Title: Demons, Demons, Demons: Information-Theoretic Statistical Mechanics & the 2nd Law

Date: May 29, 2008 04:00 PM

URL: http://pirsa.org/08050045

Abstract: TBA

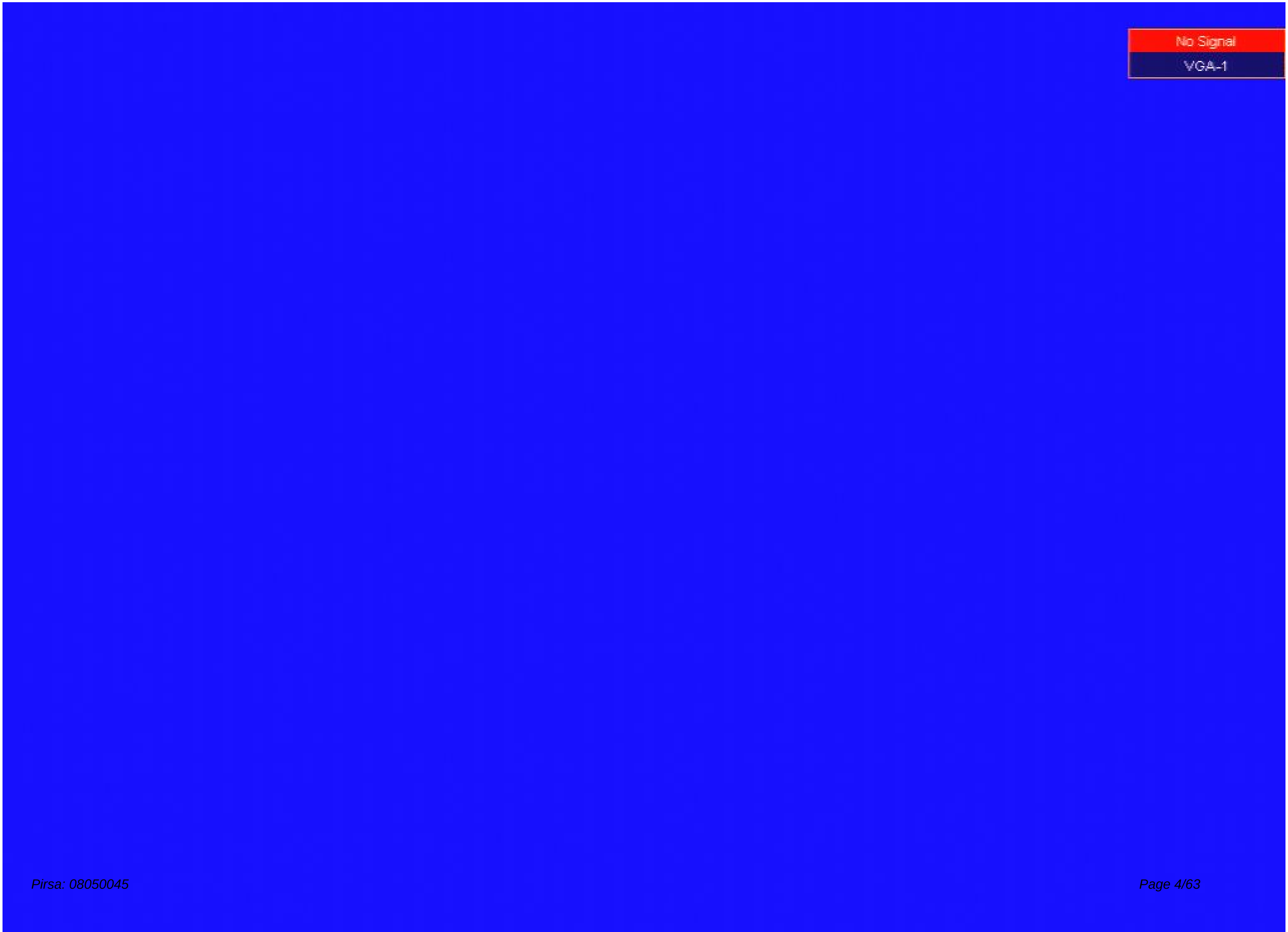
Demons Demons Demons: Information Theoretic Statistical Mechanics and the 2nd Law.

Daniel Parker

Virginia Tech

29/5/08

No Signal
VGA-1


No Signal
VGA-1

Demons Demons Demons: Information Theoretic Statistical Mechanics and the 2nd Law.

Daniel Parker

Virginia Tech

29/5/08

Outline

- What is the challenge that Maxwell's demon poses for Statistical Mechanics?
- How are we to understand the 2nd law?
  - An example based on the Szilard engine.
  - Boltzmannian SM as a model
- What is an information-theoretic 2nd law?
  - Two proposals.

Motivation

- Maxwell's demon challenges us to determine whether the 2nd law of thermodynamics is violable, given the statistical nature of its truth.
- Most present exorcisms of Maxwell's demon adopt an information-theoretic approach to statistical mechanics, coupled with Landauer's principle, to show such a demon is impossible.
- However, What does one mean by "the statistical 2nd law"?

Interpretation and SM

- What the second law permits or excludes is a matter of interpretation. How we understand the question posed by a Maxwellian Demon depends on 3 factors:
  - How we understand the 2nd law in TD
  - How we interpret SM
  - How we interpret the reducing relations between SM and TD

Some general notes on Reduction

- Typically, we expect our understanding of the reduced theory to be altered in light of the proposed reduction. The meanings and application of the terms appearing in the theory often need to be revised, and we might expect ambiguities in the formulation of the reduced theory to be resolved and/or clarified.
- Recognising the statistical nature of TD phenomena forces us to revise our understanding of what the 2nd law forbids, and Maxwell's demon is a challenge to provide just such an understanding.
- But how this understanding is to be effected will depend on how one understands both SM, and its connection to the terms that appear in TD, such as entropy and temperature.

Some general notes on Reduction

- Typically, we expect our understanding of the reduced theory to be altered in light of the proposed reduction. The meanings and application of the terms appearing in the theory often need to be revised, and we might expect ambiguities in the formulation of the reduced theory to be resolved and/or clarified.
- Recognising the statistical nature of TD phenomena forces us to revise our understanding of what the 2nd law forbids, and Maxwell's demon is a challenge to provide just such an understanding.
- But how this understanding is to be effected will depend on how one understands both SM, and its connection to the terms that appear in TD, such as entropy and temperature.

Some general notes on Reduction

- Typically, we expect our understanding of the reduced theory to be altered in light of the proposed reduction. The meanings and application of the terms appearing in the theory often need to be revised, and we might expect ambiguities in the formulation of the reduced theory to be resolved and/or clarified.
- Recognising the statistical nature of TD phenomena forces us to revise our understanding of what the 2nd law forbids, and Maxwell's demon is a challenge to provide just such an understanding.
- But how this understanding is to be effected will depend on how one understands both SM, and its connection to the terms that appear in TD, such as entropy and temperature.

Information-theoretic SM?

- Given the preponderance of information-theoretic exorcisms, the first question one should ask before engaging in such an exorcism is:
- How is TD to be reconceived in light of an information-theoretic interpretation?
- More specifically: How does one state the second law in information-theoretic terms?

Sound and Profound

- Earman and Norton (1999) offer a framework for analysing such approaches, which they pose in the form of a dilemma:
  - Sound horn: The demon is itself a canonical thermal system, and thus itself subject to the laws of thermodynamics.
  - Profound horn: Maxwell's demon is not a canonical thermal system, but the possibility of a violation of the laws of TD is blocked by introducing a link between the principles of statistical mechanics and some information-theoretic principle (such as Landauer's).

Sound and Profound (cont.)

- This is a false dilemma: *no system in nature is a classical thermodynamic system*, and the whole point of investigating the possibility of a Maxwellian Demon is to determine whether it is possible to violate a statistical version of the 2nd law.
- In effect, we're refusing to answer Maxwell's challenge on the sound horn.

Sound and Profound (cont.)

- If the dilemma is to retain its bite, we must understand the sound horn to reflect choosing to represent the demon as a canonical thermal system in the statistical sense; that is, as a canonical SM system.
- But this leaves the question open: what is a canonical thermal system, and what is the correct statement of the statistical 2nd law?
- Note: Earman and Norton demand overly restrictive conditions on what constitutes such a system.

Sound and Profound (cont.)

- If the dilemma is to retain its bite, we must understand the sound horn to reflect choosing to represent the demon as a canonical thermal system in the statistical sense; that is, as a canonical SM system.
- But this leaves the question open: what is a canonical thermal system, and what is the correct statement of the statistical 2nd law?
- Note: Earman and Norton demand overly restrictive conditions on what constitutes such a system.

The Second law in TD

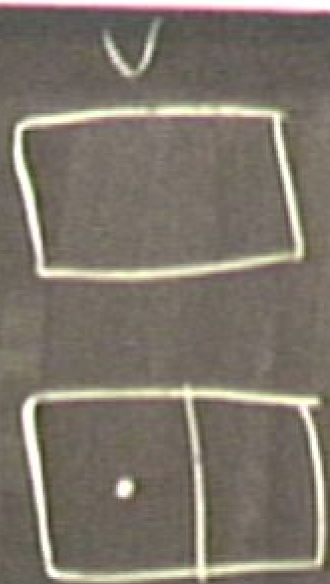
- As Uffink has explored, the meaning of the 2nd law is not clear even in classical TD.
- Some standard formulations:
  - A perpetuum mobile, where heat is continually extracted from the environment and converted into net mechanical work, is impossible.
  - It is impossible to perform a cyclic process with no other result than that heat is absorbed from a reservoir, and work is performed.
  - A thermodynamic system, operating in a cycle, will satisfy the inequality $\oint \frac{dQ}{T} \leq 0$

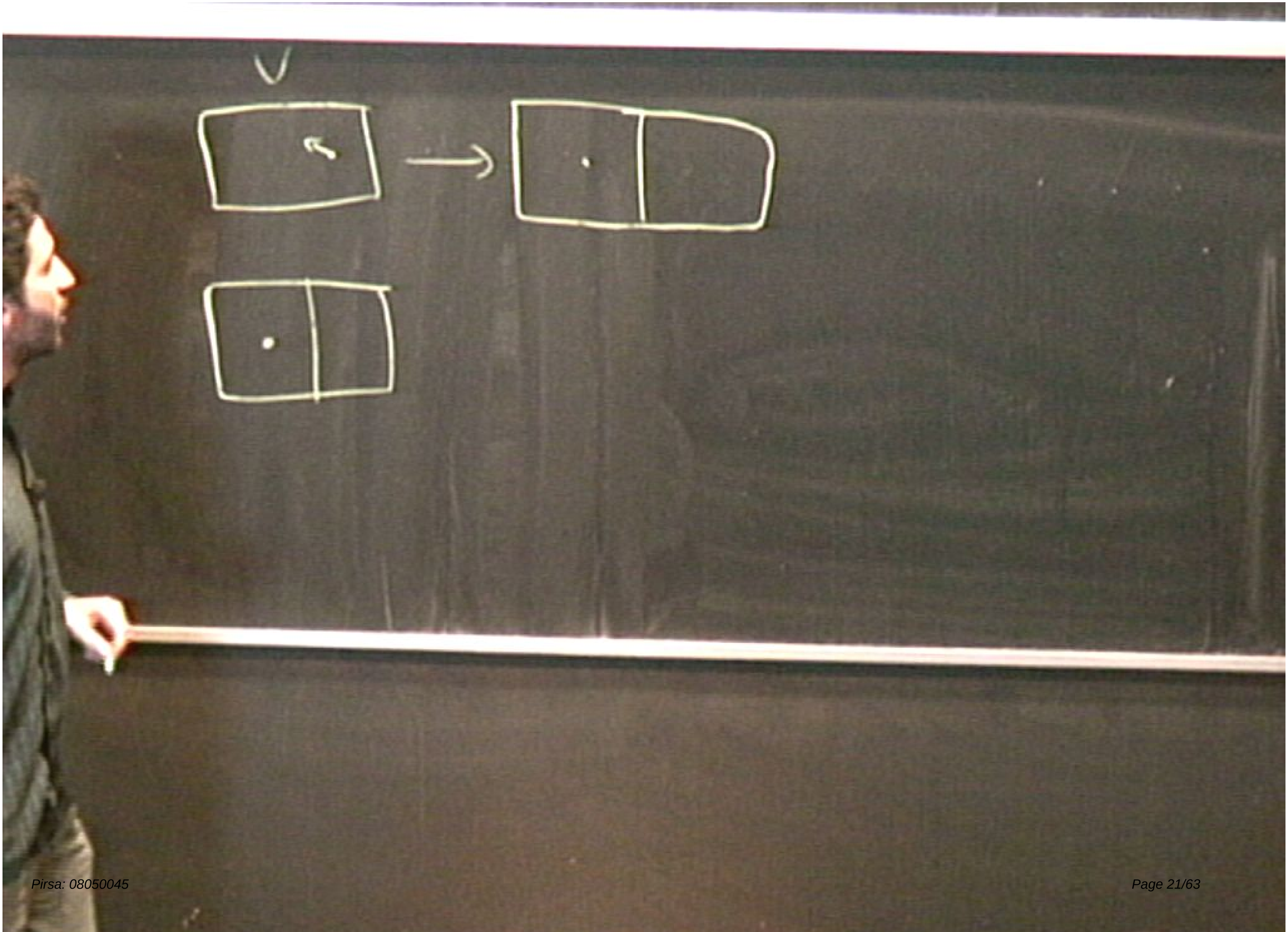
The Second Law (cont.)

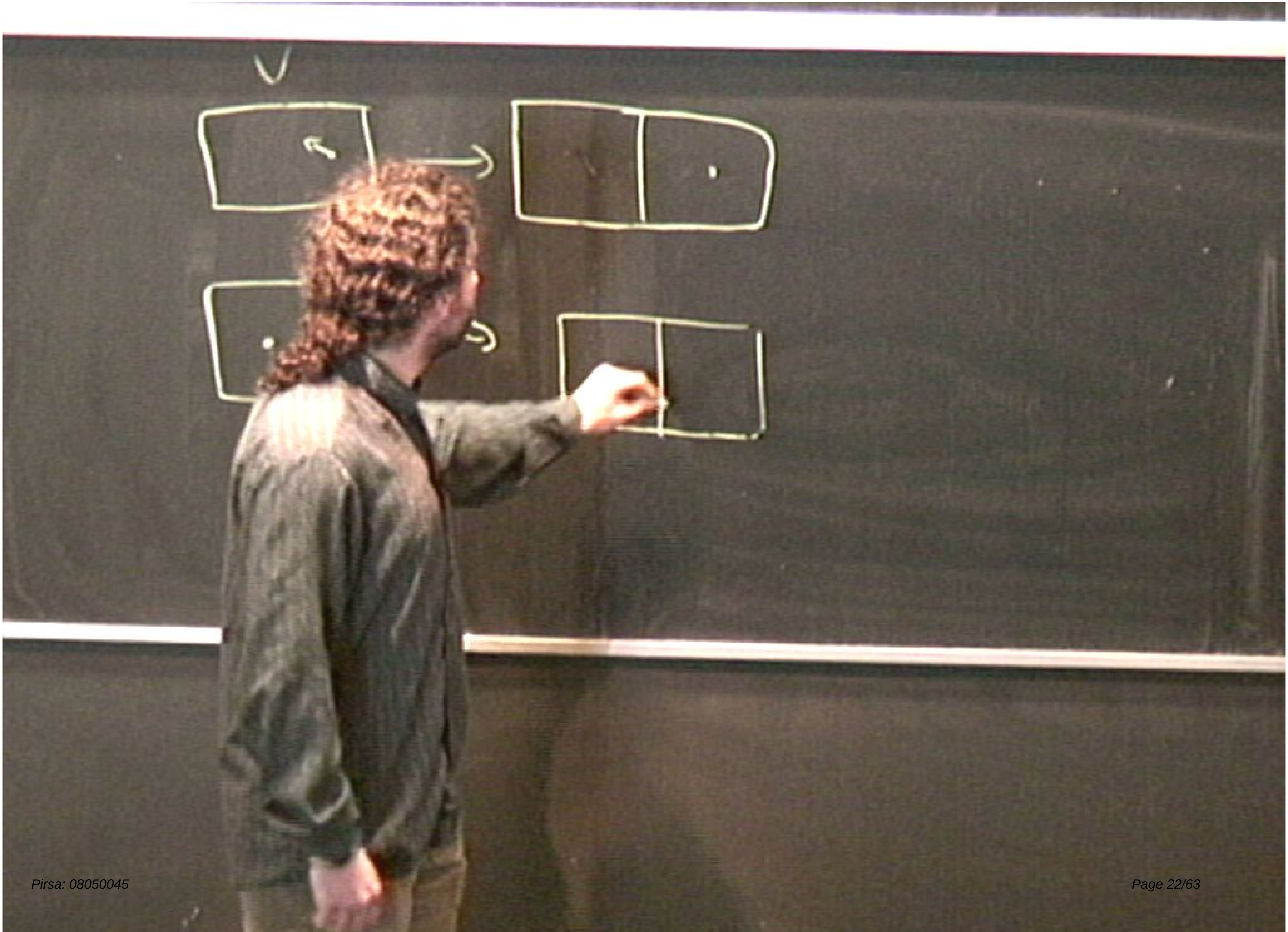
- The key term that appear in the latter two formulations is that of a ‘cycle’.
- Two notions of a cycle:
  - Strong sense: The system must return to its exact initial state, ready to run anew.
  - Weak sense: The system must return to its initial thermodynamic state (or at least one with equal entropy), though complete recovery of the initial state is not required.

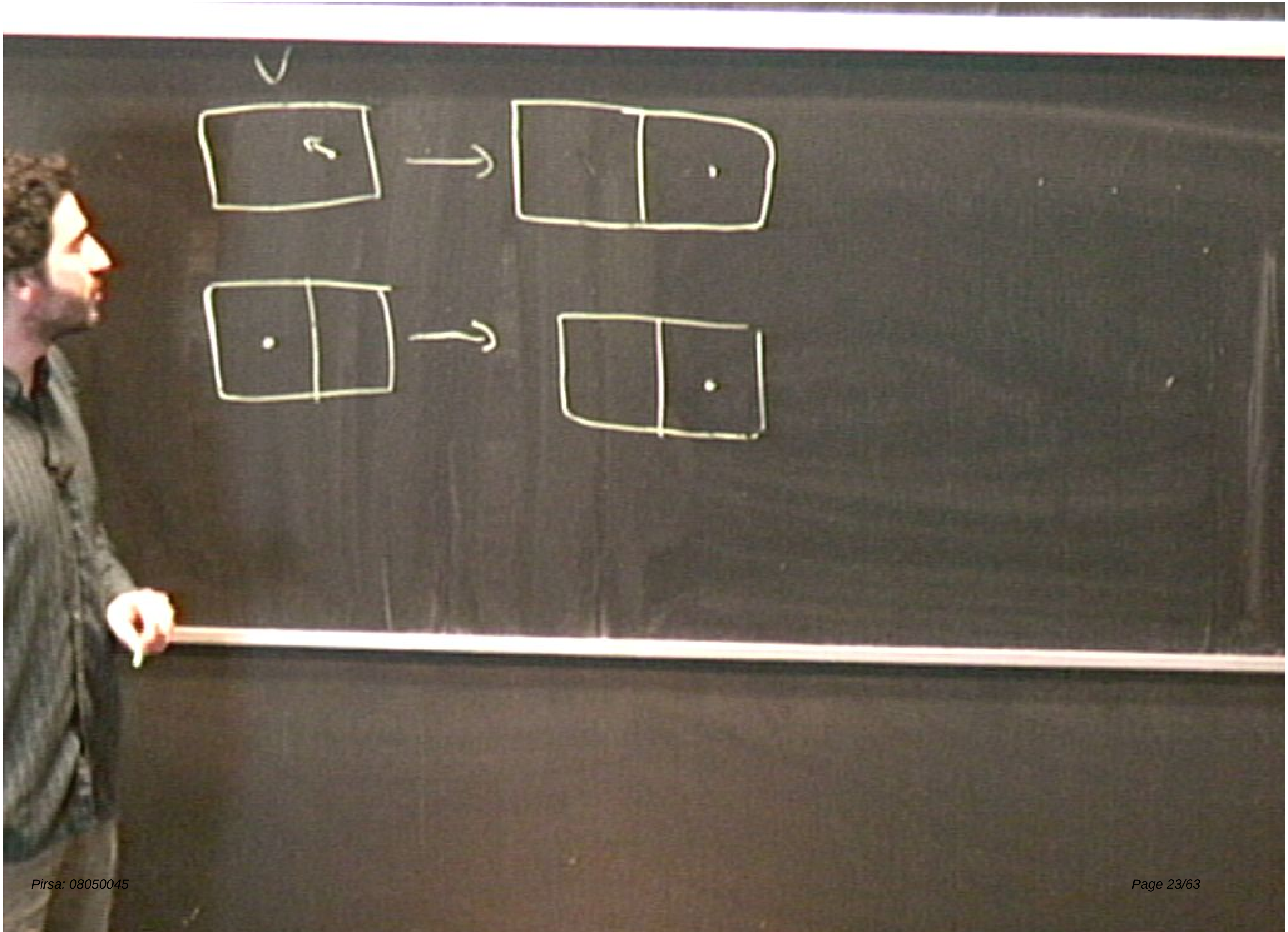
An Example from the Szilard Engine

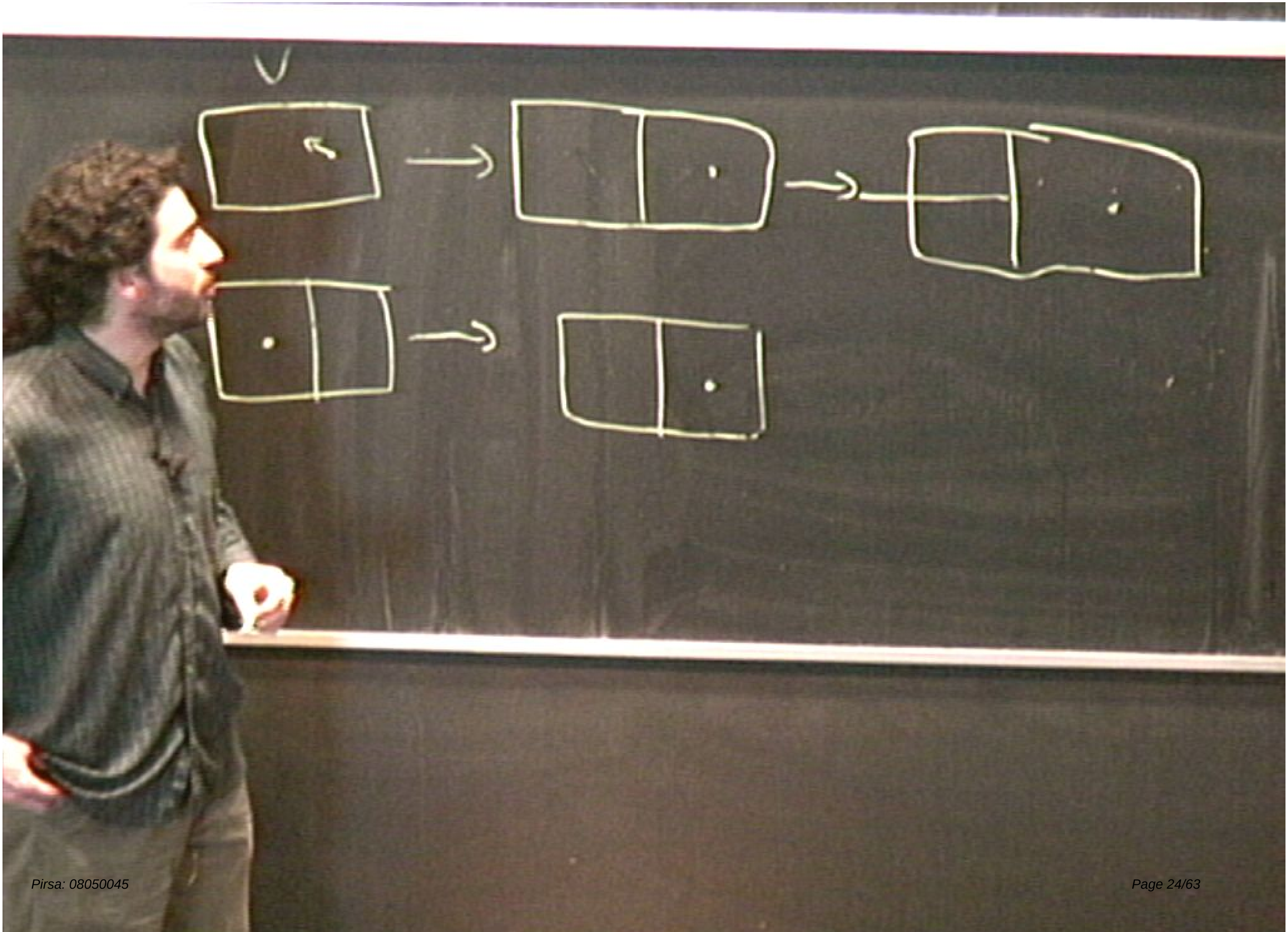
- We consider a standard setup of the Szilard engine, where the engine is a 1-particle ideal gas in a volume V in contact with a reservoir at temperature T.
- The demon is also represented as a one particle ideal gas.
- Initially, the particle in the engine has access to the entire volume of the box, and the demon's memory is in a 'ready state', such that it is restricted by a partition to, say, the left side of the memory cell.

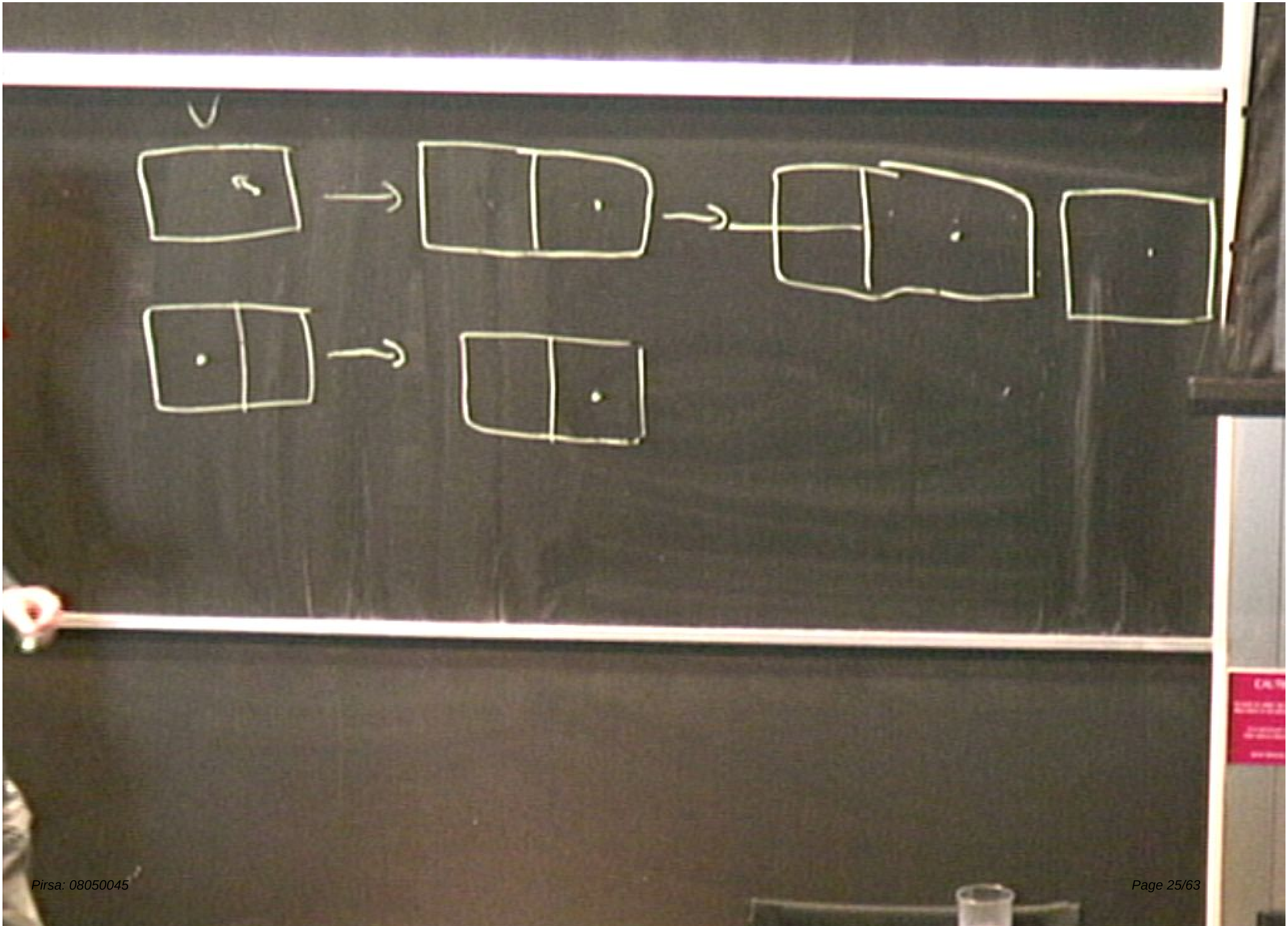

V








V



An Example from the Szilard Engine

- We consider a standard setup of the Szilard engine, where the engine is a 1-particle ideal gas in a volume V in contact with a reservoir at temperature T.
- The demon is also represented as a one particle ideal gas.
- Initially, the particle in the engine has access to the entire volume of the box, and the demon's memory is in a 'ready state', such that it is restricted by a partition to, say, the left side of the memory cell.

The Szilard Engine

- A partition is introduced into the engine, trapping the particle on one side of the engine.
- A dissipationless measurement is performed by the demon on the engine, and correlates the state of his memory to the state of the engine. The demon then replaces the partition in the engine by a piston, and the engine is isothermally expanded, extracting $kT\ln 2$ units of work from the engine.

The Szilard Engine

- This procedure completes a cycle in the weak sense: the initial thermodynamic state has been recovered.
- This does not complete a cycle in the strong sense: the procedure cannot be run again unless the demon's memory is reset.
- For Albert (2000), this is sufficient to claim that Maxwell's Demon is a distinct possibility. Why?

Boltzmann's Approach

- The above analysis assumes that the entropy is a state function of the thermodynamic constraints operating on the system. The approach Albert, Hemmo and Shenker favour couples this analysis of the Szilard engine to a Boltzmannian interpretation of SM, where:
  - The entropy is a property of individual physical systems.
  - The SM analogue of TD entropy is taken to be a relational property of the microstate and the macrostate of the system, where the entropy is calculated by ascertaining the phase volume of the macrostate to which the microstate belongs

Boltzmann's approach

- This analysis serves to provide an unusually clear and straightforward response to Maxwell's demon:
  - It gives a clear interpretation of the 2nd law.
  - It offers an unambiguous statement of the relation between the SM analysis of the situation and its import for the 2nd law.
  - It asserts that a demon is possible.
- However, this is not to say that the interpretation is entirely unproblematic, as it might be beset by other problems, and there may be reasons to reject their interpretation of the 2nd law.

A different interpretation?

- The Albert style answer thus can serve as a model for what an answer to the problem of Maxwell's demon might look like. But other responses are possible.
- The most common response is that the cycle has not been completed, even in the weak sense, in the example above.

A different Interpretation?

- Since we do not know what state the demon's memory cell is in after work has been extracted from the engine, the probability distribution describing the demon's state is distributed over both sides of the demon's memory, whereas initially it covered only the phase volume compatible with the ready state. So the cycle is not complete.
- Further, we can deny that the entropy is defined as a state function of the thermodynamic constraints, but rather is a function of this probability distribution, so unless the probability distribution is compressed, the cycle has not been completed.

An information-theoretic 2nd law?

- Unfortunately, this is as far as most proponents of information-theoretic statistical mechanics go. But this alternative interpretation only suggests a formulation of the 2nd law, and one that is clearly false!
- Elements of a "standard" information-theoretic approach:
  - The probability distribution employed is to be understood epistemically: it describes one's knowledge (or lack thereof) of the system's actual microstate.
  - The relevant entropy to consider is a function of this probability distribution.

An information-theoretic 2nd law?

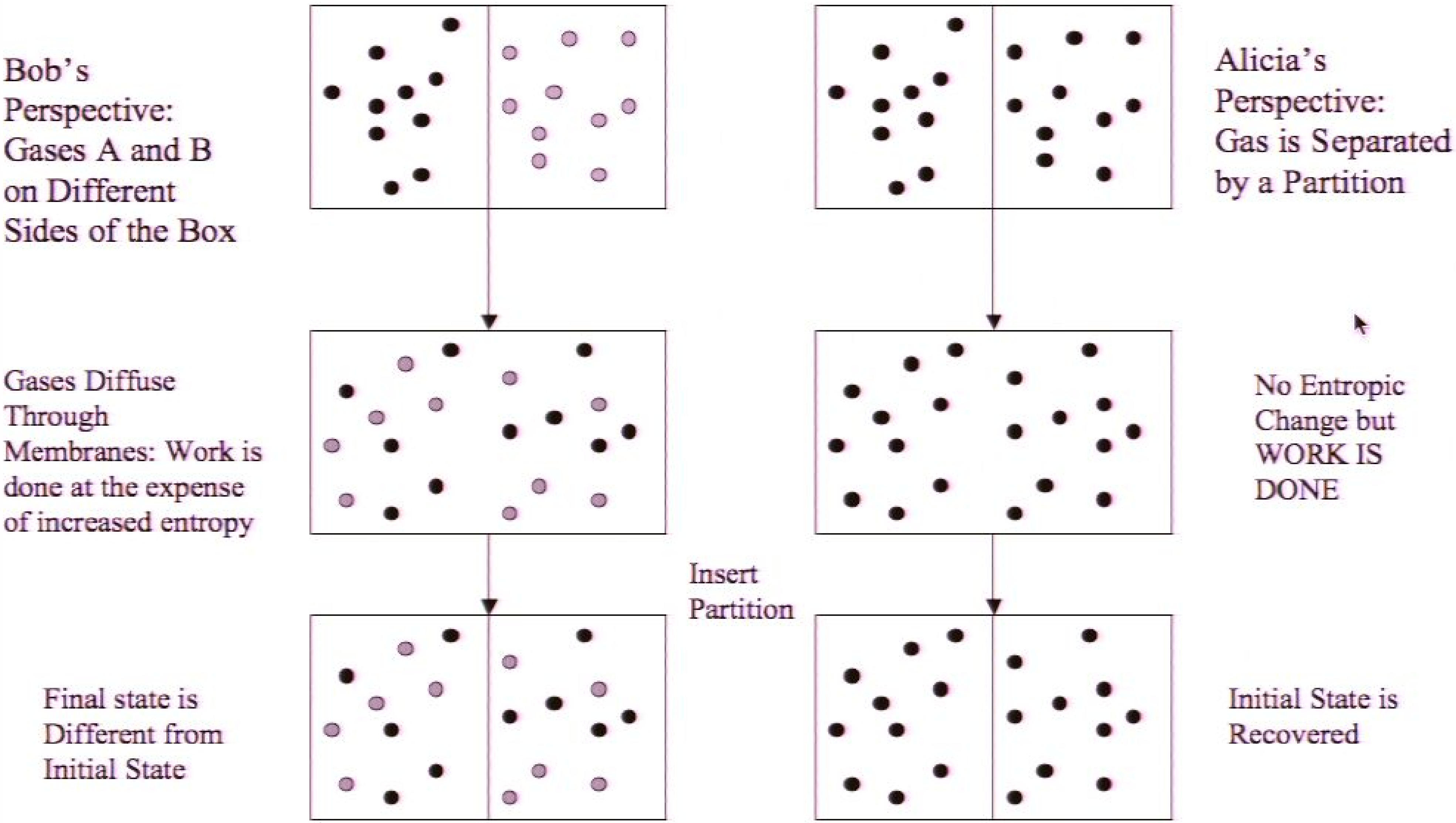
- The Kelvin reading of the second law:
- It is impossible to perform a cyclic process with no other result than that heat is absorbed from a reservoir, and work is performed.
  - The natural interpretation of a cycle here is where one's knowledge of the system returns to its initial state.
  - A counterexample to this second law is easy to furnish

An information-theoretic 2nd law?

- Suppose I am in possession of a black box. The box has an on/off switch, and a weight attached to a light rope. Upon flipping the switch, the weight is lifted, and heat is extracted from the environment. Once I've flipped the switch off:
  - Heat has been extracted from the environment.
  - Work has been performed.
  - My knowledge of the box's state is identical to the initial situation.
- Voila!

An information-theoretic 2nd law?

- So what is an information-theoretic 2nd law? Such a statement is noticeably lacking in the literature.
- However, there are two proposals one can distill from the work of E.T. Jaynes.
- First:"an entropy decrease…cannot be achieved reproducibly by manipulating the macrovariables that we have chosen to define our macrostate" (1992 p.10)


Bob's Perspective: Gases A and B on Different Sides of the Box
Alicia's Perspective: Gas is Separated by a Partition
Gases Diffuse Through Membranes: Work is done at the expense of increased entropy
No Entropic Change but WORK IS DONE
Insert Partition
Final state is Different from Initial State
Initial State is Recovered

Jaynes #1

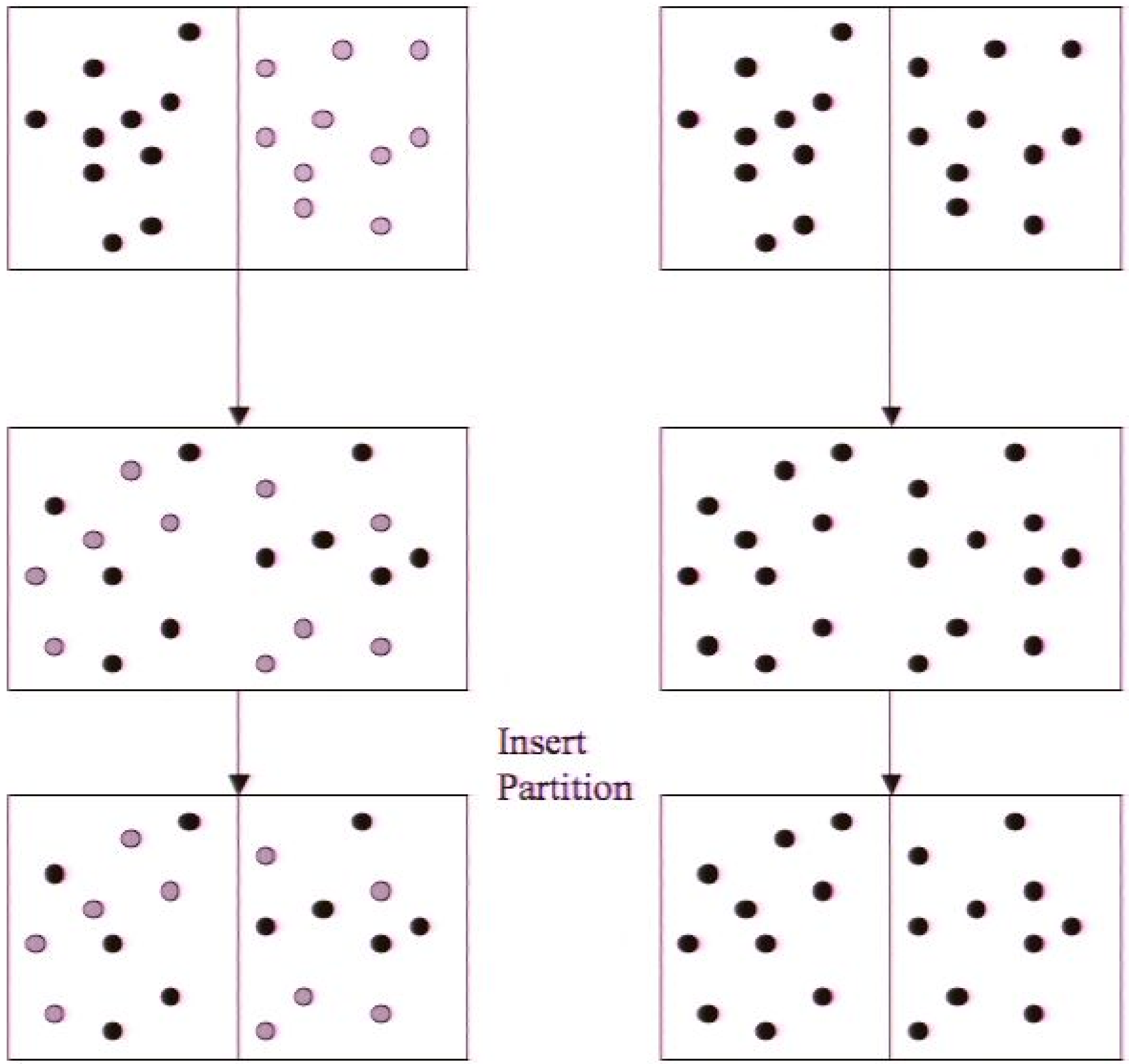
- The suggestion here is that we have some freedom in choosing how we partition a system into sets of macroscopic constraints. One cannot violate the second law by manipulating *those* macroscopic variables, though it might appear that the 2nd law is violated if other macrovariables are manipulated.
- This accords with Hemmo and Shenker's analysis.

An information-theoretic 2nd law?

- So what is an information-theoretic 2nd law? Such a statement is noticeably lacking in the literature.
- However, there are two proposals one can distill from the work of E.T. Jaynes.
- First:"an entropy decrease…cannot be achieved reproducibly by manipulating the macrovariables that we have chosen to define our macrostate" (1992 p.10)

Bob's Perspective: Gases A and B on Different Sides of the Box

Alicia's Perspective: Gas is Separated by a Partition



Gases Diffuse Through Membranes: Work is done at the expense of increased entropy

No Entropic Change but WORK IS DONE

Final state is Different from Initial State

Initial State is Recovered

Jaynes #1

- What makes this information-theoretic?
  - The entropy is relativised to a mode of description of a system, not an intrinsic feature of the system itself.
  - In the Szilard example, the position of the particle is among the macrovariables being manipulated, but the entropy calculation is effected independently of the position of the particle.
  - However, it is not clear how to connect this to the SM formalism.

Jaynes #2

- Jaynes (1963) offers an "almost unbelievably short" proof of the 2nd law.
- The interest here is not in the details of the proof, but what it can tell us about the status of the 2nd law on an IT account.
- This version of the proof is from Lavis & Milligan (1985), drawing on an expanded version of the proof found in Robertson (1966) and Hobson & Loomis (1968)

Jaynes #2

- For a system undergoing an adiabatic expansion:
- For some set of observables $\{\Omega_1(t), \Omega_2(t), \Omega_3(t), \Omega_4(t), \Omega_5(t)\ldots\}$ and measurements made on these observables at time t_0, $\{\omega_1(t_0), \omega_2(t_0), \omega_3(t_0), \omega_4(t_0)\ldots\}$, we find the density matrix operator $\rho_0(t_0)$, that maximises the information-theoretic entropy subject to the constraints

$$\omega_k(t_0) = Tr\left[\hat{\rho}_0(t_0)\hat{\Omega}(t_0)\right], k = 1\ldots m$$

Jaynes #2

- And is taken to be equal to the thermodynamic entropy
  $$S_e^{(0)}(t_0) = -kTr\left[\hat{\rho}_0(t_0)\ln\hat{\rho}_0(t_0)\right]$$
- As the system evolves, at some later time t, we *predict* the values of the observables $\{\omega_1(t), \omega_2(t), \omega_3(t), \omega_4(t), \omega_5(t)\ldots\}$, using
  $$\omega_k(t) = Tr\left[\hat{\rho}_0(t)\hat{\Omega}(t)\right], k = 1\ldots m$$
- Given these predicted values for the observables, we can define a new density matrix according to these constraints, and also subject to maximising the IT entropy

$$\omega_k(t) = Tr\left[\hat{\rho}(t)\hat{\Omega}(t)\right], k = 1\ldots m$$

Jaynes #2

- The new thermodynamic entropy is then given by the new density operator

  $$S_e(t) = -kTr\left[\hat{\rho}(t)\ln\hat{\rho}(t)\right]$$

- Since
  - The initial entropy is invariant under unitary evolution
  - both $\rho_0(t)$ and $\rho(t)$ satisfy the constraints provided by the values of the predicted observables.
  - $\rho(t)$, but not necessarily $\rho_0(t)$, maximises the information-theoretic entropy for the predicted values of the observables.

- It follows that: $S_e^{(0)}(t_0) \leq S_e(t)$

Worries

- Of the standard variety (Lavis & Milligan, Earman, Sklar, Frigg)
  - This "proof" cannot give us any of the details we want, from details of the dynamical evolution to, say, transport coefficients.
  - The inequality only shows that the entropy at later times is greater than the initial entropy, but there is no proof of monotonic increase.
  - The measurements taken at t_0 are privileged, in that this is the only time at which the density matrix makes contact with the actual values of the observables. We could use the updated density matrix used to define S(t) at later times to get a monotonic increase, but this makes the entropy curve dependent on the times chosen.

Worries

- These are not good criticisms, in a certain sense.
- The reason why these criticisms have any value is for two reasons:
  - All that is required of the density matrix is that it evolves unitarily, but otherwise is open about the dynamics of the system. If we strengthened the description of the evolution these concerns might be answered.
  - The predicted entropy at later times is assumed to be the *Thermodynamic Entropy*. If we were to relax the assumption that what is being predicted is not the thermodynamic entropy, but some other function, these worries seem moot.

Jaynes #2

- The new thermodynamic entropy is then given by the new density operator

  $S_e(t) = -kTr\left[\hat{\rho}(t)\ln\hat{\rho}(t)\right]$

- Since
  - The initial entropy is invariant under unitary evolution
  - both $\rho_0(t)$ and $\rho(t)$ satisfy the constraints provided by the values of the predicted observables.
  - $\rho(t)$, but not necessarily $\rho_0(t)$, maximises the information-theoretic entropy for the predicted values of the observables.

- It follows that: $S_e^{(0)}(t_0) \leq S_e(t)$

Worries

- These are not good criticisms, in a certain sense.
- The reason why these criticisms have any value is for two reasons:
  - All that is required of the density matrix is that it evolves unitarily, but otherwise is open about the dynamics of the system. If we strengthened the description of the evolution these concerns might be answered.
  - The predicted entropy at later times is assumed to be the *Thermodynamic Entropy*. If we were to relax the assumption that what is being predicted is not the thermodynamic entropy, but some other function, these worries seem moot.

Going the other way

- In the information-theoretic vein, we would like the probabilities to be interpreted epistemically, not physically.
  - Why should we assume that our *knowledge* of the system evolves unitarily, let alone according to the dynamics of the system? This is almost surely a false assumption, unless we are Laplacian demons.
  - The values of $\omega_k(t)$ are the *predicted values for the observables*, not the actual values. The thermodynamic entropy is then calculated for the predicted values, not the values that actually obtain. In what sense is this the *thermodynamic entropy?* What is the connection between this entropy and the experimental entropy?

Worries

- These are not good criticisms, in a certain sense.
- The reason why these criticisms have any value is for two reasons:
  - All that is required of the density matrix is that it evolves unitarily, but otherwise is open about the dynamics of the system. If we strengthened the description of the evolution these concerns might be answered.
  - The predicted entropy at later times is assumed to be the *Thermodynamic Entropy*. If we were to relax the assumption that what is being predicted is not the thermodynamic entropy, but some other function, these worries seem moot.

Jaynes #2

- The new thermodynamic entropy is then given by the new density operator

$$S_e(t) = -kTr\left[\hat{\rho}(t)\ln\hat{\rho}(t)\right]$$

- Since
  - The initial entropy is invariant under unitary evolution
  - both $\rho_0(t)$ and $\rho(t)$ satisfy the constraints provided by the values of the predicted observables.
  - $\rho(t)$, but not necessarily $\rho_0(t)$, maximises the information-theoretic entropy for the predicted values of the observables.

- It follows that: $S_e^{(0)}(t_0) \leq S_e(t)$

Jaynes #2

- And is taken to be equal to the thermodynamic entropy
  $$S_e^{(0)}(t_0) = -kTr\left[\hat{\rho}_0(t_0)\ln\hat{\rho}_0(t_0)\right]$$
- As the system evolves, at some later time t, we *predict* the values of the observables $\{\omega_1(t), \omega_2(t), \omega_3(t), \omega_4(t), \omega_5(t)\ldots\}$, using
  $$\omega_k(t) = Tr\left[\hat{\rho}_0(t)\hat{\Omega}(t)\right], k = 1\ldots m$$
- Given these predicted values for the observables, we can define a new density matrix according to these constraints, and also subject to maximising the IT entropy

$$\omega_k(t) = Tr\left[\hat{\rho}(t)\hat{\Omega}(t)\right], k = 1\ldots m$$

Jaynes #2

- The new thermodynamic entropy is then given by the new density operator

$$S_e(t) = -kTr\left[\hat{\rho}(t)\ln\hat{\rho}(t)\right]$$

- Since
  - The initial entropy is invariant under unitary evolution
  - both $\rho_0(t)$ and $\rho(t)$ satisfy the constraints provided by the values of the predicted observables.
  - $\rho(t)$, but not necessarily $\rho_0(t)$, maximises the information-theoretic entropy for the predicted values of the observables.

- It follows that: $S_e^{(0)}(t_0) \leq S_e(t)$

Worries

- These are not good criticisms, in a certain sense.
- The reason why these criticisms have any value is for two reasons:
  - All that is required of the density matrix is that it evolves unitarily, but otherwise is open about the dynamics of the system. If we strengthened the description of the evolution these concerns might be answered.
  - The predicted entropy at later times is assumed to be the *Thermodynamic Entropy*. If we were to relax the assumption that what is being predicted is not the thermodynamic entropy, but some other function, these worries seem moot.

Going the other way

- In the information-theoretic vein, we would like the probabilities to be interpreted epistemically, not physically.
  - Why should we assume that our *knowledge* of the system evolves unitarily, let alone according to the dynamics of the system? This is almost surely a false assumption, unless we are Laplacian demons.
  - The values of $\omega_k(t)$ are the *predicted values for the observables*, not the actual values. The thermodynamic entropy is then calculated for the predicted values, not the values that actually obtain. In what sense is this the *thermodynamic entropy?* What is the connection between this entropy and the experimental entropy?

Going the other way

- Generally, we are not Laplacian demons, and our knowledge of the evolution of the system does not proceed in a unitary fashion. How does this affect the result?
- Define two sets of entropies, S_e and S_I, corresponding to the thermodynamic entropy and the information-theoretic entropy, respectively.
- By construction, $S_e(t_0)=S_I(t_0)$
- At some later time t, we predict the expected values of the observables as before, though we make no assumption regarding the dynamics:

$$\omega_k(t) = Tr\left[\hat{\rho}_0(t)\hat{\Omega}(t)\right], k = 1 \ldots m$$

Going the other way

- We calculate $S_I(t)$ as before, by defining the density matrix as the one that satisfies these constraints and simultaneously maximises S_I
- Note that this will not generally correspond to an accurate prediction of $S_e(t)$, since this involves considerations of the dynamics.

Going the other way

- However, we can make a statement about the relation between $S_I(t_0)$ and $S_I(t)$.
- If we assume that one cannot compress the probability distribution on the basis of the knowledge we possess, it will still be the case that $S_I(t_0) \leq S_I(t)$
- This appears to be a form of an information-theoretic second law.

Remaining worries

- This result borders on trivial, for it roughly says that we don't magically acquire knowledge but are apt to lose or forget it (though we may later remember).
- The criticisms of the original proof are evaded by moving to an explicitly epistemic context, but does so at the cost of relinquishing any contact with the entropy of traditional thermodynamics.
- We need to establish the conditions under which these two frameworks coincide, and this is a highly non-trivial task.

On information and entropy

- Jaynes (1963) notes that
- "*for some distributions* and *in some physical situations*, has long been recognised as representing entropy. However, we have to emphasise that the "information-theory entropy" S_I and the experimental thermodynamic entropy S_e are entirely different concepts. Our job cannot be to *postulate* any relation between them; it is rather to *deduce* whatever relations we can from known mathematical and physical facts." (187, emphasis original)

Information and Entropy

- At best, this remains a promissory note, but it is on the right track.
- We only have motions towards what one might mean by an information-theoretic 2nd law.
- We only have motions as to what the relation between an information-theoretic approach to SM and its relation to TD might mean.
- What is needed is a framework that models the clarity of the Boltzmannian interpretation offered by Albert, Hemmo and Shenker.

Information and Entropy

- At best, this remains a promissory note, but it is on the right track.
- We only have motions towards what one might mean by an information-theoretic 2nd law.
- We only have motions as to what the relation between an information-theoretic approach to SM and its relation to TD might mean.
- What is needed is a framework that models the clarity of the Boltzmannian interpretation offered by Albert, Hemmo and Shenker.