Denison University

Denison Digital Commons

2018

Axiomatic Information Thermodynamics

Austin Hulse

Benjamin Schumacher

Mike Westmoreland
Denison University

Recommended Citation

Axiomatic Information Thermodynamics

Austin Hulse [1], Benjamin Schumacher [1,*] and Michael D. Westmoreland [2]

[1] Department of Physics, Kenyon College, Gambier, OH 43022, USA; hulsea@kenyon.edu
[2] Department of Mathematics, Denison University, Granville, OH 43023, USA; westmoreland@denison.edu
[*] Correspondence: schumacherb@kenyon.edu; Tel.: +1-740-427-5832

Abstract: We present an axiomatic framework for thermodynamics that incorporates information as a fundamental concept. The axioms describe both ordinary thermodynamic processes and those in which information is acquired, used and erased, as in the operation of Maxwell's demon. This system, similar to previous axiomatic systems for thermodynamics, supports the construction of conserved quantities and an entropy function governing state changes. Here, however, the entropy exhibits both information and thermodynamic aspects. Although our axioms are not based upon probabilistic concepts, a natural and highly useful concept of probability emerges from the entropy function itself. Our abstract system has many models, including both classical and quantum examples.

Keywords: axiomatic thermodynamics; information; entropy; Maxwell's demon

1. Introduction

Axiomatic approaches to physics are useful for exploring the conceptual basis and logical structure of a theory. One classic example was presented by Robin Giles over fifty years ago in his monograph *Mathematical Foundations of Thermodynamics* [1]. His theory is constructed upon three phenomenological concepts: thermodynamic states, an operation (+) that combines states into composites, and a relation (→) describing possible state transformations. From a small number of basic axioms, Giles derives a remarkable amount of thermodynamic machinery, including conserved quantities ("components of content"), the existence of an entropy function that characterizes irreversibility for possible processes, and so on.

Alternative axiomatic developments of thermodynamics have been constructed by others along different lines. One notable example is the framework of Lieb and Yngvason [2,3] (which has recently been used by Thess as the basis for a textbook [4]). Giles's abstract system, meanwhile, has found application beyond the realm of classical thermodynamics, e.g., in the theory of quantum entanglement [5].

Other work on the foundations of thermodynamics has focused on the concept of information. Much of this has been inspired by Maxwell's famous thought-experiment of the demon [6]. The demon is an entity that can acquire and use information about the microscopic state of a thermodynamic system, producing apparent violations of the Second Law of Thermodynamics. These violations are only "apparent" because the demon is itself a physical system, and its information processes are also governed by the underlying dynamical laws.

Let us examine a highly simplified example of the demon at work. Our thermodynamic system is a one-particle gas enclosed in a container (a simple system also used in [7]). The gas may be allowed to freely expand into a larger volume, but this process is irreversible. A "free compression" process that took the gas from a larger to a smaller volume with no other net change would decrease the entropy of the system and contradict the Second Law. See Figure 1.

Figure 1. A one-particle gas may freely expand to occupy a larger volume, but the reverse process would violate the Second Law.

Now, we introduce the demon, which is a machine that can interact with the gas particle and acquire information about its location. The demon contains a one-bit memory register, initially in the state 0. First, a partition is introduced in the container, so that the gas particle is confined to the upper or lower half (*U* or *L*). The demon measures the gas particle and records its location in memory, with 0 standing for *U* and 1 for *L*. On the basis of this memory bit value, the demon moves the volume containing the particle. In the end, the gas particle is confined to one particular smaller volume, apparently reducing the entropy of the system. This is illustrated in Figure 2.
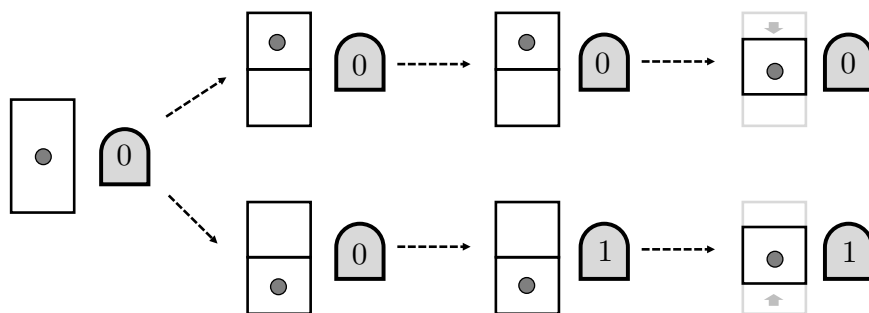

Figure 2. A Maxwell's demon device acquires one bit of information about the location of a gas particle, allowing it to contract the gas to a smaller volume.

As Bennett pointed out [8], every demon operation we have described can be carried out reversibly, with no increase in global entropy. Even the "measurement" process can be described by a reversible interaction between the particle location and the demon's memory bit b, changing the states according to

$$(\text{U}, b) \longleftrightarrow (\text{U}, b) \qquad (\text{L}, b) \longleftrightarrow (\text{L}, \bar{b}), \tag{1}$$

where $\bar{b}$ is the binary negation of b. However, the demon process as described leaves the demon in a new situation, since an initially blank memory register now stores a bit of information. To operate in a cycle (and thus unambiguously violate the Second Law), the demon must erase this bit and recover the blank memory. However, as Landauer showed [9], the erasure of a bit of information is always accompanied by an entropy increase of at least $k_B \ln 2$ in the surroundings, just enough to ensure that there is no overall entropy decrease in the demon's operation. The Second Law remains valid.

Information erasure is a physical process. The demon we have described can also erase a stored bit simply by reversing the steps outlined. Depending on the value of the bit register, the demon moves the gas particle to one of two corresponding regions separated by a partition. The same interaction that accomplished the "measurement" process (which takes $(U, 0) \dashrightarrow (U, 0)$ and $(L, 0) \dashrightarrow (L, 1)$) can now be used to reset the bit value to 0 (taking $(U, 0) \dashrightarrow (U, 0)$ and $(L, 1) \dashrightarrow (L, 0)$). In other words, the information stored redundantly in both the gas and the register is "uncopied", so that it remains only in the gas. Finally, the partition is removed and the gas expands to the larger volume. The net effect is to erase the memory register while increasing the entropy of the gas by an amount $k_B \ln 2$, in conventional units.

Another link between thermodynamics and information comes from statistical mechanics. As Jaynes has shown [10,11], the concepts and measures of information devised by Shannon for

communication theory [12] can be used in the statistical derivation of macroscopic thermodynamic properties. In a macroscopic system, we typically possess only a small amount of information about a few large-scale parameters (total energy, volume, etc.). According to Jaynes, we should therefore choose a probability distribution over microstates that maximizes the Shannon entropy, consistent with our data. That is to say, the rational probability assignment includes no information (in the Shannon sense) except that found in the macroscopic state of the system. This prescription yields the usual distributions used in statistical mechanics, from which thermodynamic properties may be derived.

　　Axiomatic theories and information analyses each provide important insights into the meaning of thermodynamics. The purpose of this paper is to synthesize these two approaches. We will present an axiomatic basis for thermodynamics that uses information ideas from the very outset. In such a theory, Maxwell's demon, which accomplishes state changes by acquiring information, is neither a paradox nor a sideshow curiosity. Instead, it is a central conceptual tool for understanding the transformations between thermodynamic states. In our view, thermodynamics is essentially a theory of the descriptions of systems possessed by agents that are themselves (like the demon) physical systems. These descriptions may change as the external systems undergo various processes; however, they also may change when the agent acquires, uses or discards information.

　　Our work is thus similar in spirit to that of Weilenmann et al. [13], though our approach is very different. They essentially take the Lieb-Yngvason axiomatic framework and apply it to various quantum resource theories. In particular, for the resource theory of non-uniformity [14], the Lieb-Yngvason entropy function coincides with the von Neumann entropy of the quantum density operator, which is a measure of quantum information. We, on the other hand, seek to modify axiomatic thermodynamics itself to describe processes involving "information engines" such as Maxwell's demon. We do not rely on any particular microphysics and have both classical and quantum models for our axioms. The connections we find between information and thermodynamic entropy are thus as general as the axioms themselves. Furthermore, to make our concept of information as clear as possible, we will seek to base our development on the most elementary ideas of state and process.

　　We therefore take as our prototype the axiomatic theory of Giles [1]. In fact, Giles's monograph contains two different axiomatic developments. (The first system is presented in Chapters 1–6 of Giles's book, and the second is introduced beginning in Chapter 7.) The first, which we might designate Giles I, is based on straightforward postulates about the properties of states and processes. The second, Giles II, is more sophisticated and powerful. The axioms are less transparent in meaning (e.g., the assumed existence of "anti-equilibrium" and "internal" thermodynamic states), but they support stronger theorems. This difference can be illustrated by the status of entropy functions in the two developments. In Giles I, it is shown that an entropy function exists; in fact, there may be many such functions. In Giles II, it is shown that an *absolute* entropy function, one that is zero for every anti-equilibrium state, exists and is unique. Giles himself regarded the second system as "The Formal Theory", which he summarizes in an Appendix of that title.

　　The information-based system we present here, on the other hand, is closer to the more elementary framework of Giles I. We have taken some care to use notation and concepts that are as analogous as possible to that theory. We derive many of Giles's propositions, including some that he uses as axioms. Despite the similarities, however, even the most elementary ideas (such as the combination of states represented by the $+$ operation) will require some modifications. These changes highlight the new ideas embodied in axiomatic information thermodynamics, and so we will take some care to discuss them as they arise.

　　Section 2 introduces the fundamental idea of an eidostate, including how two or more eidostates may be combined. In Section 3, we introduce the $\rightarrow$ relation between eidostates: $A \rightarrow B$ means that eidostate A can be transformed into eidostate B, with no other net change in the apparatus that accomplishes the transformation. The collection of processes has an algebraic structure, which we present in Section 4. We introduce our concept of information in Section 5 and show that the structure of information processes imposes a unique entropy measure on pure information states.

Section 6 introduces axioms inspired by Maxwell's demon and shows their implications for processes involving thermodynamic states. Sections 7 and 8 outline how our information axioms also yield Giles's central results about thermodynamic processes, including the existence of an entropy function, conserved components of content, and mechanical states. In Section 10, we see that an entropy function can be extended in a unique way to a wider class of "uniform" eidostates. This, rather remarkably, gives rise to a unique probability measure within uniform eidostates, as we describe in Section 11.

Sections 12 and 13 present two detailed models of our axioms, each one highlighting a different aspect of the theory. We conclude in Section 14 with some remarks on the connection between information and thermodynamic entropy, a connection that emerges necessarily from the structure of processes in our theory.

2. Eidostates

Giles [1] bases his development on a set $\mathscr{S}$ of *states*. The term "state" is undefined in the formal theory, but heuristically it represents an equilibrium macrostate of some thermodynamic system. (Giles, in fact, identifies $a \in \mathscr{S}$ with some method of preparation; the concept of a "system" does not appear in his theory, or in ours.) States can be combined using the $+$ operation, so that if $a, b \in \mathscr{S}$ then $a + b \in \mathscr{S}$ also. The new state $a + b$ is understood as the state of affairs arising from simultaneous and independent preparations of states a and b. For Giles, this operation is commutative and associative; for example, $a + b$ is exactly same state as $b + a$.

We also have a set $\mathscr{S}$. However, our theory differs in two major respects. First, the combination of states is not assumed to be commutative and associative. On the contrary, we regard $a + b$ and $b + a$ as entirely distinct states for any $a \neq b$. The motivation for this is that the way that a composite state is assembled encodes information about its preparation, and we want to be able to keep track of this information. On the states in $\mathscr{S}$, the operation $+$ is simply that of forming a Cartesian pair of states: $a + b = (a, b)$, and nothing more.

A second and more far-reaching difference is that we cannot confine our theory to individual elements of $\mathscr{S}$. A theory that keeps track of information in a physical system must admit non-deterministic processes. For instance, if a measurement is made, the single state a before the measurement may result in any one of a collection of possible states $\{a_0, a_1, \ldots, a_n\}$ corresponding to the various results.

Therefore, our basic notion is the *eidostate*, which is a finite nonempty set of states. The term "eidostate" derives from the Greek word *eidos*, meaning "to see". The eidostate is a collection of states that may be regarded as possible from the point of view of some agent. The set of eidostates is designated by $\mathscr{E}$. When we combine two eidostates $A, B \in \mathscr{E}$ into the composite $A + B$, we mean the set of all combinations $a + b$ of elements of these sets. That is, $A + B$ is exactly the Cartesian product (commonly denoted $A \times B$) for the sets. Thus, our state combination operation $+$ is very different from that envisioned by Giles. His operation is an additional structure on $\mathscr{S}$ that requires an axiom to fix its properties, whereas we simply use the Cartesian product provided by standard set theory, which we of course assume.

Some eidostates in $\mathscr{E}$ are Cartesian products of other sets (which are also eidostates); some eidostates are not, and are thus "prime". We assume that each eidostate has a "prime factorization" into a finite number of components. Here, is our first axiom:

Axiom 1. *Eidostates: $\mathscr{E}$ is a collection of sets called* eidostates *such that:*

(a) Every $A \in \mathscr{E}$ is a finite nonempty set with a finite prime Cartesian factorization.
(b) $A + B \in \mathscr{E}$ if and only if $A, B \in \mathscr{E}$.
(c) Every nonempty subset of an eidostate is also an eidostate.

Part (c) of this axiom ensures, among other things, that every element a of an eidostate can be associated with a singleton eidostate $\{a\}$. Without too much confusion, we can simply denote any singleton eidostate $\{a\}$ by the element it contains, writing a instead. We can therefore regard the set of states $\mathscr{S}$ in two ways. Either we may think of $\mathscr{S}$ as the collection of singleton eidostates in $\mathscr{E}$, or we may say that $\mathscr{S}$ is the collection of all elements of all eidostates: $\mathscr{S} = \bigcup\limits_{A \in \mathscr{E}} A$. Either way, the set $\mathscr{E}$ characterized by Axiom 1 is the more fundamental object.

Our "eidostates" are very similar to the "specifications" introduced by del Rio et al. [15] in their general framework for resource theories of knowledge. The two ideas, however, are not quite identical. To begin with, a specification V may be any subset of a state space Ω, whereas the eidostates in $\mathscr{E}$ are required to be finite nonempty subsets of $\mathscr{S}$—and indeed, not every such subset need be an eidostate. For example, the union $A \cup B$ of two eidostates is not necessarily an eidostate. The set $\mathscr{E}$ therefore does not form a Boolean lattice. Specifications are a general concept applicable to state spaces of many different kinds, and are used in [15] to analyze different resource theories and to express general notions of approximation and locality. We, however, will be restricting our attention to an $\mathscr{E}$ (and $\mathscr{S}$) with very particular properties, expressed by Axiom 1 and our later axioms, that are designed to model thermodynamics in the presence of information engines such as Maxwell's demon.

Composite eidostates are formed by combining eidostates with the $+$ operation (Cartesian product). The same pieces may be combined in different ways. We say that $A, B \in \mathscr{E}$ are *similar* (written $A \sim B$) if they are made up of the same components. That is, $A \sim B$ provided there are eidostates $E_1, \dots, E_n$ such that $A = F_A(E_1, \dots, E_n)$ and $B = F_B(E_1, \dots, E_n)$ for two Cartesian product formulas F_A and F_B. Thus,

$$(E_1 + E_2) + E_3 \sim E_2 + (E_1 + E_3), \tag{2}$$

and so on. The similarity relation $\sim$ is an equivalence relation on $\mathscr{E}$.

We sometimes wish to combine an eidostate with itself several times. For the integer $n \geq 1$, we denote by nA the eidostate $A + (A + \dots)$, where A appears n times in the nested Cartesian product. This is one particular way to combine the n instances of A, though of course all such ways are similar. Thus, we may assert equality in $A + nA = (n+1)A$, but only similarity in $nA + A \sim (n+1)A$.

Finally, we note that we have introduced as yet no probabilistic ideas. An eidostate is a simple enumeration of possible states, without any indication that some are more or less likely than others. However, as we will see in Section 11, in certain contexts, a natural probability measure for states does emerge from our axioms.

3. Processes

In the axiomatic thermodynamics of Giles, the $\rightarrow$ relation describes state transformations. The relation $a \rightarrow b$ means that there exists another state z and a definite time interval $\tau \geq 0$ so that

$$a + z \overset{\tau}{\rightsquigarrow} b + z, \tag{3}$$

where $\overset{\tau}{\rightsquigarrow}$ indicates time evolution over the period τ. The pair (z, τ) is the "apparatus" that accomplishes the transformation from a to b. This dynamical evolution is a deterministic process; that is, in the presence of the apparatus (z, τ) the initial state a is guaranteed to evolve to the final state b. This rule of interpretation for $\rightarrow$ motivates the properties assumed for the relation in the axioms.

Our version of the arrow relation $\rightarrow$ is slightly different, in that it encompasses non-deterministic processes. Again, we envision an apparatus (z, τ) and we write

$$a + z \overset{\tau}{\dashrightarrow} b + z, \tag{4}$$

to mean that the initial state $a + z$ may possibly evolve to $b + z$ over the stated interval. Then, for eidostates A and B, the relation $A \rightarrow B$ means that, if $a \in A$ and $b \in \mathscr{S}$, then there exists an

apparatus (z, τ) such that $a + z \xrightarrow{\tau} b + z$ only if $b \in B$. Each possible initial state a in A might evolve to one or more final states, but all of the possible final states are contained in B. For singleton eidostates a and b, the relation $a \to b$ represents a deterministic process, as in Giles's theory.

Again, we use our heuristic interpretation of $\to$ to motivate the essential properties specified in an axiom:

Axiom 2. *Processes: Let eidostates* $A, B, C \in \mathscr{E}$, *and* $s \in \mathscr{S}$.

(a) If $A \sim B$, then $A \to B$.
(b) If $A \to B$ and $B \to C$, then $A \to C$.
(c) If $A \to B$, then $A + C \to B + C$.
(d) If $A + s \to B + s$, then $A \to B$.

Part (a) of the axiom asserts that it is always possible to "rearrange the pieces" of a composite state. Thus, $A + B \to B + A$ and so on. Of course, since $\sim$ is a symmetric relation, $A \sim B$ implies both $A \to B$ and $B \to A$, which we might write as $A \leftrightarrow B$.

Part (b) says that a process from A to C may proceed via an intermediate eidostate B. Parts (c) and (d) establish the relationship between $\to$ and $+$. We can always append a "bystander" state C to any process $A \to B$, and a bystander singleton state s in $A + s \to B + s$ can be viewed as part of the apparatus that accomplishes the transformation $A \to B$.

We use the $\to$ relation to characterize various conceivable processes. A *formal process* is simply a pair of eidostates $\langle A, B \rangle$. Following Giles, we may say that $\langle A, B \rangle$ is:

- *natural* if $A \to B$;
- *antinatural* if $B \to A$;
- *possible* if $A \to B$ or $B \to A$ (which may be written $A \rightleftharpoons B$);
- *impossible* if it is not possible;
- *reversible* if $A \leftrightarrow B$; and
- *irreversible* if it is possible but not reversible.

Thus, any formal process must be one of four *types*: reversible, natural irreversible, antinatural irreversible, or impossible.

Any nonempty subset of an eidostate is an eidostate. If the eidostate A is an enumeration of possible states, the proper subset eidostate $B \subsetneq A$ may be regarded as an enumeration with some additional condition present that eliminates one or more of the possibilities. What can we say about processes involving these "conditional" eidostates? To answer this question, we introduce two further axioms. The first is this:

Axiom 3. *If $A, B \in \mathscr{E}$ and B is a proper subset of A, then $A \nrightarrow B$.*

This expresses the idea that no natural process can simply eliminate a state from a list of possibilities. This is a deeper principle than it first appears. Indeed, as we will find, in our theory, it is the essential ingredient in the Second Law of Thermodynamics.

To express the second new axiom, we must introduce the notion of a *uniform* eidostate. The eidostate A is said to be uniform provided, for every $a, b \in A$, we have $a \rightleftharpoons b$. That is, every pair of states in A is connected by a possible process. This means that all of the A states are "comparable" in some way involving the $\to$ relation.

All singleton eidostates are uniform. Are there any non-uniform eidostates? The axioms in fact do not tell us. We will find models of the axioms that contain non-uniform eidostates and others that contain none. Even if there are states $a, b \in \mathscr{S}$ for which $a \nrightarrow b$ and $b \nrightarrow a$, nothing in our axioms guarantees the existence of an eidostate that contains both a and b as elements. We denote the set of uniform eidostates by $\mathcal{U} \subseteq \mathscr{E}$.

Now, we may state the second axiom about conditional processes.

Axiom 4. *Conditional processes:*

(a) Suppose $A, A' \in \mathscr{E}$ and $b \in \mathscr{S}$. If $A \to b$ and $A' \subseteq A$ then $A' \to b$.
(b) Suppose A and B are uniform eidostates that are each disjoint unions of eidostates: $A = A_1 \cup A_2$ and $B = B_1 \cup B_2$. If $A_1 \to B_1$ and $A_2 \to B_2$ then $A \to B$.

The first part makes sense given our interpretation of the $\to$ relation in terms of an apparatus (z, τ). If every $a \in A$ satisfies $a + z \xrightarrow{\ \tau\ } b + z$ for only one state b, then the same will be true for every $a \in A' \subseteq A$. Part (b) of the axiom posits that, if we can find an apparatus whose dynamics transforms A_1 states into B_1 states, and another whose dynamics transforms A_2 states into B_2 states, then we can devise a apparatus with "conditional dynamics" that does both tasks, taking A to B. This is a rather strong proposition, and we limit its scope by restricting it to the special class of uniform eidostates.

We can as a corollary extend Part (b) of Axiom 4 to more than two subsets. That is, suppose $A, B \in \mathcal{U}$ and the sets A and B are each partitioned into n mutually disjoint, nonempty subsets: $A = A_1 \cup \cdots \cup A_n$ and $B = B_1 \cup \cdots B_n$. In addition, suppose $A_k \to B_k$ for all $k = 1, \ldots, n$. Then, $A \to B$.

4. Process Algebra and Irreversibility

Following Giles, we now explore the algebraic structure of formal processes and describe how the type of a possible process may be characterized by a single real-valued function. Although the broad outlines of our development will follow that of Giles [1], there are significant differences. (For example, in our theory, the unrestricted set $\mathcal{P}$ of eidostate processes does not form a group.)

There is an equivalence relation among formal processes, based on the similarity relation $\sim$ among eidostates. We say that $\langle A, B \rangle \doteq \langle C, D \rangle$ if there exist singletons $x, y \in \mathscr{S}$ such that $A + x \sim C + y$ and $B + x \sim D + y$. That is, if we augment the eidostates in $\langle A, B \rangle$ by x and those in $\langle C, D \rangle$ by y, the corresponding eidostates in the two processes are mere rearrangements of each other. It is not hard to establish that this is an equivalence relation. Furthermore, equivalent processes (under $\doteq$) are always of the same type. To see this, suppose that if $\langle A, B \rangle \doteq \langle C, D \rangle$ and $A \to B$. Then, $A + x \sim C + y$, etc., and

$$C + y \to A + x \to B + x \to D + y. \tag{5}$$

Hence, $C \to D$. It is a straightforward corollary that $A \to B$ if and only if $C \to D$.

Processes (or more strictly, equivalence classes of processes under $\doteq$) have an algebraic structure. We define the *sum* of two processes as

$$\langle A, B \rangle + \langle C, D \rangle = \langle A + C, B + D \rangle, \tag{6}$$

and the *negation* of a process as $- \langle A, B \rangle = \langle B, A \rangle$. We call the $-$ operation "negation" even though in general $- \langle A, B \rangle$ is not the additive inverse of $\langle B, A \rangle$. Such an inverse may not exist. As we will see, however, the negation does yield an additive inverse process in some important special contexts.

The sum and negation operations on processes are both compatible with the equivalence $\doteq$. That is, first, if $\langle A, B \rangle \doteq \langle A', B' \rangle$ then $- \langle A, B \rangle \doteq - \langle A', B' \rangle$. Furthermore, if $\langle A, B \rangle \doteq \langle A', B' \rangle$ and $\langle C, D \rangle \doteq \langle C', D' \rangle$, then

$$\langle A, B \rangle + \langle C, D \rangle \doteq \langle A', B' \rangle + \langle C', D' \rangle. \tag{7}$$

This means that the sum and negation operations are well-defined on equivalence classes of processes. If $[\langle A, B \rangle]$ and $[\langle C, D \rangle]$ denote two equivalence classes represented by $\langle A, B \rangle$ and $\langle C, D \rangle$, then

$$[\langle A, B \rangle] + [\langle C, D \rangle] = [\langle A + C, B + D \rangle], \tag{8}$$

which is the same regardless of the particular representatives chosen to indicate the two classes being added.

We let $\mathcal{P}$ denote the collection of all equivalence classes of formal processes. Then, the sum operation both is associative and commutative on $\mathcal{P}$. Furthermore, it contains a zero element $\mathbf{0} = [\langle s, s \rangle]$ for some singleton s. That is, if $\Gamma = [\langle A, B \rangle]$,

$$\Gamma + \mathbf{0} = [\langle A, B \rangle] + [\langle s, s \rangle] = [\langle A + s, B + s \rangle] = [\langle A, B \rangle] = \Gamma. \tag{9}$$

The set $\mathcal{P}$ is thus a monoid (a semigroup with identity) under the operation $+$. Moreover, the subset $\mathcal{P}_S$ of singleton processes—equivalence classes of formal processes with singleton initial and final states—is actually an Abelian group, since

$$[\langle a, b \rangle] + [\langle b, a \rangle] = [\langle a + b, b + a \rangle] = [\langle a + b, a + b \rangle] = \mathbf{0}. \tag{10}$$

In $\mathcal{P}_S$, the negation operation does yield the additive inverse of an element.

The $\rightarrow$ relation on states induces a corresponding relation on processes. If $\Gamma, \Delta \in \mathcal{P}$, we say that $\Gamma \rightarrow \Delta$ provided the process $\Gamma - \Delta$ is natural (a condition we might formally write as $\Gamma - \Delta \rightarrow \mathbf{0}$). Intuitively, this means that process Γ can "drive process Δ backward", so that Γ and the opposite of Δ together form a natural process.

Now, suppose that $\hat{\mathcal{P}}$ is a collection of equivalence classes of possible processes that is closed under addition and negation. An *irreversibility function* $\mathbb{I}$ is a real-valued function on $\hat{\mathcal{P}}$ such that

1.　If $\Gamma, \Delta \in \hat{\mathcal{P}}$, then $\mathbb{I}(\Gamma + \Delta) = \mathbb{I}(\Gamma) + \mathbb{I}(\Delta)$.
2.　The value of $\mathbb{I}$ determines the type of the processes in $\hat{\mathcal{P}}$:

  - $\mathbb{I}(\Gamma) > 0$ whenever Γ is natural irreversible.
  - $\mathbb{I}(\Gamma) = 0$ whenever Γ is reversible.
  - $\mathbb{I}(\Gamma) < 0$ whenever Γ is antinatural irreversible.

An irreversibility function, if it exists, has a number of elementary properties. For instance, since the process $\Gamma + (-\Gamma)$ is always reversible for any $\Gamma \in \hat{\mathcal{P}}$, it follows that $\mathbb{I}(-\Gamma) = -\mathbb{I}(\Gamma)$. In fact, we can show that all irreversibility functions on $\hat{\mathcal{P}}$ are essentially the same:

Theorem 1. *An irreversibility function on $\hat{\mathcal{P}}$ is unique up to an overall positive factor.*

Proof. Suppose $\mathbb{I}_1$ and $\mathbb{I}_2$ are irreversibility functions on the set $\hat{\mathcal{P}}$. If Γ is reversible, then $\mathbb{I}_1(\Gamma) = \mathbb{I}_2(\Gamma) = 0$. If there are no irreversible processes in $\hat{\mathcal{P}}$, then $\mathbb{I}_1 = \mathbb{I}_2$.

Now, suppose that $\hat{\mathcal{P}}$ contains at least one irreversible process Γ, which we may suppose is natural irreversible. Thus, both $\mathbb{I}_1(\Gamma) > 0$ and $\mathbb{I}_2(\Gamma) > 0$. Consider some other process $\Delta \in \hat{\mathcal{P}}$. We must show that

$$\frac{\mathbb{I}_1(\Delta)}{\mathbb{I}_1(\Gamma)} = \frac{\mathbb{I}_2(\Delta)}{\mathbb{I}_2(\Gamma)}. \tag{11}$$

We proceed by contradiction, imagining that the two ratios are not equal and (without loss of generality) that the second one is larger. Then, there exists a rational number m/n (with $n > 0$) such that

$$\frac{\mathbb{I}_1(\Delta)}{\mathbb{I}_1(\Gamma)} < \frac{m}{n} < \frac{\mathbb{I}_2(\Delta)}{\mathbb{I}_2(\Gamma)}. \tag{12}$$

The first inequality yields $m\mathbb{I}_1(\Gamma) - n\mathbb{I}_1(\Delta) > 0$, so that the process $m\Gamma - n\Delta$ (that is, $m\Gamma + n(-\Delta)$) must be natural irreversible. The second inequality yields $m\mathbb{I}_2(\Gamma) - n\mathbb{I}_2(\Delta) < 0$, so that the process $m\Gamma - n\Delta$ must be antinatural irreversible. These cannot both be true, so the original ratios must be equal.　□

The additive irreversibility function $\mathbb{I}$ is only defined on a set $\hat{\mathcal{P}}$ of possible processes. However, we can under some circumstances extend an additive function to a wider domain. It is convenient to state here the general mathematical result we will use later for this purpose:

Theorem 2 ("Hahn–Banach theorem" for Abelian groups). *Let $\mathcal{G}$ be an Abelian group. Let ϕ be a real-valued function defined and additive on a subgroup $\mathcal{G}_0$ of $\mathcal{G}$. Then, there exists an additive function ϕ' defined on $\mathcal{G}$ such that $\phi'(x) = \phi(x)$ for all x in $\mathcal{G}_0$.*

Note that the extension ϕ' is not necessarily unique—that is, there may be many different extensions of a single additive function ϕ.

5. Information and Entropy

In our theory, *information* resides in the distinction among possible states. Thus, the eidostate $A = \{a_1, a_2, a_3\}$ represents information in the distinction among it elements. However, thermodynamic states such as the a_ks may also have other properties such as energy, particle content, and so on. To disentangle the concept of information from the other properties of these states, we introduce a notion of a "pure" information state.

The intuitive idea is this. We imagine that the world contains freely available memory devices. Different configurations of these memories—different memory records—are distinct states that are degenerate in energy and every other conserved quantity. Any particular memory record can thus be freely created from or reset to some null value.

We therefore define a *record state* to be an element $r \in \mathscr{S}$ such that there exists $a \in \mathscr{S}$ so that $a \leftrightarrow a + r$. The state a can be thought of as part of the apparatus that reversibly exchanges the particular record r with a null value. In fact, if A is any eidostate at all, we find that

$$A + a \leftrightarrow A + (a + r) \leftrightarrow (A + r) + a, \tag{13}$$

and so by cancellation of the singleton state a, $A \leftrightarrow A + r$. We denote the set of record states by $\mathscr{R}$. If $r, s \in \mathscr{R}$, then $r + s \in \mathscr{R}$, and furthermore $r \leftrightarrow s$. Any particular record state can be reversibly transformed into any other, a fact that expresses the arbitrariness of the "code" used to represent information in a memory device.

An *information state* is an eidostate whose elements are all record states, and the set of such eidostates is denoted $\mathscr{I}$. All information states are uniform eidostates. An *information process* is one that is equivalent to a process $\langle I, J \rangle$, where $I, J \in \mathscr{I}$. (Of course, information processes also include processes of the form $\langle I + x, J + x \rangle$ for a non-record state $x \in \mathscr{S}$, as well as more complex combinations of record and non-record states.) Roughly speaking, an information process is a kind of computation performed on information states.

It is convenient at this point to define a *bit state* (denoted $I_{\rm b}$) as an information state containing exactly two record states. That is, $I_{\rm b} = \{r_0, r_1\}$. A *bit process* is an information process of the form $\Theta_{\rm b} = \langle r, I_{\rm b} \rangle$ for some $r \in \mathscr{R}$—that is, a process by which a bit state is created from a single record state.

We can illustrate a bit process in a thermodynamics context by considering a thought-experiment involving Maxwell's demon. We imagine that the demon operates on a one-particle gas confined to a volume, a situation described by gas state v (see Figure 3). The demon, whose memory initially has a record state r, inserts a partition into the container, dividing the volume in two halves labeled 0 and 1. The gas molecule is certainly in one sub-volume or the other. The demon then records in memory which half the particle occupies. Finally, the partition is removed and the particle wanders freely around the whole volume of the container. In thermodynamic terms, the gas state relaxes to the original state v. The overall process establishes the following relations:

$$v + r \rightarrow \{v_0, v_1\} + r \rightarrow \{v_0 + r_0, v_1 + r_1\} \rightarrow v + \{r_0, r_1\}, \tag{14}$$

and so $r \rightarrow \{r_0, r_1\}$. In our example, the bit process $\Theta_{\rm b} = \langle r, I_{\rm b} \rangle$ is natural one.
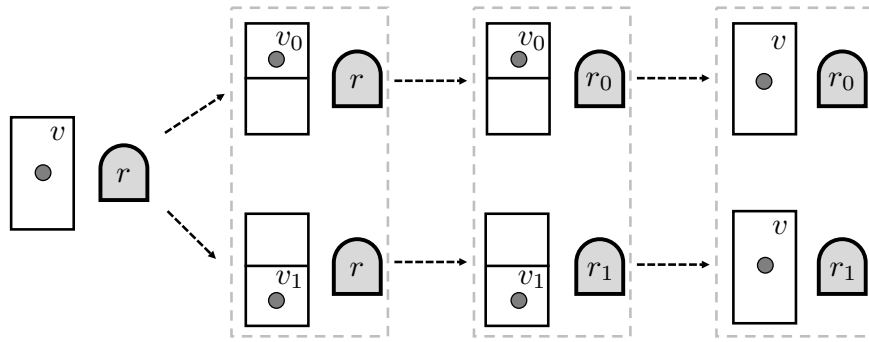
Figure 3. Maxwell's demon interacting with a one-particle gas, illustrating a bit process $\langle r, \{r_0, r_1\} \rangle$.

However, we do not yet know that there actually *are* information states and information processes in our theory. We address this by a new axiom.

Axiom 5. *Information: There exist a bit state and a possible bit process.*

Axiom 5 has a wealth of consequences. Since an information state exists, record states necessarily also exist. There are infinitely many record states, since for any $r \in \mathscr{R}$ we also have distinct states $r + r, r + (r + r), \ldots, nr, \ldots$, all in $\mathscr{R}$. We may have information states in $\mathscr{I}$ that contain arbitrarily many record states, because $n I_b$ contains 2^n elements. Furthermore, since every nonempty subset of an information state is also an information state, for any integer $k \geq 1$ there exists $I \in \mathscr{I}$ so that $\#(I) = k$.

Any two bit states can be reversibly transformed into one another. Consider $I_b = \{r_0, r_1\}$ and $I_b' = \{s_0, s_1\}$. Since $r_0 \to s_0$ and $r_1 \to s_1$, it follows by Axiom 4 that $I_b \to I_b'$ (and hence $I_b \leftrightarrow I_b'$.)

If any bit process is possible, then every bit process is possible. Furthermore, every such process is of the natural irreversible type. To see why, consider the bit state $I_b = \{r_0, r_1\}$ and suppose $I_b \to r$ for a record state r. Since $r \to r_0$, this implies that $\{r_0, r_1\} \to r_0$, which is a violation of Axiom 3. Therefore, it must be that $r \to I_b$ but $I_b \nrightarrow r$; and this is true for any choice of r and I_b.

Now, we can prove that every information process is possible, and that the $\to$ relation is determined solely by the relative sizes of the initial and final information states.

Theorem 3. *Suppose $I, J \in \mathscr{I}$. Then, $I \to J$ if and only if $\#(I) \leq \#(J)$.*

Proof. To begin with, we can see that $\#(I) = \#(J)$ implies $I \to J$. This is because we can write $I = \{r_1, \ldots, r_n\}$ and $J = \{s_1, \ldots, s_n\}$. Since $r_k \to s_k$ for every $k = 1, \ldots, n$, the finite extension of Axiom 4 tells us that $I \to J$.

Now, imagine that $\#(I) > \#(J)$. There exists a proper subset I' of I so that $\#(I') = \#(J)$, and thus $J \to I'$. If it happened that $I \to J$, it would follow that $I \to I'$, a contradiction of Axiom 3. Hence, $\#(I) > \#(J)$ implies $I \nrightarrow J$.

It remains to show that if $\#(I) < \#(J)$, it must be that $I \to J$. As a first case, suppose that $\#(I) = n$ and $\#(J) = n + 1$. Letting $I = \{r_1, \ldots, r_n\}$ and $J = \{s_1, \ldots, s_n, s_{n+1}\}$, we note that $r_1 \to s_1, r_2 \to s_2, \ldots, r_n \to \{s_n, s_{n+1}\}$ (the last being a bit process. It follows that $I \to J$.

We now proceed inductively. Given $\#(J) = \#(I) + n$, we can imagine a sequence of information states K_m with successive numbers of elements between $\#(I)$ and $\#(J)$, so that $\#(K_m) = \#(I) + m$. From what we have already proved,

$$I \to K_1 \to \cdots \to K_{n-1} \to J, \tag{15}$$

and by transitivity of the $\to$ relation we may conclude that $I \to J$. Therefore, $I \to J$ if and only if $\#(I) \leq \#(J)$. $\square$

Intuitively, the greater the number of distinct record states in an information state, the more information it represents. Thus, under our axioms, a natural information process may maintain or increase the amount of information, but never decrease it.

We can sharpen this intuition considerably. Let $\mathcal{P}_I$ be the set of information processes, which is closed under addition and negation, and in which every process is possible. We can define a function $\mathbb{I}$ on $\mathcal{P}_I$ as follows: For $\Gamma = [\langle I, J \rangle] \in \mathcal{P}_I$,

$$\mathbb{I}(\Gamma) = \log\left(\frac{\#(J)}{\#(I)}\right) = \log \#(J) - \log \#(I). \tag{16}$$

The logarithm function guarantees the additivity of $\mathbb{I}$ when two processes are combined, and the sign of $\mathbb{I}$ exactly determines whether $I \to J$. Thus, $\mathbb{I}$ is an irreversibility function on $\mathcal{P}_I$, as the notation suggests. This function is unique up to a positive constant factor—i.e., the choice of logarithm base. If we choose to use base-2 logarithms, so that a bit process has $\mathbb{I}(\Theta_b) = 1$, then $\mathbb{I}$ is uniquely determined.

The irreversibility function on information processes can be expressed in terms of a function on information states. Suppose we have a collection of eidostates $\mathcal{K} \subseteq \mathcal{E}$ that is closed under the $+$ operation. With Giles, we define a *quasi-entropy* on $\mathcal{K}$ to be a real-valued function $\mathbb{S}$ such that, for $A, B \in \mathcal{K}$:

1. $\mathbb{S}(A + B) = \mathbb{S}(A) + \mathbb{S}(B)$.
2. If $\langle A, B \rangle$ is natural irreversible, then $\mathbb{S}(A) < \mathbb{S}(B)$.
3. If $\langle A, B \rangle$ is reversible, then $\mathbb{S}(A) = \mathbb{S}(B)$.

(A full "entropy" function satisfies one additional requirement, which we will address in Section 8 below.) Given a quasi-entropy $\mathbb{S}$, we can derive an irreversibility function $\mathbb{I}$ on possible $\mathcal{K}$-processes by $\mathbb{I}(\langle A, B \rangle) = \mathbb{S}(B) - \mathbb{S}(A)$.

Obviously, $\mathbb{S}(I) = \log \#(I)$ is a quasi-entropy function on $\mathcal{I}$ that yields the irreversibility function on $\mathcal{P}_I$. We recognize it as the Hartley–Shannon entropy of an information source with $\#(I)$ possible outputs [16]. In fact, this is the only possible quasi-entropy on $\mathcal{I}$. Since every information process is possible, two different quasi-entropy functions $\mathbb{S}$ and $\mathbb{S}'$ can only differ by an additive constant. However, since $I + r \leftrightarrow I$ for $I \in \mathcal{I}$ and $r \in \mathcal{R}$, we know that $\mathbb{S}(r) = 0$ for any quasi-entropy. Therefore, $\mathbb{S}(I) = \log \#(I)$ is the unique quasi-entropy function for information states. The quasi-entropy of a bit state is $\mathbb{S}(I_b) = 1$.

6. Demons

Maxwell's demon accomplishes changes in thermodynamic states by acquiring and using information. For example, a demon that operates a trapdoor between two containers of gas can arrange for all of the gas molecules to end up in one of the containers, "compressing" the gas without work. As we have seen, it is also possible to imagine a reversible demon, which acquires and manipulates information in a completely reversible way. If such a demon produces a transformation from state x to state y by acquiring k bits of information in its memory, it can accomplish the reverse transformation (from y to x) while erasing k bits from its memory.

Maxwell's demon is a key concept in axiomatic information thermodynamics, and we introduce a new axiom to describe "demonic" processes.

Axiom 6. *Demons: Suppose $a, b \in \mathcal{S}$ and $J \in \mathcal{I}$ such that $a \to b + J$.*

(a) There exists $I \in \mathcal{I}$ such that $b \to a + I$.
(b) For any $I \in \mathcal{I}$, either $a \to b + I$ or $b + I \to a$.

In Part (a), we assert that what one demon can do (transforming a to b by acquiring information in J), another demon can undo (transforming b to a by acquiring information in I). Part (b) envisions a

reversible demon. Any amount of information in I is either large enough that we can turn a to b by acquiring I, or small enough that we can erase the information by turning b to a.

A process $\langle A, B \rangle$ is said to be *demonically possible* if one of two equivalent conditions hold:

- There exists an information state $J \in \mathscr{I}$ such that either $A \to B + J$ or $B \to A + J$.
- There exists an information process $\langle I, J \rangle \in \mathcal{P}_I$ such that $\langle A, B \rangle + \langle I, J \rangle$ is possible; that is, either $A + I \to B + J$ or $B + J \to A + I$.

It is not hard to see that these are equivalent. Suppose we have $J \in \mathscr{I}$ such that $A \to B + J$. Then, for any $I \in \mathscr{I}$, $A + I \to B + (I + J)$. Conversely, we note that $A \to A + I$ for any $I \in \mathscr{I}$. Thus, if $A + I \to B + J$ then $A \to B + J$ as well.

If a process is possible, then it is also demonically possible, since if $A \to B$ it is also true that $A + I \to B + I$. For singleton processes in $\mathcal{P}_S$, moreover, the converse is also true. Suppose $a, b \in \mathscr{S}$ and $\langle a, b \rangle$ is demonically possible. Then, there exists $I \in \mathscr{I}$ such that either $a \to b + I$ or $b \to a + I$. Either way, either trivially or by an application of Axiom 6, there must be $J \in \mathscr{I}$ so that $a \to b + J$. A single record state $r \in \mathscr{R}$ is a singleton information state in $\mathscr{I}$. Thus, by Axiom 6 it must be that either $a \to b + r$ or $b + r \to a$. Since $b + r \leftrightarrow b$, we find that $a \rightleftharpoons b$, and so $\langle a, b \rangle$ is possible.

A singleton process is demonically possible if and only if it is possible. This means that we can use processes involving demons to understand processes that do not. In fact, we can use Axiom 6 to prove a highly significant fact about the $\to$ relation on singleton eidostates.

Theorem 4. *Suppose $a, b, c \in \mathscr{S}$. If $\langle a, b \rangle$ and $\langle a, c \rangle$ are possible, then $\langle b, c \rangle$ is possible.*

Proof. First, we note the general fact that, if $\langle x, y \rangle$ is a possible singleton process, then there exist $I, J \in \mathscr{I}$ so that $x \to y + I$ and $y \to x + J$.

Given our hypothesis, therefore, there must be $I, J \in \mathscr{I}$ such that $b \to a + I$ and $a \to c + J$. Then

$$b \to a + I \to c + (I + J). \tag{17}$$

That is, $\langle b, c \rangle$ is demonically possible, and hence possible. $\square$

This fact is so fundamental that Giles made it an axiom in his theory. For us, it is a straightforward consequence of the axiom about processes involving demons. It tells us that the set of singleton states $\mathscr{S}$ is partitioned into equivalence classes, within each of which all states are related by $\rightleftharpoons$.

This statement is more primitive than, but closely related to, a well-known principle called the Comparison Hypothesis. The Comparison Hypothesis deals with a state relation called *adiabatic accessibility* (denoted $\prec$) which is definable in Giles's theory (and ours) but is taken as an undefined relation in some other axiomatic developments. According to the Comparison Hypothesis, if X and Y are states in a given thermodynamic space, either $X \prec Y$ or $Y \prec X$. Lieb and Yngvason, for instance, show that the Comparison Hypothesis can emerge as a consequence of certain axioms for thermodynamic states, spaces, and equilibrium [2,3].

Theorem 4 also sheds light on our axiom about conditional processes, Axiom 4. In Part (a) of this axiom, we suppose that $A \to b$ for some $A \in \mathscr{E}$ and $b \in \mathscr{S}$. The axiom itself allows us to infer that $a \to b$ for every $a \in A$. However, Theorem 4 now implies that, for every $a, a' \in A$, either $a \to a'$ or $a' \to a$. In other words, Part (a) of Axiom 4, similar to Part (b) of the same axiom, only applies to uniform eidostates.

Finally, we introduce one further "demonic" axiom.

Axiom 7. *Stability: Suppose $A, B \in \mathscr{E}$ and $J \in \mathscr{I}$. If $nA \to nB + J$ for arbitrarily large values of n, then $A \to B$.*

According to the Stability Axiom, if a demon can transform arbitrarily many copies of eidostate A into arbitrarily many copies of B while acquiring a bounded amount of information, then we may say

that $A \to B$. This can be viewed as a kind of "asymptotic regularization" of the $\to$ relation. The form of the Stability Axiom that we have chosen is a particularly simple one, and it suffices for our purposes in this paper. However, more sophisticated axiomatic developments might require a refinement of the axiom. Compare, for instance, Axiom 2.1.3 in Giles to its refinement in Axiom 7.2.1 [1].

To illustrate the use of the Stability Axiom, suppose that $A, B \in \mathscr{E}$ and $I \in \mathscr{I}$ such that $A + I \to B + I$. Our axioms do not provide a "cancellation law" for information states, so we cannot immediately conclude that $A \to B$. However, we can show that $nA \to nB + I$ for all positive integers n. The case $n = 1$ holds since $A \to A + I \to B + I$. Now, we proceed inductively, assuming that $nA \to nB + I$ for some n. Then,

$$
\begin{aligned}
(n+1)A \quad &\to \quad nA + A \\
&\to \quad (nB + I) + A \\
&\to \quad nB + (A + I) \\
&\to \quad nB + (B + I) \to (n+1)B + I.
\end{aligned} \tag{18}
$$

Thus, $nA \to nB + I$ for arbitrarily large (and indeed all) values of n. By the Stability Axiom, we see that $A \to B$. Thus, there is after all a general cancellation law information states that appear on both sides of the $\to$ relation.

7. Irreversibility for Singleton Processes

From our two "demonic" axioms (Axioms 6 and 7), we can use the properties of information states to derive an irreversibility function on singleton processes. Let $\hat{\mathcal{P}}_S$ denote the set of possible singleton processes. This is a subgroup of the Abelian group $\mathcal{P}_S$. Thus, if we can find an irreversibility function $\mathbb{I}$ on $\hat{\mathcal{P}}_S$, we will be able to extend it to all of $\mathcal{P}_S$.

We begin by proving a useful fact about possible singleton processes:

Theorem 5. *Suppose $a, b \in \mathscr{S}$ so that $\alpha = \langle a, b \rangle$ is possible. Then, for any integers $n, m \geq 0$, either $a + mI_b \to b + nI_b$ or $b + nI_b \to a + mI_b$.*

Proof. If $m = n$, the result is easy. Suppose that $n > m$, so that $n = m + k$ for positive integer k. By Axiom 6, either $a \to b + kI_b$ or $b + kI_b \to a$. We can then append the information state mI_b to both sides and rearrange the components. The argument for $m > n$ is exactly similar. $\square$

This fact has a corollary that we may state using the $\to$ relation on processes. Suppose $\alpha = \langle a, b \rangle$ is a possible singleton process, and let q, p be integers with $q > 0$. Then, either $q\alpha \to p\Theta_b$ or $q\alpha \leftarrow p\Theta_b$ (so that $q\alpha - p\Theta_b$ is either natural or antinatural) for the bit process $\Theta_b = \langle r, I_b \rangle$.

Given α, therefore, we can define two sets of rational numbers:

$$
L_\alpha = \{ p/q : q\alpha \to p\Theta_b \} \qquad U_\alpha = \{ p/q : q\alpha \leftarrow p\Theta_b \} \tag{19}
$$

where $q > 0$. Both sets are nonempty and every rational number is in at least one of these sets. Furthermore, if $p/q \in U_\alpha$ and $p'/q' \in L_\alpha$, we have that $q\alpha \leftarrow p\Theta_b$ and $q'\alpha \to p'\Theta_b$, and so

$$
pq'\Theta_b \to qq'\alpha \to p'q\Theta_b. \tag{20}
$$

Hence, $(pq' - p'q)\Theta_b \to \mathbf{0}$. Since Θ_b is itself a natural irreversible process, it follows that $pq' - p'q \geq 0$, and so

$$
\frac{p}{q} \geq \frac{p'}{q'}. \tag{21}
$$

Every element of U_α is an upper bound for L_α. It follows that U_α and L_α form a Dedekind cut of the rationals, which leads us to the following important result.

Theorem 6. *For $\alpha \in \hat{\mathcal{P}}_S$, define $\mathbb{I}(\alpha) = \inf U_\alpha = \sup L_\alpha$. Then, $\mathbb{I}$ is an irreversibility function on $\hat{\mathcal{P}}_S$.*

Proof. First, we must show that $\mathbb{I}$ is additive. Suppose $\alpha, \beta \in \hat{\mathcal{P}}_S$. If $p/q \in U_\alpha$ and $p'/q \in U_\beta$, then $q(\alpha + \beta) \to (p + p')\Theta_b$, and so $(p + p')/q \in U_{\alpha+\beta}$. It follows that $\mathbb{I}(\alpha + \beta) \leq \mathbb{I}(\alpha) + \mathbb{I}(\beta)$. The corresponding argument involving L_α and L_β proves that $\mathbb{I}(\alpha + \beta) \geq \mathbb{I}(\alpha) + \mathbb{I}(\beta)$. Thus, $\mathbb{I}$ must be additive.

Next, we must show that the value of $\mathbb{I}(\alpha)$ tells us the type of the process α. If $\mathbb{I}(\alpha) > 0$, then $0 \in L_\alpha$ but $0 \notin U_\alpha$, and so $\alpha \to \mathbf{0}$ but $\alpha \nleftarrow \mathbf{0}$. That is, α is natural irreversible. Likewise, if $\mathbb{I}(\alpha) < 0$, then $0 \notin L_\alpha$ but $0 \in U_\alpha$, from which we find that α must be antinatural irreversible. Finally, if $\mathbb{I}(\alpha) = 0$, we find that $q\alpha - \Theta_b \to \mathbf{0}$ and $q\alpha + \Theta_b \gets \mathbf{0}$ for arbitrarily large values of q. From Axiom 7, we may conclude that $\alpha \leftrightarrow \mathbf{0}$, and so α is reversible. $\square$

Notice that we have arrived at an irreversibility function for possible singleton processes—those most analogous to the ordinary processes in Giles or any text on classical thermodynamics—from axioms about information and processes involving demons (Axioms 5–7). In our view, such ideas are not "extras" to be appended onto a thermodynamic theory, but are instead central concepts throughout. In ordinary thermodynamics, the possibility of a reversible heat engine can have implications for processes that do not involve any heat engines at all. In the information thermodynamics whose axiomatic foundations we are exploring, the possibility of a Maxwell's demon has implications even for situations in which no demon acts.

We now have irreversibility functions for both information processes and singleton processes. These are closely related. In fact, it is possible to prove the following general result:

Theorem 7. *If $\alpha \in \hat{\mathcal{P}}_S$ and $\Gamma \in \mathcal{P}_I$, then the combined process $\alpha + \Gamma$ is natural if and only if $\mathbb{I}(\alpha) + \mathbb{I}(\Gamma) \geq 0$.*

Since the set $\hat{\mathcal{P}}_S$ of possible singleton processes is a subgroup of the Abelian group $\mathcal{P}_S$ of all singleton processes, we can extend the additive irreversiblity function $\mathbb{I}$ to all of $\mathcal{P}_S$. Though $\mathbb{I}$ is unique on the possible set $\hat{\mathcal{P}}_S$, its extension to $\mathcal{P}_S$ is generally not unique.

8. Components of Content and Entropy

Our axiomatic theory of information thermodynamics is fundamentally about the set of eidostates $\mathscr{E}$. However, the part of that theory dealing with the set $\mathscr{S}$ of singleton eidostates includes many of the concepts and results of ordinary axiomatic thermodynamics [1]. We have a group of singleton processes $\mathcal{P}_S$ containing a subgroup $\hat{\mathcal{P}}_S$ of possible processes, and we have constructed an irreversibility function $\mathbb{I}$ on $\hat{\mathcal{P}}_S$ that may be extended to all of $\mathcal{P}_S$. From these we can establish several facts.

- We can construct *components of content*, which are the abstract versions of conserved quantities. A component of content Q is an additive function on $\mathscr{S}$ such that, if the singleton process $\langle a, b \rangle$ is possible, then $Q(a) = Q(b)$. (In conventional thermodynamics, components of content include energy, particle number, etc.)
- We can find a *sufficient set* of components of content. The singleton process $\langle a, b \rangle$ is possible if and only if $Q(a) = Q(b)$ for all Q in the sufficient set.
- We can use $\mathbb{I}$ to define a *quasi-entropy* $\mathbb{S}$ on $\mathscr{S}$ as follows: $\mathbb{S}(a) = \mathbb{I}(\langle a, 2a \rangle)$. This is an additive function on states in $\mathscr{S}$ such that $\mathbb{I}(\langle a, b \rangle) = \mathbb{S}(b) - \mathbb{S}(a)$.

Because the extension of the irreversibility function $\mathbb{I}$ from $\hat{\mathcal{P}}_S$ to all of $\mathcal{P}_S$ is not unique, the quasi-entropy $\mathbb{S}$ is not unique either. How could various quasi-entropies differ? Suppose $\mathbb{I}_1$

and $\mathbb{I}_2$ are two different extensions of the same original $\mathbb{I}$, leading to two quasi-entropy functions $\mathbb{S}_1$ and $\mathbb{S}_2$ on $\mathscr{S}$. Then, the difference $Q = \mathbb{S}_1 - \mathbb{S}_2$ is a component of content. That is, if $\langle a, b \rangle$ is possible,

$$
\begin{aligned}
Q(b) - Q(a) &= \mathbb{S}_1(b) - \mathbb{S}_2(b) - \mathbb{S}_1(a) + \mathbb{S}_2(a) \\
&= \mathbb{I}_1(\langle a, b \rangle) - \mathbb{I}_2(\langle a, b \rangle) \\
&= 0,
\end{aligned}
\tag{22}
$$

since $\mathbb{I}_1$ and $\mathbb{I}_2$ agree on $\hat{\mathcal{P}}_S$, which contains $\langle a, b \rangle$.

Another idea that we can inherit without alteration is the concept of a *mechanical state*. A mechanical state is a singleton state that reversibly stores one or more components of content, in much the same way that we can store energy reversibly as the work done to lift or lower a weight. The mechanical state in this example is the height of the weight. In Giles's theory [1], mechanical states are the subject of an axiom, which we also adopt:

Axiom 8. *Mechanical states: There exists a subset $\mathscr{M} \subseteq \mathscr{S}$ of* mechanical states *such that:*

(a) *If $l, m \in \mathscr{M}$, then $l + m \in \mathscr{M}$.*
(b) *For $l, m \in \mathscr{M}$, if $l \to m$ then $m \to l$.*

Nothing in this axiom asserts the actual *existence* of any mechanical state. It might be that $\mathscr{M} = \varnothing$. Furthermore, the choice of the designated set $\mathscr{M}$ is not determined solely by the $\to$ relations among the states. For instance, the set $\mathscr{R}$ of record states might be included in $\mathscr{M}$, or not. This explains why the introduction of mechanical states must be phrased as an axiom, rather than a definition: a complete specification of the system must include the choice of which set is to be designated as $\mathscr{M}$. Whatever choice is made for $\mathscr{M}$, the set $\mathcal{P}_M$ of *mechanical processes* (i.e., those equivalent to $\langle l, m \rangle$ for $l, m \in \mathscr{M}$) will form a subgroup of $\mathcal{P}_S$.

A mechanical state may "reversibly store" a component of content Q, but it need not be true that every Q can be stored like this. We say that a component of content Q is *non-mechanical* if $Q(m) = 0$ for all $m \in \mathscr{M}$. For example, we might store energy by lifting or lowering a weight, but we cannot store particle number in this way.

Once we have mechanical states and processes, we can give a new classification of processes. A process $\Gamma \in \mathcal{P}$ is said to be *adiabatically natural* (*possible, reversible, antinatural*) if there exists a mechanical process $\mu \in \mathcal{P}_M$ such that $\Gamma + \mu$ is natural (possible, reversible, antinatural). We can also define the "adiabatic accessibility" relation for states in $\mathscr{S}$, as mentioned in Section 6: $a \prec b$ whenever the process $\langle a, b \rangle$ is adiabatically natural.

The set $\mathscr{M}$ of mechanical states allows us to refine the idea of a quasi-entropy into an *entropy*, which is a quasi-entropy $\mathbb{S}$ that takes the value $\mathbb{S}(m) = 0$ for any mechanical state m. Such a function is guaranteed to exist. We end up with a characterization theorem, identical to a result of Giles [1], that summarizes the general thermodynamics of singleton eidostates in our axiomatic theory.

Theorem 8. *There exist an entropy function $\mathbb{S}$ and a set of components of content Q on $\mathscr{S}$ with the following properties:*

(a) *For any $a, b \in \mathscr{S}$, $\mathbb{S}(a + b) = \mathbb{S}(a) + \mathbb{S}(b)$.*
(b) *For any $a, b \in \mathscr{S}$ and component of content Q, $Q(a + b) = Q(a) + Q(b)$.*
(c) *For any $a, b \in \mathscr{S}$, $a \to b$ if and only if $\mathbb{S}(a) \leq \mathbb{S}(b)$ and $Q(a) = Q(b)$ for every component of content Q.*
(d) *$\mathbb{S}(m) = 0$ for all $m \in \mathscr{M}$.*

An entropy function is not unique. Two entropy functions may differ by a non-mechanical component of content.

The entropy function on $\mathscr{S}$ is related to the information entropy function we found for information states in $\mathscr{I}$. Suppose we have $a, b \in \mathscr{S}$ and $I, J \in \mathscr{I}$. Then, Theorem 7 tells us that $a + I \rightarrow b + J$ only if

$$\mathbb{S}(a) + \log \#(I) \leq \mathbb{S}(b) + \log \#(J). \tag{23}$$

Let $\mathscr{E}_{SI}$ represent the set of eidostates that are similar to a singleton state combined with an information state. Then, $\mathbb{S}(a + I) = \mathbb{S}(a) + \log \#(I)$ is an entropy function on $\mathscr{E}_{SI}$. In the next section, we will extend the domain of the entropy function even further, to the set $\mathcal{U}$ of all uniform eidostates.

9. State Equivalence

Consider a thought-experiment (illustrated in Figure 4) in which a one-particle gas starts out in a volume v_0 and a second thermodynamic system starts out in one of three states e_1, e_2 or e_3. We assume that all conserved quantities are the same for these three states, but they may differ in entropy. We can formally describe the overall situation by the eidostate $E + v_0$, where $E = \{e_1, e_2, e_3\}$ is uniform.
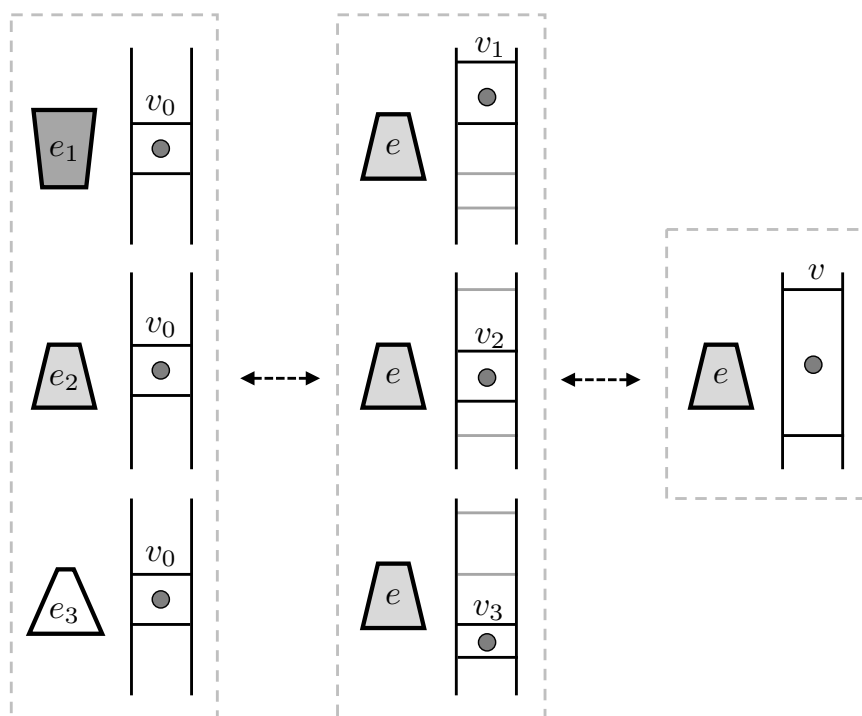

Figure 4. A state equivalence thought-experiment involving states of an arbitrary thermodynamic system and a one-particle gas.

We can reversibly transform each of the e_k states to the same state e, compensating for the various changes in entropy by expanding or contracting the volume occupied by the gas. That is, we can have $v_0 + e_k \leftrightarrow v_k + e$ for adjacent but non-overlapping volumes v_k. Axiom 4 indicates that we can write $E + v_0 \leftrightarrow e + V$, where $V = \{v_1, v_2, v_3\}$.

Now, we note that V can itself be reversibly transformed into a singleton eidostate v. A gas molecule in one of the sub-volumes (eidostate V) can be turned into a molecule in the whole volume (eidostate e) by removing internal partitions between the sub-volumes; and when we re-insert these partitions the particle is once again in just one of sub-volumes. Thus, $V \leftrightarrow v$. To summarize, we have

$$E + v_0 \leftrightarrow e + v. \tag{24}$$

The uniform eidostate E, taken together with the gas state v_0, can be reversibly transformed into the singleton state $v + e$. We call this a *state equivalence* for the uniform eidostate E. By choosing the

state e properly (say, by letting $e = e_k$ for some k), we can also guarantee that the volume v is larger than v_0, so that $v_0 \to v$ by free expansion.

This discussion motivates our final axiom, which states that this type of reversible transformation is always possible for a uniform eidostate.

Axiom 9. *State equivalence: If E is a uniform eidostate then there exist states $e, x, y \in \mathscr{S}$ such that $x \to y$ and $E + x \leftrightarrow e + y$.*

Axiom 9 closes a number of gaps in our theory. For example, our previous axioms (Axioms 1–8) do not by themselves guarantee than *any* state in $\mathscr{S}$ has a nonzero entropy. With the new axiom, however, we can prove that such states exist. The bit state I_b, which is uniform, has some state equivalence given by $I_b + x \leftrightarrow e + y$. Since the eidostates on each side of this relation are in $\mathscr{E}_{SI}$, we can determine the entropies on each side. We find that

$$1 + \mathbb{S}(x) = \mathbb{S}(e) + \mathbb{S}(y). \tag{25}$$

It follows that at least one of the states e, x, y must have $\mathbb{S} \neq 0$.

State equivalence also allows us to define the entropy of any uniform eidostate $E \in \mathcal{U}$. If $E + x \leftrightarrow e + y$ then we let

$$\mathbb{S}(E) = \mathbb{S}(e) + \mathbb{S}(y) - \mathbb{S}(x). \tag{26}$$

We must first establish that this expression is well-defined. If we have two state equivalences for the same E, so that $E + x \leftrightarrow e + y$ and $E + x' \leftrightarrow e' + y'$, then

$$(e + y) + x' \leftrightarrow E + (x + x') \leftrightarrow (e' + y') + x, \tag{27}$$

from which it follows that

$$\mathbb{S}(e) + \mathbb{S}(y) - \mathbb{S}(x) = \mathbb{S}(e') + \mathbb{S}(y') - \mathbb{S}(x'). \tag{28}$$

Thus, our definition for $\mathbb{S}(E)$ does not depend on our choice of state equivalence for E.

Is $\mathbb{S}$ an entropy function on $\mathcal{U}$? It is straightforward to show that $\mathbb{S}$ is additive on $\mathcal{U}$ and that $\mathbb{S}(m) = 0$ for any mechanical state m. It remains to show that, for any $E, F \in \mathcal{U}$ with $\langle E, F \rangle$ possible, $E \to F$ if and only if $\mathbb{S}(E) \leq \mathbb{S}(F)$. We will use the state equivalences $E + x \leftrightarrow e + y$ and $F + w \leftrightarrow f + z$.

Suppose first that $E \to F$. Then, $E + (x + w) \to F + (x + w)$ and so

$$(e + y) + w \leftrightarrow (E + x) + w \to (F + w) + x \leftrightarrow (f + z) + x. \tag{29}$$

From this, it follows that

$$\mathbb{S}(e) + \mathbb{S}(y) - \mathbb{S}(x) \leq \mathbb{S}(f) + \mathbb{S}(z) - \mathbb{S}(w), \tag{30}$$

and hence $\mathbb{S}(E) \leq \mathbb{S}(F)$.

We can actually extract one more fact from this argument. If we assume that $\langle E, F \rangle$ is possible, it must also be true that $\langle (e + y) + w, (f + z) + x \rangle$ is a possible singleton process. If we now suppose that $\mathbb{S}(E) \leq \mathbb{S}(F)$, we know that $\mathbb{S}((e + y) + w) \leq \mathbb{S}((f + z) + x)$ and thus

$$E + (x + w) \leftrightarrow (e + y) + w \to (f + z) + x \leftrightarrow F + (x + w). \tag{31}$$

Therefore, $E \to F$, as desired.

We have extended the entropy $\mathbb{S}$ to uniform eidostates. It is even easier to extend any component of content function Q to these states. If $E \in \mathcal{U}$, then any $e_1, e_2 \in E$ must have $Q(e_1) = Q(e_2)$, since $e_1 \rightleftharpoons e_2$. Thus, we can define $Q(E) = Q(e_k)$ for any $e_k \in E$. This is additive because the elements of $E + F$ are combinations $e_k + f_j$ of states in E and F. Furthermore, suppose we have a state equivalence

$E + x \leftrightarrow e + y$. Since we assume $x \to y$ in a state equivalence, $Q(x) = Q(y)$. By the conditional process axiom (Axiom 4) we know that $e_k + x \to e + y$ for any $e_k \in E$. It follows that $Q(e) = Q(e_k) = Q(E)$.

Now, let E_1 and E_2 be uniform eidostates with state equivalences $E_k + x_k \leftrightarrow e_k + y_k$. Suppose further that $Q(E_1) = Q(E_2)$ for every component of content Q. We know that $Q(x_1) = Q(y_1)$, $Q(x_2) = Q(y_2)$ and $Q(e_1) = Q(e_2)$ for every component of content. Thus

$$E_1 + (x_1 + x_2) \leftrightarrow (e_1 + y_1) + x_2 \rightleftharpoons (e_2 + y_2) + x_1 \leftrightarrow E_2 + (x_1 + x_2). \tag{32}$$

It follows that $E_1 \rightleftharpoons E_2$, i.e., that $\langle E_1, E_2 \rangle$ is a possible eidostate process.

We have therefore extended Theorem 8 to all uniform eidostates. We state the new result here.

Theorem 9 (Uniform eidostate thermodynamics). *There exist an entropy function $\mathbb{S}$ and a set of components of content Q on $\mathcal{U}$ with the following properties:*

(a) For any $E, F \in \mathcal{U}$, $\mathbb{S}(E + F) = \mathbb{S}(E) + \mathbb{S}(F)$.
(b) For any $E, F \in \mathcal{U}$ and component of content Q, $Q(E + F) = Q(E) + Q(F)$.
(c) For any $E, F \in \mathcal{U}$, $E \to F$ if and only if $\mathbb{S}(E) \leq \mathbb{S}(F)$ and $Q(E) = Q(F)$ for every component of content Q.
(d) $\mathbb{S}(m) = 0$ for all $m \in \mathcal{M}$.

The set $\mathcal{U}$ of uniform eidostates includes the singleton states in $\mathscr{S}$, the information states in $\mathscr{I}$, all combinations of these, and perhaps many other states as well. (Non-uniform eidostates in $\mathscr{E}$ *might* exist, as we will see in the model we discuss in Section 12, but their existence cannot be proved from our axioms.) The type of every process involving uniform eidostates can be determined by a single entropy function (which must not decrease) and a set of components of content (which must be conserved).

10. Entropy for Uniform Eidostates

We have extended the entropy function from singleton states and information states to all uniform eidostates. It turns out that this extension is unique. The following theorem and its corollaries actually allow us to compute the entropy of any $E \in \mathcal{U}$ from the entropies of the states contained in E.

Theorem 10. *Suppose $E \in \mathcal{U}$ is a disjoint union of uniform eidostates E_1 and E_2. Then*

$$\mathbb{S}(E) = \mathbb{S}(E_1 \cup E_2) = \log\left(2^{\mathbb{S}(E_1)} + 2^{\mathbb{S}(E_2)}\right). \tag{33}$$

Proof. E and E_k $(k = 1, 2)$ all have equal components of content. Define

$$\Delta(E_1, E_2) = \mathbb{S}(E) - \log\left(2^{\mathbb{S}(E_1)} + 2^{\mathbb{S}(E_2)}\right). \tag{34}$$

Note that, if we replace E_k by $E_k' = E_k + I$ for $I \in \mathscr{I}$, then these new eidostates are still disjoint and

$$E' = E_1' \cup E_2' = (E_1 + I) \cup (E_2 + I) = (E + I). \tag{35}$$

These eidostates have the same components of content as the original E. Furthermore,

$$
\begin{aligned}
\Delta(E_1', E_2') &= \Delta(E_1 + I, E_2 + I) \\
&= \mathbb{S}(E + I) - \log\left(2^{\mathbb{S}(E_1 + I)} + 2^{\mathbb{S}(E_2 + I)}\right) \\
&= \mathbb{S}(E) + \log \#(I) - \log\left(2^{\log \#(I)}\left(2^{\mathbb{S}(E_1)} + 2^{\mathbb{S}(E_2)}\right)\right) \\
&= \Delta(E_1, E_2).
\end{aligned}
\tag{36}
$$

We can find a uniform eidostate E_0 with the same components of content such that $S_0 = \mathbb{S}(E_0)$ is less than or equal to $\mathbb{S}(E_1)$, $\mathbb{S}(E_2)$ and $\mathbb{S}(E)$. (It suffices to pick E_0 to be the state of smallest entropy among E_1, E_2 and E.) Then, there exist integers $m_k \geq 1$ such that

- There are disjoint information states J_k containing m_k record states.
- There are disjoint information states J_k^* containing $m_k + 1$ record states.
- $J = J_1 \cup J_2$ and $J^* = J_1^* \cup J_2^*$ have $m_1 + m_2$ and $m_1 + m_2 + 2$ record states, respectively.
- We have

$$S_0 + \log m_k \leq \mathbb{S}(E_k) < S_0 + \log(m_k + 1). \tag{37}$$

That is, we choose m_k so that $\mathbb{S}(E_k) - S_0 \geq 0$ is between $\log(m_k)$ and $\log(m_k + 1)$. To put it more simply, $m_k = \lfloor 2^{\mathbb{S}(E_k) - S_0} \rfloor$.

Once we have m_k, we can write that

$$2^{S_0} \cdot m_k \leq 2^{\mathbb{S}(E_k)} < 2^{S_0} \cdot (m_k + 1). \tag{38}$$

Adding these inequalities for $k = 1, 2$ and taking the logarithm yields

$$S_0 + \log(m_1 + m_2) \leq \log\left(2^{\mathbb{S}(E_1)} + 2^{\mathbb{S}(E_2)}\right) < S_0 + \log(m_1 + m_2 + 2). \tag{39}$$

How far apart are the two ends of this chain of inequalities? Here, is a useful fact about base-2 logarithms: If $n \geq 1$, then $\log(n + 2) < \log(n) + 2/n$. This implies

$$S_0 + \log(m_1 + m_2) \leq \log\left(2^{\mathbb{S}(E_1)} + 2^{\mathbb{S}(E_2)}\right) < S_0 + \log(m_1 + m_2) + \frac{2}{m_1 + m_2}. \tag{40}$$

The two ends of the inequality differ by less than $2/(m_1 + m_2)$.

We can get another chain of inequalities by applying Axiom 4 about conditional processes. Since all of our uniform eidostates have the same components of content, we know that

$$E_0 + J_k \to E_k \to E_0 + J_k^*. \tag{41}$$

From the axiom, we can therefore say

$$E_0 + J \to E \to E_0 + J^* \tag{42}$$

which implies that

$$\begin{aligned} S_0 + \log(m_1 + m_2) \leq \mathbb{S}(E) &\leq S_0 + \log(m_1 + m_2 + 2) \\ &< S_0 + \log(m_1 + m_2) + \frac{2}{m_1 + m_2}. \end{aligned} \tag{43}$$

We have two quantities that lie in the same interval. Their separation is therefore bounded by the interval width—i.e., less than $2/(m_1 + m_2)$. Therefore,

$$\begin{aligned} |\Delta(E_1, E_2)| &= \left| \mathbb{S}(E) - \log\left(2^{\mathbb{S}(E_1)} + 2^{\mathbb{S}(E_2)}\right) \right| \\ &< \frac{2}{m_1 + m_2}. \end{aligned} \tag{44}$$

How big are the numbers m_k? We can make such numbers as large as we like by considering instead the eidostates $E_k' = E_k + I$, where I is an information state. As we have seen, $\Delta(E_1', E_2') = \Delta(E_1, E_2)$. Given any $\epsilon > 0$, we can choose I so that

$$m_k' = \left\lfloor 2^{\mathbb{S}(E_k') - S_0} \right\rfloor = \left\lfloor \#(I) \cdot 2^{\mathbb{S}(E_k) - S_0} \right\rfloor > \frac{1}{\epsilon}. \tag{45}$$

Then, $2/(m'_1 + m'_2) < \epsilon$, and so

$$|\Delta(E_1, E_2)| = |\Delta(E'_1, E'_2)| < \epsilon. \tag{46}$$

Since this is true for any $\epsilon > 0$, we must have $\Delta(E_1, E_2) = 0$, and so

$$\mathbb{S}(E) = \log\left(2^{\mathbb{S}(E_1)} + 2^{\mathbb{S}(E_2)}\right), \tag{47}$$

as desired.   □

Theorem 10 has a corollary, which we obtain by applying the theorem inductively:

Theorem 11. *If E is a uniform eidostate,*

$$\mathbb{S}(E) = \log\left(\sum_{e_k \in E} 2^{\mathbb{S}(e_k)}\right). \tag{48}$$

The entropy of any uniform eidostate is a straightforward function of the entropies of the states contained therein. If E contains more than one state, we notice that $\mathbb{S}(E) > \mathbb{S}(e_k)$ for any $e_k \in E$. It follows that $\langle e_k, E \rangle$ is a natural irreversible process.

To take a simple example of Theorem 11, consider the entropy of an information state. Every record state r has $\mathbb{S}(r) = 0$. Thus, for $I \in \mathscr{I}$,

$$\mathbb{S}(I) = \log\left(\sum_{r_k \in I} 2^0\right) = \log \#(I), \tag{49}$$

as we have already seen.

11. Probability

An eidostate in $\mathscr{E}$ represents a state of knowledge of a thermodynamic agent. It is, as we have said, a simple list of possible states, without any assignment of probabilities to them. If the agent is to use probabilistic reasoning, then it needs to assign conditional probabilities of the form $P(A|B)$ where $A, B \in \mathscr{E}$.

The entropy formula in Theorem 11 allows us to make such an assignment based on the entropy itself, provided the eidostate conditioned upon is uniform. Suppose $E \in \mathcal{U}$ and let $a \in \mathscr{S}$. Then, the *entropic probability* of a conditioned on E is

$$P(a|E) = \begin{cases} \dfrac{2^{\mathbb{S}(a)}}{2^{\mathbb{S}(E)}} = 2^{\mathbb{S}(a) - \mathbb{S}(E)} & a \in E \\[2mm] 0 & a \notin E \end{cases}. \tag{50}$$

Clearly, $0 \le P(a|E) \le 1$ and $\sum_a P(a|E) = 1$. It is worth noting that, although E is "uniform" (in the sense that all $a \in E$ have exactly the same conserved components of content), the probability distribution $P(a|E)$ is *not* uniform, but assigns a higher probability to states of higher entropy.

We can generalize entropic probabilities a bit further. Let A be any subset of $\mathscr{S}$ (be it an eidostate or not) and $E \in \mathcal{U}$. Then, $A \cap E$ is either a uniform eidostate or the empty set $\emptyset$. If we formally assign $\mathbb{S}(\emptyset) = -\infty$, then both E and $A \cap E$ have well-defined entropies. Then, we define

$$p(A|E) = \sum_{a \in A} P(a|E) = \frac{2^{\mathbb{S}(A \cap E)}}{2^{\mathbb{S}(E)}} = 2^{\mathbb{S}(A \cap E) - \mathbb{S}(E)}. \tag{51}$$

Obviously, $P(E|E) = 1$. Now, consider two disjoint sets A and B along with $E \in \mathcal{U}$. The set $(A \cup B) \cap E$ is a disjoint union of uniform eidostates (or empty sets) $A \cap E$ and $B \cap E$. Thus,

$$
\begin{aligned}
P(A \cup B|E) &= \frac{2^{\mathbb{S}((A \cup B) \cap E)}}{2^{\mathbb{S}(E)}} \\
&= \frac{2^{\mathbb{S}(A \cap E)} + 2^{\mathbb{S}(B \cap E)}}{2^{\mathbb{S}(E)}} \\
&= P(A|E) + P(B|E),
\end{aligned}
\tag{52}
$$

in accordance with the rules of probability. We also have the usual rule for conditional probabilities. Suppose $A, B \subseteq \mathscr{S}$ and $E \in \mathcal{U}$ such that $A \cap E \neq \varnothing$. Then

$$
P(B|A \cap E) = \frac{2^{\mathbb{S}(B \cap (A \cap E))}}{2^{\mathbb{S}(A \cap E)}} = \frac{P(B \cap A|E)}{P(A|E)}.
\tag{53}
$$

The entropy formula in Theorem 11 and the probability assignment in Equation (50) call to mind familiar ideas from statistical mechanics. According to Boltzmann's formula, the entropy is $\mathbb{S} = \log \Omega$, where Ω is (depending on the context) the number (or phase space volume or Hilbert space dimension) of the microstates consistent with macroscopic data about a system. In the microcanonical ensemble, a uniform probability distribution is assigned to these microstates. In an eidostate E comprising non-overlapping macrostates states $\{e_1, e_2, \ldots\}$, we would therefore expect $\Omega(E) = \Omega(e_1) + \Omega(e_2) + \ldots$, and the probability of state e_k should be proportional to $\Omega(e_k)$. However, in our axiomatic system Theorem 11 and Equation (50) do not arise from any assumptions about microstates, but solely from the "phenomenological" $\rightarrow$ relation among eidostates in $\mathscr{E}$.

Even though the entropy function $\mathbb{S}$ is not unique, the entropic probability assignment is unique. Suppose $\mathbb{S}_1$ and $\mathbb{S}_2$ are two entropy functions for the same states and processes. Then, as we have seen, the difference $\mathbb{S}_1 - \mathbb{S}_2$ is a component of content. All of the states within a uniform eidostate E (as well as E itself) have the same values for all components of content. That is, $\mathbb{S}_1(a) - \mathbb{S}_2(a) = \mathbb{S}_1(E) - \mathbb{S}_2(E)$ for any $a \in E$. Thus,

$$
P_1(a|E) = 2^{\mathbb{S}_1(a) - \mathbb{S}_1(E)} = 2^{\mathbb{S}_2(a) - \mathbb{S}_2(E)} = P_2(a|E).
\tag{54}
$$

(Both probabilities are zero for $a \notin E$, of course.)

The entropic probability assignment is not the only possible probability assignment, but it does have a number of remarkable properties. For instance, suppose E and F are two uniform eidostates. If we prepare them independently, we have the combined eidostate $E + F$ and the probability of some particular state $x + y$ is

$$
P(x + y|E + F) = \frac{2^{\mathbb{S}(x) + \mathbb{S}(y)}}{2^{\mathbb{S}(E) + \mathbb{S}(F)}} = P(x|E) \, P(y|F),
\tag{55}
$$

as we would expect for independent events.

The entropic probability also yields an elegant expression for the entropy $\mathbb{S}(E)$ of a uniform eidostate E.

Theorem 12. *Suppose E is a uniform eidostate, and $P(a|E)$ is the entropic probability for state $a \in E$. Then*

$$
\mathbb{S}(E) = \left\langle \mathbb{S}(a) \right\rangle + H(\vec{P}),
\tag{56}
$$

where $\langle \mathbb{S}(a) \rangle$ is the average state entropy in E and $H(\vec{P})$ is the Shannon entropy of the $P(a|E)$ distribution.

Proof. We first note that, for any $a \in E$, $\log P(a|E) = \mathbb{S}(a) - \mathbb{S}(E)$. We rewrite this as $\mathbb{S}(E) = \mathbb{S}(a) - \log P(a|E)$ and take the mean value with respect to the $P(a|E)$ probabilities:

$$
\begin{aligned}
\mathbb{S}(E) &= \langle \mathbb{S}(E) \rangle \\
&= \sum_{a \in E} P(a|E) \left(\mathbb{S}(a) - \log P(a|E) \right) \\
&= \sum_{a \in E} P(a|E) \mathbb{S}(a) - \sum_{a \in E} P(a|E) \log P(a|E),
\end{aligned}
\tag{57}
$$

Therefore, $\mathbb{S}(E) = \langle \mathbb{S}(a) \rangle + H(\vec{P})$, as desired. $\square$

We previously said that the list of possible states in an eidostate represents a kind of information. Theorem 12 puts this intuition on a quantitative footing. The entropy of a uniform eidostate E can be decomposed into two parts: the average entropy of the states, and an additional term representing the information contained in the distinction among the possible states. For a singleton state, the entropy is all of the first sort. For a pure information state $I \in \mathscr{I}$, it is all of the second.

In fact, the decomposition itself uniquely picks out the entropic probability assignment. Suppose $P(a|E)$ is the entropic probability of a given E, and $P'(a|E)$ is some other probability distribution over states in E. By Gibbs's inequality [17],

$$
0 \leq \sum_{a \in E} P'(a|E) \log \left(\frac{P'(a|E)}{P(a|E)} \right)
\tag{58}
$$

with equality if and only if $P'(a|E) = P(a|E)$ for all $a \in E$. We find that

$$
\begin{aligned}
0 &\leq \sum_{a \in E} P'(a|E) \log \left(\frac{P'(a|E)\, 2^{\mathbb{S}(E)}}{2^{\mathbb{S}(a)}} \right) \\
&= \sum_{a \in E} P'(a|E) \log P'(a|E) + \mathbb{S}(E) - \sum_{a \in E} P'(a|E) \mathbb{S}(a)
\end{aligned}
\tag{59}
$$

and so

$$
\mathbb{S}(E) \geq \left\langle \mathbb{S}(a) \right\rangle_{P'} + H(\vec{P}'),
\tag{60}
$$

with equality if and only if the P' distribution is the entropic one.

An agent that employs entropic probabilities will regard the entropy of a uniform eidostate E as the sum of two parts, one the average entropy of the possible states and the other the Shannon entropy of the distribution. For an agent that employs some other probability distribution, things will not be so simple. Besides the average state entropy and the Shannon entropy of the distribution, $\mathbb{S}(E)$ will include an extra, otherwise unexplained term. Thus, for a given collection of eidostates connected by the $\rightarrow$ relation, the entropic probability provides a uniquely simple account of the entropy of any uniform eidostate.

It is in this sense we say that the entropic probability "emerges" from the entropy function $\mathbb{S}$, just as that function itself emerges from the $\rightarrow$ relation among eidostates.

Some further remarks about probability are in order. Every formal basis for probability emphasizes a distinct idea about it. In Kolmogorov's axioms [18], probability is simply a measure on a sample space. High-measure subsets are more probable. In the Bayesian approach of Cox [19], probability is a rational measure of confidence in a proposition. Propositions in which a rational agent is more confident are also more probable. Laplace's early discussion [20] is based on symmetry. Symmetrically equivalent events—two different orderings of a shuffled deck of cards, for instance—are equally probable. (Zurek [21] has used a similar principle of "envariance" to discuss the origin of quantum

probabilities.) In algorithmic information theory [16], the algorithmic probability of a bit string is related to its complexity. Simpler bit strings are more probable.

In a similar way, entropic probabilities express facts about state transformations. In a uniform eidostate E, any two states $a, b \in E$ are related by a possible process. If $a \to b$, then the output state is at least as probable as the input state: $P(a|E) \le P(b|E)$.

12. A Model for the Axioms

A model for an axiomatic system may serve several purposes. The existence of a model establishes that the axioms are self-consistent. A model may also demonstrate that the system can describe an actual realistic physical situation. If the axioms have a variety of interesting models, then the axiomatic theory is widely applicable. We may almost say that the entire significance of an axiomatic system lies in the range of models for that system.

Terms that are undefined in an abstract axiomatic system are defined within a model as particular mathematical structures. The axioms of the system are provable properties of those structures. Therefore, a model for axiomatic information thermodynamics must include several elements:

- A set $\mathscr{S}$ of states and a collection $\mathscr{E}$ of finite nonempty subsets of $\mathscr{S}$ to be designated as eidostates.
- A rule for interpreting the combination of states $(+)$ in $\mathscr{S}$.
- A relation $\to$ on $\mathscr{E}$.
- A designated set $\mathscr{M} \subseteq \mathscr{S}$ of mechanical states (which might be empty).
- Proofs of Axioms 1–9 within the model, including the general properties of $\to$, the existence of record states and information states, etc.

The model will therefore involve specific meanings for $\mathscr{S}, \mathscr{E}, +, \to$ and so forth. It will also yield interpretations of derived concepts and results, such as entropy functions and conserved components of content. In the abstract theory, the combination $a + b$ is simply the Cartesian pair (a, b). This definition may suffice for the model, or the model may have a different interpretation of $+$. In any case, it must be true in the model that $a + b = a' + b'$ implies $a = a'$ and $b = b'$.

Our first model for axiomatic thermodynamics is based on a set $\mathscr{A}$ of "atomic" states, from which all states in $\mathscr{S}$ and all eidostates in $\mathscr{E}$ are constructed. We assign entropies and components of content to these states, which extend to composite states by additivity. Let us consider a simple but non-trivial example that has just one component of content Q. A suitable set $\mathscr{A}$ of atomic states is shown in Figure 5. It includes a special state r (with $\mathbb{S}(r) = 0$ and $Q(r) = 0$) and a continuous set of states s_λ with $Q(s_\lambda) = 1$ and $\mathbb{S}(s_\lambda) = \lambda$. The parameter λ ranges over the closed interval $[0, 1]$.
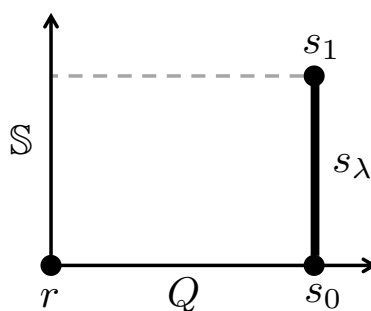

Figure 5. Atomic states in our simple "macrostate" model.

The set of states $\mathscr{S}$ includes everything that can be constructed from $\mathscr{A}$ by finite application of the pairing operation. In this way, we can build up $a \in \mathscr{S}$ with any non-negative integer value of the component of content $Q(a)$ and any entropy value $0 \le \mathbb{S}(a) \le Q(a)$. Indeed, there will typically be many different ways to create given $(Q, \mathbb{S})$ values. To obtain $Q(a) = 2$ and $\mathbb{S}(a) = 1$, for example, we might have $a = s_1 + s_0, s_0 + s_1, (s_{1/2} + s_{1/2}) + r, \dots$.

Anticipating somewhat, we call an eidostate *uniform* if all of its elements have the same Q-value. We calculate the entropy of a uniform eidostate by applying Theorem 11 to it.

Our model allows any finite nonempty set of states to play the role of an eidostate. Hence, we have both uniform and non-uniform eidostates in $\mathscr{E}$. Each eidostate A has a finite Cartesian factorization

$$A = F_A(E_1, \ldots, E_n). \tag{61}$$

If A is uniform, then all of its factors are also uniform. If none of its factors are uniform, we say that A is *completely non-uniform*. More generally, we can write down an *NU-decomposition* for any eidostate A:

$$A \sim N_A + U_A, \tag{62}$$

where N_A is a completely non-uniform eidostate and U_A is a uniform eidostate. Of course, if A itself is either completely non-uniform or uniform, one or the other of these eidostates may be absent from the decomposition. The NU-decomposition is unique up to similarity: If $N_A + U_A \sim N_A' + U_A'$ for completely non-uniform Ns and uniform Us, then $N_A \sim N_A'$ and $U_A \sim U_A'$.

We can now define the $\rightarrow$ relation on $\mathscr{E}$ in our model. If $A, B \in \mathscr{E}$, we first write down NU-decompositions $A \sim N_A + U_A$ and $B \sim N_B + U_B$. We say that $A \rightarrow B$ provided three conditions hold:

1. Either N_A and N_B both do not exist, or $N_A \sim N_B$.
2. Either U_A and U_B both do not exist, or only one exists and its Q-value is 0, or both exist and $Q(U_A) = Q(U_B)$.
3. Either U_A and U_B both do not exist, or only U_A exists and $\mathbb{S}(U_A) = 0$, or only U_B exists and $\mathbb{S}(U_B) \geq 0$, or both exist and $\mathbb{S}(U_A) \leq \mathbb{S}(U_B)$.

We may call these the N-criterion, Q-criterion, and $\mathbb{S}$-criterion, and summarize their meaning as follows: $A \rightarrow B$ provided we can transform A to B by: (a) rearranging the non-uniform factors; and (b) transforming the uniform factors in a way that conserves Q and does not decrease $\mathbb{S}$.

Now, let us examine each of the axioms in turn.

Axiom 1 The basic properties of eidostates follow by construction.

Axiom 2 Part (a) holds because $A \sim B$ implies that A and B can have the same NU-decomposition. Part (b) holds because similarity, equality (for Q) and inequality (for $\mathbb{S}$) are all transitive. Parts (c) and (d) make use of the general facts that $N_{A+B} \sim N_A + N_B$ and $U_{A+B} \sim U_A + U_B$.

Axiom 3 If $N_A \not\sim N_B$, then $A \not\rightarrow B$. If $N_A \sim N_B$, then it must be true that $U_B \subsetneq U_A$, and so $\mathbb{S}(U_A) > \mathbb{S}(U_B)$. The $\mathbb{S}$-criterion fails, so $A \not\rightarrow B$ in this case as well.

Axiom 4 For Part (a), we note that A must be uniform, and so $A' \subseteq A$ is also uniform. The statement follows from the $\mathbb{S}$-criterion. Part (b) also follows from the $\mathbb{S}$-criterion.

Axiom 5 The atomic state r with $Q(r) = 0$ and $\mathbb{S}(r) = 0$ is a record state, as is $r + r$, etc. We can take our bit state to be $I_b = \{r, r + r\}$. Since every information state $I \in \mathscr{I}$ is uniform with $Q(I) = 0$, every information process (including $\Theta_b = \langle r, I_b \rangle$) is possible.

Axiom 6 Since all of the states of the form $a + I$ are uniform, the statements in this axiom follow from the $\mathbb{S}$-criterion.

Axiom 7 Suppose $nA \rightarrow nB + J$. Since J is uniform, it must be that $nN_A \sim nN_B$, from which it follows that $N_A \sim N_B$. The $\mathbb{S}$-criterion for U_A and U_B follows from a typical stability argument—that is, if $nx \leq ny + z$ for arbitrarily large values of n, then it must be true $x \leq y$.

Axiom 8 The set $\mathscr{M}$ of mechanical states may be defined to include all states that can be constructed from the zero-entropy atomic state s_0 ($s_0 + s_0$, $s_0 + (s_0 + s_0)$, etc.). The required properties of $\mathscr{M}$ follow.

Axiom 9 The uniform eidostate E has $Q(E) = q \geq 0$ and $\mathbb{S}(E) = \sigma \geq 0$. Choose an integer $n > \sigma$. Now, let $e = qs_0$ (or $e = r$ if $q = 0$), $x = ns_0$, and $y = ns_\lambda$ where $\lambda = \sigma/n$. We find that $Q(e) = q$, $Q(x) = Q(y) = n$, $\mathbb{S}(e) = \mathbb{S}(x) = 0$ and $\mathbb{S}(y) = n(\sigma/n) = \sigma$. It follows that $x \rightarrow y$ and $E + x \leftrightarrow e + y$.

This model based on a simple set of atomic states has several sophisticated characteristics, including a non-trivial component of content Q and possible processes involving non-uniform eidostates. It is not difficult to create models of this type that are even richer and more complex. However, it may be objected that this type of model obscures one of the key features of axiomatic information thermodynamics. Here, the entropy function $\mathbb{S}$ does not *emerge from* the $\rightarrow$ relation among eidostates, but instead is imposed by hand to *define* $\rightarrow$ within the model. We address this deficiency in our next model.

13. A Simple Quantum Model

Now, we present a model for the axioms in which the entropy function does emerge from the underlying structure. The model is a simple one without mechanical states or non-trivial components of content. Every eidostate is uniform and every process is possible. On the other hand, the model is based on quantum mechanics, and so is not devoid of features of interest.

Consider an agent A that can act upon an external qubit system Q having Hilbert space $\mathcal{Q}$. Based on the information the agent possesses, it may assign the states $|\psi_1\rangle$ or $|\psi_2\rangle$ to the qubit. These state vectors need not be orthogonal. That is, it may be that no measurement of Q can perfectly distinguish which of the two states is actually present. The states, however, correspond to states of knowledge of agent A, and the agent is able to perform different operations on Q depending on whether it judges the qubit to be in one state or the other. Our notion of information possessed by the agent is thus similar to Zurek's concept of *actionable information* [22]. Roughly speaking, information is actionable if it can be used as a control variable for conditional unitary dynamics. This means that the two states of the agent's memory ($|\mu_1\rangle$ and $|\mu_2\rangle$ in a Hilbert space $\mathcal{A}$) must be distinguishable. Hence, if we include the agent in our description of the entire system, the states $|\mu_1\rangle \otimes |\psi_1\rangle$ and $|\mu_2\rangle \otimes |\psi_2\rangle$ are orthogonal, even if $|\psi_1\rangle$ and $|\psi_2\rangle$ are not.

Our model for axiomatic information thermodynamics envisions a world consisting of an agent A and an unbounded number of external qubits. Nothing essential in our model would be altered if the external systems had dim $\mathcal{Q} = d$ —"qudits" instead of qubits. The thermodynamic states of the qubit systems are actually states of knowledge of the agent, and so we must include the corresponding state of the agent's memory in our physical description. The quantum state space for our model is of the form:

$$\mathcal{H} = \mathcal{A} \otimes \mathcal{Q} \otimes \mathcal{Q} \otimes \cdots \tag{63}$$

To be a bit more rigorous, we restrict $\mathcal{H}$ to vectors of the form $|\Psi\rangle \otimes |0\rangle \otimes |0\rangle \otimes \cdots$, where $|\Psi\rangle \in \mathcal{A} \otimes \mathcal{Q}^{\otimes n}$ for some finite n, and $|0\rangle$ is a designated "zero" ket in $\mathcal{Q}$. Physical states in $\mathcal{H}$ have $\langle\Psi|\Psi\rangle = 1$. (The space $\mathcal{H}$ is not quite a Hilbert space, since it is not topologically complete, but this mathematical nicety will not affect our discussion.)

Since the Qs are qubits, dim $\mathcal{Q} = 2$. The agent space $\mathcal{A}$, however, must be infinite-dimensional, so that it contains a countably infinite set of orthogonal quantum states. These are to be identified as distinct records of the agent's memory.

In our thermodynamic model, the elements of $\mathscr{S}$ (the thermodynamic "states") are projection operators on $\mathcal{H}$. For any $a \in \mathscr{S}$, we have a projection on $\mathcal{H}$ of the form

$$\mathbf{\Pi}_a = |a\rangle\langle a| \otimes \boldsymbol{\pi}_a \otimes |0\rangle\langle 0| \otimes \cdots \tag{64}$$

where $|a\rangle$ is an agent state in $\mathcal{A}$ and $\boldsymbol{\pi}_a$ is a non-null projection in $\mathcal{Q}^{\otimes n}$ for some finite $n \geq 1$. The value of n is determined by a specified integer function $L(a)$, which we call the *length* of the state a. Heuristically, the thermodynamic state a means that the state of the world lies in the subspace $\mathcal{S}_a$ onto which $\mathbf{\Pi}_a$ projects. The agent's memory is in the state $|a\rangle$ and the the quantum state of the first $L(a)$ external qubits lies somewhere in the subspace onto which $\boldsymbol{\pi}_a$ projects. (All of the subsequent qubits are in the state $|0\rangle$.)

Two distinct thermodynamic states correspond to orthogonal states of the agent's memory. If $a, b \in \mathscr{S}$ with $a \neq b$, then $\langle a | b \rangle = 0$. The projections $\mathbf{\Pi}_a$ and $\mathbf{\Pi}_b$ are orthogonal to each other (so that $\mathbf{\Pi}_a \mathbf{\Pi}_b = 0$). However, it need not be the case that $\boldsymbol{\pi}_a$ and $\boldsymbol{\pi}_b$ are orthogonal.

Given a, the projection $\mathbf{\Pi}_a$ projects onto the subspace $\mathcal{S}_a$ The dimension of this subspace is $d_a = \dim \mathcal{S}_a = \operatorname{Tr} \mathbf{\Pi}_a = \operatorname{Tr} \boldsymbol{\pi}_a$. Note that $d_a \leq 2^{L(a)}$. We will assume that there are $\mathcal{S}_a$ subspaces of every finite dimension: For any integer $n \geq 1$, there exists $a \in \mathscr{S}$ with $d_a = n$.

Suppose $a, b \in \mathscr{S}$ correspond to $\mathbf{\Pi}_a = |a\rangle\langle a| \otimes \boldsymbol{\pi}_a \otimes \cdots$ and $\mathbf{\Pi}_b = |b\rangle\langle b| \otimes \boldsymbol{\pi}_b \otimes \cdots$. Then, we will specify that the combined state $a + b$ corresponds to

$$\mathbf{\Pi}_{a+b} = |a+b\rangle\langle a+b| \otimes \boldsymbol{\pi}_a \otimes \boldsymbol{\pi}_b \otimes \cdots . \tag{65}$$

Since the state $a + b$ entails a distinct state of the agent's knowledge, the agent state vector $|a + b\rangle$ is orthogonal to both $|a\rangle$ and $|b\rangle$. We also note that $L(a + b) = L(a) + L(b)$.

Here is a clarifying example. Suppose $a, b, c \in \mathscr{S}$. Then,

$$\mathbf{\Pi}_{(a+b)+c} = |(a+b)+c\rangle\langle (a+b)+c| \otimes \boldsymbol{\pi}_a \otimes \boldsymbol{\pi}_b \otimes \boldsymbol{\pi}_c \otimes \cdots \tag{66}$$

$$\mathbf{\Pi}_{a+(b+c)} = |a+(b+c)\rangle\langle a+(b+c)| \otimes \boldsymbol{\pi}_a \otimes \boldsymbol{\pi}_b \otimes \boldsymbol{\pi}_c \otimes \cdots \tag{67}$$

are distinct thermodynamic states and hence orthogonal projections in $\mathcal{H}$, even though they correspond to exactly the same qubit states. The difference between $(a + b) + c$ and $a + (b + c)$ entirely lies in the distinct representations of the states in the agent's memory.

The eidostates in our model are the finite nonempty collections of states in $\mathscr{S}$. We can associate each eidostate with a projection operator as well. Let $E = \{a, \ldots\}$ be an eidostate. We define

$$\mathbf{\Pi}_E = \sum_{a \in E} \mathbf{\Pi}_a = \sum_{a \in E} |a\rangle\langle a| \otimes \boldsymbol{\pi}_a \otimes \cdots . \tag{68}$$

This is a projection operator because the $\mathbf{\Pi}_a$ projections are orthogonal to one another. $\mathbf{\Pi}_E$ projects onto a subspace $\mathcal{S}_E$, which is the linear span of the collection of subspaces $\{\mathcal{S}_a, \ldots\}$.

Interestingly, this subspace $\mathcal{S}_E$ might contain quantum states in which the agent A is entangled with one or more external qubits. Suppose $a, b \in \mathscr{S}$ are associated with single-qubit projections onto distinct states $|\psi_a\rangle$ and $|\psi_b\rangle$. The eidostate $E = \{a, b\}$ is associated with the projection

$$\mathbf{\Pi}_E = (|a\rangle\langle a| \otimes |\psi_a\rangle\langle\psi_a| + |b\rangle\langle b| \otimes |\psi_b\rangle\langle\psi_b|) \otimes \cdots , \tag{69}$$

which projects onto a subspace $\mathcal{S}_E$ that contains the quantum state

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|a\rangle \otimes |\psi_a\rangle + |b\rangle \otimes |\psi_b\rangle) \otimes \cdots . \tag{70}$$

In this state, the agent does not have a definite memory record state. However, if a measurement is performed on the agent (perhaps by asking it a question), then the resulting memory record a or b would certainly be found to be consistent with the state of the qubit system, $|\psi_a\rangle$ or $|\psi_b\rangle$.

Suppose we combine two eidostates $A = \{a, \ldots\}$ and $B = \{b, \ldots\}$. Then, the quantum state lies in a subspace of dimension

$$
\begin{aligned}
d_{A+B} = \text{Tr}\,\mathbf{\Pi}_{A+B} &= \text{Tr}\sum_{a,b} \mathbf{\Pi}_{a+b} \\
&= \sum_{a,b} \text{Tr}\left(|a+b\rangle\langle a+b| \otimes \boldsymbol{\pi}_a \otimes \boldsymbol{\pi}_b\right) \\
&= \sum_{a,b} (\text{Tr}\,\boldsymbol{\pi}_a)(\text{Tr}\,\boldsymbol{\pi}_b) \\
&= \left(\sum_a \text{Tr}\,\boldsymbol{\pi}_a\right)\left(\sum_b \text{Tr}\,\boldsymbol{\pi}_b\right) \\
&= d_A \cdot d_B.
\end{aligned}
\tag{71}
$$

When eidostates combine, subspace dimension is multiplicative.

It remains to define the $\to$ relation in our quantum model. We say that $A \to B$ if there exists a unitary time evolution operator $\boldsymbol{U}$ on $\mathcal{H}$ such that $|\Psi\rangle \in \mathcal{S}_A$ implies that $\boldsymbol{U}\,|\Psi\rangle \in \mathcal{S}_B$. That is, every quantum state consistent with A evolves to one consistent with B under the time evolution $\boldsymbol{U}$. (Note that the evolution includes a suitable updating of the agent's own memory state.) This requirement is easily expressed as a subspace dimension criterion: $A \to B$ if and only if $d_A \leq d_B$.

We are now ready to verify our axioms.

Axiom 1 This follows from our construction of the eidostates $\mathcal{E}$.

Axiom 2 All of these basic properties of the $\to$ relation follow from the subspace dimension criterion.

Axiom 3 If $B \subsetneq A$, then $d_A > d_B$, and so $A \not\to B$.

Axiom 4 Again, both parts of this axiom follow from the subspace dimension criterion. If eidostate A is a disjoint union of eidostates A_1 and A_2, then $d_A = d_{A_1} + d_{A_2}$.

Axiom 5 Any state r with $d_r = 1$ functions as a record state. We have assumed that such a state exists. We can take $I_b = \{r, r+r\}$. The bit process $\Theta_b = \langle r, I_b\rangle$ is natural ($r \to I_b$) by the subspace dimension criterion. Notice that, for any information state I, $d_I = \#(I)$.

Axiom 6 For any $b \in \mathscr{S}$ and $I \in \mathscr{I}$, we have $d_{b+I} = d_b \cdot \#(I)$. For Part (a), we can always find a large enough information state so that $d_b \leq d_a \cdot \#(I)$. For Part (b), either $d_a \leq d_{b+I}$ or $d_{b+I} \leq d_a$.

Axiom 7 If $(d_A)^n \leq (d_B)^n \cdot \#(J)$ for arbitrarily large values of n, then $d_A \leq d_B$.

Axiom 8 It is consistent to take $\mathscr{M} = \varnothing$.

Axiom 9 All of our eidostates are uniform. For any eidostate E, we can choose e so that $d_e = d_E$. (Recall that we have assumed states with every positive subspace dimension.) If we chose $x = y$ to be any state, then $E + x \leftrightarrow e + y$.

Our quantum mechanical model is therefore a model of the axioms of information thermodynamics. It is a relatively simple model, of course, having no non-trivial conserved components of content and no mechanical states.

In the quantum model, the entropy of any eidostate is simply the logarithm of the dimension of the corresponding subspace: $\mathbb{S}(E) = \log d_E$. This is the von Neumann entropy of a uniform density operator $\boldsymbol{\rho}_E = \frac{1}{d_E}\mathbf{\Pi}_E$. We can, in fact, recast our entire discussion in terms of these mixed states, and this approach does yield some insights. For example, we find that the density operator $\boldsymbol{\rho}_E$ for eidostate E is a mixture of the density operators for its constituent states:

$$
\boldsymbol{\rho}_E = \sum_{a \in E} P(a|E)\boldsymbol{\rho}_a,
\tag{72}
$$

where $P(a|E)$ is the entropic probability

$$
P(a|E) = \frac{2^{\mathbb{S}(a)}}{2^{\mathbb{S}(E)}} = \frac{d_a}{d_E}.
\tag{73}
$$

There are, of course, many quantum states of the agent and its qubit world that do not lie within any eidostate subspace $\mathcal{S}_E$. For example, consider a state associated with a pure state projection $\boldsymbol{\Pi}_a = |a\rangle\langle a| \otimes |\psi_a\rangle\langle\psi_a| \otimes \cdots$. Let $|\chi\rangle$ be a state orthogonal to $|\psi_a\rangle$. Then, $|a\rangle \otimes |\chi\rangle \otimes \cdots$ is a perfectly legitimate quantum state that is orthogonal to $\mathcal{S}_a$ and every other eidostate subspace. This state represents a situation in which the agent's memory record indicates that the first $L(a)$ external qubits are in state $|\psi_a\rangle$, but *the agent is wrong.*

The exclusion of such physically possible but incongruous quantum states tells us something significant about our theory of axiomatic information thermodynamics. The set $\mathscr{E}$ does not necessarily include all possible physical situations; the arrow relations $\rightarrow$ between eidostates do not necessarily represent all possible time evolutions. Our axiomatic system is simply a theory of what transformations are possible among a collection of allowable states. In this, it is similar to ordinary classical thermodynamics, which is designed to consider processes that begin and end with states in internal thermodynamic equilibrium.

14. Remarks

The emergence of the entropy $\mathbb{S}$, a state function that determines the irreversibility of processes, is a key benchmark for any axiomatic system of thermodynamics. Our axiomatic system does not yield a unique entropy on $\mathscr{S}$, since it is based on the extension of an irreversibility function to impossible processes. However, many of our results and formulas for entropy are uniquely determined by our axioms. The entropy measure for information states, the Hartley–Shannon entropy $\log\#(I)$, is unique up to the choice of logarithm base. This in turn uniquely determines the irreversibility function on possible singleton processes, since this is defined in terms of the creation and erasure of bit states. There is a unique relationship between the entropy of a uniform eidostate and the entropy of the possible states it contains. Finally, the entropic probability distribution on a uniform eidostate, which might appear at first to depend on the singleton state entropy, is nonetheless unique. It remains to be seen how the axiomatic system developed here for state transformations is related to the axiomatic system, similar in some respects, given by Knuth and Skilling for considering problems of inference [23]. There, symmetry axioms in a lattice of states give rise to probability and entropy measures. In a similar way, the "entropy first, probability after" idea presented here is reminiscent of Caticha's "entropic inference" [24], in which the probabilistic Bayes rule is derived (along with the maximum entropy method) from a single relative entropy functional.

The entropy in information thermodynamics has two aspects: a measure of the information in a pure information state and an irreversibility measure for a singleton eidostate (the kind most closely analogous to a conventional thermodynamic state). The most general expression for the entropy of a uniform eidostate in Theorem 12 exhibits this twofold character. However, the two aspects of entropy are not really distinct in our axiomatic system. Both are based on the structure of the $\rightarrow$ relation among eidostates, which tells how one eidostate may be transformed into another by processes that may include demons.

The connection between information and thermodynamic entropy is nowhere more clearly stated than in Landauer's principle, the minimum thermodynamic cost of information erasure. It is easy to state a theorem of our system corresponding to Landauer's principle. Suppose $a, b \in \mathscr{S}$ and I_b is a bit state. If $a + I_b \rightarrow b$, then it follows that $\mathbb{S}(b) \geq \mathbb{S}(a) + 1$. Erasing a bit state is necessarily accompanied by an increase of at least one unit in the thermodynamic entropy. More generally, suppose A and B are uniform eidostates with $A \rightarrow B$. If we use Theorem 12 to write the entropies of the two states as

$$\mathbb{S}(A) = \langle\mathbb{S}\rangle_A + H_A \qquad \text{and} \qquad \mathbb{S}(B) = \langle\mathbb{S}\rangle_B + H_B, \tag{74}$$

then it follows that $\Delta\langle\mathbb{S}\rangle \geq -\Delta H$ for the process taking A to B. In other words, any decrease of the "Shannon information" part of the eidostate entropy $(-\Delta H)$ must be accompanied by an increase, on average, in the thermodynamic entropy of the possible states $(\Delta\langle\mathbb{S}\rangle)$.

These theorems do not really constitute a "proof" of Landauer's principle, because they depend upon the physical applicability of our axioms. On the other hand, our axiomatic framework does obviate some of the objections that have been raised to existing derivations of Landauer's principle. Norton [25], for example, has argued that many "proofs" of the principle improperly mix together two different kinds of probability distribution, the microstate distributions associated with thermodynamic equilibrium states and the probability assignments to memory records that represent information. He calls this the problem of "illicit ensembles". Our approach, by contrast, is not based on concepts of probability, except for those that emerge naturally from the structure of the $\rightarrow$ relation. We do not resolve thermodynamic states into their microstates at all, and we represent information by a simple enumeration of possible memory records.

The Second Law of Thermodynamics, the law of non-decrease of entropy, is the canonical example of the "arrow of time" in physics. Time asymmetry in our axiomatic theory is found in the direction of the $\rightarrow$ relation. It is an enlightening exercise to consider the axioms of information thermodynamics with the arrows reversed. Some axioms (e.g., Axioms 1 and 2) are actually unchanged by this. Others may be modified but still remain true statements in the theory. A few become false. The most striking example of the last is Axiom 3, which states that no eidostate may be transformed into a proper subset of itself. That is, no process can deterministically delete one of a list of possible states. This is the principle that leads to irreversbility in information processes, and through them, to more general irreversibility and entropy measures.

Author Contributions: This work began as a student research project by A.H. under the supervision of B.S., with M.D.W. acting as a consultant. That project established many of the basic concepts, axioms and proofs (particularly the elegant formulation of Axiom 3). B.S. and M.D.W. refined and extended the axiomatic system and wrote the present paper. All three authors have contributed editorial input.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

This appendix provides a convenient summary of axioms of our theory.

First, we remind ourselves of a few essential definitions:

An *eidostate* is a set whose elements are called *states*. The collection of eidostates is $\mathscr{E}$ and the collection of states is $\mathscr{S}$. An element $a \in \mathscr{S}$ may be identified with the singleton eidostate $\{a\} \in \mathscr{E}$.

Eidostates are combined by the Cartesian product, which we denote by the symbol $+$. When we combine an eidostate with itself n times, we use nA to denote $A + (A + (A + \ldots))$. Two eidostates are *similar* (written $A \sim B$) if they are made up of the same Cartesian factors, perhaps combined in a different way.

There is a relation $\rightarrow$ on $\mathscr{E}$, and thus also on the singletons in $\mathscr{S}$. An eidostate A is *uniform* if, for all $a, b \in A$, either $a \rightarrow b$ or $b \rightarrow a$. A formal *process* is a pair of eidostates $\langle A, B \rangle$. We say that a process $\langle A, B \rangle$ is *possible* if either $A \rightarrow B$ or $B \rightarrow A$.

A *record state* r is a state for which there exists another state a such that $a \rightarrow a + r$ and $a + r \rightarrow a$ (denoted $a \leftrightarrow a + r$). An *information state* is an eidostate containing only record states, and the set of information states is called $\mathscr{I}$. A *bit state* I_b is an information state containing exactly two distinct record states. A *bit process* is a formal process $\langle r, I_b \rangle$, where r is a record state and I_b is a bit state.

Now, we may state our axioms:

Axiom 1 (Eidostates.) $\mathscr{E}$ is a collection of sets called *eidostates* such that:

(a) Every $A \in \mathscr{E}$ is a finite nonempty set with a finite prime Cartesian factorization.
(b) $A + B \in \mathscr{E}$ if and only if $A, B \in \mathscr{E}$.

(c) Every nonempty subset of an eidostate is also an eidostate.

Axiom 2 (Processes.) Let eidostates $A, B, C \in \mathscr{E}$, and $s \in \mathscr{S}$.

(a) If $A \sim B$, then $A \to B$.
(b) If $A \to B$ and $B \to C$, then $A \to C$.
(c) If $A \to B$, then $A + C \to B + C$.
(d) If $A + s \to B + s$, then $A \to B$.

Axiom 3 If $A, B \in \mathscr{E}$ and B is a proper subset of A, then $A \nrightarrow B$.

Axiom 4 (Conditional processes.)

(a) Suppose $A, A' \in \mathscr{E}$ and $b \in \mathscr{S}$. If $A \to b$ and $A' \subseteq A$ then $A' \to b$.
(b) Suppose A and B are uniform eidostates that are each disjoint unions of eidostates: $A = A_1 \cup A_2$ and $B = B_1 \cup B_2$. If $A_1 \to B_1$ and $A_2 \to B_2$ then $A \to B$.

Axiom 5 (Information.) There exist a bit state and a possible bit process.

Axiom 6 (Demons.) Suppose $a, b \in \mathscr{S}$ and $J \in \mathscr{I}$ such that $a \to b + J$.

(a) There exists $I \in \mathscr{I}$ such that $b \to a + I$.
(b) For any $I \in \mathscr{I}$, either $a \to b + I$ or $b + I \to a$.

Axiom 7 (Stability.) Suppose $A, B \in \mathscr{E}$ and $J \in \mathscr{I}$. If $nA \to nB + J$ for arbitrarily large values of n, then $A \to B$.

Axiom 8 (Mechanical states.) There exists a subset $\mathscr{M} \subseteq \mathscr{S}$ of *mechanical states* such that:

(a) If $l, m \in \mathscr{M}$, then $l + m \in \mathscr{M}$.
(b) For $l, m \in \mathscr{M}$, if $l \to m$ then $m \to l$.

Axiom 9 (State equivalence.) If E is a uniform eidostate then there exist states $e, x, y \in \mathscr{S}$ such that $x \to y$ and $E + x \leftrightarrow e + y$.

References

1. Giles, R. *Mathematical Foundations of Thermodynamics*; Pergamon Press Ltd.: Oxford, UK, 1964.
2. Lieb, E.H.; Yngvason, J. The physics and mathematics of the second law of thermodynamics. *Phys. Rep.* **1999**, *310*, 1–96.
3. Lieb, E.H.; Yngvason, J. A guide to entropy and the second law of thermodynamics. *Not. Am. Math. Soc.* **1998**, *45*, 571–581.
4. Thess, A. *The Entropy Principle: Thermodynamics for the Unsatisfied*; Springer: Berlin, Germany, 1964.
5. Vedral, V.; Kashefi, E. Uniqueness of the entanglement measure for bipartite pure states. *Phys. Rev. Lett.* **2002**, *89*, 037903.
6. Leff, H.S.; Rex, A.F. *Maxwell's Demon 2: Entropy, Classical and Quantum Information, Computing*; IOP Publishing Ltd.: Bristol, UK, 2003.
7. Szilard, L. On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings. *Z. Phys.* **1929**, *53*, 840–856. (In German)
8. Bennett, C.H. The thermodynamics of computation—A review. *Int. J. Theor. Phys.* **1982**, *21*, 905–940.
9. Landauer, R. Irreversibility and heat generation in the computing process. *IBM J. Res. Dev.* **1961**, *5*, 183–191.
10. Jaynes, E.T. Information theory and statistical mechanics. *Phys. Rev.* **1957**, *106*, 620–630.
11. Jaynes, E.T. Information theory and statistical mechanics II. *Phys. Rev.* **1957**, *108*, 171–190.
12. Shannon, C.E. A mathematical theory of communication. *Bell Sys. Tech. J.* **1948**, *27*, 379–423, 623–656.
13. Weilenmann, M.; Kraemer, L.; Faist, P.; Renner, R. Axiomatic relation between thermodynamic and information-theoretic entropies. *Phys. Rev. Lett.* **2016**, *117*, 260601.
14. Gour, G.; Müller, M.P.; Narasimhachar, V.; Spekkens, R.W.; Halpern, N.Y. Resource theory of informational nonequilibrium in thermodynamics. *Phys. Rep.* **2015**, *583*, 1–58.
15. Del Rio, L.; Kraemer, L.; Renner, R. Resource theories of knowledge. *arXiv* **2015**, arXiv:1511.08818.
16. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*, 2nd ed.; John Wiley and Sons: Hoboken, NJ, USA, 2006.
17. MacKay, D.J.C. *Information Theory, Inference, and Learning Algorithms*; Cambridge University Press: Cambridge, UK, 2003.

18. Kolmogorov, A.N. *Grundbegriffe der Wahrscheinlichkeitsrechnung*; Springer: Berlin, Germany, 1933. (In German)

19. Cox, R.T. *The Algebra of Probable Inference*; Johns Hopkins Press: Baltimore, MD, USA, 1961.

20. De Laplace, P.S. *A Philosophical Essay on Probabilities*; Dover: New York, NY, USA, 1951.

21. Zurek, W.H. Probabilities from entanglement, Born's rule $p_k = |\psi_k|^2$ from envariance. *Phys. Rev. A* **2005**, *71*, 052105.

22. Zurek, W.H. Wave-packet collapse and the core quantum postulates: Discreteness of quantum jumps from unitarity, repeatability, and actionable information. *Phys. Rev. A* **2013**, *87*, 052111.

23. Knuth, K.H.; Skilling, J. Foundations of inference. *Axioms* **2012**, *1*, 38–73.

24. Caticha, A. Entropic inference. *AIP Conf. Proc.* **2011**, *1305*, 20–29.

25. Norton, J.D. Eaters of the lotus: Landauer's principle and the return of Maxwell's demon. *Stud. Hist. Phil. Mod. Phys.* **2005**, *36*, 375–411.