systems will undergo time-dilation in accordance with equation (16) due to the presence of kinetic energy within the system that codes for the overall motion of the system. For example, the biological processes of a living organism with some overall velocity would therefore progress slower, since the quanta that code for its biological processes will be active less often due to the presence of the kinetic quanta that code for its overall velocity.

Finally, note that it is ultimately the number of kinetic quanta that code for the overall velocity of a system that causes the system to experience time-dilation, not its velocity, which is, in our model, a consequence of that kinetic energy in the first instance. For example, even if two particles have the same mass energy and kinetic energy, the particles can nonetheless have different velocities if they have different values for μ. However, because both particles have the same mass energy and kinetic energy, the value given by $\gamma = \frac{E_T}{E_M}$ is the same for both particles. Thus, both particles will experience the same amount of time-dilation, despite traveling at different velocities. Therefore, as a general matter, we cannot determine γ based upon the velocity of a system alone, without making additional assumptions about the value of μ.

3.5   The Doppler Effect

Because we do not need to assume that the measured velocity of light is constant in all inertial frames in order to achieve time-dilation, we do not make that assumption.[38] As a result, the classical Doppler equations, which distinguish between the motions of a light source and a detector, would apply to our model, as adjusted to account for time-dilation, which we discuss in the paragraphs below. But rather than present the classical explanation for the Doppler effect, which is in no way inconsistent with our model, we present an explanation that is unique to our model of energy.

Specifically, because we view photons as consisting solely of kinetic quanta, we can view a light source as ejecting kinetic quanta. As such, over any interval of time, we can consider the average distance between quanta ejected by a stationary, monochromatic light source $\mu_s = \frac{1}{N} \sum_{i=1}^{N} l_i$, where each l_i is the distance between a given quantum and the quantum that is closest to it in the direction the light is traveling, and N is the number of such pairs of quanta

---

[38]We present an analysis of the velocity of light as measured within a moving inertial frame in Section 5.2 below.

traveling between the source and the detector at a given moment in time. We assume that μ_s is approximately constant over time. Thus, at any point along the detector in Figure 4, the expected time between quanta is $t_s = \frac{\mu_s}{c}$. As such, when the source and the detector are both stationary, we can express the expected total amount of energy incident upon any single point along the detector over an interval of time t as $E_s = E_0 \frac{t}{t_s}$.
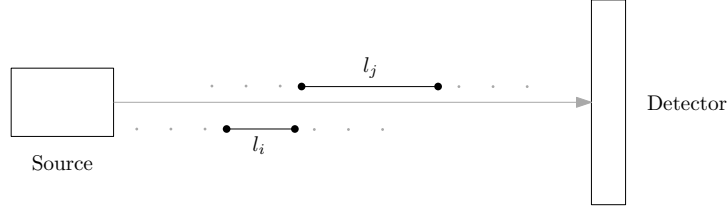

Figure 4: A light source ejecting quanta towards a detector.

As the velocity of the source changes, the average distance between each such pair of quanta will also change, just as the distance between each wavefront will change when we view the light source as generating a wave.[39] Thus, the amount of energy that is incident upon a given point along the detector over any interval of time will vary with the velocity of the source, increasing when the source moves towards the detector, and decreasing when the source moves away from the detector. Specifically, if the distance between two quanta is l_i when the source is stationary, then when the source has a velocity of v towards the detector, the distance between the two quanta will be $l_i - vt_s$. Thus, the expected distance between any two quanta will be $\mu_v = \frac{1}{N} \sum_{i=1}^{N} (l_i - vt_s) = \mu_s - vt_s$. As such, the expected time between quanta at any point along the detector is $t_v = \frac{\mu_s - vt_s}{c}$. Thus, the amount of energy incident upon any point along the detector over an interval of t seconds is $E_v = E_0 \frac{t}{t_v} = E_s \frac{c}{c-v}$.

Note that if the source is moving away from the detector, then the same analysis would apply, but v would in that case be negative. Further, if the source is moving at an angle of θ relative to the direction of its light, then $v = \cos(\theta)v'$, where v' is the velocity of the source.[40] Thus, as a source moves in the direction of its light, the ejected quanta become more concentrated, whereas when a source moves away from the direction of its light, the ejected quanta become more diffuse. If we assume that E_v and E_s are partitioned among an equal number of M photons, then $E_v = Mhf_v$, and $E_s = Mhf_s$, for some frequencies f_v and f_s. Under this view, the frequency of the light incident upon the detector is determined by the rate at which energy is incident upon a single point along the detector. In contrast, we view the luminosity of the light as

---

[39]For the moment, we will ignore any red-shift due to time-dilation, which we will address in the paragraphs below.

[40]That is, a given quantum is always "paired" with another quantum that lies on the same line connecting the source to the detector for purposes of calculating each l_i. As such, if the source is moving perpendicular to the direction of its light, that motion will not impact the average distance between the quanta.

determined by the rate at which energy is incident upon the entire surface of the detector.[41] Since $E_v = E_s \frac{c}{c-v}$, it follows that,

$$f_v = f_s \frac{c}{c-v}. \tag{17}$$

Now consider the case of a stationary source, and a detector moving towards that source. Since we do not assume that the velocity of light is constant in every inertial frame, it follows that an analogous effect would occur when the detector moves towards or away from a stationary source, causing the number of quanta incident at any point along the detector over any interval of time to increase as the detector moves towards the source, and decrease as the detector moves away from the source. However, the actual average distance between quanta is unchanged, since the source is stationary. Instead, the average amount of time between quanta at any point along the detector will vary with the velocity of the detector. Specifically, the expected time between quanta when the detector is moving towards a stationary source is $t_v = \frac{\mu_s}{c+v}$. In this case, it follows that,

$$f_v = f_s \frac{c+v}{c}. \tag{18}$$

Note that if the detector is moving away from the source, then the same analysis would apply, but v would in that case be negative. Further, if the detector is moving at an angle of θ relative to the direction of the light, then $v = \cos(\theta)v'$, where v' is the velocity of the detector. Moreover, note that in all cases, we simply change the time at which the quanta from the source arrive at the detector, and thus, total energy is conserved. That is, the quanta ejected by a moving source will be more concentrated than that of a stationary source if the source is traveling in the direction of its light, and more diffuse if the source is traveling away from the direction of its light. Similarly, a detector that is moving towards a stationary source will experience a more concentrated field of quanta, whereas a detector that is moving away from a stationary source will experience a more diffuse field of quanta, though in this case the actual distance between quanta is unchanged. Thus, in all cases, total energy is conserved, since the number of quanta ejected by the source over any interval of time is not affected by the motion of the source or the detector, except for the effects of time-dilation, which we now address in the paragraphs below.

We can account for the red-shift associated with time-dilation, while still conserving total energy, by assuming that the rate at which energy is ejected or absorbed by a system is affected by time-dilation in the same way that all internal processes of a moving system affected by the kinetic energy of the system experience time-dilation. Specifically, assume that a stationary source ejects quanta at a rate of $E_s = \bar{t}NE_0$ Joules every $\bar{t}$ seconds of absolute time, for some N. We assume that when the source has some overall velocity, it will eject quanta at a rate of $E_v = \frac{\bar{t}}{\gamma}NE_0 = \frac{1}{\gamma}E_s$ every $\bar{t}$ seconds. That is, it will take longer for the source to eject the same amount of energy as a consequence of time-dilation. If we assume that this energy is again in the form of M photons,

---

[41] Thus, a more luminous source will leave fewer "dark spots" on the surface of the detector.

then $Mhf_v = \frac{1}{\gamma}Mhf_s$, where f_s is the frequency of the source when stationary. Thus, $f_v = \frac{1}{\gamma}f_s$, causing a red-shift.

We assume that the same arguments in the paragraph above apply to a moving detector. Specifically, while the energy incident upon the surface of the detector is of course not affected by time-dilation's effects on the detector, we assume that the rate at which energy is exchanged by the atoms within the detector is affected by time-dilation, in the same way that the atoms within a source of light eject energy at a slower rate due to time-dilation. That is, we assume that time-dilation causes the atoms within the detector to exchange the energy of the photons incident upon the detector at a lower rate of energy per unit of time, just like a moving source will eject less energy per unit of time due to time-dilation. Thus, the atoms within a moving detector will exchange less energy per unit of time, ultimately resulting in a lower detected frequency of $\frac{1}{\gamma}f_s$.

If we assume that $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$, where v is the velocity of the source towards the detector, then combining the Doppler effect with the red-shift due to time-dilation yields the following equation for the frequency of light measured by a stationary detector:

$$f = f_s \frac{c}{c-v}\sqrt{1 - \frac{v^2}{c^2}}. \tag{19}$$

In the case of a stationary source, and a detector moving with a velocity of v towards the source, we have the following equation:

$$f = f_s \frac{c+v}{c}\sqrt{1 - \frac{v^2}{c^2}}. \tag{20}$$

Note that equation (19) is consistent in form with the special theory of relativity. However, equation (20) is not, since we have distinguished between the motions of the detector and the source. Applying the special theory of relativity, we would not distinguish between the motion of the detector and the motion of the source, but would instead only calculate their relative velocity. As such, we would treat the case of a detector moving towards a stationary source with a velocity of v in the same manner that we would treat the case of a source moving towards a stationary detector with a velocity of v. Thus, we can compare the energy per photon as predicted in that case by our model, with that predicted by the special theory of relativity, by taking the difference between equations (19) and (20), and multiplying by Planck's constant h:

$$\Delta = \frac{1}{\gamma}hf_s\left(\frac{c}{c-v} - \frac{c+v}{c}\right).$$

As such, Δ will be very small for even substantial values of v and f_s. In order to test the validity of equation (20), a detector could simply be accelerated

towards a stationary source, with the resultant detected frequencies measured and compared to the frequencies predicted by equations (19) and (20).[42]

Now assume that each quantum of energy ejected by a stationary, monochromatic source generates a wavefront. It follows that μ_s is the expected distance between the wavefronts generated by the source. Thus, $\mu_s = \lambda$, where λ is the wavelength of the light generated by the source. That is, we can think of μ_s as the average wavelength of the light generated by a source. Therefore, $M\frac{hc}{\lambda} = E_0\frac{t}{t_s} = E_0\frac{ct}{\lambda}$, for some interval of time t. It follows that $Mh = E_0 t$. Now assume that $M = 1$, and that as a result, $h = E_0 t_p$, for some interval of time t_p. As such, the total amount of energy incident upon a single point along the detector over an interval of t_p seconds is $E_0\frac{t_p}{t_s}$. Note that this is also the energy of a single photon generated by the source, which is given by $\frac{hc}{\lambda} = hf$. Because $hf = E_0\frac{t_p}{t_s}$, it follows that $f = \frac{1}{t_s}$, which is also the number of quanta ejected by the source that will be incident upon a single point in space over an interval of 1 second. As such, we can interpret the frequency of a light source f as the number of quanta ejected by the source that will be incident upon a single point in space over an interval of 1 second. Further, we can interpret Planck's equation hf as the total amount of energy incident upon a single point in space over an interval of t_p seconds.
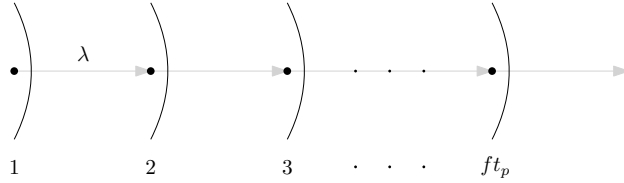

Figure 5: A photon as a discrete series of wavefronts.

Under this view, a photon consists of a discrete series of wavefronts, generated by a discrete series of quanta, that are incident upon a single point in space over an interval of t_p seconds, as set forth in Figure 5. It follows that the number of wavefronts and the number of quanta within a single photon is simply ft_p. Note that this does not explain single photon interference, since a single photon would, for example, still traverse a single path.[43] However, under this view, a group of photons would consist of multiple, discrete sets of wavefronts, which could interfere with each other like a classical wave.[44] Thus, our model could explain how light can be viewed as a wave, while nonetheless always producing energy in multiples of hf. As such, upon each click, exactly one quantum within

---

[42]Note that experiments such as [15] involving a rotating Mössbauer detector would not constitute tests of the validity of equation (20), since the detector is not moving towards the source in these experiments, but is instead rotating around the source.

[43]We are of course not suggesting that photons necessarily traverse a single path, but rather, suggesting that our model is necessarily incomplete, and that perhaps another model like Quantum Mechanics could explain single photon interference in conjunction with the model of time-dilation that we present herein.

[44]We discuss wave interference in Section 3.8 below.

a photon will become active, which will cause each quantum, and therefore each wavefront, within the entire photon to change positions. The amount of displacement and direction of motion will be determined by the displacement and direction coded for by the active quantum, which will be the same for all quanta within the photon. Because the notion of a cluster makes use of discrete graphs, it does not matter for this purpose that the photon is a "horizontal" particle, or that the quanta are separated by what could be substantial distances. However, because we allow for the possibility of a single quantum in isolation to be ejected by an atom, we assume that the first quantum within a photon is the "central" quantum of the photon. For example, in Figure 5, the quantum labeled ft_p would be the central quantum of the photon.

Note that because $E_0 = \frac{h}{t_p}$, it follows that if $\lambda > ct_p$, then $\frac{hc}{\lambda} < E_0$. As such, because we assume that E_0 is the minimum energy of any particle, it follows that we would view quanta separated by a distance of greater than ct_p meters as discrete particles, as opposed to a single photon. Thus, as a theoretical matter, we can view a single quantum of kinetic energy in isolation as a photon with an infinite wavelength, since there is no "next" quantum in that case. That is, $t_s = \infty$, and as such, $E_0 \frac{t_p}{t_s} = 0$. This is consistent with Planck's equation, which would give an energy of 0 for a photon with a wavelength of $\lambda = \infty$. Nonetheless, the energy of that single quantum is simply E_0. While it might seem odd that a single quantum in isolation, which has a definite non-zero energy of E_0, would produce a zero energy when viewed as a photon, it is consistent with our interpretation of a photon as an amount of energy incident at a single point over a period of t_p seconds. Thus, if the distance between the quanta ejected by the source is greater than ct_p, then the number of quanta that are incident upon a single point every t_p seconds will be less than 1. If there is no "next" quantum, then the number of quanta incident upon a single point per t_p seconds is 0, which is distinguishable, for example, from a source that ejects 1 quanta every t_p seconds, which would have a wavelength of exactly ct_p, and an energy of exactly E_0. If this model of the photon is correct, and t_p is not an immeasurably small interval of time, then t_p could serve as a measure of absolute time. Finally, this implies that Planck's constant h is equal to the product of the minimum energy E_0, multiplied by the amount of time it takes for the energy of a single photon to be incident upon a single point in space, t_p.

3.6 Matter as a Wave

As noted in Section 3.1 above, if m is the mass of a stationary elementary particle, then the number of quanta within the particle is given by $\frac{E}{E_0} = t_p \frac{mc^2}{h}$. In the case of a photon, the number of quanta is given by $t_p f$. If we assume that the particle is a "horizontal" particle with a length of $t_p c$, then we can view $f_c = \frac{mc^2}{h}$, which has units of $\frac{1}{seconds}$, as the "frequency" of the particle. Note that $f_c = \frac{c}{\lambda_c}$, where $\lambda_c = \frac{h}{mc}$ is the Compton wavelength of the particle. Thus, we can view λ_c as the actual distance between mass quanta within the particle. Under this view, a stationary elementary particle would consist of

34

a discrete series of standing wavefronts, each separated by a length of λ_c. If the particle has some velocity, then we assume that the kinetic quanta in that case would not generate wavefronts, but would instead be "attached" to the mass quanta within the particle, driving its motion. Thus, the wavelength of a particle with mass does not change as a result of a change in its kinetic energy. As noted above, this view of an elementary particle as a discrete series of wavefronts would not explain single particle interference. Thus, our model is not in any way a substitute for quantum mechanics. Rather, our model allows all elementary particles to be described in terms of a discrete series of wavefronts. Interestingly, it also shows that the Compton wavelength is not necessarily a product of quantum mechanics, but can viewed as the wavelength of a quasi-classical particle that consists of a discrete series of wavefronts.

Now imagine that the particle is traveling with a velocity of v. It follows that the rate per second at which the mass quanta within the particle will be incident upon a single point in space is given by $f_v = \frac{1}{t_v} = \frac{v}{\lambda_c}$, where t_v is the time between mass quanta, and λ_c is the Compton wavelength of the particle, which we view as the actual distance between the mass quanta within the particle. Note that $f_v = \frac{mcv}{h}$, which is also $\frac{c}{\lambda_v}$, where $\lambda_v = \frac{h}{mv}$ is the de Broglie wavelength of the particle. It follows that the amount of energy contained within the mass quanta that will be incident upon a single point in space over a period of t_p seconds is given by $E = E_0 t_p f_v = mcv$. If we assume that $v = c$, it follows that $E = mc^2$, and $f_v = f_c$. Thus, we can interpret Einstein's equation $E = mc^2$ as the amount of energy contained within the mass quanta that will be incident upon a single point in space over a period of t_p seconds, assuming the particle has a velocity of c.[45] This is the same amount of energy that would be incident upon a single point in space if all mass quanta within the particle transitioned from a mass state to a kinetic state, and formed a single photon. In that case, the wavelength of the resultant photon would equal the Compton wavelength of the particle, since both particles would contain the same number of quanta, and both particles would be horizontal particles with a length of ct_p. Note that this also follows from assuming that $\frac{hc}{\lambda} = mc^2$.

Finally, note that the classical momentum of the particle is given by $\rho = f_v \frac{h}{c} = mv$. Generalizing upon this observation, it follows that the momentum of a photon is given by $f \frac{h}{c} = \frac{h}{\lambda}$, which is in fact the momentum of a photon. Thus, the classical momentum of a particle is proportional to an amount of energy contained within the particle that is incident upon a single point in space over an interval of t_p seconds. In contrast, the special theory of relativity would express the momentum of a particle as γmv. Note that each cluster within the particle has an energy of $E_0(1 + \frac{E_K}{E_M})$. That is, each cluster consists of one quantum of mass energy, and the cluster's share of the total kinetic energy of the particle. If we assume that the kinetic quanta within each cluster occupy the same point in space occupied by the central mass quantum, then the total

---

[45]Though presented as a hypothetical, as noted in Section 3.3 above, it is possible in our model for a particle with mass to have a velocity of c.

energy incident upon a single point in space over a period of t_p seconds is given by $E = E_0 t_p f_v (1 + \frac{E_K}{E_M}) = \gamma mcv$. As such, the momentum of an elementary particle, as expressed in the special theory of relativity, is proportional to the total amount of energy within the particle that is incident upon a single point in space over an interval of t_p seconds. Because $\gamma = 1$ for a photon,[46] it follows that the momentum of a photon is the same in either case. This notion of momentum, as proportional to an amount of energy incident upon a single point in space over a period of t_p seconds, implies that an elementary particle must have a finite wavelength in order for the particle to have momentum or mass. However, we can still define the momentum and mass of a single quantum in isolation, by viewing a single quantum in isolation as a particle itself, even though a single quantum, as a technical matter, has an infinite wavelength. The momentum of a single quantum in a mass state is zero, simply because a single quantum in a mass state has a velocity of zero, and will therefore never be incident upon any point in space other than its position. In contrast, the momentum of a single quantum in a kinetic state is given by $\frac{h}{ct_p}$, which is proportional to the total amount of energy that will be incident upon any point in space along its path. The mass of a single quantum in a mass state is given by $\frac{h}{t_p c^2}$, which follows from the equation $E = mc^2$. We assume that the mass of a single quantum in a kinetic state is zero. Finally, note that assuming that $v = c$ implies that $E = \gamma mc^2$.

3.7   The Photoelectric Effect

Our description of the structure of a photon in Section 3.5 above is consistent with the photoelectric effect, since the number of photons ejected by a source will of course not impact the amount of energy contained within the individual photons ejected by the source. However, since we view the photon as a horizontal particle, we would interpret the threshold frequency of a material as the minimum rate at which energy must be incident upon a single point on the surface of the material in order to cause electrons to be ejected from the material. As such, it is the frequency of a source, and not its luminosity, that will determine whether or not electrons are ejected by a material when light from the source is incident upon the surface of the material. Thus, if light from a source with a frequency greater than or equal to the threshold frequency of a material is incident upon the material, then electrons will be ejected by the material.[47]

---

[46]Although the equation $\frac{E_T}{E_M}$ is undefined for a photon, there are no internal processes or machinations within a photon, other than the movement of the photon itself, which occurs uniformly upon each click. Moreover, we assume that the quanta within a photon are all arranged horizontally, and thus, exactly one quantum is incident upon a given point in space along its path at all times.

[47]Though we clearly have not presented any meaningful analysis of the dynamics of electrons within a material subject to incident light, we note that it is not necessarily the case that the threshold frequency of a material would be impacted by time-dilation, at least within our model, since only those interactions that are affected by the presence of kinetic energy would be affected by time-dilation. Though due to distinct processes in the nucleus of an

3.8   Wave Interference With Quantized Energy

Because we assume that energy is quantized, we cannot allow for wave interference to reduce the energy of a wavefront to below E_0, or cause a transfer of energy that is not an integer multiple of E_0. Instead, we assume that wave interference reduces the probability that the quanta generating two interfering waves will interact with a given system or particle, but does not change the energy contained in each wavefront, which we assume to be E_0. For example, consider a series of wavefronts that pass through a point in space, and assume that if a particle were present at that point in space, then the probability of an interaction between a single quantum generating a wavefront and the particle is p_1. That is, if a particle is stationary at a point in space though which the wave passes, then each time a wavefront crosses that point, the probability of an interaction between any such wavefront and the particle is p_1. Now assume that there is a second wave that also passes through that same point, and that in the absence of the first wave, the probability of an interaction between any given wavefront in the second wave and the particle would be p_2.
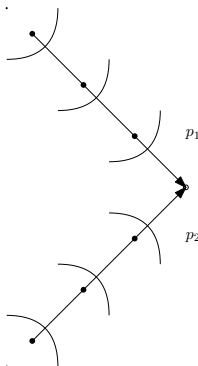

Figure 6: Two waves interfering at a point.

If the two waves perfectly destructively interfere, then we assume that the probability that either wave will interact with a particle at the point of interfer-

ence is zero. Similarly, if the two waves perfectly constructively interfere, then we assume that the probability that the first wave will interact with a particle at the point of interference is p_1, and the probability that the second wave will interact with a particle at the point of interference is p_2. That is, if the two waves perfectly constructively interfere, then we assume that the probabilities of interaction remain unchanged at the point of interference. If the two waves partially interfere with each other, then we assume that the probabilities of interaction will be reduced, but greater than zero. In each case, any such interaction will involve an integer number of quanta, and the transference of an amount of energy that is an integer multiple of E_0, though the expected energy involved in each such interaction could of course be any real number.

4 Time-Dilation Due to Gravity

4.1 Gravitational Quanta

Consider a single quantum in true isolation that is in a mass state. It follows that a completely empty vacuum extends in every direction from the point in space occupied by the mass quantum. We assume that upon each click, the mass quantum emits a single quantum of energy in each direction of motion,[48] each of which we call a **gravitational quantum**.
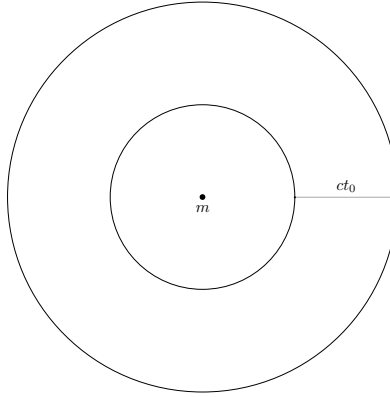

Figure 7: The gravitational field of a single mass quantum.

We assume that each such quantum travels in a fixed direction of motion, and at a fixed velocity of c, neither of which changes over time, in the absence of an interaction with another system or particle.[49] Further, we assume that the

---

[48]Recall that we assumed that direction of motion is quantized. However, we assume that K is so large that direction of motion can be treated as effectively continuous.

[49]We assume that gravitational quanta are emitted upon each click in order to allow for the possibility that a gravitational quantum passes through a given point in space upon each click. We assume that gravitational quanta travel at a velocity of c given experimental evidence that the speed of gravity is approximately c [1].

gravitational quanta emitted by the mass quantum "power" the gravitational field generated by the single mass quantum, by interacting with other systems and particles. Along any line emanating from the single mass quantum, we can view the gravitational quanta emitted by the mass quantum as generating a series of wavefronts, each separated by a distance of ct_0.
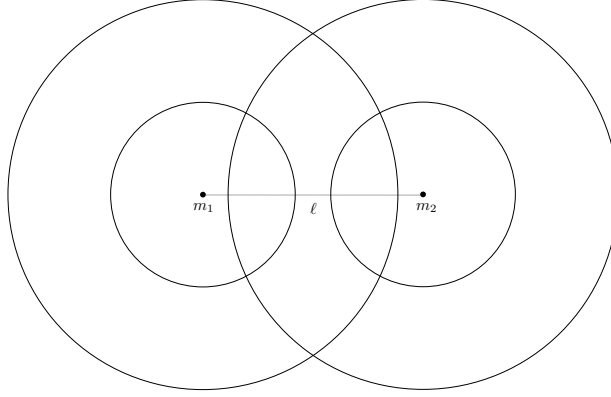

Figure 8: The gravitational field of two mass quanta.

Now consider a system S comprised of two otherwise isolated quanta of energy, both in a mass state, separated by a distance of ℓ.[50] Because each quantum is in a mass state, we assume that each quantum will emit gravitational quanta upon each click. Because the two streams of gravitational wavefronts emitted by S are separated by some distance ℓ, they will interfere with each other, both constructively and destructively, at different points in space, depending upon ℓ, thereby altering the probabilities of an interaction occurring at different points. Further, we assume that, upon any given click, the gravitational quanta emitted by S can interact with the individual clusters within a particle. If an interaction occurs between a given gravitational quantum and a given cluster, then we assume that no quanta within the cluster will be active upon that click. That is, the interaction between the gravitational field and the cluster occurs upon a given click in lieu of any quanta within the cluster becoming active. Therefore, if a particle interacts with a gravitational field, it will experience time-dilation.

Now consider a point in space that is at a distance of r from the center of mass of a system S with a mass of M. Let $p_G(r)$ denote the probability of an interaction between the gravitational field emitted by S and a cluster that is within a particle located at that point.[51] This probability will account for any interference that might occur between the gravitational wavefronts emitted by S at that point in space. Assume for simplicity that the particle in question has no kinetic energy. In this case, a given mass quantum within the particle will

---

[50]For simplicity, we assume that their positions are fixed, and thus, ignore any gravitational force of attraction between the two quanta.

[51]For simplicity, we treat the particle as a point particle, and as such, treat all clusters within the particle as located at a single point in space.
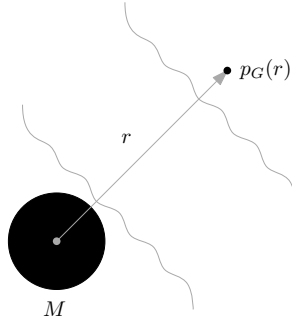
Figure 9: The gravitational field of a system with a mass of M.

be active so long as it does not interact with the gravitational field. It follows that the probability that the mass quantum within a given cluster within the particle is active upon any given click is simply $1 - p_G(r)$. If we again view the mass quanta of the particle as the "clock" of the particle, it follows that if $\bar{t}$ is the amount of absolute time that has elapsed, then the amount of time that has elapsed as measured by the particle is $t = \bar{t}(1 - p_G(r))$.[52] If we assume that $1 - p_G(r) = \sqrt{1 - \frac{2GM}{rc^2}}$, where G is the gravitational constant, then our model will be consistent with the general theory of relativity's prediction for the amount of time-dilation that will occur within the particle. That is, we can view $p_G(r) = 1 - \sqrt{1 - \frac{2GM}{rc^2}}$ as the probability of an interaction upon any given click between the gravitational field and a cluster within a particle that is at a distance of r from the mass generating the field, that takes into account all interference between the gravitational wavefronts.[53]

Though we assume gravitational quanta are emitted upon each click by each of the mass quanta within a system, we allow for interference between the individual gravitational wavefronts, which, as noted above, will ultimately determine the probability of an interaction between the resultant gravitational field and a particle at a particular point in space. Thus, regardless of the actual number of gravitational quanta that cross a given point in space per second, the expected number of interactions per second between a gravitational field and a particle at a particular point in space is given by $\frac{p_G(r)}{t_0}$. As such, we can interpret $f_G = \frac{p_G(r)}{t_0}$ as the frequency of the gravitational field at a distance

---

[52]Because we have assumed the particle is stationary, with no kinetic energy, this would again imply a binomial distribution, where the probability of "success" is $p = 1 - p_G(r)$. Using the analysis in Section 3.4 above, a stationary clock in a gravitational field at a distance of r from the mass generating the field will, after $\bar{k}$ clicks, report a total change in time given by $t = \bar{k}pN_M\tau$, where N_M is the number of mass quanta within the clock. Because $\bar{k}N_M\tau = \bar{k}t_0 = \bar{t}$, it follows that $t = \bar{t}p = \bar{t}(1 - p_G(r))$.

[53]For simplicity, we are assuming that the system generates a gravitational field that is identical at all points at a distance of r from its center of mass. This may not always be the case, and will instead depend upon the distribution of mass quanta within the system, which, as a practical matter, will depend upon the distribution of its mass.

of r from its center of mass, since it is by definition the expected number of interactions per second between the gravitational field and a particle that is at a distance of r from its center of mass. We can, therefore, also consider the wavelength of the gravitational field at a point in space, given by $\lambda_G = \frac{ct_0}{p_G(r)}$.

4.2   Potential Energy and the Conservation of Energy

Obviously, the emission of an effectively unlimited number of quanta from a single quantum does not conserve energy. However, fields by their very nature do not conserve energy, in the sense that systems can accelerate "for free" if they interact with a field. We can argue that the concept of potential energy allows for the conservation of total energy, even in the presence of fields that cause acceleration, but this begs the question of where the potential energy came from in the first instance. For example, two particles with mass that have no kinetic energy will accumulate kinetic energy if they are brought into proximity with each other due to gravitational attraction. These interactions cause the particles to have kinetic energy that did not exist beforehand. As such, the bottom line is that fields, such as gravitational fields, can act as net contributors of energy to the systems with which they interact. As a consequence, fields, by their very nature, do not conserve energy, but are instead a free source of energy for the systems and particles with which they interact.

Moreover, throughout this paper, we have disregarded the potential energy of systems and particles. This is because the potential energy of a system, in our model, plays no role in time-dilation, as we do not view potential energy as an amount of energy that is physically contained within a system. This view follows from the fact that mass energy generates a gravitational field that can be measured, and kinetic energy generates motion that can be measured, whereas the potential energy of a system generates no measurable effects at all. As such, we view the potential energy of a system or particle as an amount of energy that potentially could be contained in the system at some future time, assuming some future state of affairs. For example, a book on a shelf near the surface of the Earth is generally presumed *to have* potential energy due to the gravitational field of the Earth. Instead, we view the book as being constantly bombarded by gravitational quanta that have a certain probability of interacting with the book. This is distinct from the potential energy of the book, which we view as an amount of kinetic energy the book *would have* if it were to fall from the shelf. That is, the realization of potential energy requires the realization of some future state of a system that is different from its present state. This is in contrast to the kinetic energy and mass energy of a system, which is always measurably manifested in the present state of the system. Thus, by its nature, potential energy is not an amount of energy that is contained within the present state of a system. Returning to our example, if the book falls off the shelf, then we would argue that the book accumulates kinetic energy during its fall through its interactions with the gravitational quanta emitted by the Earth, as opposed to assuming that some quantity of potential energy that was always present in the book is converted into kinetic energy.

Similarly, if we were to instead throw the book upwards into the air, eventually it will reach a point where its velocity is zero, before falling back to the ground. We view the gradual reduction in the book's upward velocity as a result of the gravitational quanta interacting with the book, stripping the book of kinetic energy, and absorbing that kinetic energy, until the book completely stops moving, at which point the book begins to absorb gravitational quanta, causing it to fall. As a general matter, when a system loses kinetic energy in a gravitational field, we assume that the kinetic energy lost by the system becomes "attached" to the gravitational quanta within the field. In contrast, when a system gains kinetic energy in a gravitational field, we assume that the gravitational quanta become "attached" to the system. This is consistent with the fact that photons are "blue-shifted" as they fall into gravitational fields, and "red-shifted" as they escape gravitational fields.[54]

4.2.1 Rotational Kinetic Energy and Time-Dilation

We note that a rotating system traveling at a velocity of v should experience more time-dilation than an otherwise identical system traveling at a velocity of v in a straight line, since the rotational motion of the former system implies the repeated occurrence of interactions that change the direction of motion of the system, which we assume to cause additional time-dilation.

5 Inertial Frames and the Velocity of Light

5.1 Inertial Frames

We assume that a single quantum of mass energy in isolation would not just appear to be stationary, but rather, would in fact be stationary, since it has no kinetic energy with which it can move. As such, we can define a truly stationary coordinate system by reference to a set of orthogonal vectors whose distances from a single mass quantum are fixed. We call any such coordinate system an **absolute coordinate system**. The **absolute distance** between any two points, measured in meters, can then be determined by reference to any absolute coordinate system, and thus, we assume that the absolute distance between any two points is equal in all absolute coordinate systems. We define the change in **absolute time** as the number of clicks that have elapsed multiplied by t_0, which we assume to be equal to the change in time reported by any zero energy clock over that same number of clicks. We define **absolute displacement** as

---

[54]Though we have not presented a fulsome theory of gravity, there is nothing in our model that would require the curvature of space in order to have photons travel along a curved path. That is, there is nothing in our model that prevents the energy of a photon from being affected by forces such as gravity. Moreover, the simple phenomena of reflection and dispersion unambiguously demonstrate that the path of a light wave can be affected through interactions. It would therefore be perfectly consistent with our model to assume that a light wave repeatedly interacts with a gravitational field, thereby giving the wave the overall appearance of a curved path, when that path is actually the result of a discrete set of interactions that incrementally alter the path of the wave, much like a series of extremely small mirrors.

the absolute distance between the initial and final positions of a system over some interval of absolute time. We define **absolute velocity** as the ratio of the absolute displacement of a system over a given interval of absolute time divided by the length of the interval of time.

We define a **zero energy inertial frame** as any set of systems whose positions are constant in an absolute coordinate system. In general, we define an **inertial frame** as any set of systems whose absolute distances from each other are constant. We define the **velocity of an inertial frame** as the absolute velocity of any system within the inertial frame, which will be equal for all such systems. Thus, we can define a coordinate system by reference to a set of orthogonal vectors whose absolute distances from any system within a given inertial frame are constant. We call any such coordinate system an **inertial coordinate system**. We assume that the distance between any two points in a given inertial coordinate system is equal to the absolute distance between those two points. Thus, we reject the notion of length contraction implied by the special theory of relativity. However, we do not assume that measurements of displacement are constant in all inertial frames. Rather, in general, we define **displacement** as the distance, measured in meters, between the initial and final positions of a system in some inertial coordinate system over some interval of **time**, which we define as the change in time reported by a clock within the applicable inertial frame. Finally, we define **velocity** as the ratio of the displacement of a system over a given interval of time divided by the length of the interval of time.

Consider a simple example of a bus traveling with a constant absolute velocity of v meters per second. All seated passengers on the bus are part of the same inertial frame that includes the bus itself, and all other objects that are traveling at the same absolute velocity as the bus, which, therefore, appear stationary relative to the bus. However, if a passenger runs from one end of the bus to the other, that passenger is not part of the inertial frame of the bus for so long as the passenger is running, and is instead an object whose velocity can be measured from within the inertial frame of the bus. Assume that the bus is 12 meters long, that the bus is traveling forwards, and that the passenger runs from the rear of the bus to the front of the bus in 5 seconds of absolute time. Because we assume measurements of distance are uniform in all inertial frames, the length of the bus is uniform in both the inertial frame of the bus and any zero energy inertial frame. However, the displacement of the running passenger is not uniform in all inertial frames. Specifically, the displacement of the running passenger as measured from the inertial frame of the bus is 12 meters. In contrast, the absolute displacement of the running passenger is $5v + 12$ meters. Similarly, the amount of time that has elapsed is also not uniform across all inertial frames. Specifically, the absolute change in time is $\bar{t} = 5$ seconds, but a clock on the bus will report a change in time of $t = \frac{5}{\gamma}$ seconds. Thus, the absolute velocity of the passenger is $\frac{5v+12}{5}$ meters per second, whereas the velocity of the passenger as measured from within the inertial frame of the bus is $\gamma \frac{12}{5}$ meters per second.

Now assume that the passenger runs from the front of the bus to the back of the bus in 5 seconds. The absolute velocity of the passenger is therefore $\frac{5v-12}{5}$ meters per second, whereas the velocity of the passenger as measured in the inertial frame of the bus is again given by $\gamma\frac{12}{5}$. We assume that the reason the absolute velocity of the passenger is less than v while running is because the kinetic energy of the passenger decreases as the passenger runs from the front of the bus to the back of the bus. Because we assume that total energy is conserved, we assume that the kinetic energy lost by the passenger is gained by the bus. However, because the bus has a very large mass compared to the passenger, the impact of this additional kinetic energy on the velocity of the bus is negligible.

5.2 Measuring the Velocity of Light

In some sense, the measurements made in the inertial frame of the bus are simply incorrect, since they are based upon limited information. That said, historically, it has generally been assumed to be impossible for an observer to determine their absolute velocity. This assumption is not specific to the special theory of relativity, and in fact dates back to at least as early as 1632, when Galileo asserted that all motion is relative [11]. However, because of our model of energy, all motion is by definition absolute, since it is always driven by the presence of some definite number of kinetic quanta, even though measurements of displacement and time are not uniform across all inertial frames. Because our model implies a constant, and universal absolute velocity of light, but nonetheless allows for time-dilation and non-uniform measurements of displacement within inertial frames, we argue below that it is possible for an observer to determine their absolute velocity by measuring the velocity of light using a clock within the observer's inertial frame. Thus, our model implies that the measured velocity of light is not uniform in every inertial frame, though our model implies that the absolute velocity of light is always exactly c. This initially appears contrary to over a century's worth of experimental evidence demonstrating that the measured velocity of light is always extremely close to the exact value of c [10]. However, even though high precision measurements of the velocity of light on the surface of the Earth show little deviation from the exact value of c, these results simply imply that the measured deviations are small enough to disregard in the absence of a deliberate effort to search for the predictable, minute deviations we discuss below.

As discussed above, our model of the photon necessarily implies that the absolute velocity of a photon is constant regardless of the velocity of its source. Thus, a photon will always have an absolute velocity of c, regardless of the velocity of its source. However, we also assume that the direction of motion of a photon is independent of the velocity of its source.

Assumption 5.1. *If a source ejects photons at an angle of θ relative to some absolute coordinate system when stationary, then that source will eject photons at an angle of θ relative to that same absolute coordinate system when moving*

in any direction at any velocity.

As such, we assume that a source will always eject photons at the same angles, regardless of the velocity of the source.[55] Thus, a photon does not participate in the velocity of its source in any manner whatsoever, but is instead ejected at a constant velocity and in a direction that does not depend upon the direction of the motion of its source. For example, assume that an observer is part of an inertial frame with a velocity of v. Further, assume that the observer has an apparatus that consists of a clock, a light source, a mirror, and a detector located within the source, as set forth in Figure 10. Thus, the apparatus can measure the two-way velocity of light in that inertial frame. Further, assume that the distance from the source to the mirror is L, and thus, the round-trip displacement of a photon through the apparatus as measured within the inertial frame is $2L$. As noted above, we assume that the distance between the source and the mirror is uniform in all inertial frames, whereas the displacement of the photon over a round-trip, and the amount of time it takes for the photon to complete a round-trip, are not uniform in all inertial frames.
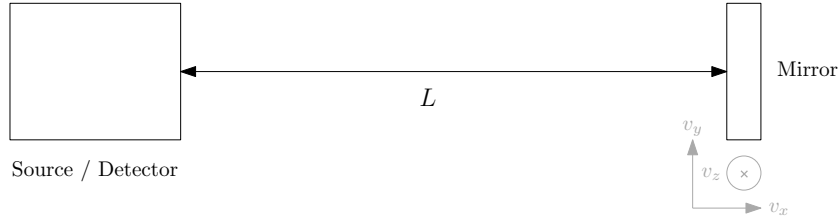

Figure 10: A light source with a detector.

Thus, the velocity of light as measured in a given inertial frame is not uniform across all inertial frames. Specifically, assume that the velocity of the inertial frame is $v = ||(v_x, v_y, v_z)||$, where v_x is the velocity of the inertial frame in the direction parallel to the line that connects the surface of the source to the surface of the mirror. It follows that the absolute displacement of the photon as it travels from the source to the mirror is $L + \bar{t}_1 v_x$, where $\bar{t}_1$ is the absolute time it takes for the photon to travel from the source to the mirror. That is, Assumption 4.1 implies that the velocity of the inertial frame in the y and z directions simply changes the location that the photon is incident upon the mirror and the detector, and does not alter the absolute displacement of the photon. Since the absolute velocity of light is c, it follows that the absolute displacement of the photon is also given by $c\bar{t}_1 = L + \bar{t}_1 v_x$, and thus $\bar{t}_1 = \frac{L}{c - v_x}$. Similarly, the absolute displacement of the photon over the return path from the mirror back to the source is $L - v_x \bar{t}_2$, and thus, $\bar{t}_2 = \frac{L}{c + v_x}$. Thus, the total

---

[55]We are assuming that the orientation of the source does not change as a result of its velocity. Further, note that we are of course not suggesting that aberration will not occur, though we do not address aberration in this article. Rather, we are assuming that the angle at which a photon is ejected is completely independent of the motion of its source.

absolute time and absolute displacement over a round-trip are, respectively, given by,

$$\bar{t} = L(\frac{1}{c - v_x} + \frac{1}{c + v_x}),$$ (21)

and,

$$\bar{d} = 2L + Lv_x(\frac{1}{c - v_x} - \frac{1}{c + v_x}).$$ (22)

As such, the absolute velocity over the round-trip is simply c. In contrast, the round-trip displacement as measured within the inertial frame is $2L$, and the total change in time as measured within the inertial frame is $t = \frac{\bar{t}}{\gamma}$. Note that γ is a function of the kinetic energy of the apparatus, which, as discussed in Section 3.4 above, cannot be determined by reference to the absolute velocity of the apparatus alone, without making additional assumptions about the value of μ. In contrast, the absolute displacement $\bar{d}$ is a function of only the absolute velocity of the apparatus in the x direction. Thus, the velocity of light, as measured within in the inertial frame using the apparatus set forth in Figure 10, is given by,

$$c_v = \gamma(c - \frac{v_x^2}{c}).$$ (23)

Now assume that we rotate the apparatus. It follows that v_x, which is a measure of absolute velocity, will increase or decrease, and thus, our model implies that the velocity of light is not necessarily independent of the direction in which it is measured. However, note that equation (23) is specific to two-way measurements of the velocity of light using an apparatus of the type set forth in Figure 10, and thus, our model could imply different predicted variations from the exact value of c using other methods of measuring the velocity of light.[56] Experiments have shown that it is possible for a moving clock to run slower than a clock that is stationary relative to the surface of the Earth (see [13]), and thus, the absolute velocity of the Earth cannot be 0.[57] However, it could be the case that the absolute velocity of the Earth is substantially less than its orbital velocity around the Sun. For example, if we assume that we can reliably measure displacements of one nanometer $(10^{-9}m)$,[58] and we assume that $L = 1$ meter, then in order for the difference in displacement given by $\bar{d} - 2L$ to be ignored or undetected, it follows that v_x must be less than approximately 6703.56 meters per second. While substantially less than the orbital velocity of the Earth, which is roughly $30,000$ meters per second, it is approximately 14.57

---

[56]See [10] for examples of other methods of measuring the velocity of light.

[57]If that were the case, then any clock on the surface of the Earth would constitute a zero energy clock

[58]For example, an interferometer could be used to measure displacements on this scale. However, there could be other methods that are even better suited for measuring minute deviations from the exact value of c (see [10] generally). We have chosen an apparatus of the type in Figure 10 for purposes of illustration because of its simplicity.

times greater than the rotational velocity of the Earth, which is approximately 460 meters per second.

Now assume that we repeatedly rotate the apparatus, measuring the value of c_v in each such orientation. Because the kinetic energy of the apparatus will not depend upon its orientation, it follows that γ will be constant in any given inertial frame, regardless of the orientation of the apparatus. Further, if we find some orientation for which our measurement of c_v is minimized, then equation (23) implies that we have in that case pointed the apparatus in a direction that maximizes v_x, which must be the actual direction of motion of the inertial frame. If we assume that $\gamma = \frac{1}{\sqrt{1-\frac{\bar{v}^2}{c^2}}}$, where $\bar{v}$ is the absolute velocity of the inertial frame, then in that case $v_x = \bar{v}$.[59] In short, we can rotate the apparatus, and eventually find some orientation which minimizes c_v, our measured velocity of light, and because $\bar{v}$ does not depend upon the orientation of the apparatus, in that case, equation (23) implies that $v_x = \bar{v}$. Thus, we can in that case solve for the absolute velocity of our inertial frame, which will be,

$$\bar{v} = \sqrt{c^2 - c_v^2}. \tag{24}$$

Further, recall that time as measured within any inertial frame is given by $t = \frac{\bar{t}}{\gamma}$. Since we can measure $\bar{v}$ by maximizing v_x, and thereby derive γ, then for any change in time t measured within an inertial frame, we can solve for the absolute change in time $\bar{t} = t\gamma$.[60]

6 A Discrete Model of an Infinite Universe

In this section, we examine the consequences of assuming that time, and the amount of energy in the universe, are infinite. We do so by applying concepts from set theory and graph theory to physical systems.

6.1 Infinite Sets

The concept of infinity was poorly understood until the work of German mathematician Georg Cantor, who provided a rigorous, mathematical concept of infinity. The core insight of his work was to define scale by reference to abstract mathematical functions, rather than make use of a physically intuitive concept of scale determined by measurement or counting. Consider two sets, $A = \{a_1, \ldots, a_n\}$ and $B = \{b_1, \ldots, b_m\}$. We define the **cardinality** of the set A as the number of elements within the set, denoted $|A| = n$. Now assume that there is a function f, such that for all $a_i \in A$, there exists exactly one b_i such that $f(a_i) = b_i$, and for all $a_i, a_k \in A$, $f(a_i) \neq f(a_k)$. As such, every element

---

[59]For simplicity, we are ignoring any time-dilation due to gravity.

[60]The results presented thus far in this paper, if true, would clearly have significant consequences for other areas of physics, particularly theories of electrostatic charge and electromagnetism. We concede that a proper treatment of these topics would require a significant undertaking, and thus, we view these topics as beyond the scope of this paper.

of A is associated with exactly one unique element of B, and therefore, it must be the case that $|A| \leq |B|$. Now assume that there is a function g such that for all $b_i \in B$, there exists exactly one a_i such that $g(b_i) = a_i$, and for all b_i, b_k, $g(b_i) \neq g(b_k)$. It follows that $|B| \leq |A|$, and therefore, $|A| = |B|$. The functions f and g together constitute a **one-to-one correspondence** between the sets A and B. Thus, in general, given any two sets A and B, if there is a one-to-one correspondence between A and B, then $|A| = |B|$.

Note that the existence of f and g together imply that A and B contain the same number of elements, regardless of the number of elements in A and B. Using this insight, Cantor demonstrated that the cardinalities of infinite sets could be compared using functions of this type. For example, consider the set of natural numbers $\mathbb{N}$, and a finite subset of the natural numbers $A = \{1, 2, \ldots, m\} \subset \mathbb{N}$. We can construct a function f that maps each element of A to a unique element of $\mathbb{N}$ by simply assuming that $f(a) = a$, for all $a \in A$. As such, whatever the cardinality of $\mathbb{N}$ might be, it must be the case that $|A| \leq |\mathbb{N}|$. Similarly, we can show that there is no analogous function from $\mathbb{N}$ to A. Let C be a subset of $\mathbb{N}$ such that $|C| = m + 1$. As such, $|A| < |C|$. Since C is a subset of $\mathbb{N}$, it follows that $|C| \leq |\mathbb{N}|$. Thus, $|A| < |\mathbb{N}|$, and therefore, there is no function that maps each element of $\mathbb{N}$ to a unique element of A.

We define a set A as **countable** if there is a one-to-one correspondence between A and $\mathbb{N}$. The intuition for the label "countable" comes from the fact that the elements of any countable set can be enumerated, since, by definition, there is a mapping from that set to the natural numbers $\{1, 2, 3, \ldots\}$. For example, let $A = \{2, 4, 6, \ldots\}$ be the set of all even numbers, let f be the function $f(n) = \frac{n}{2}$, and let g be the function $g(n) = 2n$. As such, f maps each element of A to a unique element of $\mathbb{N}$. Similarly, g maps each element of $\mathbb{N}$ to a unique element of A. As such, together, f and g constitute a one-to-one correspondence between A and $\mathbb{N}$. Therefore, it must be the case that $|A| = |\mathbb{N}|$. Note that $A \subset \mathbb{N}$, yet at the same time, they are "the same size", and thus, A is a countable set. Similar arguments apply to all infinite subsets of the natural numbers, all of which are countable. In his seminal work [6], Cantor began his paper by showing that the algebraic real numbers[61] are also countable. Thus, we might be tempted to say that there is one "size" of infinite set, given by $|\mathbb{N}|$, and that all infinite sets are countable. However, Cantor showed that there is no one-to-one correspondence between the real numbers and the natural numbers, and thus, not all infinite sets have the same cardinality. The method of proof Cantor presented in [6] is known as "diagonalization", which we now summarize.

Let $\mathbb{R}$ denote the set of all real numbers, and let A be the set of all real numbers in the interval $(0, 1)$. Note that every $r \in A$ can be expressed as a decimal number of the form $r = 0.x_1x_2x_3x_4\cdots$, where each x_i is an integer from 0 to 9. If we assume that A is countable, then we can assign each $r \in A$ an integer $i \in \mathbb{N}$. Let $x_{i,j}$ denote the j-th digit of $r_i \in A$. For example, $x_{1,2}$ is the second digit of r_1. As such, we can arrange the elements of A as follows:

---

[61]The algebraic real numbers are the set of all real number solutions to polynomial equations of the form $a_nx^n + a_{n-1}x^{n-1} + \ldots + a_0$, where each a_i is an integer.

$$
\begin{aligned}
r_1 &= \quad 0.x_{1,1}x_{1,2}x_{1,3}x_{1,4}\cdots \\
r_2 &= \quad 0.x_{2,1}x_{2,2}x_{2,3}x_{2,4}\cdots \\
r_3 &= \quad 0.x_{3,1}x_{3,2}x_{3,3}x_{3,4}\cdots \\
r_4 &= \quad 0.x_{4,1}x_{4,2}x_{4,3}x_{4,4}\cdots
\end{aligned}
$$

$$
\begin{array}{cc}
. & . \\
. & . \\
. & .
\end{array}
$$

We define $\hat{r} = 0.x_1x_2x_3\cdots$ as any real number in the interval $(0,1)$ for which $x_i \neq x_{i,i}$ for all $i \in \mathbb{N}$. That is, $\hat{r}$ can be any real number between 0 and 1 so long as the i-th digit of $\hat{r}$ is not equal to the digit that appears along the diagonal in the enumeration of A above. For example, for any given $\hat{r}$, the 3rd digit of $\hat{r}$ cannot be equal to $x_{3,3}$ above. Note that our definition of $\hat{r}$ precludes any given $\hat{r}$ from being included in the enumeration of A above. For example, assume that $r_j = \hat{r}$. The j-th digit of r_j is $x_{j,j}$, and thus, the j-th digit of $\hat{r}$ is $x_{j,j}$, which contradicts our assumption that $x_j \neq x_{j,j}$. Because this would be the case regardless of our chosen enumeration, it follows that it is not possible for any $\hat{r}$ to appear in any enumeration of the elements of A. However, every $\hat{r}$ is a real number of the form $0.x_1x_2x_3x_4\ldots$, and thus, every $\hat{r}$ is a real number in the interval $(0,1)$. This leads to a contradiction, since A consists of all real numbers in the interval $(0,1)$. It follows that it is not possible to enumerate all real numbers in the interval $(0,1)$, which implies that there is no function mapping each element of A to a unique element of $\mathbb{N}$. Therefore, $|\mathbb{N}| < |A|$. Since $A \subset \mathbb{R}$, it follows that $|\mathbb{N}| < |\mathbb{R}|$. Note that it can be shown that $|A| = |\mathbb{R}|$ [6].

Cantor's work was initially poorly received, perhaps in part because it leads to such remarkable and counterintuitive conclusions. Nonetheless, his work constituted the first rigorous treatment of the peculiar properties of the real numbers, and paved the way for others to eventually axiomatize all of mathematics. Cantor's work also played a critical role in defining the boundaries of mathematical logic and computer theory, with deep connections to both Kurt Godel's incompleteness theorem, and Turing's concept of a non-computable function. For example, every finite binary string corresponds to a unique integer, and every integer corresponds to a unique finite binary string. As such, the set of all finite binary strings is countable. Thus, the set of all finite inputs to a UTM is countable. In contrast, the reals are uncountable, and as such, there must be real numbers that cannot be calculated by a UTM.[62] As a general matter, Cantor's work on infinite sets was the starting point for entire disciplines that ultimately cut through to the very nature of mathematics, computation, and knowledge itself. For purposes of this paper, we aim to give physical meaning to Cantor's work, by showing below that Cantor's concepts of infinity can be used to model both time, and the amount of energy in an infinite universe.

---

[62]These numbers are are known as "non-computable numbers".

6.2 Counting the Number of States of the Universe

We can define the energy of the universe as a whole as the set of all quanta of energy. While this may seem too curt a definition for such a sweeping concept, so long as we accept that there is no energy outside of the universe, this definition is a practical one, since it includes the energy of every system within the entire universe. We call the set of all such quanta U.[63] Because we assume that energy is conserved, it must be the case that $|U|$ is constant over time.[64] That is, because energy is conserved, the quanta of energy within the universe at a given moment in time will always be present in the universe, though they may of course change states and positions over time. By $S_U(t)$, we denote a set containing the state and position of every quanta of energy within U at time t. As such, the information contained in $S_U(t)$ completely characterizes the state of the entire universe at time t.

We assume that there are only a finite number of possible next states for a quantum of energy given its present state. Because each quantum of energy can occupy only a finite number of intrinsic states, this is equivalent to assuming that the set of next possible positions for any given quantum is always finite. This does not imply that the position of a quantum is bounded over time, but rather, that upon any given click, the number of possible next positions for any given quantum is finite. Assume that $|U| = N$, for some finite N, and let n_i denote the number of possible next states that the i-th quantum can occupy at time t. That is, N is the number of quanta of energy within the entire universe, and n_i is the number of states that the i-th quantum can occupy at time $t + t_0$. Because $|U|$ is finite, there is some integer n such that $n \geq n_i$ for all i. Thus, $|R(S_U(t))| \leq n^N$, and as such, in this case, the set of all possible next states for the entire universe is always finite.[65] Note that this is the maximum number of possible next states of S_U at time t, and as such, if the state and position of each quantum are not independent of those of other quanta, which could very well be the case, then the number of possible next states of S_U will be less than n^N. Further, note that this does not imply that $|R(S_U(t))|$ is constant for all t, but rather, that upon any given click, there are only a finite number of possible next states for the universe.

For a closed system, $|S|$ is constant. However, because space is presumably unbounded, it follows that the universe is not a closed system in terms of its volume. Thus, it is reasonable to assume that $|S_U|$ is not constant, but will instead depend upon the interval of time under consideration. For example, imagine a single quantum of kinetic energy that travels indefinitely along an undisturbed, rectilinear path. Upon each click, that quantum will occupy a new

---

[63]Note that U has nothing to do with the concept of a universal set.

[64]We are considering only the energy contained within systems, and not any energy contained within fields.

[65]Although U is presumably not a closed system in terms of its volume, it is a closed system in terms of its energy, and thus, its information. That is, since U contains all of the energy in the entire universe, there is no additional information that we could provide R that would change the set of next states generated by R when applied to U, since by definition, no such additional information exists.

position, and therefore, a new state. As such, while the number of possible positions the quantum can occupy over any finite number of clicks is finite, there is no maximum number of positions that the quantum can occupy over all periods of time, since its position is unbounded. Thus, there is no maximum number of states that quantum can occupy over all periods of time. Because $|R(S_U(t))|$ is finite for all t, $|S_U|$ is necessarily finite over any finite period of time. However, because the position of every quantum in the universe is theoretically unbounded over time, it is reasonable to assume that $|S_U|$ is similarly unbounded over time.

Now assume that $|U| = |\mathbb{N}|$. Under this assumption, the universe contains a countably infinite number of quanta, and thus, an infinite amount of energy. Note that in this case, it is possible that $R(S_U(t))$ is not computable, since there might not be any finite input to a UTM that generates $R(S_U(t))$.[66] Nonetheless, just like we used $K(x)$, which is not computable, to derive meaningful results, we can use $R(S_U(t))$ to derive meaningful results, even if $R(S_U(t))$ is not computable. Because U is a countable set, we can still assign each quantum within U an integer i. As such, we can still let n_i denote the number of states that the i-th quantum can occupy at time $t + t_0$, which will again be finite. Since all quanta are always in some state, each $n_i \geq 1$. As such, we can represent the state of the i-th quantum at a given moment in time with an integer $1 \leq x_i \leq n_i$. Thus, we can represent each possible next state of S_U as an infinite sequence of integers $x_1, x_2, \ldots$, where x_i represents the state of the i-th quantum at time $t + t_0$. If we assume that each $n_i = 1$, then each x_i can take on only one value, meaning only one such sequence is possible. Thus, if each $n_i = 1$, then $|R(S_U(t))| = 1$, and thus, there is exactly one possible state of the universe at time $t + t_0$. Under this assumption, the universe would be deterministic at time t.

Now assume that for some finite subset $A \subset \mathbb{N}$, $n_i > 1$ for all $i \in A$, and for all $i \notin A$, $n_i = 1$. We can arrange the sequence $x_1, x_2, \ldots$ so that the first $|A|$ numbers represent the state of the $|A|$ quanta for which $n_i > 1$. The remaining x_i would be fixed, since they can take on only a single value. Thus, the maximum number of possible sequences is given by $\prod_{i=1}^{|A|} n_i$, which is finite. Therefore, if upon a given click, only a finite number of quanta are non-deterministic, then $|R(S_U(t))|$ is finite.

Now assume that $n_i > 1$ for all $i \in \mathbb{N}$. Because each n_i is at least 2, we can set each x_i to 1 for all i, which would represent a possible state of S_U given by the sequence $1, 1, 1, \ldots$, and then iteratively generate other unique states beginning with $2, 1, 1, \ldots$, followed by $1, 2, 1, \ldots$, and so on. There are countably many such unique states, and thus, the maximum value of $|R(S_U(t))|$ is not finite. Assume that the maximum value of $|R(S_U(t))|$ is countable, and that as such, we can enumerate all possible states of S_U at $t + t_0$ as follows:

---

[66]While it is possible for a UTM to generate an infinite string, a UTM cannot generate all infinite strings. Thus, it is possible that there is no encoding of $R(S_U(t))$ that can be generated by a UTM.

$$
\begin{aligned}
S_1 &= & x_{1,1}x_{1,2}x_{1,3}x_{1,4}\cdots \\
S_2 &= & x_{2,1}x_{2,2}x_{2,3}x_{2,4}\cdots \\
S_3 &= & x_{3,1}x_{3,2}x_{3,3}x_{3,4}\cdots \\
S_4 &= & x_{4,1}x_{4,2}x_{4,3}x_{4,4}\cdots
\end{aligned}
$$

$$
\begin{aligned}
&\quad. \qquad\qquad . \\
&\quad. \qquad\qquad . \\
&\quad. \qquad\qquad .
\end{aligned}
$$

In the enumeration above, $x_{i,j}$ represents the state of the j-th quantum in the i-th possible state of S_U. We define $\hat{S} = x_1, x_2, \ldots$ as any sequence for which $x_i \neq x_{i,i}$ for all i in the sequence above. Because each x_i in $\hat{S}$ is a possible state of the i-th quantum, each $\hat{S}$ represents a possible state of S_U, assuming that the state and position of a given quantum are independent from those of all other quanta. Thus, the maximum number of possible next states of S_U at time t is uncountable. As noted above, this is the maximum number of possible next states of S_U, and thus, the actual number of possible next states of S_U could be countable, and could even be finite.

6.3 Sequences of States of the Universe

In Section 6.2 above, we showed that the number of possible next states of the universe could be finite, countably infinite, or uncountably infinite, depending upon our assumptions. In this section, we will analyze the probabilities of arbitrarily long sequences of states of the universe in each of these three cases.

6.3.1 The Set of Possible Next States is Finite

We begin by assuming that $|R(S_U(t))|$ is finite for some t. As such, at time t, there are only a finite number of possible next states for the entire universe. Thus, the application of R to $S_U(t)$ will generate some finite set of possible states of $S_U(t + t_0)$. As discussed in Section 1.4 above, by application of R to each such next state, we can generate the set of all possible sequences of states that begin at $S_U(t)$ and end at some $S_U(t + 2t_0)$. In general, by $\{t, t_f\}_{S_U}$ we denote the set of all possible sequences of states of S beginning at time t and ending at t_f. That is, each element $s \in \{t, t_f\}_{S_U}$ is a sequence of states of S_U that begin with $S(t)$ and end with some $S(t_f)$. For example, in Figure 11, $\{t, t + 2t_0\}_{S_U}$ consists of 3 sequences of states.

We can represent every possible sequence of states in $\{t, t_f\}_{S_U}$ as a graph, a simple example of which is given in Figure 11. That is, we begin with a vertex that represents the state of S_U at time t, and then apply R to that state. This will generate a set of next of next possible states of S_U, and we represent each such possible next state in $R(S_U(t))$ with a unique vertex. We then apply R to each state in $R(S_U(t))$, and represent each resultant possible state with a unique vertex. This process is carried out over the entire interval $[t, t_f]$, which will produce a graph of the type in Figure 11. We call this graph the **state graph** of S_U over $[t, t_f]$, which we denote by $G\{t, t_f\}_{S_U}$. If the vertex
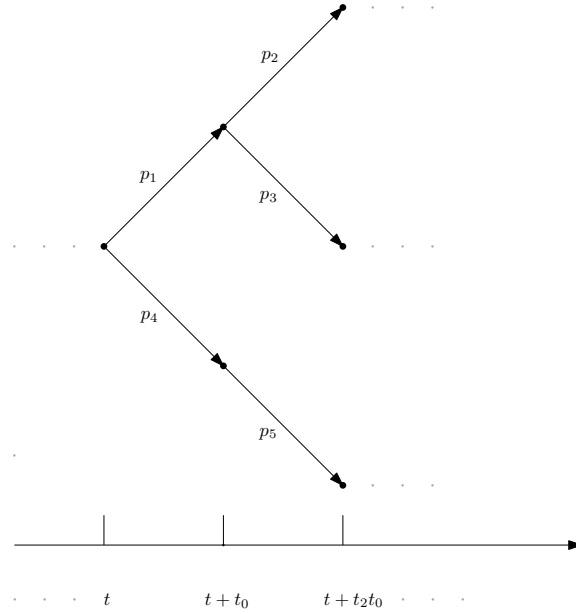
Figure 11: Possible sequences of states of S_U over two clicks.

that represents some state S_i is adjacent to the vertex that represents S_j, then we label the edge that connects them with the probability of that transition. That is, the label of every edge in the state graph represents the probability of the transition from one state to the next state represented by that edge. Note that each such probability represents the conditional probability of transitioning from one state to the next, assuming that the system is in the prior state. Thus, the probability assigned to a particular edge does not give us the probability of being in a particular state, but rather, give us the probability that the universe will transition into some next state assuming the universe is in some prior state. Thus, every possible sequence of states of the universe corresponds to some path, and each such path has a probability that is equal to the product of the probabilities of the edges along that path. For example, in Figure 11, the probability of the uppermost path is p_1p_2. That is, assuming the universe is in the state represented by the vertex listed at time t, the probability that the universe will transition through the states represented by the uppermost path is p_1p_2.

Assume that $|\{t, t_f\}_{S_U}|$ is finite. That is, the set of all possible sequences of states over the interval of time $[t, t_f]$ is finite. It follows that, for any given sequence, at each state in the sequence, there are only a finite number of possible next states. Assuming otherwise implies that there is some state S along some sequence s for which there are an infinite number of possible next states. Since each such possible next state is part of a distinct sequence of states, this implies that the set of all possible sequences of S_U over $[t, t_f]$ is infinite, which contra-

dicts our assumption that $|\{t, t_f\}_{S_U}|$ is finite. As such, regardless of the sequence chosen, there will always be a finite number of possible next states for any given state along any such path. Thus, assuming $|\{t, t_f\}_{S_U}|$ is finite implies that the set of possible next states of any state in the state graph $G\{t, t_f\}_{S_U}$ is also finite. Moreover, because $|\{t, t_f\}_{S_U}|$ is finite, and each sequence $s \in \{t, t_f\}_{S_U}$ contains a finite number of $\frac{t_f - t}{t_0}$ states, the state graph $G\{t, t_f\}_{S_U}$ will contain a finite number of vertices. Since all edges must emanate from some vertex, it follows that any given vertex in $G\{t, t_f\}_{S_U}$ will be connected to some finite number of other vertices. It follows that the **in-degree** of each vertex, which is the number of edges pointing into a given vertex, and the **out-degree** of each vertex, which is the number of edges emanating from a given vertex, is finite for all vertices in $G\{t, t_f\}_{S_U}$.[67] Thus, assuming $|\{t, t_f\}_{S_U}|$ is finite implies that the in-degree and out-degree of every vertex in $G\{t, t_f\}_{S_U}$ is also finite.

We assume that an event that actually occurs cannot have a probability of zero. Thus, if a particular sequence of states of the universe actually occurs, then we assume that the probability of that sequence cannot be zero. Because the universe exists, there is some sequence of states that describes its progression over every period of time up to the present moment. For example, if t is the present, then the probability of the sequence of states that connects the state of the universe at some prior time to the present state at time t is equal to the product of the probabilities of the edges along the corresponding path to the present state. For example, in Figure 12, we present a sequence of states that leads to the present state of the universe $S_U(t)$. Because multiplication is commutative, we can calculate the probability of this sequence "backwards", beginning with the probability of the edge that connects the current state of S_U to its prior state, and so on. Thus, the probability of the sequence that connects S_j to the present state is $\prod_{i=1}^{5} p_i$. This will be true of any sequence commencing at some prior state of the universe.
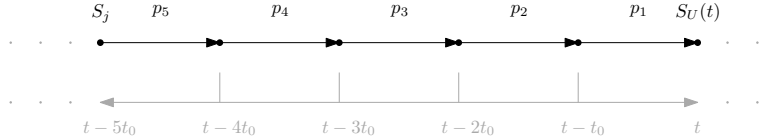

Figure 12: A possible sequence of states leading to $S(t)$.

If the history of the universe is infinite, then we assume that the sequence of states that led to the current state of the universe is some countably infinite path. Thus, the probability of that sequence is equal to the product of the probabilities of the edges along that infinite path. If there are an infinite number of edges with a probability of less than 1, then the probability of the path is 0. Since we are assuming that this path connects the present state of the universe to all of its prior states, it cannot be the case that the probability of this path is 0, since all states along this path actually, physically occurred. This implies that

---

[67]Note that any pair of vertices are connected by at most one edge.

either the history of the universe is finite, or that the history of the universe is non-deterministic at only a finite number of points in time throughout its entire infinite history.
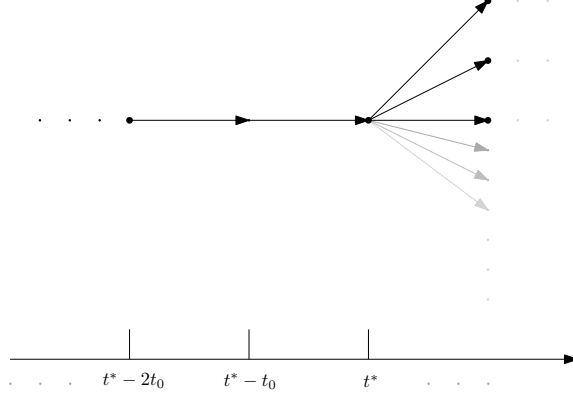

Figure 13: A non-deterministic sequence commences at time t^*.

The claim that the history of the universe is finite does not fit well with the model presented herein, since time is simply the product of quanta changing states. As such, to assume that time had a beginning is to assume that in some prior state of the universe, quanta did not change states. Alternatively, we could assume that the path connecting the prior states of the universe to the present state consists of only a finite number of non-deterministic points in time. For example, it could be the case that prior to some moment in the finite past, the universe was deterministic, and thus, the probability of each edge in the corresponding path of states prior to that moment had a probability of 1. Under this assumption, the present state of the universe is the product of an infinite period of time during which the universe was deterministic, after which there was a finite period of time during which the universe was non-deterministic, ultimately leading up to the present moment (see Figure 13).

6.3.2   The Set of Possible Next States is Countably Infinite

Now assume that $|R(S_U(t))|$ is countably infinite for some t. It follows that there are a countable number of possible next states of S_U. Thus, for any t_f, it follows that $|\{t, t_f\}_{S_U}|$ cannot be finite, since there are at least a countable number of distinct sequences that commence at $S_U(t)$. Thus, the out-degree of the vertex representing $S_U(t)$ in $G\{t, t_f\}_{S_U}$ will be countably infinite, since $S_U(t)$ can transition into any one of a countable number of next states. Now assume that the out-degree of every vertex in $G\{t, t_f\}_{S_U}$ is countably infinite. Thus, any possible state of S_U over $[t, t_f]$ can transition into any one of a countable number of possible next states. We can represent $G\{t, t_f\}_{S_U}$ as set forth in Figure 14.
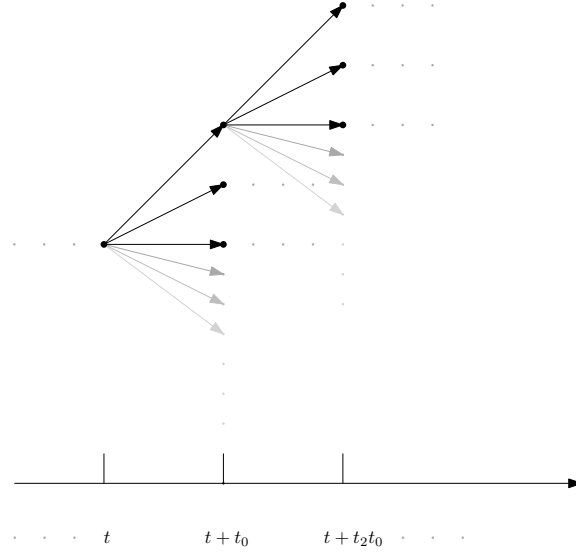
Figure 14: A state graph whose vertices all have a countably infinite out-degree.

Because the vertex representing $S(t)$ has a countable out-degree, we can enumerate each possible state of $S(t + t_0)$, and assign each such state a unique integer $n \in \mathbb{N}$. The same would be true of any vertex in $G\{t, t_f\}_{S_U}$, since each has a countably infinite out-degree. Let $k = \frac{t_f - t}{t_0}$ be the number of clicks that elapse over $[t, t_f]$. It follows that every possible sequence of k integers $n_1, \ldots, n_k$ corresponds to a unique sequence of states of S_U, and therefore, a unique path in $G\{t, t_f\}_{S_U}$. Note that each n_i is unbounded, and can take on any finite value in $\mathbb{N}$. Thus, the number of such sequences cannot be finite. Note that each n_i can be expressed as some base 10 numeral. If we concatenate the numerical expressions of each number in the sequence $n_1, \ldots, n_k$, the result will also be the numerical expression of some integer. For example, if $k = 2$, and our sequence is $(10, 25)$, then concatenating these two numbers yields another integer 1025. Thus, every possible sequence corresponds to some unique integer. Because each such sequence corresponds to some unique positive integer, they can be ordered, and thus, there is a first such sequence, a second, and so on. Thus, the set of all such sequences can be enumerated, and it is therefore a countable set.

Because each vertex has a countably infinite out-degree, our ability to assign probabilities to the edges emanating from each vertex will be restricted, assuming we make use of standard probabilities. For example, it cannot be the case that all possible states of $S_U(t + t_0)$ have an equal probability of occurring, because there is no uniform distribution on $\mathbb{N}$ using standard probabilities. Nonetheless, we can still assign probabilities to each edge in $G\{t, t_f\}_{S_U}$, and as before, the probability of any sequence of states over $[t, t_f]$ is the probability of the corresponding path in $G\{t, t_f\}_{S_U}$. Thus, over any infinite path, if there are an infinite number of edges with a probability of less than 1, then the probabil-

ity of the path is 0. This again implies that either the history of the universe is finite, or that the history of the universe is non-deterministic at only a finite number of points in time throughout its entire infinite history.

6.3.3 The Set of Possible Next States is Uncountably Infinite

Now assume that $|R(S_U(t))|$ is uncountably infinite for some t. It follows that there are an uncountable number of possible next states of S_U. Thus, for any t_f, it follows that $|\{t, t_f\}_{S_U}|$ cannot be finite, since there are at least an uncountable number of distinct sequences that commence at $S_U(t)$. Because $|R(S_U(t))|$ is uncountable, we cannot enumerate the vertices of $G\{t, t_f\}_{S_U}$. However, we can still meaningfully describe the out-degree of the vertex representing $S_U(t)$ in $G\{t, t_f\}_{S_U}$, by viewing the edges emanating from that vertex as a mapping from that vertex to some uncountable set, such as the set of all real numbers in $(0, 1)$. As noted above, the set of all real numbers in $(0, 1)$ is uncountable. Thus, we can think of the vertex representing $S_U(t)$ in $G\{t, t_f\}_{S_U}$ as being mapped to each real number in $(0, 1)$. As such, for every real number in $(0, 1)$, there is some possible state $S(t + t_0)$. Though we cannot iteratively apply R the way we can to a finite or countable set of states, we can nonetheless define $R(S(t + t_0))$ for all such possible states. If we assume that each such application again yields an uncountable set of possible next states, and assume that $k = \frac{t_f - t}{t_0}$ is the number of clicks that elapse, then we can represent each possible sequence of states as a sequence of k real numbers $r_1, r_2, \ldots, r_k$, where each $r_i \in (0, 1)$. Because each r_i can take on any value in $(0, 1)$, the set of all such sequences cannot be finite. Because k is finite, it can be shown that the cardinality of the set of all such sequences is equal to $|\mathbb{R}|$.[68]
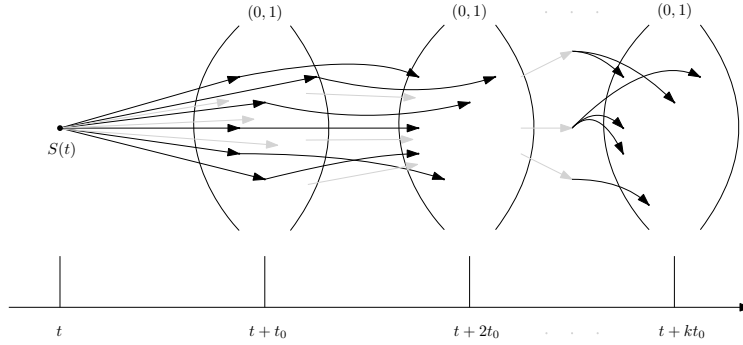

Figure 15: A representation of $G\{t, t_f\}_{S_U}$.

Because each vertex in $G\{t, t_f\}_{S_U}$ has an uncountably infinite out-degree, our ability to assign probabilities to the edges emanating from each vertex will again be restricted, assuming we make use of standard probabilities. For example, it cannot be the case that all possible states of $S_U(t + t_0)$ have an equal

---

[68]In general, it can be shown that $|\mathbb{R}^n| = |\mathbb{R}|$ for all finite n.

probability of occurring, because there is no uniform distribution on $\mathbb{R}$ using standard probabilities. Nonetheless, we can still assign probabilities to each edge in $G\{t, t_f\}_{S_U}$, and as before, the probability of any sequence of states over $[t, t_f]$ is the probability of the corresponding path in $G\{t, t_f\}_{S_U}$. Thus, over any infinite path, if there are an infinite number of edges with a probability of less than 1, then the probability of that path is again 0. Therefore, either the history of the universe is finite, or the history of the universe is non-deterministic at only a finite number of points in time throughout its entire infinite history.

6.4 Infinite Energy Systems

If the universe contains a countably infinite number of quanta, then it would be possible for other systems within the universe to contain an infinite number of quanta as well, since, as noted above, any countable set can be subdivided into countable subsets. For example, the set of all natural numbers can be subdivided into the set of even numbers and the set of odd numbers, both of which are countably infinite sets. As such, if $|U|$ is countable, then $|U|$ can be subdivided into countable subsets. Therefore, if $|U|$ is countable, then it would be possible for a system to contain a countably infinite number of quanta, and thus, an infinite amount of energy. Assume that a system S contains a countably infinite number of quanta, and assume that there is some drain on the energy of S that extracts quanta at some rate per click. If the rate per click at which energy is extracted from S is finite, then the cardinality of the set of all quanta within S will be unchanged over any finite interval of time.[69] Any such system would be capable of providing an unlimited amount of energy over any finite period of time. Therefore, if a system transitions from an infinite amount of energy to a finite amount of energy, this transition must be discontinuous in nature.

Imagine a source that contributes quanta at some rate per click to some system S that initially contains some finite number of quanta. If the rate per click at which energy is contributed to S by the source is finite, then it is impossible for the source to contribute an infinite number of quanta to S in any finite amount of time. This is true regardless of the rate at which the source contributes energy to S, and regardless of the amount of time that has elapsed. Thus, if the rate at which energy is contributed by the source to S is finite, and the number of clicks that have elapsed is finite, then there is always some next click for which the energy of S is greater than it was upon the previous click, but nonetheless finite. Therefore, if a system transitions from a finite amount of energy to an infinite amount of energy, this transition must be discontinuous in nature.

Now assume that S consists of a countably infinite number of kinetic quanta, and a finite number of mass quanta. Because this kinetic energy must be contained within the clusters of S, and there are only a finite number of clusters

---

[69]For example, if we remove the first n integers from $\mathbb{N}$, the remaining set will still be countably infinite. If we continue to iteratively remove a finite number of integers, each resulting set will still be countably infinite.

within S, it follows that there is at least one cluster within S that consists of a single mass quantum, and a countably infinite number of kinetic quanta. In Section 3, we assumed that the probability that any given quantum within a cluster is active upon a given click is equal for all quanta within the cluster. In the case of S, this assumption fails using standard real number probabilities, since if we assign any single real number probability to each quantum becoming active, then the sum of those probabilities will be infinite, and therefore, greater than 1. We can, however, generalize upon the notion of a uniform distribution if we allow for an infinite amount of time to elapse. That is, we can assume that over any infinite sequence of active quanta, the number of times each quantum appears in the sequence is equal for all quanta. Specifically, because the set of quanta within the cluster is countable, we can assign each such quantum a number $i \in \mathbb{N}$. Assume that t is the first moment in time we consider, and that a countably infinite number of clicks follow. As such, each moment in time $t + jt_0$ will be associated with exactly one active quantum, and therefore, one integer i, as set forth in Figure 16. Thus, we can describe the sequence of active quanta over time as a map from time to the integers. Note that it is not necessarily the case that all i are associated with exactly one moment in time. That is, a given quantum could be active more than once over that infinite sequence.
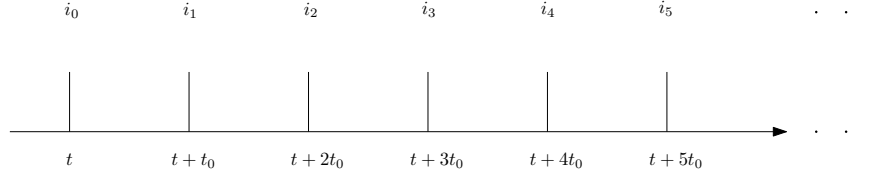

Figure 16: A mapping from time to the set of integers.

As a result, it is perhaps more meaningful to represent this mapping as a graph, where each moment in time is mapped to a vertex that is labeled with some integer i, and the in-degree of a vertex in the range of the mapping represents the number of times the quantum represented by that vertex is active over the infinite sequence. In Figure 17, each vertex in the range of the mapping has an in-degree of one, and thus, each quantum appears exactly once in the infinite sequence of active quanta. As noted above, we cannot assign a real number probability to each quantum in this sequence, but if we generalize upon the concept of a uniform distribution to include any infinite sequence of events in which each possible event occurs an equal number of times, then the mapping represented in Figure 17 would constitute a uniform distribution. Moreover, this definition allows the in-degree of each vertex to take on any value, so long as all such vertices have an equal in-degree. As such, there are an infinite number of "uniform distributions" of this type.

Note that the sequence of integers generated by such a mapping might not be random, in the sense that the Kolmogorov complexity of any resultant finite sequence could be low relative to the length of that sequence. For example, a
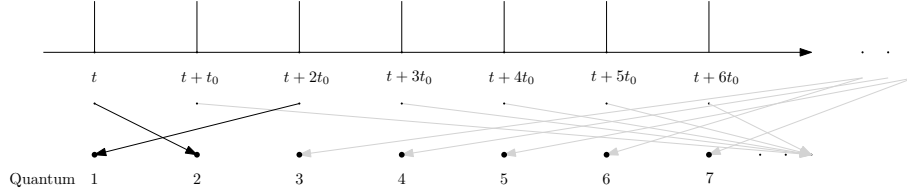
Figure 17: A mapping from time to the set of active quanta.

mapping that simply lists the integers in numerical order would certainly not be random. As a general matter, the sequences generated by these mappings are not the product of a random variable, since they are in fact fixed, and therefore, deterministic mappings from time to the natural numbers. However, in the case where the in-degree of each vertex in the range is one, we can generate any infinite permutation of the natural numbers using such a mapping. In contrast, a UTM cannot generate all permutations of the natural numbers, but rather, can generate only those permutations of the natural numbers that are computable. Moreover, since the in-degree of the vertices in the range in such a mapping is unbounded, we can also generate any finite sequence of natural numbers using such a mapping. For example, assume a random variable produces the sequence of integers $s = n_1, n_2, n_3, \ldots, n_k$. Each integer within s will appear at most k times within s. As such, we can generate s using a mapping where the in-degree of each vertex in the range is k, by mapping the first k clicks to the integers that appear in the sequence s, in the same order in which they appear in s. This would be true for any sequence, regardless of k. Thus, any finite sequence of events generated by a random variable would be indistinguishable from a finite sequence of events generated by a deterministic mapping of this type. Therefore, we can generate finite sequences of arbitrarily high Kolmogorov complexity using mappings of this type. However, a mapping of this type is not restricted by the standard axioms of probability, which would preclude a uniform distribution on an infinite set. Thus, by making use of these mappings, we can conceive of uniform distributions on infinite sets, and describe the emergence of random phenomena over finite periods of time.

We are of course not suggesting that all apparently random phenomena are actually the product of a deterministic mapping of this type. Rather, we are arguing that mappings of this type allow us to generalize the notion of a uniform distribution to allow for uniform distributions on infinite sets, which would be impossible to do using standard probabilities. Moreover, they allow for apparently random phenomena to persist over an infinite period of time. For example, assume that the current state of a system $S(t)$ is the product of an infinite number of non-deterministic events. As discussed above, the probability of $S(t)$ is therefore zero, which contradicts the assertion that $S(t)$ actually occurred. In contrast, if we assume that $S(t)$ is the product of an infinite number of deterministic events, then the probability of $S(t)$ occurring is one. If the sequence of states of S are determined using a mapping of the type described above, then

that sequence of states could nonetheless be random, despite being deterministic. However, because a mapping of the type described above could generate a non-computable sequence of states of S, we cannot in this case assume that R is computable. Thus, if we want the ability to describe any possible sequence of events over a countably infinite period of time, then we cannot assume that R is computable. As a practical matter, if it is actually the case that R is non-computable, then the laws of physics would be non-computable, which would imply that they cannot be calculated by using any known mechanical process. This would not, however, preclude the existence of computable approximations to the "true", non-computable laws of physics.
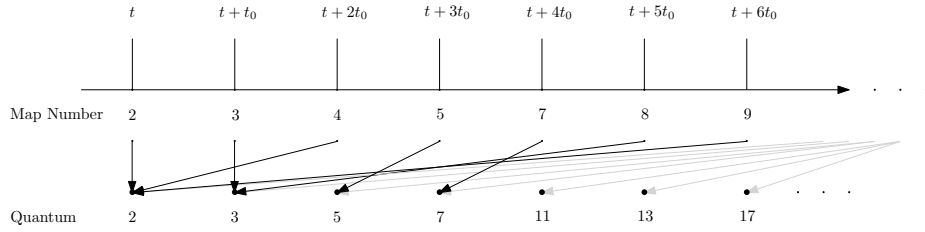

Figure 18: A mapping from time to the set of prime numbers.

Because there are an infinite number of clicks in any such sequence, we can also construct a mapping in which each vertex in the range has a countably infinite in-degree. Constructing such a mapping requires subdividing the natural numbers into a countable number of countable subsets. This can be accomplished by subdividing the natural numbers into sets of powers of the prime numbers. For example, the first set would be the set of all integers of the form 2^n, the next, the set of all integers of the form 3^n, and so on. All such sets have empty intersections due to the nature of prime numbers, and thus, each constitutes a unique countable subset of the natural numbers. As set forth in Figure 18, we would label each quantum with a unique prime number. We would then map each moment in time to a "map number", which is a unique power of some base prime number, with the map numbers listed in ascending order over time. Each map number would then map to the quantum labeled with its base prime number. For example, time t would map to 2, which would in turn map to the quantum labeled 2. Time $t + 6t_0$ would map to 9, which would in turn map to quantum 3. Since there are a countably infinite number of powers of each prime number, using this approach, all quanta in the range would appear a countably infinite number of times, and therefore, each quantum would be active an equal number of times over a countably infinite period of time.

6.5 Time-Dilation in Infinite Energy Systems

The amount of time-dilation experienced by a system with a countably infinite number of quanta will depend upon our assumptions regarding the sequences of active quanta within each cluster. For simplicity, assume that all clusters

within a system S consist of a single mass quantum, and a countable number of kinetic quanta. Further, assume that S has a finite mass, and thus, a finite number of clusters. If we assume that all quanta within each cluster are active exactly once over an infinite number of clicks, then every mass quantum within the system will be active exactly once over that infinite number of clicks. This can be described by a mapping of the type set forth in Figure 17. As such, each quantum within each cluster is associated with exactly one moment in time. If we again assume that time-dilation can be approximated by the rate at which the mass quanta within the system are active, it follows that after some finite period of time, all mass quanta within the system will have been active once, after which, "time will stop", since the mass quanta will not become active again. Note that the number of clicks required for all mass quanta within the system to become active once is unbounded, because we cannot say which moment in time each mass quantum within the system is associated with, but rather, that it must be associated with some moment in time. Further, note that the number of clicks between any two mass quanta within the system becoming active is also unbounded, meaning that each cluster could "age" at its own rate. In general, similar reasoning would apply if we assume that each quantum within each cluster is active $k \in \mathbb{N}$ times over an infinite number of clicks, since there has to be a "last" moment in time with which each mass quantum is associated.

If instead we assume that the quanta within the cluster become active in accordance with the distribution set forth in Figure 18, then the mass quantum within a given cluster can be thought of as being assigned some finite prime number q. Note that we are of course not arguing that every moment in time is actually associated with some map number. Rather, we are assuming that the sequence of active quanta generated by each cluster can be described by making use of a mapping of this type. Thus, at some point, over some finite period of time, we assume that the cluster will behave as if there is some map number associated with a given click that is a power of q, at which time, the mass quantum within the cluster will become active. Note that because each mass quantum could be assigned a different prime number, each cluster could again "age" at its own rate. This implies that even if a system has an infinite amount of kinetic energy, time as measured within that system does not necessarily "stop", but could instead proceed at an arbitrarily slow pace, since q could be any prime number. Moreover, because the difference $q^n - q^{n-1} = q^n(1 - \frac{1}{q})$ grows exponentially with n, it follows that the amount of objective time between clicks upon which the mass quanta are active will grow exponentially with objective time. Therefore, the amount of time-dilation experienced by an infinite energy system could increase as a function of objective time.

Finally, note that in all cases, a system with an infinite number of kinetic quanta would nonetheless have a finite velocity, since each kinetic quantum codes for a finite displacement per click.

References

[1] ABBOTT, B., ET AL. Gravitational waves and gamma-rays from a binary neutron star merger: Gw170817 and grb 170817a. *The Astrophysical Journal Letters 848*, 2 (2017), 1–27.

[2] ADAMSON, P., ET AL. Measurement of neutrino velocity with the minos detectors and numi neutrino beam. *Physical Review 76*, 7 (2007).

[3] BENNETT, C. H. Demons, engines and the second law. *Scientific American 257*, 5 (1987), 108–116.

[4] BETHE, H. A. On the annihilation radiation of positrons. *Proceedings of the Royal Society of London, Series A 150*, 869 (1935), 129–141.

[5] BURKE, D. L., ET AL. Positron production in multiphoton light-by-light scattering. *Physical Review Letters 79*, 9 (1997), 1626–1629.

[6] CANTOR, G. Ueber eine eigenschaft des inbegriffs aller reellen algebraischen zahlen. *Journal fr die reine und angewandte Mathematik / Zeitschriftenband* (1874), 259–262.

[7] COLGATE, S. A., AND GILBERT, F. C. Electron-positron annihilation in flight. *Physical Review 89*, 4 (1953), 790–792.

[8] COMPTON, A. H. A quantum theory of the scattering of x-rays by light elements. *Physical Review 21*, 5 (1923), 483 – 502.

[9] DREXLIN, G., HANNEN, V., MERTENS, S., AND WEINHEIMER, C. Current direct neutrino mass experiments. *Advances in High Energy Physics 2013* (2013).

[10] EVENSON, K., AND PETERSEN, F. Laser frequency measurements, the speed of light, and the meter. *Topics in Applied Physics 2* (1976), 349 – 368.

[11] GALILEI, G. *Dialogue Concerning the Two Chief World Systems*. 1632.

[12] GRUNWALD, P., AND VITANYI, P. Shannon information and kolmogorov complexity. *eprint arXiv:cs/0410002* (2004).

[13] HAFELE, J. C., AND KEATING, R. E. Around-the-world atomic clocks: Observed relativistic time gains. *Science 177*, 4044 (1972), 168 – 170.

[14] HUFFMAN, D. A. A method for the construction of minimum-redundancy codes. *Proceedings of the IRE 40*, 9 (1952), 1098 – 1101.

[15] KUNDIG, W. Measurement of the transverse doppler effect in an accelerated system. *Physical Review 129*, 6 (1963), 2371– 2375.

[16] MARUYAMA, K., NORI, F., AND VEDRAL, V. The physics of maxwell's demon and information. *Reviews of Modern Physics 81*, 1 (2009), 1–23.

[17] MILLIKAN, R. A. A new modification of the cloud method of measuring the elementary electrical charge, and the most probable value of that charge. *Physical Review 29*, 6 (1909), 560–561.

[18] MILLIKAN, R. A. A direct photoelectric determination of planck's h. *Physical Review 7*, 3 (1916), 355–388.

[19] SCHULTZE, M., ET AL. Delay in photoemission. *Science 328*, 5986 (2010), 1658–1662.

[20] SHANNON, C. E. A mathematical theory of communication. *The Bell System Technical Journal 27*, 3 (1948), 379–423.

[21] TURING, A. M. On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society s2-42*, 1 (1937), 230–265.

[22] TURING, A. M. Systems of logic based on ordinals. *Proceedings of the London Mathematical Society s2-45*, 1 (1939), 161–228.

[23] VANDERLOOY, S., AND SPRINKHUIZEN-KUYPER, I. G. An overview of algorithmic randomness and its application to reliable instance classification. *Technical Report 07-02, Universiteit Maastricht* (2007).