Article

On the arrow of spacetime

Alexandre Harvey-Tremblay 1*

1 Independent researcher
* Correspondence: aht@protonmail.ch

Abstract: Consistent with special relativity and statistical physics, here we construct a partition function of space-time events. The union of these two theories resolves longstanding problems regarding time. We will argue that it augments the standard description of time given by the (non-relativistic) arrow of time to one able to describe the past, the present and the future in a manner consistent with our macroscopic experience of such. First, using Fermi-Dirac statistics, we find that the system essentially describes a "waterfall" of space-time events. This "waterfall" recedes in space-time at the speed of light towards the direction of the future as it "floods" local space with events that it depletes from the past. In this union, an observer O will perceive two horizons that can be interpreted as hiding events behind them. The first is an event horizon and its entropy hides events in the regions that O cannot see. The second is a time horizon, and its entropy "shields" events from O's causal influence. As only past events are "shielded" and not future events, an asymmetry in time is thus created. Finally, future events are hidden by an entropy prohibiting O from knowing the future before the present catches on.

Keywords: Statistical Physics; Special Relativity; Arrow of Time

1. Introduction

1.1. Time

The connection between time in statistical physics and time in special relativity are both well established (Popper [1], Layzer [2], Davies [3], Zeh and Page [44], Callender [5], Price [6], Landsberg [7], Halliwell et al. [8], Carroll [9]). On the one hand, we have the statistical emergence of a macroscopic arrow of time, and on the other, we have a causal relationship between space-time events limited by the speed of light. As most other theories consider time to be little more than some reversible variable, the claims made regarding time by these two theories pack quite a lot of heat (figuratively). Thus, to further investigate the nature of time, a promising avenue might be to ask: what about time in "statistical physics ∪ special relativity" — is it possible to combine causality (in the sense given by special relativity) with the arrow of time? As it turns out, in doing so (achieving the combination) we are rewarded with quite a lot of insight. First, let's clarify the motivation for this union.

1.2. Problem 1: The direction of time

We recall the well-known thought experiment in which Maxwell’s demon (Maxwell [10], Knott [11], Thomson [12], Leff and Rex [13]) consumes good information to return a macroscopic system to some initial low-entropy state. An observer might naively conclude that the (macroscopic) arrow of time (Weinert [14]) is reversed, but inspection of the global system reveals that Maxwell’s demon produced entropy in amounts at least equal to the information consumed (Szilard [15], Landauer [16], Bennett [17]).

Taking the universe to be the global system, and sub-systems to be local, equivalent thought experiments can be produced. Borrowing terms from special relativity, we might say that the arrow of time holds globally, but not locally. Gotcha! A global property in a non-relativistic theory is bound to violate causality.
Or as George F R Ellis (Ellis [18]) states:

“The arrow of time locality issue: If there is a purely local process for determining the arrow of time, why does it give the same result everywhere? [...] Local determination has to arbitrarily choose one of the two directions of time as the positive direction indicating the future; but as this decision is made locally, there is no reason whatever why it should be consistent globally. If it emerges locally, opposite arrows may be expected to occur in different places.”

Resolving this issue is, of course, an interesting result of the model. In the model, the arrow of space-time cannot be reversed by Maxwell’s demon, even locally, because the space-like entropy of the sub-system, associated with the system’s past, grows throughout the cycle proportionally to the speed of light. The causal influence of the system (which propagates at the speed of light in all directions), is, in this model, the macroscopic system that would need to be reversed by Maxwell’s demon to reverse the arrow of time. Thus, Maxwell’s demon cannot succeed unless it violates the speed of light.

1.3. Problem 2: Macroscopic time

If we were to only rely on time-symmetric differential equations (e.g. classical mechanics, unitary evolution, etc.), there would be no meaningful difference between the past, the present, and the future. Basically, time would be a point on a line and we could solve our equations as far into the future or into the past as we wanted. This is quite far from our day-to-day experience of time.

Consistent with our macroscopic experience of time, the arrow of time improves the connection between "time in the equations" and "time in real life" by giving it a direction favored by entropy. This is better, but, as we will argue, not yet complete.

We will now scope our work by formulating some contemplation questions to elucidate what is missing from the arrow of time, along with some leading questions to guide the discussion towards the solution presented in this paper. Let’s start with two questions specifically related to the arrow of time:

1. It is very unusual in science to have a quantity which sets the direction of another, but whose magnitudes are not proportional. For instance, we can imagine an experiment which attains thermodynamic equilibrium very quickly (e.g. an explosion), and one which takes a long time (e.g. Earth’s core cooling off to room temperature). But in each case, the clocks are ticking at the same rate. Why is the rate of passage of time not proportional to the magnitude of the arrow of time, if its direction is correlated to it?

2. Special relativity sets the magnitude for the time evolution of an observer \( O \) in its own frame of reference to be proportional to the speed of light. Leading question: Why is the direction of time set by statistical physics, but its magnitude is set by special relativity, two seemingly unrelated theories?

Furthermore, there is a noticeable difference between the past, the present, and the future that is not captured by the arrow of time. An image that I like to use is that of an observer regretting the past, worrying about the future, and acting (or failing to act) in the present. Consistent with causality, \( O \) ought not to experience the other permutations, such as regretting the future or worrying about the past.

Or, again, as George F R Ellis (Ellis [18]) states:

"As the nature of existence is different in the past and in the future - Becoming has meaning. Different ontologies apply in the past and future, as well as different epistemologies"

With this image in mind, a further series of questions arises:

1. Suppose \( O \) goes bird watching. Say the birds are observed for 20 years by \( O \). Then \( O \) should know a lot more about those birds today than \( O \) did 20 years ago. \( O \)'s conclusion that time
has passed is evidenced by the existence of a "log of observations" maintained by $\mathcal{O}$, and not
so much on the fact that the birds of today are ever so slightly closer to final cosmological
thermodynamic equilibrium than the birds of 20 years ago. Let’s emphasize the dichotomy:
the arrow of time points towards higher entropy (more hidden information), but $\mathcal{O}$ concluded
that time moves forward by building up a log of events (gaining information). Is $\mathcal{O}$ making a
mistake by using a log of events to conclude as to the passage of time, instead of estimating the
system’s overall entropy as per the arrow of time? As George F R Ellis puts it, “we are aware of
the flow of time because of the existence of records of the past” (Ellis [18]). Leading question: Is
$\mathcal{O}$’s construction of a log of events contributing to increasing the entropy of the system? What if
$\mathcal{O}$’s log is complete (it lists all space-time events) — in this case, is $\mathcal{O}$’s behavior of producing
information by increasing entropy the opposite of Maxwell’s demon? Perhaps it is “Maxwell’s
angel”? (Touchette and Lloyd [19], Parrondo al. [20])?

2. Generally, if the universe started from a low entropy Big Bang and is evolving into a high entropy
state and that entropy is associated with a loss of information, what sort of compensation is there
to offset this loss? Leading question: Can the information gained by observing a system over time
be accumulated and is it enough to offset the information lost by the action of the arrow of time?
Can a statistical system aware of events reconcile both notions? From Landauer’s aphorism,
since “information is physical” (Landauer et al. [21]) where and how would his information be
stored(Goold et al. [22], Streltsov et al. [23], Adesso et al. [24])?

3. Why is the causal connection between past, present and future asymmetric (Halliwell et al. [8])?
For example, $\mathcal{O}$ opens the fridge and notices that there is milk in it. $\mathcal{O}$ wonders why the milk
is there. It would make no sense to appeal to the future to justify the presence of milk in the
present. Thus saying, "the milk is in the fridge because $\mathcal{O}$ will drink it tomorrow and therefore
it must be here today for $\mathcal{O}$ to be able to drink it tomorrow — that’s why it is there!” makes no
sense because $\mathcal{O}$ could simply elect not to drink the milk and the milk will still be in the fridge
today. It’s similar to how some people feel absolutely certain that they will win the lottery at the
next draw, and thus will buy a ticket today in an attempt to set up the present to be as consistent
as possible with the anticipated future winnings, only to witness the attempt fail. However, the
reverse explanation easily makes sense: the milk is in the fridge today because $\mathcal{O}$ bought it at
the supermarket yesterday and put it in the fridge this morning. Why does the “algorithmic
reconstruction” of the past based on present evidence work, but the same reconstruction fails
when applied to the future? Using purely time-symmetric differential equations we would get a
different story: the present would fix both the past and the future! We could even appeal to the
future to justify the present. The “arrow of causality” that we experience fails to emerge from
exactly solvable time-symmetric differential equations.

4. From $\mathcal{O}$’s present, the future appears to have multiple outcomes that could occur (e.g: $\mathcal{O}$ could
elect to drink or not to drink the milk — we won’t know until tomorrow) (Eddington [25]).
Leading question: Is there an entropy that we could have missed, but is nonetheless associated
with $\mathcal{O}$’s possible futures that somehow gets reduced as $\mathcal{O}$ travels forward in time, as the present
unravels? Does this entropy prohibit $\mathcal{O}$’s knowledge about the future?

Combining special relativity with statistical physics will allow us to account for the past, the
present, and the future as three distinct regimes of time in a manner consistent with experience and
with the points raised in these questions. Let us now describe in greater detail how this will be done.

1.4. Outline

We consider two approaches that we could use to combine statistical physics with special relativity.
One approach is to make the standard quantities of statistical physics (e.g., entropy, temperature)
Lorentz-invariant. Using this methodology, we are rewarded with a frame-independent partition
function. The result of this approach yields non-relativistic statistical physics in the limit $c \to \infty$. This
is the approach taken by Giorgio Kaniadakis (Kaniadakis [26]).
Another approach is to seek a statistical ensemble whose macroscopic description relates to special relativity. As special relativity is concerned with the relationship between space-time events, this approach elects to make special relativity emergent (via an equation of state) from a statistical ensemble of space-time events. In this case, special relativity is an emergent behavior caused by the random statistics of space-time events.

The latter approach is the one that will be taken in this paper. Not only does it paint a seductive picture of time and space, but it is also attractive because of its ability to explain the nature of macroscopic time.

Using Fermi-Dirac statistics over an ensemble of space-time events, we show (section 3.3) that the space-time counterpart to the arrow of time is a "waterfall of events". Like the arrow of time, it too has a direction: it recedes in space-time at the speed of light towards the direction of the future (section 3.4). Furthermore, as it recedes, the waterfall of events floods the present with immediate events and depletes the past of events. As we will see, this behavior connects to the three regimes of time (section 3.7 and 4).

1.5. Improvements upon prior work

The prior work closest to ours are probably the works of (Jacobson [27]) and the works of (Verlinde [28]), in which general relativity, classical gravity and the law of inertia are derived from statistical physics. To derive the law of inertia, Verlinde [28] considers a length conjugated with an entropic force, then under injection of the Unruh temperature, finds $F = ma$. To derive general relativity, Jacobson [27] considers an area conjugated with an entropic surface tension, then under multiple reasonable assumptions (Raychaud-Huri equation, local energy conservation, Null-horizons, etc.) recovers Einstein’s fields equations. Thus, a pattern is beginning to emerge in the literature in which some laws of physics are consistent when formulated as having their origin in entropy. Here, we extend their work by placing time on equal entropic footing as these authors have done for length and area. Specifically, we will consider a time conjugated to an entropic power in the context of emergent special relativity.

2. A statistical ensemble of events

Our goal for this section will be to recover the “features” of special relativity strictly using the facilities of statistical physics. In this case, we would say that special relativity is an emergent property of the constructed statistical ensemble. Even the speed of light will not be taken as an axiom, but it will instead be a property emergent from the construction. How will we do that? First, we have to interpret the speed of light as a tool to hide information. Specifically, the speed of light hides information regarding events whose intervals to the observer are space-like. Interpreted as such, we can then use the entropy in statistical physics to achieve the same purpose as the speed of light (hide information), provided that we “place” the entropy appropriately in the system.

For instance, we can imagine an observer $O$ whose event horizon is defined by the usual light cone of special relativity. We can then describe this light cone entirely using notions of statistical physics by analyzing the number of configurations of events outside the horizon and associating it to an entropy.

We deduce that, to prevent faster-than-light communication, all possible configurations of events outside the event horizon must be of maximal entropy (e.g., equally likely within the priors) so as to be void of information from the perspective of $O$. This entropy thus hides events outside the horizon. Attributing an entropy to events separated by a horizon in order to connect to thermodynamics has been done since at least 1973 by J.D. Bekenstein (Bekenstein [29]). Furthermore, from G. W. Gibbons and S. W. Hawking’s 1977 article (Gibbons and Hawking [30]), I quote:

"An observer in these models will have an event horizon whose area can be interpreted as the entropy or lack of information of the observer about the regions which he cannot see."
The part missing from the picture of time, we suggest, is to apply the same line of reasoning to configurations of future events. We ask, how many configurations of future events are compatible with the present state? First, we pose this assumption: since the present is caused by the past (not the future), there exists strictly more than one configuration of future states compatible with the current present. The use of the quantifier strictly more than one is important here as if it were not the case (i.e. there is only one configuration of future events), then the present would be "caused/determined" by the future as much as by the past. Causality would be time-symmetric. This assumption is the foundation of the past/future asymmetry we recover in this union. For instance, an observer trying to pre-solve time-symmetric exact equations will, due to this entropy, fail to predict which of the possible futures will actually occur. However, once the future has become present, solving those same equations backward will go as expected. We, of course, assume that an observer has no prior knowledge of his future (no foresight). Thus, configurations of compatible future events must be of maximum entropy otherwise information about the future would be available to $O$ (foresight).

Using this strategy, we can obtain a system of statistical physics over space-time events that follows the rules of special relativity.

Other emergent features we would like to have (briefly listed and to be expanded in section 3) are: a) recovering the fundamental equation of special relativity, linking time to space $dx = cdt$ from the equation of state; b) having the entropy at the horizon corresponding to the Bekenstein-Hawking entropy $S = k_B A/(4 L_p^2)$; c) having the speed of light defined as an emergent property of the system, which is constant and that cannot be exceeded; d) connecting time-like separated events with a "time-like entropy" and space-like separated events with a "space-like entropy"; e) having an emergent arrow of space-time that generalizes the arrow of time to special relativity. This list of requirements might sound like the statistical system would be complicated to describe, but all it takes is maximizing the entropy on average $t$ (system age) and $x$ (system size), and everything we need will emerge out of it. Let's get started.

2.1. Background

We suppose a 3+1 space-time $M_4$ in spherical coordinates $\{r, \theta, \phi, t\}$. Under isotropic assumptions $\{d\theta = 0, d\phi = 0\}$ the space-time is simplified to a 1+1 space-time $M$ with coordinates $\{r, t\}$, and it enforces the constraint that $r \in \mathbb{R}_{\geq 0}$. Additionally, we pose that $t = 0$ is the origin of the system, and thus $t \in \mathbb{R}_{\geq 0}$. Finally, we denote an observer by $O$.

2.2. Events

Let $Q$ be the set of events in $M$. We define the functions $r$ and $t$ as mapping each event $q \in Q$ to its space-time position in $M$ as:

$$r: Q \rightarrow \mathbb{R}_{\geq 0} \text{ [meters]}$$

$$t: Q \rightarrow \mathbb{R}_{\geq 0} \text{ [seconds]}$$

2.3. Macroscopic system of events

We consider a macroscopic system defined for a set of events $Q$ and two macroscopic quantities (the priors): an average event-time $\bar{t} \in \mathbb{R}_{\geq 0}$, and an average event-distance $\bar{r} \in \mathbb{R}_{\geq 0}$. We offer two justifications for these priors.

Mathematical justification: First, to treat the events as microscopic elements of a macroscopic system of statistical physics, we are required to take the averages of the quantities $r(q)$ and $t(q)$ as the priors $\bar{r}$ and $\bar{t}$ of the macroscopic system when we derive the Gibbs ensemble under an appropriate notion of equilibrium. This is analogous to taking the average energy $\bar{E}$ as the macroscopic description of a system microscopically described by energy levels $E(q)$, under the assumptions of thermodynamic...
equilibrium. Another example would be taking an average volume \( V \) over the possible volumes \( V(q) \), etc.

Physical justification:

Math notwithstanding, this is a physics paper and thus, the priors must be physically justified. First, consider that in usual special relativity, the observer is conceptually understood to be the origin of the system, and as such its light cone extends to infinity both in the past and in the future. Empirically however, this is not the quite the case. The past light cone of any observer that we find in nature (e.g. of real observers) does not extend to infinity in the past, but terminates at the Big Bang. Thus, such real observers come with a prior light cone of a certain size and age. For instance, the light cone of a present-day observer will come with priors regarding the size of its light cone and the origin is not the observer, but instead coincide with the Big Bang at \((0,0)\). In this context, the priors are essentially the initial conditions regarding the size of the light cone of real observers respective to the origin of the system.

This accounts for the priors being non-null, but it doesn’t yet answer why these priors are average values instead of exact values. To address this, let us first give a classical example, then we will justify it more fundamentally.

1. If we were to request five hundred independent measurements of the size of the present-day particle horizon, we would likely get five hundred slightly different values. Thus, in the case where the values of the priors are empirically derived, we would end up with an average experimental value with some expected fluctuations over the set of measurements. This argument is purely classical, but we can make it fundamental by evoking notions of quantum mechanics.

2. In the formalism of special relativity, the observer \( O \) is point-like and the size of the light cone expanding away from \( O \) can be measured to arbitrary precision. Ergo, the theory is not aware of the quantum mechanical restrictions on the precision of measurements. Assume that instead of being point-like, \( O \) is of a size described by its Compton wavelength. In this case, since the observer is not point-like, the dimension of the light cone expanding from it can no longer be considered arbitrarily precisely, even fundamentally. Thus, in this context we would expect that any attempt at measuring the dimensions the surface of the light cone will exhibit a certain statistical character preventing its measurement to a precision exceeding the Compton wavelength of the object of reference.

Referencing the Compton wavelength has previously and successfully been done before in an analogous context regarding the Black Hole information paradox, general relativity, black hole entropy and statistical physics in (Hawking [31]) and to entropy over positional information in (Verlinde [28]). Here, we will soon recover the Compton wavelength explicitly from the equation of state (equation 25), and by doing so we will be able to show that the system exactly connects to a gamut of related physical laws including the law of inertia, the Unruh temperature, the Bekenstein-Hawking entropy (equation 27). For these reasons, interpreting the Compton wavelength as a limitation on the size of the observer does appear to be the proper physical relation to limit the precision of the point-like observer in the context of statistical physics \( \cup \) special relativity, as it will make the union nicely fit into many prior results of physics.

2.4. Gibbs ensemble of events

Under the principle of maximum entropy, we seek the probability distribution \( \rho: Q \rightarrow \{ p \in \mathbb{R} | 0 \leq p \leq 1 \} \) and \( \sum_{q \in Q} \rho(q) = 1 \) which maximizes the entropy \( S \):

\[
S = -k_B \sum_{q \in Q} \rho(q) \ln \rho(q)
\]  

(3)

and subject to the priors \( t \) and \( \tau \)
\[ T = \sum_{q \in Q} \rho(q) t(q) \]  
\[ \tau = \sum_{q \in Q} \rho(q) r(q) \]  

We maximize the entropy using the well-known method of the Lagrange multipliers.

\[ L = \left( -k_B \sum_{q \in Q} \rho(q) \ln \rho(q) \right) + \lambda_0 \left( \sum_{q \in Q} \rho(q) - 1 \right) \]
\[ + \lambda_1 \left( \sum_{q \in Q} \rho(q) t(q) - T \right) + \lambda_2 \left( \sum_{q \in Q} \rho(q) r(q) - \tau \right) \]  

Maximizing \( L \) with respect to \( \rho(q) \) is done by taking its derivative and posing it equal to zero:

\[ \frac{\partial L}{\partial \rho(q)} = -k_B \ln \rho(q) - k_B + \lambda_0 + \lambda_1 t(q) + \lambda_2 r(q) = 0 \]

Solving for \( \rho(q) \) we obtain:

\[ \rho(q) = \exp \left( \frac{-k_B + \lambda_0 + \lambda_1 t(q) + \lambda_2 r(q)}{k_B} \right) \]

From the constraint \( 1 = \sum_{q \in Q} \rho(q) \), we can find the value for \( \lambda_0 \):

\[ \sum_{q \in Q} \rho(q) = 1 = \sum_{q \in Q} \exp \left( \frac{-k_B + \lambda_0 + \lambda_1 t(q) + \lambda_2 r(q)}{k_B} \right) \]
\[ = \exp \left( \frac{-k_B + \lambda_0}{k_B} \right) \sum_{q \in Q} \exp (\lambda_1 t(q) + \lambda_2 r(q)) \]

We define the partition function \( Z \) to be

\[ Z := \sum_{q \in Q} \exp (\lambda_1 t(q) + \lambda_2 r(q)) \]

Then, we rewrite \( \rho(q) \) using \( Z \), we pose \( \lambda_1 := 1/t_0 \) and \( \lambda_2 := -1/r_0 \) and we obtain the probability distribution:

\[ \rho(q) = \frac{1}{Z} \exp \left( \frac{1}{t_0} t(q) - \frac{1}{r_0} r(q) \right) \]

where \( 1/t_0 \) with units \([1/\text{seconds}]\) and \(-1/r_0 \) with units \([1/\text{meters}]\) are the Lagrange multipliers (a justification for the choice of signs will be provided after the results in section 4.3). Finally, we obtain the equation of state:
which represents the macroscopic evolution of the system, and where \(k_B\) is Boltzmann’s constant. The reason why the author has elected to produce the explicit derivation of the Gibbs ensemble for this system is to show clearly that a Gibbs ensemble of statistical physics (such as the one here) can legitimately be constructed without the introduction of an emergent temperature (as a Lagrange multiplier) associated with thermodynamic equilibrium. Thus, the present system holds outside of thermodynamic equilibrium, although another type of equilibrium is required on the \(1/t_0\) and \(-1/r_0\) Lagrange multipliers. Perhaps the name “tempo-dynamic equilibrium” is fitting? In this case, the Lagrange multipliers \(1/t_0\) and \(-1/r_0\) would be the "tempo-ture" of the system.

Definition: We will define "light-like" entropy, "time-like" entropy, and "space-like" entropy. Each is obtained by solving \(dS\) in (14) but under different conditions. The first refers to \(dS\) at tempo-dynamic equilibrium \((1/t_0dt = 1/r_0d\tau)\). The second refers to the case where \(1/t_0dt > 1/r_0d\tau\), and the third to \(1/t_0dt < 1/r_0d\tau\).

Remark: Since the universe is not at uniform temperature, cosmological thermodynamics has been focused on the study of event horizons, which admits temperatures (Hawking [31], Unruh [32]). Resisting the temptation to include an average \(T\), and thus relaxing the requirement that the system be at thermodynamic equilibrium with a uniform temperature \(t\) was a key insight which opened the door to apply statistical physics away from the surface of horizons, and within the volume of the enclosing surface. This is possible at "tempo-dynamic" equilibrium. As we will see, instead of admitting a uniform temperature (as in the thermodynamic equilibrium case), a system at tempo-dynamic equilibrium admits a uniform maximum speed.

3. Results

3.1. Tempo-dynamic quantities

Consistent with the standard interpretation of statistical physics, physical quantities that are extensive are conjugated with an intensive quantity. For instance the volume \(V\) in statistical physics is extensive and combining the volume of two sub-systems increases the volume by the sum \(V_1 + V_2 = V\) (extensive), but the pressure is not added: \(p_1 = p_2 = p\) (intensive). Experimentally, two systems with different pressures \(p_1 \neq p_2\) can be joined together, but this breaks the thermodynamic equilibrium until the pressures equalize. Likewise, a process which consumes \(t_1\) seconds followed by a process consuming \(t_2\) seconds consumes a total of \(t_1 + t_2 = t\) seconds (extensive) while its conjugate time is intensive \((1/t_0\) is a constant Lagrange multiplier). Same goes with position: adding two meters end-to-end doubles the length (extensive) whereas its conjugate position, as it is a Lagrange multiplier, remains the same (intensive). These quantities are summarized in the Table (1).

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Name</th>
<th>Units</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>(r(q))</td>
<td>position</td>
<td>(m)</td>
<td>extensive</td>
</tr>
<tr>
<td>(k_B/r_0)</td>
<td>conjugate to position</td>
<td>((J/K)m^{-1})</td>
<td>intensive</td>
</tr>
<tr>
<td>(\tau)</td>
<td>average position</td>
<td>(m)</td>
<td>macroscopic</td>
</tr>
<tr>
<td>(t(q))</td>
<td>time</td>
<td>(s)</td>
<td>extensive</td>
</tr>
<tr>
<td>(k_B/t_0)</td>
<td>conjugate to time</td>
<td>((J/K)s^{-1})</td>
<td>intensive</td>
</tr>
<tr>
<td>(\overline{t})</td>
<td>average time</td>
<td>(s)</td>
<td>macroscopic</td>
</tr>
</tbody>
</table>

Table 1. The relevant tempo-dynamic quantities
3.2. Tempo-dynamic cycles in time and space

Equation (14), as it references a time $t$ and a distance $r$, imposes requirements on the entropy of the system as the light cone expands. To investigate the behavior, let’s us take Figure (1) as an example of a simple tempo-dynamic cycle involving time and space. The cycle is reminiscent of a light cone, but it makes additional claims about the entropy. The transitions along the cycle are:

1. isospatial process: While transiting from $O$ to $A$ and by keeping the distance constant ($dD = 0$), the system decreases its entropy: $dS/dt = -k_B/t_0$.

2. isotemporal process: While transiting from $A$ to $B$ and by keeping the time constant ($dD = 0$), the system increases its entropy: $dS/dL = k_B/r_0$.

3. isentropic process: While transiting from $B$ to $O$ and by keeping the entropy constant ($dS = 0$), the system’s size increases at a characteristic constant speed: $dD/dt = r_0/t_0$.

3.3. Fermi-Dirac statistics of events

We consider that an event can occur at most once (whatever happens to Schrödinger’s cat, for sure, it doesn’t die twice), and thus we will use Fermi-Dirac statistics to study the occupancy distribution of events. In the case of (14), its Fermi-Dirac distribution under the assumption that $\mu = 0$ is:

$$n(q, t_0, r_0) = \frac{1}{\exp \left( \frac{1}{t_0} t(q) - \frac{1}{r_0} r(q) \right) + 1}$$  \hspace{1cm} (15)

To better understand what is going on with this equation (Fermi-Dirac statistics over space and time quantities), it helps to first illustrate the Fermi-Dirac statistics of both $r(q)$ and $t(q)$ in isolation. Therefore, consider these Fermi-Dirac distributions applicable to $r(q)$ and $t(q)$, respectively. In each case, one of the two quantities has been made constant for the purposes of simplification. The two distributions (16) and (17) are illustrated in Figure 2a and 2b, respectively. The equations are:

$$n(q, x_0) = \frac{1}{\exp \left( T_x - \frac{1}{x_0} r(q) \right) + 1}$$  \hspace{1cm} (16)

$$n(q, t_0) = \frac{1}{\exp \left( \frac{1}{t_0} t(q) - t_t \right) + 1}$$  \hspace{1cm} (17)

Now we are ready to investigate the Fermi-Dirac distribution given in (15) as illustrated in Figure 3.
Figure 2. a) Fermi-Dirac statistics (equation 16) over the occupancy of space-time events in relation to their distance from \( r = 0 \), while holding the average time \( \bar{t}_r \) constant. The shape of the curve is of the familiar shape and direction as the well-known distribution applicable to energy levels. The slope at \( t_r \) is correlated to the value of \( r_0 \) and is analogous to \( k_B T \) in the case of energy levels. b) Fermi-Dirac statistics (equation 17) over the occupancy of space-time events in relation to their time \( t \), while holding the average position \( \bar{r}_t \) constant. The shape is the same as the previous case, but its direction is mirrored at the \( \bar{r}_t \) line. c) and d): We argue at the end of section 3.3 that Bose-Einstein statistics are inapplicable for space-time events.
Figure 3. Fermi-Dirac statistics over the occupancy of space-time events (equation 15 and with \( \mu = 0 \)). Red means an occupancy rate of 100\%, whereas blue means 0\% (and with rainbow colors for intermediate values). a) The slices \( t_1 \) and \( t_2 \) have the same shape as Figure 2a. The slices \( x_1 \) and \( x_2 \) have the same shape as Figure 2b. As the system goes from \( P_1 \) to \( P_2 \), occupied past states are depleted as distance states become saturated. The slope of the line from \( P_1 \) to \( P_2 \) is the speed of light, associated with the ratio of the tempo-ture of the system. b) The image on the right is a perspective view of the image on the left.

Description: For the future, the occupancy rate of events is saturated (future events are upcoming). For the past, the occupancy rate of events is depleted (past events are gone). An observer \( O \) evolving from \( P_1 \) to \( P_2 \) will, in \( O \)'s present, experience a transfer in the saturation of the occupancy of future events to a saturation in the occupancy of events in space. This is the point at which we introduce the analogy of the waterfall of events flooding space. Along with \( O \), this "waterfall of events" recedes in space-time at the speed of light towards the direction of the future as it "floods" local space with events. \( O \) is prohibited by entropy from knowing future events until the waterfall recedes appropriately. Past events are depleted from the system following the passage of the waterfall. Since \( O \) sees no occupied past events, \( O \)'s future in space-time lies in the forward time direction, and the waterfall recedes towards the future. Let’s prove this direction in the next section, and then return to this discussion in section 4.

A note on Bose-Einstein statistics: We believe that Bose-Einstein statistics (BE) are inappropriate for the description of events for the following reasons:

1. Repeating an event that has already occurred (if such a thing is even possible) is not expected to contribute to the information of the system. For instance, re-measuring Schrödinger’s cat multiple times over does not make it more or less dead than it already is (if dead, or alive otherwise). Generally, performing a quantum measurement on a quantum system already in an eigenstate leaves the system unchanged. Thus, we would expect events to be registered only when they first occur; that is to say, their informational contribution occurs at most once.

2. The occupancy rate of events described by Bose-Einstein statistics falls quickly away from the observer and reaches near-zero well before reaching the event horizon of the light cone. Thus, the argument that the occupancy rate of events represents those events that are causally connected to the observer, does not hold under Bose-Einstein statistics.

3. Under Bose-Einstein statistics, future states are depleted and both the past and the inside of the event horizon are undefined. Figure (2c) and (2d) show that Bose-Einstein statistics are
only defined for the future and for outside of the event horizon. This is inconsistent with event
horizons in special relativity.

4. Under Bose-Einstein statistics, the present occurs at a mathematical singularity which is,
obviously, undesirable.

3.4. Arrow of time

Our goal here is to prove that the waterfall recedes in the direction of the future.

Let us first investigate the production of entropy over time in (14). To do so, we divide each side
of (14) by $\frac{dt}{d}$ then multiply it by $\frac{r_0}{k_B}$. We obtain:

$$\frac{r_0}{k_B} \frac{dS}{dt} = \frac{r_0}{t_0} - \frac{d\bar{r}}{dt}$$

(18)

First, we note that as both $-\frac{1}{r_0}$ and $\frac{1}{t_0}$ are Lagrange multipliers, then both $r_0$ and $t_0$ are
uniform in the system. The ratio $r_0/t_0$ is a speed [meters/seconds] and is also uniform.

Second, we note that (18) represents an inflection point in the production of entropy in the system
at $dS/dt = 0$. Specifically, when $dS/dt < 0$, the entropy of the system decreases as a whole, which
is prohibited by the second law of thermodynamics. This occurs if the macroscopic system ($\tau$ and $\bar{t}$)
shrinks ($d\bar{r}/dt < 0$) or grows slower than the ratio $r_0/t_0$.

Discussion: We now conclude, based on the two points mentioned, that the arrow of time points
towards the future of the macroscopic system whose growth in space-time is prohibited by the second
law of thermodynamics from being less than the ratio of the tempo-ture. Now that we know the
minimum growth rate, we might wonder, what, if anything, prevents the system from exceeding the
ratio? Answer: the observer does not see events beyond the horizon because the occupancy probability,
given by Fermi-Dirac statistics, sharply drops to zero at the horizon. This limits the growth rate of the
system as perceived by $O$ to the ratio of its tempo-ture.

3.5. Recording the passage of time

We integrate equation (14). We get:

$$\int dS = -\frac{k_B}{t_0} \int d\bar{t} + \frac{k_B}{r_0} \int d\bar{r}$$

(19)

$$\Delta S = -\frac{k_B}{t_0} \Delta \bar{t} + \frac{k_B}{r_0} \Delta \bar{r} + C$$

(20)

This relation leads to two inequalities:

$$0 \leq -\Delta S \leq \frac{k_B}{t_0} \Delta \bar{t} \quad \quad 0 \leq \Delta S \leq \frac{k_B}{r_0} \Delta \bar{r}$$

(21)

The first, involving time, relates the minimum amount of information ($-\Delta S$) that must be acquired
to prove that time has passed by a certain amount $\Delta t$. It is interpreted in the sense that the passage of
time $\Delta \bar{t}$ requires the logging of an event $-\Delta S$. The second, involving space, relates an entropy ($\Delta S$)
to the minimum increase in the size of space $\Delta \bar{r}$ required to accommodate it. Let us now study this
behavior in more detail by constructing an (abstract) thermodynamic engine that converts time to
space.

3.6. Space-time engine

We will define a thermodynamic engine that converts time to space. The engine is comprised of
a detector and a tape. The engine can write one bit on the tape at $r = 0$. It can also shift the tape to
the right by one increment \( \Delta r \) (in preparation to write the next bit). For purposes of idealization, we consider that the engine never runs out of tape. Finally, and without loss of generality, it helps for the purposes of the illustration to consider the more familiar case where the system is at thermodynamic equilibrium. Thus, we pose thermodynamic equilibrium to the inequalities with these replacements: 

\[ k_B / T_0 := P / T \] where \( P \) is a power in \([\text{Joules/seconds}]\), and \( k_B / r_0 := F / T \) where \( F \) is a force in \([\text{Joules/meters}]\), and where \( T \) is a temperature in \([\text{Kelvins}]\). We get:

\[
0 \leq -\Delta S \leq \frac{P}{T} \Delta \tau \quad 0 \leq \Delta S \leq \frac{F}{T} \Delta \tau
\] (22)

To ease the abstraction, we can picture a concrete system sharing similar characteristics (with some limitations), such as a seismograph tracing seismic data (collected over time) on a rolling tape (stored in space). The cycles of the engine occur in parallel and are completed over a time period \( \Delta t \).

Each cycle represents a logical step in the process (not chronological). They are:

1. Shifting the tape to the right by \( \Delta r \) is favored by entropy as it increases it by \( \Delta S \). Thus, an entropic force \( F \) emerges which pulls on the tape. For this engine, we associate this increase in entropy to an undefined memory address (\( \Delta S = k_B \ln 2 \)) which becomes available at position 0 of the tape when it is shifted by \( \Delta \tau \).
2. The detector clicks (it produces information). To register the click, \( \Delta \tau \) increases as \( \Delta S \) is decreased. To pay the energy cost to decrease the entropy, the detector draws a power \( P \) from the engine over the time period \( \Delta t \).
3. Finally, the engine writes the bit associated with the click in the undefined memory address of the tape. This reclaims the energy of the shift (\( F \Delta r \)) that produced the entropy and instead makes it available to power the detector (\( P \Delta t \)).

This engine has an interesting property: the further along the tape one looks, the further back in time the information on the tape refers to. At tempo-dynamic equilibrium, the tape is shifted at the speed of light towards the right, and its farthest bit refers to the very first thing that has been recorded. Ergo, this engine produces a light cone.

### 3.7. Gamut of related laws

Let us calculate the entropy at the horizon for the system then we will be in a good position to discuss these results.

We can show that the entropy at the horizon is no greater than the Bekenstein-Hawking entropy (Bekenstein [29], Hawking [31], Susskind [33]). To derive it, we must be consistent with the conditions permitting the derivation of the Bekenstein-Hawking entropy in the first place: event horizons have a temperature (Hawking [31], Unruh [32]).

First, we pose \( \frac{dT}{dt} = 0 \). Then, the first step of the proof will involve taking the system at tempo-dynamic equilibrium (14) and making it into a system that is also at thermodynamic equilibrium. To do so, we must insert a temperature into the system while keeping in mind that \( 1/t_0 \) is a Lagrange multiplier and, thus, is uniform at equilibrium. Preserving the units while injecting a temperature, we pose the relation \( k_B / t_0 := F / T \) where \( F \) is a force in \([\text{Joules/meters}]\) and \( T \) is a temperature in \([\text{Kelvin}]\), and we get:

\[
dS = \frac{F}{T} d\tau
\] (23)

For the system to be at thermodynamic equilibrium (and thus to admit a temperature), we are looking for a temperature associated with black-body radiation and applicable to an object under the action of a force. Within the context of special relativity, this is of course the Unruh effect (Unruh...
[32], Fulling [34], Davies [35]), whose characteristic temperature is given by
\[ T = \frac{\hbar}{2\pi k_B c} \]
and applicable to an object undergoing acceleration \( F := ma \). Making the replacements into \( F/T \), we get both the ratio \( F/T \) and \( T \) to remain uniform as required, and we obtain:

\[ \frac{F}{T} = 2\pi k_B \frac{mc}{\hbar} \]  

(24)

Since the mass \( m \) can change, we have to insert the ratio \( F/T \) into \( S = \frac{F}{T} r \), then take the total derivative of \( S \):

\[ dS = 2\pi k_B \frac{mc}{\hbar} dr + 2\pi k_B r \frac{c}{\hbar} dm \]  

(25)

where \( \hbar/(mc) \) is the well-known reduced Compton wavelength. Here, the first term is acting as the factor of proportionality between the entropy and the distance (Verlinde [28]). In the case of \( dr \), we find that the Compton wavelength mediates the intensity of the fluctuations around the average value \( r \), consistent with our justification of the priors in section 2.3. We notice that the higher the mass, the higher the entropy. Since the most massive object for a given \( r \) (radius) is a black hole, we pose \( r := 2Gm/c^2 \) (the Schwarzschild radius) as the upper limit on entropy. The black hole also has the benefit of being isotropic with respect to its center, consistent with our assumptions in section (2.1). We now use the Schwarzschild radius to replace \( r \) with \( m \) where appropriate to simplify (25) to:

\[ dS = 8\pi k_B \frac{Gm}{hc} dm \]  

(26)

After integrating, we get \( S = 4\pi k_B Gm^2/(hc) + C \), where \( C \) is an integration constant. Then, by posing \( A := 4\pi r_s^2 = 16\pi G^2 m^2/c^4 \) and using the Planck length \( L_p := \sqrt{\hbar G/c^3} \), we get a boundary on the size of the entropy as:

\[ S \leq k_B A \frac{1}{4L_p^2} + C \]  

(27)

which is proportional to the surface \( A \) and includes the factor \( 1/4 \). This result serves as a sanity check: The Bekenstein-Hawking entropy is recovered! Naturally, we interpret the surface as an event horizon and the ratio of tempo-ture \( r_0/t_0 \) as the speed of light.

4. Discussion

We can now understand how, precisely, recording events connects to (and even produces) the arrow of time. To investigate this, instead of considering the second law of thermodynamics as a logically independent axiom, we will inject it into an equation of state in order to study its contribution quantitatively. We recall the second law of thermodynamics which state that the entropy of a system stays constant or increases \( (dS \geq 0) \), but never decreases. In the context of the arrow of time, we are interested in the second law of thermodynamics as it pertains to time:

\[ \frac{dS}{dt} \geq 0 \]  

(28)

We can rewrite this inequality using a function \( C \) as follows:
\[ \frac{\partial S}{\partial t} \bigg|_{V,N,E} = C \quad \text{where } C \geq 0 \] (29)

Formulated as such, the arrow of time is mathematically equivalent to (28), but in this format it is more immediately recognizable as a candidate thermodynamic law. For instance, the laws of thermodynamics regarding entropy, volume and particle number are often expressed as:

\[ \frac{\partial S}{\partial E} \bigg|_{V,N} = \frac{1}{T} \quad \frac{\partial S}{\partial V} \bigg|_{E,N} = \frac{p}{T} \quad \frac{\partial S}{\partial N} \bigg|_{E,V} = -\frac{\mu}{T} \] (30)

To turn (29) into a valid thermodynamic law, it suffices to replace \( C \) with \( -P/T \) under an appropriate equilibrium situation, where \( P \) is an entropic power in [Joules/seconds] and \( T \) is the temperature in [Kelvin]. Taking these laws, along with (29), we can summarize them into an equation of state:

\[ dE = TdS - pd\text{V} + \mu d\text{N} + Pd\text{I} \] (31)

where the last term \( Pd\text{I} \) is responsible for enforcing the arrow of time explicitly within the equation of state itself.

To make sense of this equation of state, it helps to imagine a quantum gas whose molecules are under continuous measurements by the environment. First, recall that in the quantum case, the results of measurements are intrinsically random and therefore to reverse the system, contrary to the classical case, it is not sufficient to simply know, with infinite precision, the initial state of the air molecules (as per Laplace’s demon). In the quantum case each random measurements produced by the environment over time would have to be recorded in the degrees of freedom of the environment to make the history of the system complete in the information theoretic sense.

A problem however arises while doing this. The memory requirements of the log of event will outgrow any finite system, unless the system grows too. Why? Consider a gas at thermodynamic equilibrium. The quantum molecules are continuously being measured by the environment even at equilibrium and thus the system must store a perpetually growing log of events. To accommodate the growing log, the system must somehow grow its memory proportionally to the rate at which events are produced. This is why special relativity emerges as an entropic law in this context. Indeed growing, at the speed of light, an event horizon (which bears an entropy) around the system guarantees that the memory requirements for storing all such events as they occur in time are met. The entropy on the surface of this event horizon is sufficient to make the system thermodynamically-reversible in time, and thus the log of events is complete. From this, we are now ready to discuss how the past, the present and the future emerges from these properties.

4.1. Three regimes of time

Reprising our discussion over the system’s qualitatively different description of the past, the present, and the future, we now study the entropy dynamics of both time and distance in the context of entropic special relativity. As concluded in section (3.7), when time is ignored, an observer perceives the space-like entropy to be the Bekenstein-Hawking entropy. However, the situation is different at tempo-dynamic equilibrium when time is included. In this case, we find that the entropy reduction of increasing the \( \mathcal{I} \) quantity exactly compensates the entropy increase associated with increasing the \( \tau \) quantity. From this, we would find that the observer, at tempo-dynamic equilibrium, always sees an entropy of 0 in the present (plus an integration constant, to be neglected from the discussion). How do we make sense of this result?
First we note that in the quantum case, unlike Laplace’s demon in the classical case, knowing the initial conditions of the system is not sufficient to replay it. However, with both the initial conditions and the log of events, we have enough information to replay the system from the beginning even in the quantum case. As this information is sufficient to fix the present to a single solution, then any additional information beyond it is therefore necessarily redundant which is why, inevitably, we obtained the Bekenstein-Hawking entropy, as an upper bound on entropy, in this case.

The surface of the event horizon can be interpreted as keeping a record of all events relevant to O’s present that have transpired since the Big Bang, in a manner complete in the information theoretic sense such that the present state of the observer can be exactly recovered by replaying the log from the initial conditions and under an appropriate quantum theory.

We describe each regime of macroscopic time as follows:

- **Present**: We associate light-like entropy with the uniquely determined present. As stated, the present has an entropy of 0. It is uniquely determined by the initial conditions of the system plus the log of random events that have occurred since the beginning of the system up to the present time. As a result, the observer cannot be in a superposition of multiple presents and consequently, we expect the observer to “measure” Schrödinger’s cat as either dead or alive, but not as a superposition of both.

- **Past**: We associate space-like entropy with a trace of the past. As the waterfall recedes in space-time, it leaves a trace of events in the degrees of freedom of space. An observer can, by inspecting the trace, find evidence for a consistent past to account for the present. The horizon produced by the depletion of past events under Fermi-Dirac statistics prohibits O from going backward or interacting with past events (Figure 2b) directly. A similar horizon, also enforced by state depletion, is found at the edge of the event horizon in space (Figure 2a).

- **Future**: Finally, we associate time-like entropy with the future. The observer is prohibited from peeking into its future (increasing \( T \) above tempo-dynamic equilibrium) until it is imminent. Indeed, an observer peeking into the future (without first waiting for the waterfall to recede appropriately) will hit negative entropy (contradiction). The observer can hypothesize about its possible futures, but the actual future is made final no sooner than in the present when the entropy hits 0. This negative entropy prevents the macroscopic system from growing faster than allowed by tempo-dynamic equilibrium.

So why do we think there is an entropy in the gas of the room I am in, for instance, if the entropy of the present is allegedly zero? This is because I do not know all the bits of the log of events, thus, I can successfully approximate the gas in the room as an entropic system. Partial knowledge of the trace of events by O limits the uniqueness of the reconstruction of the past achievable by O based on the analysis of the trace, as it represents the number of logs of events that are compatible with the observer’s unique present under partial knowledge of the log.

What about the arrow of time? This now accounts for only half of the story. The waterfall of events, as it recedes in space-time, creates two arrows related to time. The first arrow is the familiar one. The entropy associated with the degrees of freedom of space \( ((k_B/r_0)dT) \) increases as time moves forward. The second arrow acts in the reverse direction on the degrees of freedom of time \( ((-k_B/r_0)dT) \).

With it, the entropy associated with time decreases as time moves forward because future possibilities are consumed to create a present.

### 4.2. Falsifiable prediction

From equation (14), and preserving the units, we impose thermodynamic equilibrium on the system with the replacement \( k_B/r_0 := P/T \) where \( P \) is a power [Joules/seconds] and \( T \) is a temperature [Kelvin]. By further posing \( dr = 0 \). We get:
This equation predicts that an entropic power can be made to emerge if information is produced (or consumed) over time and at constant temperature. This prediction regarding the possibility of an emergence of an entropic power should be relatively easy to show experimentally.

4.3. Note on the signs of the Lagrange multipliers

The signs of the Lagrange multipliers were chosen for the following reasons: 1) Starting with (14) and posing $dS = 0$ we get $d\tau = cd\tau$ (where $c := r_0/t_0$). This is the fundamental relation of special relativity connecting space to time. Thus, the signs of the Lagrange multipliers must be opposite. 2) In the derivation of the Bekenstein-Hawking entropy, we have used $F = ma$, and not $F = -ma$. Thus, the Lagrange multiplier of $-1/r_0$ must be negative. 3) The combination of reason one and two implies that the sign of $1/t_0$ must be positive.

5. Conclusion

We conclude that the statistical physics of space-time events admit the following:

1. The speed of light as the ratio of the tempo-ture of the system.
2. A waterfall of events receding towards the future and flooding the present with events.
3. A surface boundary to maintain the average growth of the system consistent with the speed of light and causality. This limits the entropy of the inner system proportionally to its area (Bekenstein-Hawking entropy).
4. A description of all three regimes of time (past, present, and future) distinctly from one another, and in a manner consistent with our macroscopic experience of said regimes. Specifically:
   (a) The observer perceives the macroscopic present with an entropy of 0, negating the possibility of being in a superposition of multiple possible presents,
   (b) The observer cannot peak into the future without hitting negative entropy (contradiction),
   (c) The observer cannot observe past events as their occupancy is depleted. At best, an attempt to reconstruct the past can be made based on a forensic analysis of the present. With partial knowledge of the log of events, this reconstruction is not uniquely determined as multiple logs of events lead to the same present.

Funding: This research received no external funding

Conflicts of Interest: The author declares no conflict of interest.

References


