Maxwell through the Looking Glass

From Szilard
to Landauer
and back again

"The laws of statistical mechanics
apply to conservative systems
of any number of degrees of freedom,
and are exact."

Josiah Willard Gibbs, 1902

PERIMETER INSTITUTE
FOR THEORETICAL PHYSICS

Maxwell through the looking glass

- The Szilard Engine and Landauer's Principle
  - The combined operation and its critics
- Statistical mechanical entropy
  - What should one expect of such an entropy?
- Macroscopic indeterminism
  - And when does such an entropy apply?
- Solving it all
  - Can a Maxwellian Demon exist? What is the validity of Landauer's Principle? Does understanding the Szilard Engine require understanding information theory? What is the statistical mechanical generalisation of entropy?
  - From four assumptions
    - (which are sufficient, but not necessary)
    - (and may not be true)

Maxwell's Demon
Szilard's Engine
Landauer's Principle

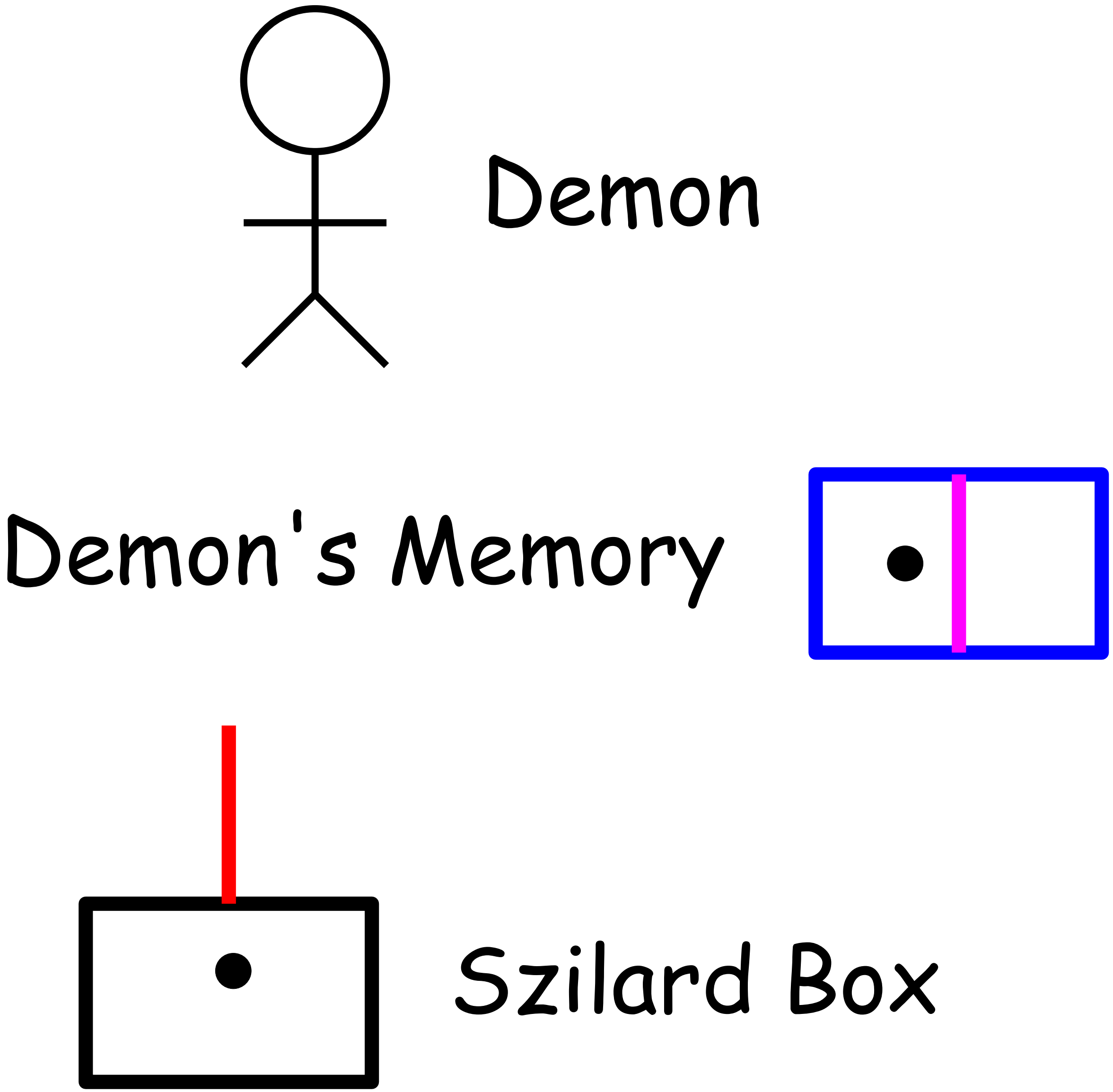
Fluctuation Phenomena and Thermal Physics

- The observability of fluctuation phenomena (since 1905) has been regarded as a challenge to the second law of thermodynamics
  - Maxwell's original demon was supposed to need to be too small
- Smoluchowski and followers show a mechanical demon goes into reverse as it is also subject to fluctuations.
  - Exorcism on a case-by-case basis.
    - Each exorcism supposedly illustrates the non-existence of Demon's, but it is less clear why one should go from the failure of a particular Demon to the assumption that all potential Demons must fail in the same way.
  - A large literature exists of continuing attempts to construct exceptions.
- It would be helpful to know: is there a *general proof*? (Yes!)

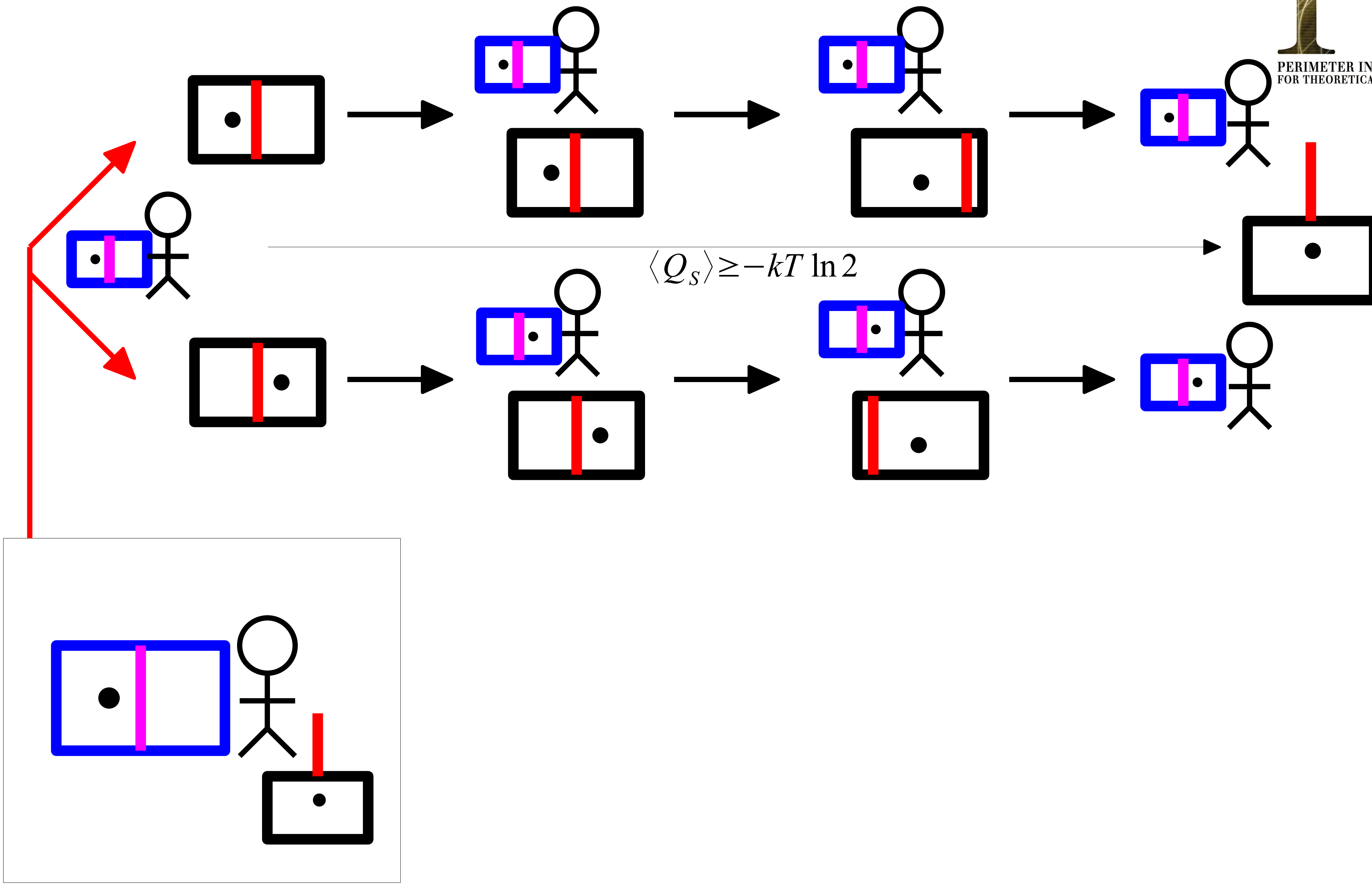
The Szilard Engine

PERIMETER INSTITUTE
FOR THEORETICAL PHYSICS

- An atom in a box, in thermal contact with a heat bath.
  - The box is separated in two by a partition, trapping the atom on one side or the other
- The fluctuation is 'guaranteed'.
  - Whichever side the atom is trapped upon, the volume available to it has decreased.
- To extract work from the fluctuation it is necessary to determine which side the atom is on.
  - Information gathering, measurement is required.
  - Information erased, Landauer's Principle is required.
- Criticism
  - Landauer's Principle not independent of second law, exorcism is circular (Earman & Norton, Norton)
  - Landauer's Principle is not sustainable, Maxwell's Demons may be possible (Shenker, Shenker & Hemmo)
  - Landauer's Principle is irrelevant, Maxwell's Demons *are* possible (Albert).


Demon
Demon's Memory
Szilard Box

Szilard's Engine

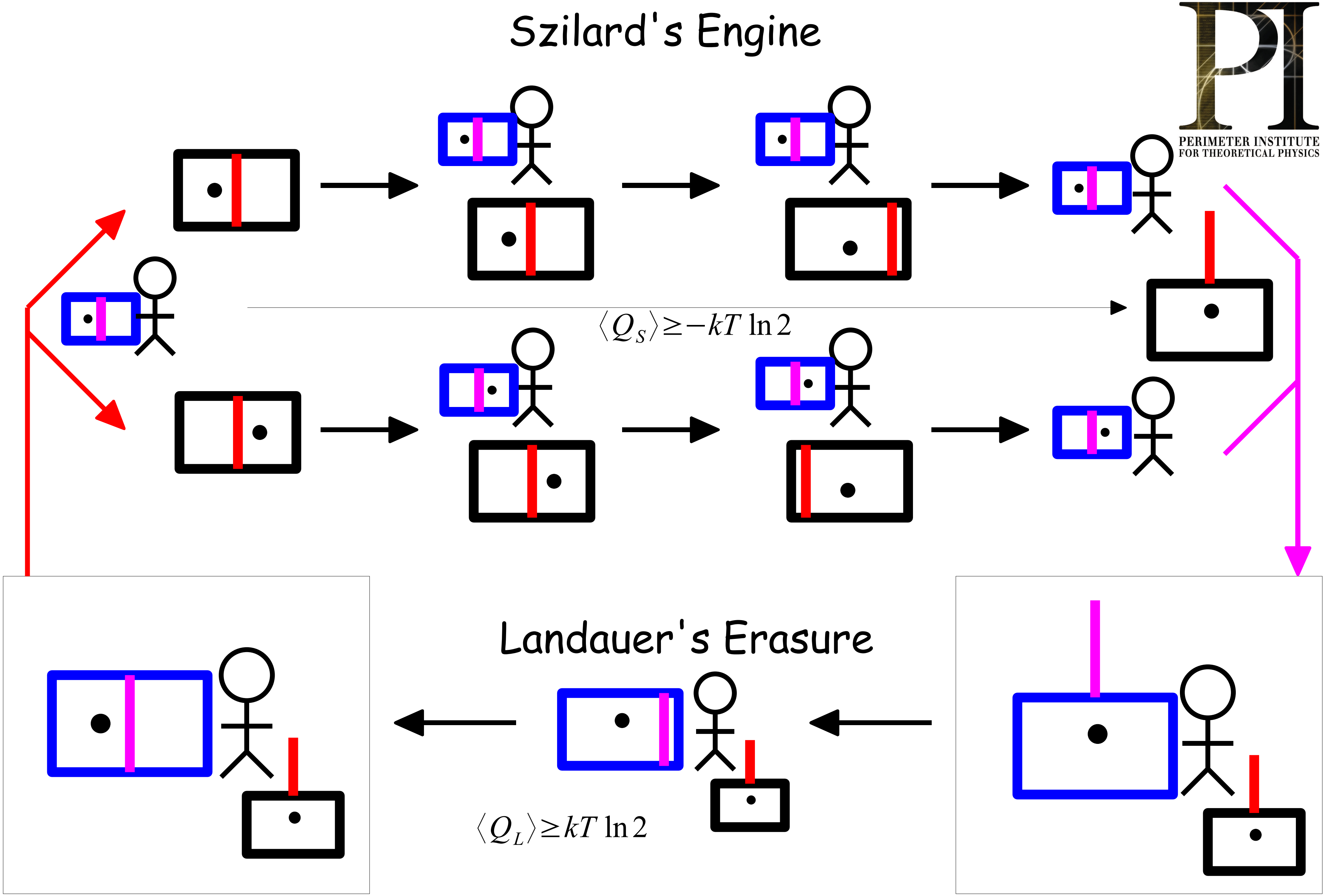


Maxwell through the Looking Glass:
From Szilard to Landauer and back again.

The Szilard-Landauer Cycle

- Work *can* be extracted from the Szilard Box.
  - What is the explanation for this? What is the origin of the work extracted?
  - Entropy is a lack of information, by performing a measurement has it been reduced?
- The Demon retains information at the end.
  - Does this compensate?
  - But each distinct state of the Demon has the same entropy.
  - Is the overall entropy higher, or lower, or the same?
- Why the need for the correlation?
  - Can we extract the work without the Demon?

PERIMETER INSTITUTE
FOR THEORETICAL PHYSICS



Maxwell through the Looking Glass:
From Szilard to Landauer and back again.

PIAF Sydney, February 2008

The Szilard-Landauer Cycle

- Where is the 'principle'?
  - And why should one believe it?
  - The circularity argument:
    - What are the independent grounds for believing there is no better return path?
- *What*, exactly, is the principle?
  - in erasing one bit . . . of information one dissipates, on average, at least kBT ln (2) of energy into the environment. [Pie00]
  - a logically irreversible operation must be implemented by a physically irreversible device, which dissipates heat into the environment [Bub01]
  - any logically irreversible manipulation of data *: : :* must be accompanied by a corresponding entropy increase in the non-information bearing degrees of freedom of the information processing apparatus or its environment. Conversely, it is generally accepted that any logically reversible transformation of information can in principle be accomplished by an appropriate physical mechanism operating in a thermodynamically reversible fashion. [Ben03]
- *Why* restore the Demon?
  - Why should one care about the Demon's memory?
    - Entropy has gone down, work has been extracted. Who cares where the shoe is?

Statistical Mechanics

Statistical Mechanics:
Assumptions

PERIMETER INSTITUTE FOR THEORETICAL PHYSICS

- Unitary evolution on density matrices
  - *(cf. wavefn. collapse; Zhang & Zhang 1992)*
- Negligible variation in interaction energies
  - *(cf. Allahverdyan & Nieuwenhuizen 2001)*
- Statistical independence between initial systems *(and no equivalent final condition)*
  - *(cf. arrow of time asymmetry; non-extensive entropies; non-markovian master equations)*
- Thermal systems are Gibbs canonical states
  - *(cf. non-extensive entropies, microcanonical entropies, objective Boltzmann entropies)*
  - NB. this can be deduced from statistical independence, with additional requirements
    - *composivity [Szilard 1925, Tisza & Quay 1963]*
    - *stable equilibrium [Hatsopoulos & Gyftopoulos 1976]*
    - *complete passivity [Pusz & Woronowicz 1978]*
    - *reservoir stability [Sewell 1980]*

Statistical Mechanics:
Assumptions

Unitary evolution on density matrices: $\rho(t)=U\rho_0 U^{\dagger}$ $\quad -i\hbar\frac{\partial U}{\partial t}=[H,U]$

$$H=H_1+H_2+V_{12} \qquad \Delta E_i=\Delta W_i+\Delta Q_i$$

$$\Delta E_i=Tr\left[H_i(t)\rho(t)-H_i(0)\rho(0)\right] \quad \Delta W_i=\int Tr\left[\frac{\partial H_i}{\partial t}\rho_i(t)\right]dt \quad \Delta Q_i=\int Tr\left[[H_i,V_{12}]\rho(t)\right]dt$$

Negligible variation in interaction energy: $\frac{\partial V_{12}(t)}{\partial t}\approx 0$ $\quad Tr\left[V_{12}(\rho(t)-\rho(0))\right]\approx 0$

$$\sum_i \Delta Q_i\approx 0$$

Statistical independence between intial systems: $\rho(0)=\prod_i \rho_i(0)$

Thermal systems are Gibbs canonical states: $\rho_i(T)=\frac{e^{-H_i/kT}}{Tr\left[e^{-H_i/kT}\right]}$

Statistical Mechanics:
Theorems

No Hamiltonian process is possible, whose sole result is to return a system to its original, marginal statistical state, while transferring mean energies to initially uncorrelated, canonical systems, with dispersions β_a, unless:

$$\sum_a \beta_a \langle \Delta Q_a \rangle \geq 0$$

If there is a process, whose sole result is to take the marginal state of a system from A to B, depositing mean energy Q_{AB} in a heat bath at temperature T, then there is no process taking the marginal state from B to A, depositing mean energy Q_{AB} in a heat bath at temperature T, unless:

$$\frac{\Delta Q_{BA}}{T} \geq \frac{-\Delta Q_{AB}}{T}$$

There exists a single valued function of the statistical states, $S[\rho]$, such that, if there exists a process from state A to B, on average depositing energy Q_{AB} in a heat bath at temperature T, then:

$$S[\rho_A] - S[\rho_B] \leq \frac{\Delta Q_{AB}}{T}$$

$$\frac{-\Delta Q_{BA}}{T} \leq S[\rho_A] - S[\rho_B] \leq \frac{\Delta Q_{AB}}{T}$$

A optimal cycle is one for which:

$$\Delta Q_{AB} + \Delta Q_{BA} = 0$$ uniquely identifying $S[\rho]$: $$S[\rho_A] = S[\rho_B] + \frac{\Delta Q_{AB}}{T}$$

There is *no* process taking state A to B, on average depositing energy Q_{AB} in a heat bath at temperature T, unless the function $S[\rho]$ has values:

$$S[\rho_A] - S[\rho_B] \leq \frac{\Delta Q_{AB}}{T}$$

Statistical Mechanics:

Theorems

$$\Delta E_i = \Delta W_i + \Delta Q_i \qquad \sum_i \Delta Q_i \approx 0 \qquad \rho_i(T) = \frac{e^{-H_i/kT}}{Tr\left[e^{-H_i/kT}\right]}$$

In the limiting case of an isothermal evolution

$$\Delta W_i = -kT \ln\left[\frac{Z_i(t)}{Z_i(0)}\right] \qquad Z_i = Tr\left[e^{-H_i/kT}\right]$$

So for an isothermal process going from A to B

$$\Delta E_{AB} = Tr[H_B \rho_B - H_A \rho_A] \qquad \Delta W_{AB} = -kT \ln\left[\frac{Z_B(t)}{Z_A(0)}\right]$$

$$\Delta Q_{AB} = k\,T\,Tr\left[\rho_B \ln \rho_B\right] - k\,T\,Tr\left[\rho_A \ln \rho_A\right]$$

This process can make an optimal cycle: $S\left[\rho_A\right] = S\left[\rho_B\right] + \frac{\Delta Q_{AB}}{T}$

$$S\left[\rho_A\right] + k\,Tr\left[\rho_A \ln \rho_A\right] = S\left[\rho_B\right] + k\,Tr\left[\rho_B \ln \rho_B\right]$$

$$S\left[\rho_A\right] = -k\,Tr\left[\rho_A \ln \rho_A\right] + c$$

Statistical Mechanics:
Reversibility

- The optimum path is not a property of a *given* unitary evolution:
  - Given initial and final, density matrices, one can, in principle, construct an optimal unitary evolution between them;
  - Given an initial density matrix and a unitary evolution, one can determine if the evolution is optimal;
    - In general an evolution optimal for one initial density matrix is not optimal for a different initial density matrix.
    - *But* it is not necessary for either initial or final density matrices to be thermal states!
- An operation is statistical mechanically reversible if, and only if, it can in principle be incorporated into a cycle for which:

$$\sum_a \frac{\Delta Q_a}{T_a} = 0$$

Statistical Mechanics:
Irreversibility

- An operation is statistical mechanically *irreversible* if, and only if, it *cannot*, even in principle, be incorporated into a cycle for which:

$$\sum_a \frac{\Delta Q_a}{T_a} = 0$$

- Source of irreversibility?
- The closed cycle requires the marginal density matrices of all (non-heat bath) systems to be returned to their initial values
  - This includes restoring *accessible* correlations between systems
    - May still have inaccessibility of newly developed *micro*correlations.
    - Irreversibility is only apparent
  - The spin-echo type experiments show that the apparent irreversibility *can* be restored.

Statistical Mechanics:

The Theorem

Given:

Unitary evolution on density matrices
Negligible variation in interaction energy
Statistical independence between initial systems
Thermal systems are Gibbs canonical states

then:

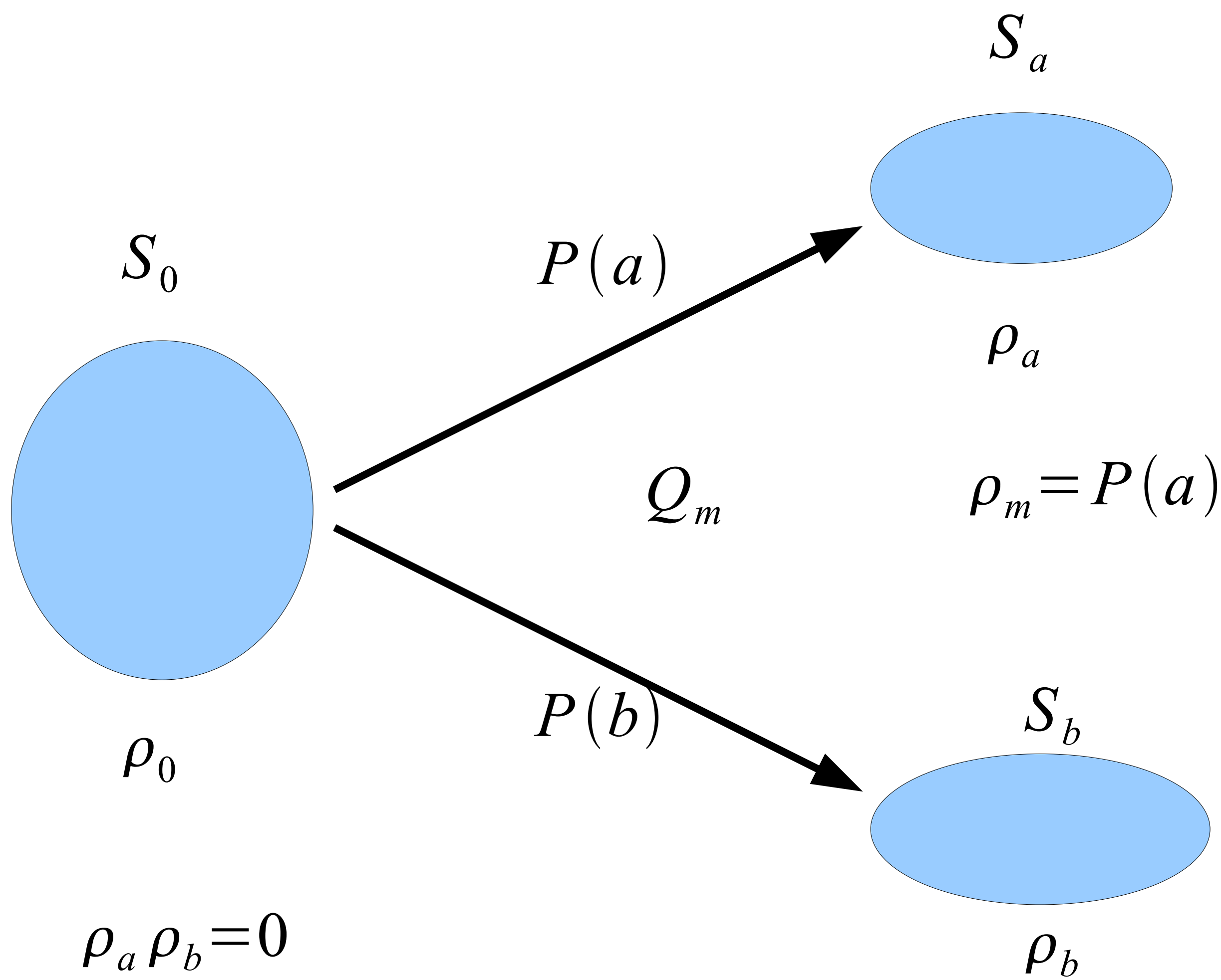
There is *no* process, whose sole result is to change the marginal state of the system from ρ_A to ρ_B, while depositing, on average, energy Q_{AB} in a heat bath at temperature T, unless:

$$Tr\left[\rho_B \ln \rho_B - \rho_A \ln \rho_A\right] \leq \frac{Q_{AB}}{k\,T}$$

Statistical Mechanics:
Macroscopic Indeterminism

- The probability distribution can be over macroscopically distinct states
  - Some explanations of *thermodynamic* entropy are based upon microscopic notions (eg. mixing, inaccessibility) that are not self-evidently applicable to macroscopic uncertainty.
  - The system is objectively within a particular region (corresponding to a particular macrostate), surely this should be the objective characterisation of the state, regardless of our uncertainty over which region it is in?
- Different possible sources of indeterminism
  - Different notions of probability may be involved
  - But the theorem still represents a limitation on the intraconvertibility of heat and work.
    - Different interpretations of what a mean value signifies
    - But provided everyone assigns the same "probability" to each particular value, everyone gets the same mean value, even if they disagree what that signifies!

Macroscopic Indeterminism



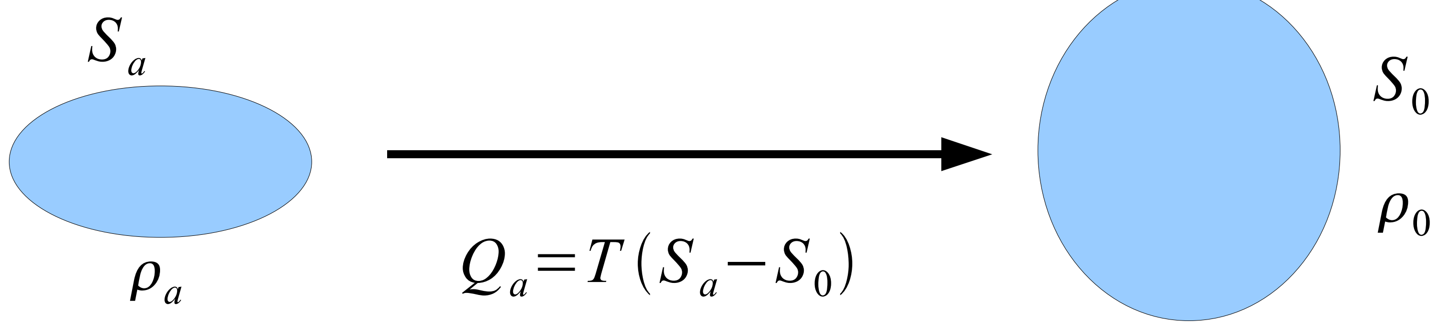
$$\rho_m = P(a)\rho_a + P(b)\rho_b$$

$$\rho_a \rho_b = 0$$

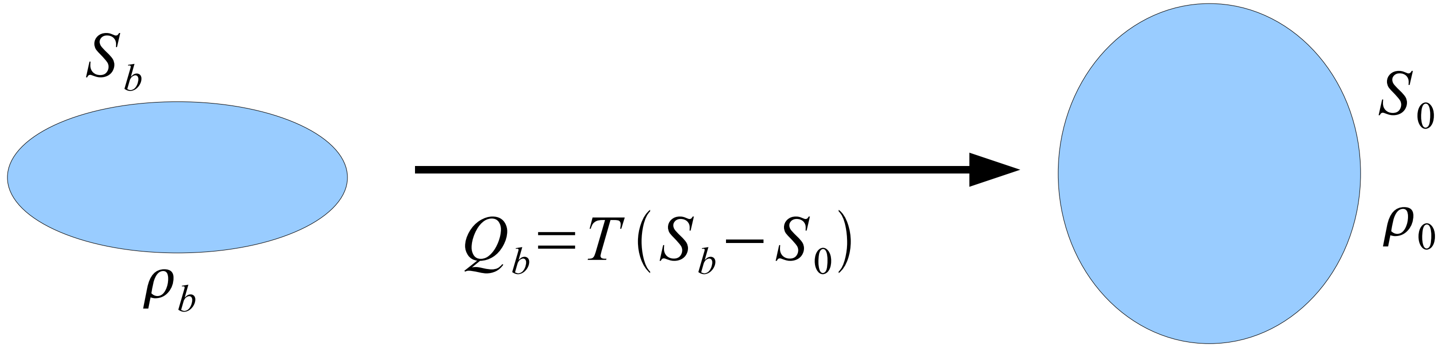
$$S_m = \sum_i P(i) S_i - k\, P(i) \ln P(i)$$

$$Q_m \geq T(S_0 - S_m)$$

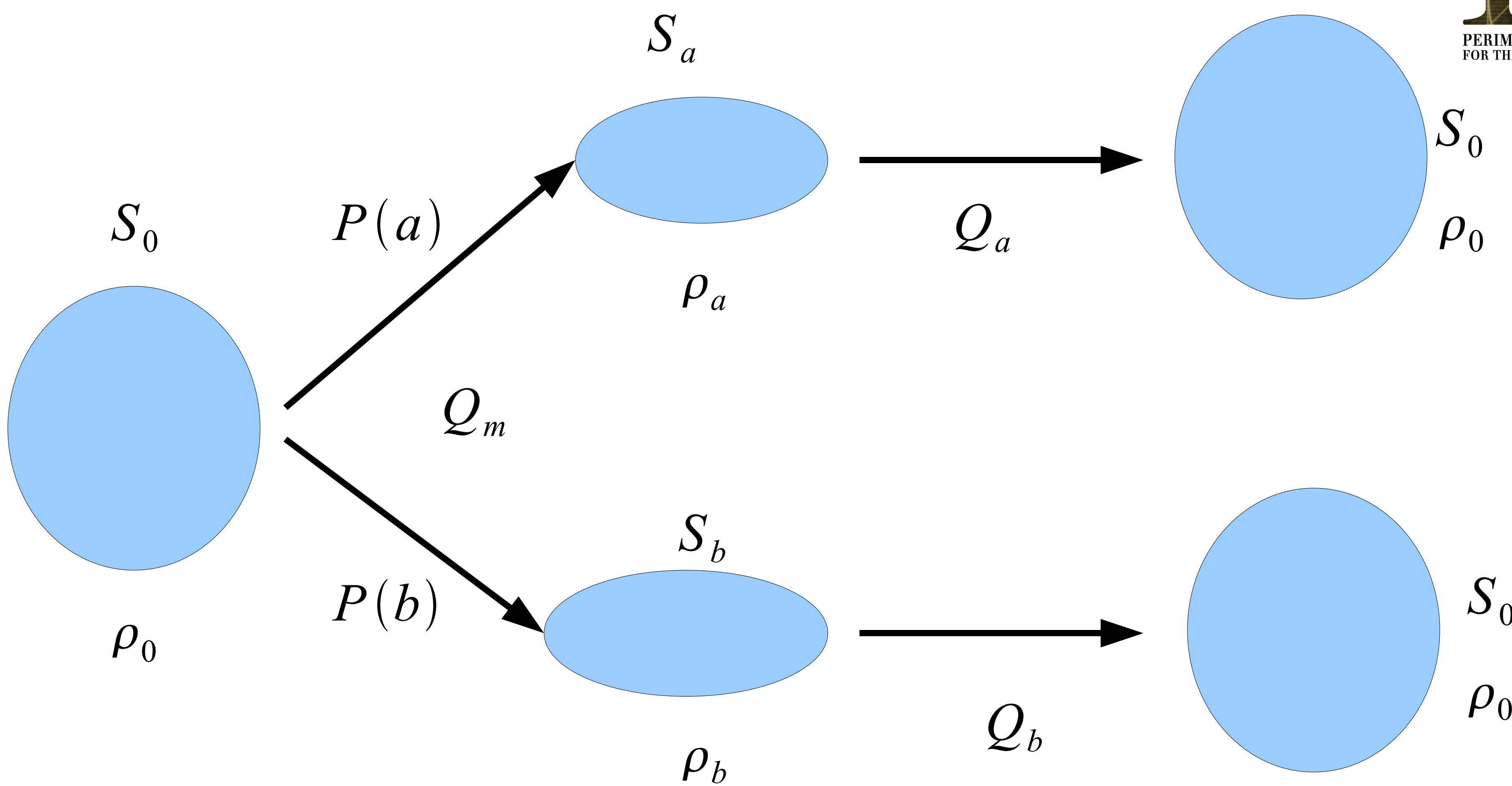
Macroscopic Indeterminism



$$\overline{Q} = \sum_i P(i) Q_i = T \sum_i P(i) S_i - T S_0 \geq kT \sum_i P(i) \ln P(i) - Q_m$$

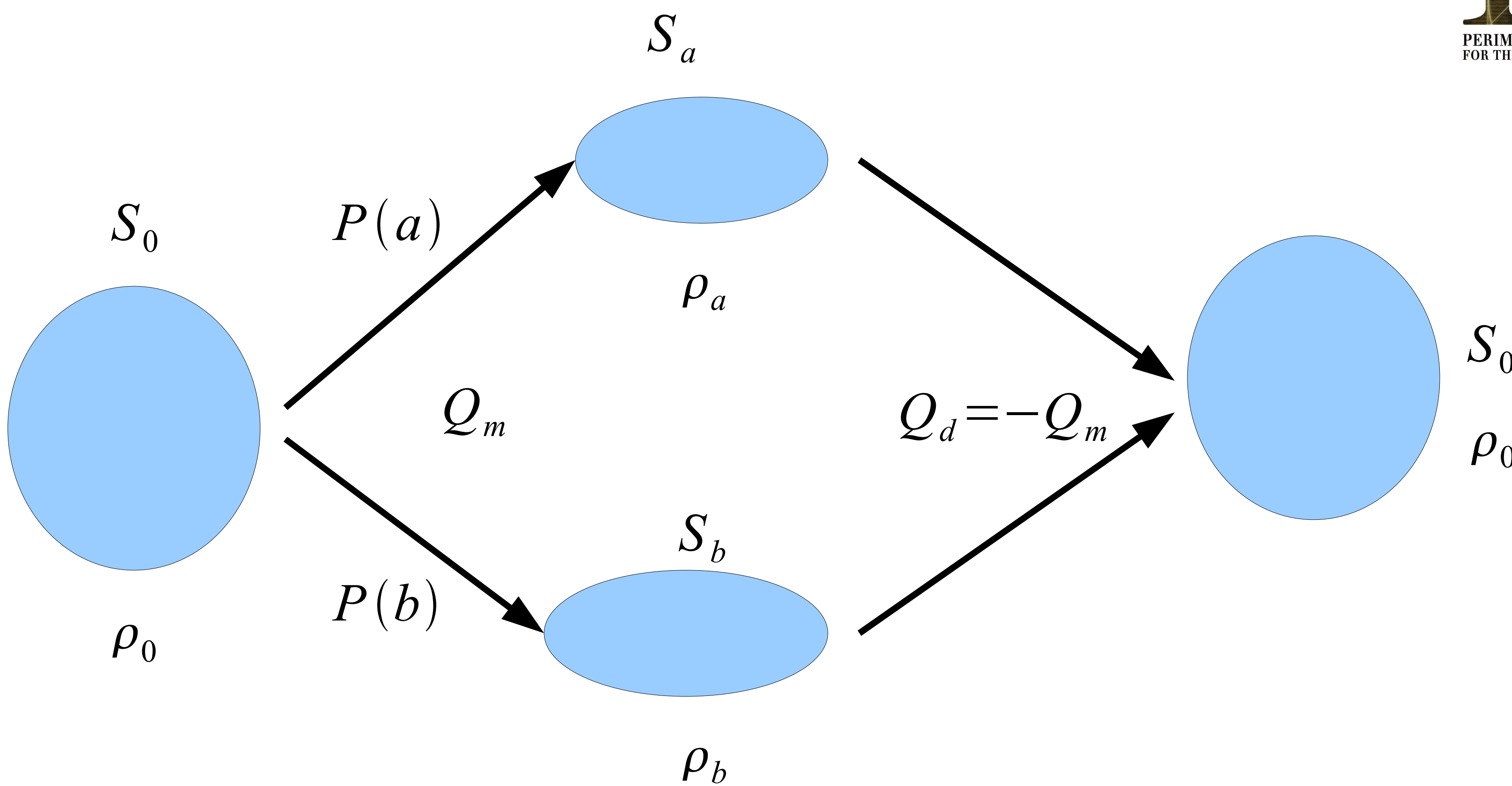


Macroscopic Indeterminism



$$\overline{Q} + Q_m \geq kT \sum_i P(i) \ln P(i)$$

Macroscopic Indeterminism



$$Q_d = \overline{Q} - kT \sum_i P(i) \ln P(i)$$

Macroscopic Indeterminism

PERIMETER INSTITUTE
FOR THEORETICAL PHYSICS

- The optimum path is not a property of a *given* unitary evolution:
  - Given initial and final, density matrices, one can, in principle, construct an optimal unitary evolution between them;
  - In general an evolution optimal for one initial density matrix is not optimal for a different initial density matrix.
- The optimal unitary evolution for extracting work from state A to state 0, and the optimal unitary evolution for extracting work from state B to state, cannot, in general, be combined into a single unitary evolution
  - The optimal unitary evolution for extracting work from a mixture of state A and state B, extracts a reduced amount of work:

$$Q_d = \overline{Q} - kT \sum_i P(i) \ln P(i)$$

Macroscopic Indeterminism

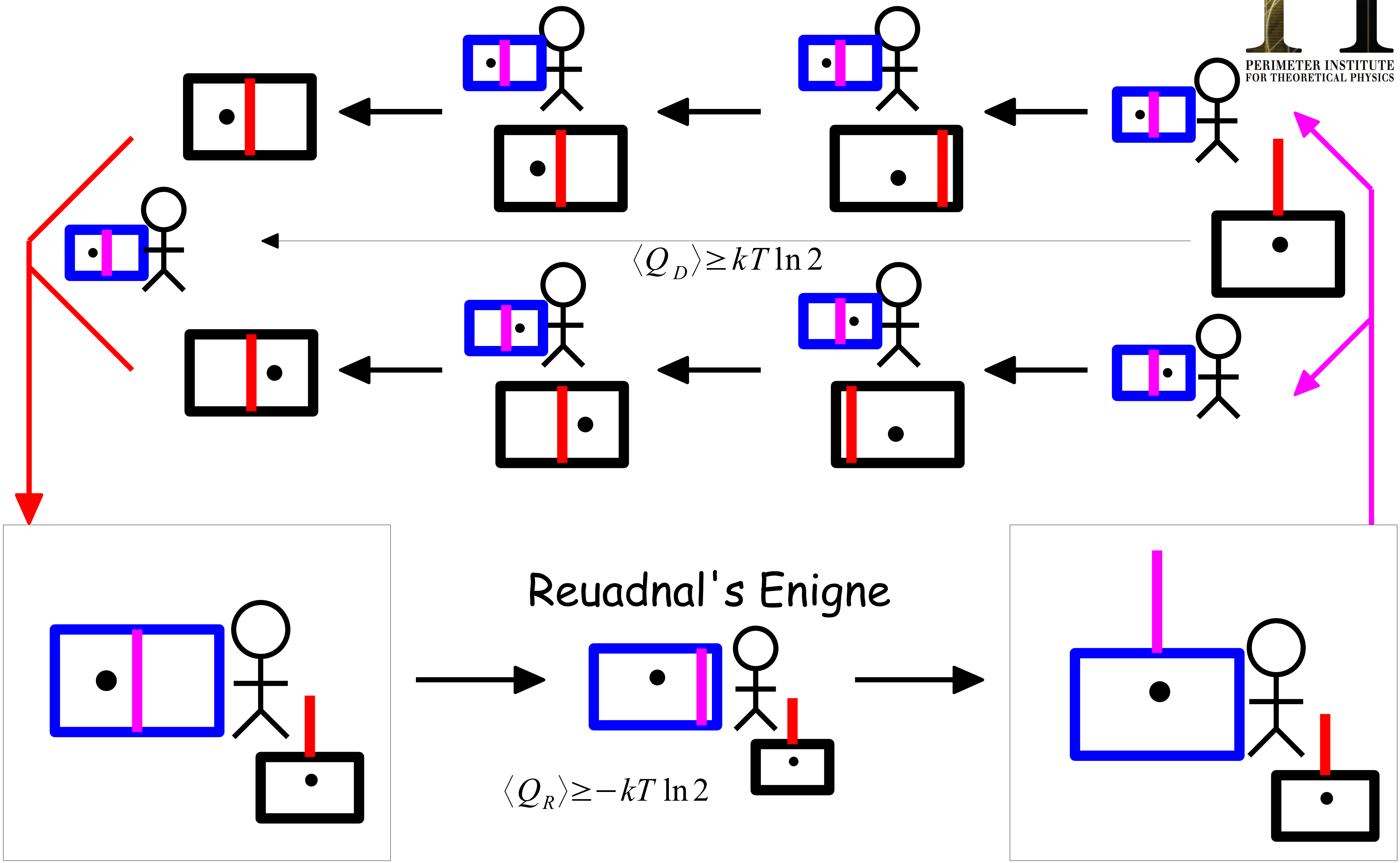
- But if I *see* the new macrostate, hasn't the entropy gone down?
  - The evolution of the system must still be unitary. Defined over the whole of the state space
    - Which now includes you, if you interact with the system, and your correlation with the system.
  - Analogy: you must place your bets before the wheel stops spinning!
- From Boltzmann to Gibbs (Penrose 1970)
  - Start by defining (objective?) Boltzmann entropies for macroscopic states
  - Then note unitary evolution allows macroscopic state entropy to fall *even on average* when macros-state evolution is indeterministic $\sum_a p_a S_a < S$
  - But also note that the fall is by the term: $-k \sum_a p_a \ln p_a$
  - Construct (or calculate) an erasure principle
    - Define optimum cost for eliminating macroscopic uncertainty $-kT \sum_a p_a \ln p_a$
  - Reconstruct Gibbs entropy! $S_{mix} = \sum_a p_a S_a - k\, p_a \ln p_a = S$

PERIMETER INSTITUTE FOR THEORETICAL PHYSICS

Through the Looking Glass

PERIMETER INSTITUTE FOR THEORETICAL PHYSICS

Dralizs's Erusare

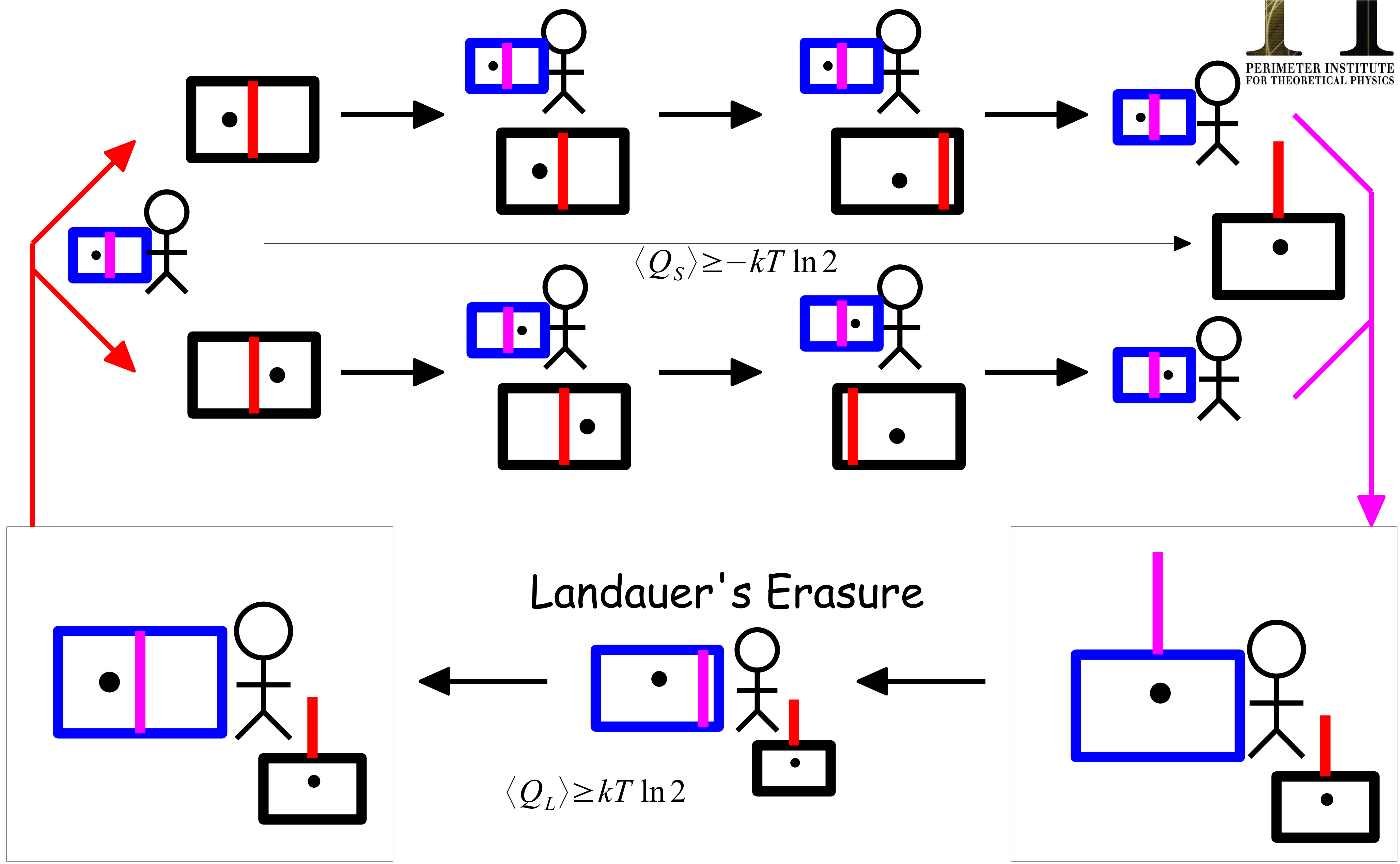


Maxwell through the Looking Glass:
From Szilard to Landauer and back again.

PIAF Sydney, February 2008

Szilard's Engine

PERIMETER INSTITUTE FOR THEORETICAL PHYSICS



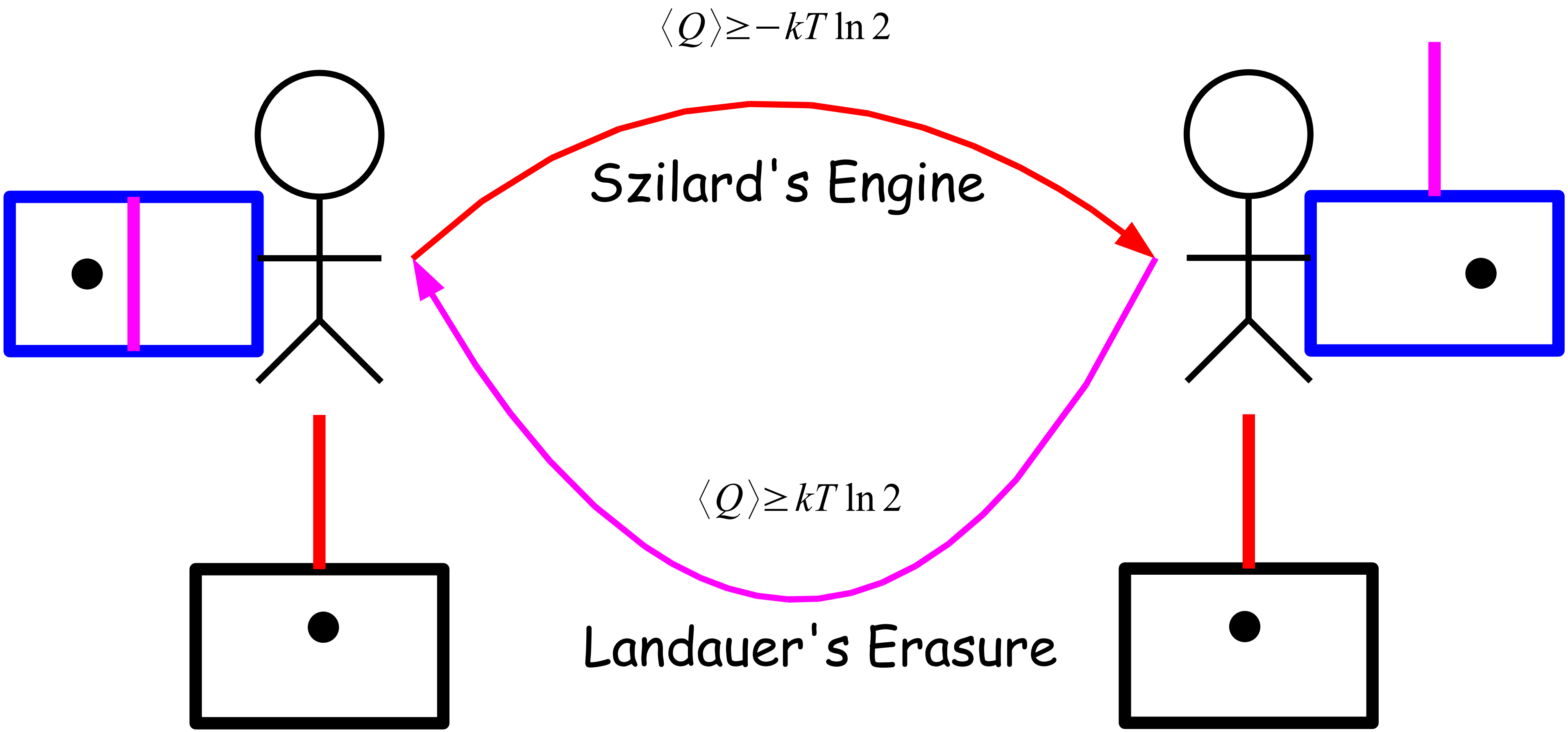
Maxwell through the Looking Glass:
From Szilard to Landauer and back again.

PIAF Sydney, February 2008

How to understand the Szilard-Landauer-Dralizs-Reuadnal cycles?

Cyclic paths

$\sum \langle Q \rangle = 0$ $\Delta S = k \ln 2$



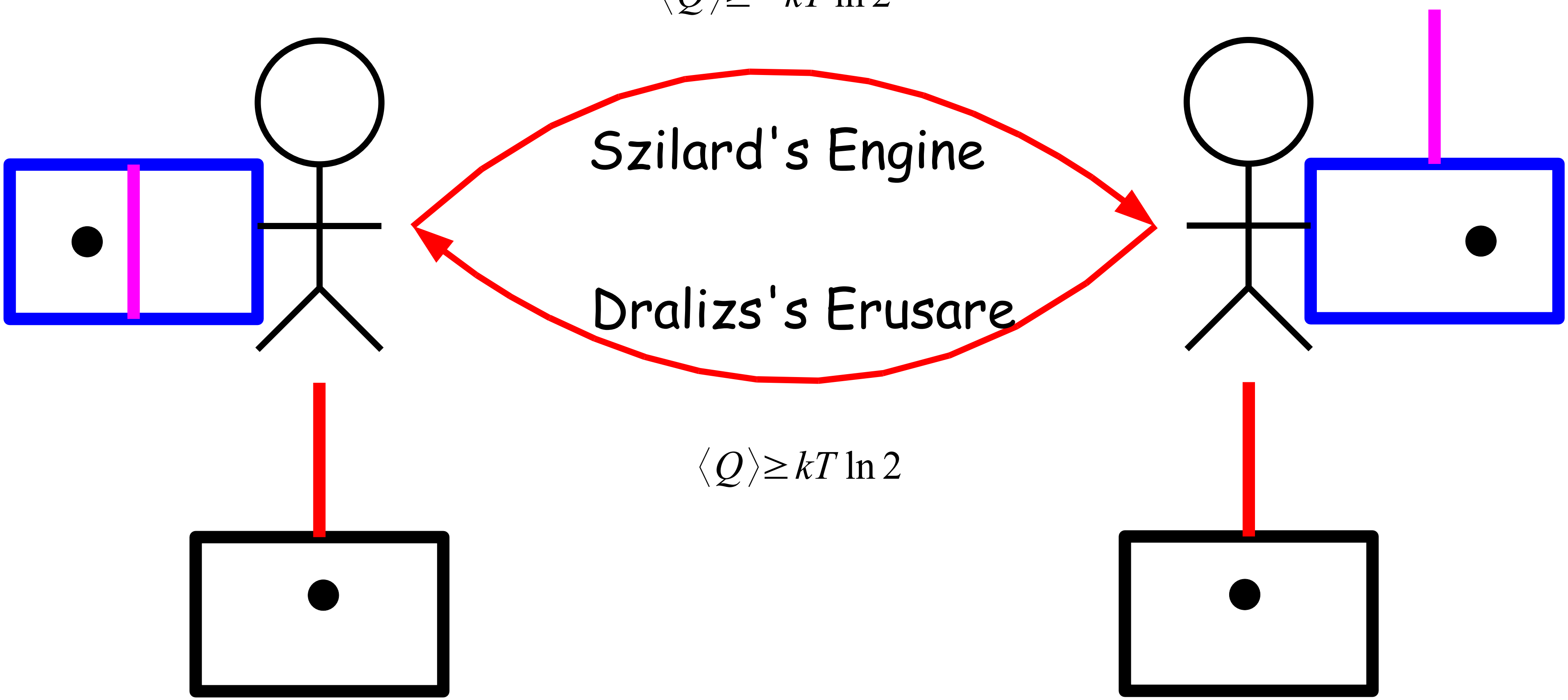
Cyclic paths

PERIMETER INSTITUTE
FOR THEORETICAL PHYSICS

$\sum \langle Q \rangle = 0$

$\Delta S = k \ln 2$

$\langle Q \rangle \geq -kT \ln 2$

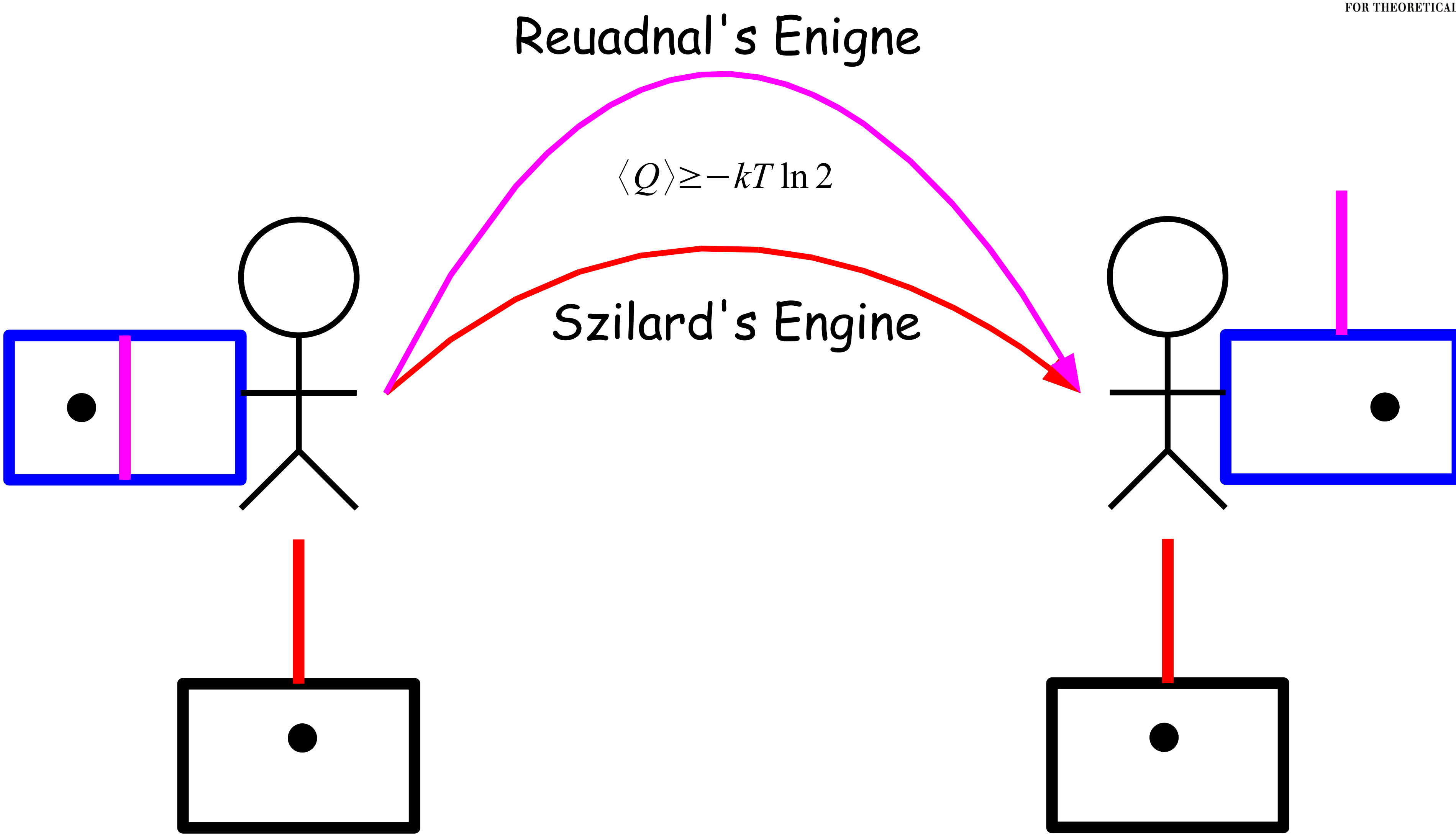


$\langle Q \rangle \geq kT \ln 2$

Szilard-Dralizs Cycle

- Landauer Erasure (or Landauer's Principle) not required for the exorcism!
  - Dralizs Erusare suffices.
  - Simply reversing the path to complete the cycle should always have been considered sufficient to establish the Szilard Engine is not a challenge to the second law.
    - *An earlier, mistaken, belief that measurement was statistically mechanically irreversible seems to have obscured this.*
- The failure is known in advance (as none of the assumptions of the statistical mechanical entropy theorem are violated)
  - The cycle merely establishes the statistical mechanical entropy difference.

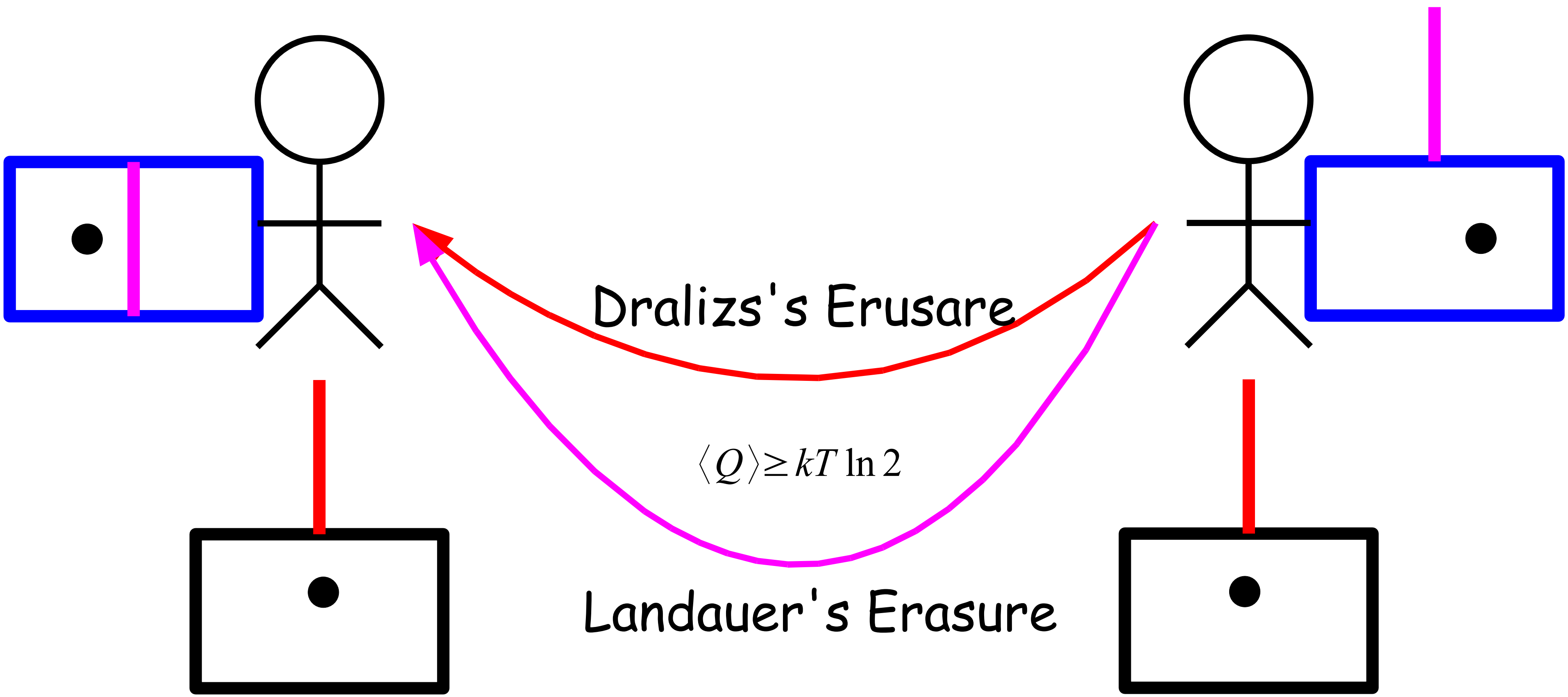
Cyclic paths



Reuadnal vs Szilard

- The extraction of the energy from the Szilard Engine is supposed to be through fluctuations, measurements, information, correlations and other complex interactions.
- In the Reuadnal Enigne it is manifestly from the isothermal expansion of the Demon state.
  - There is no need to refer to fluctuations, information, measurement, correlations to understand the source of energy extracted. The Szilard Box is *irrelevant*
- Who cares about the shoe?
  - It is the isothermal expansion of the shoe that is the source of the energy!
  - You *can* get energy from the isothermal expansion of a shoe. But you need a *really big* warehouse to hold the shoe!

Cyclic paths

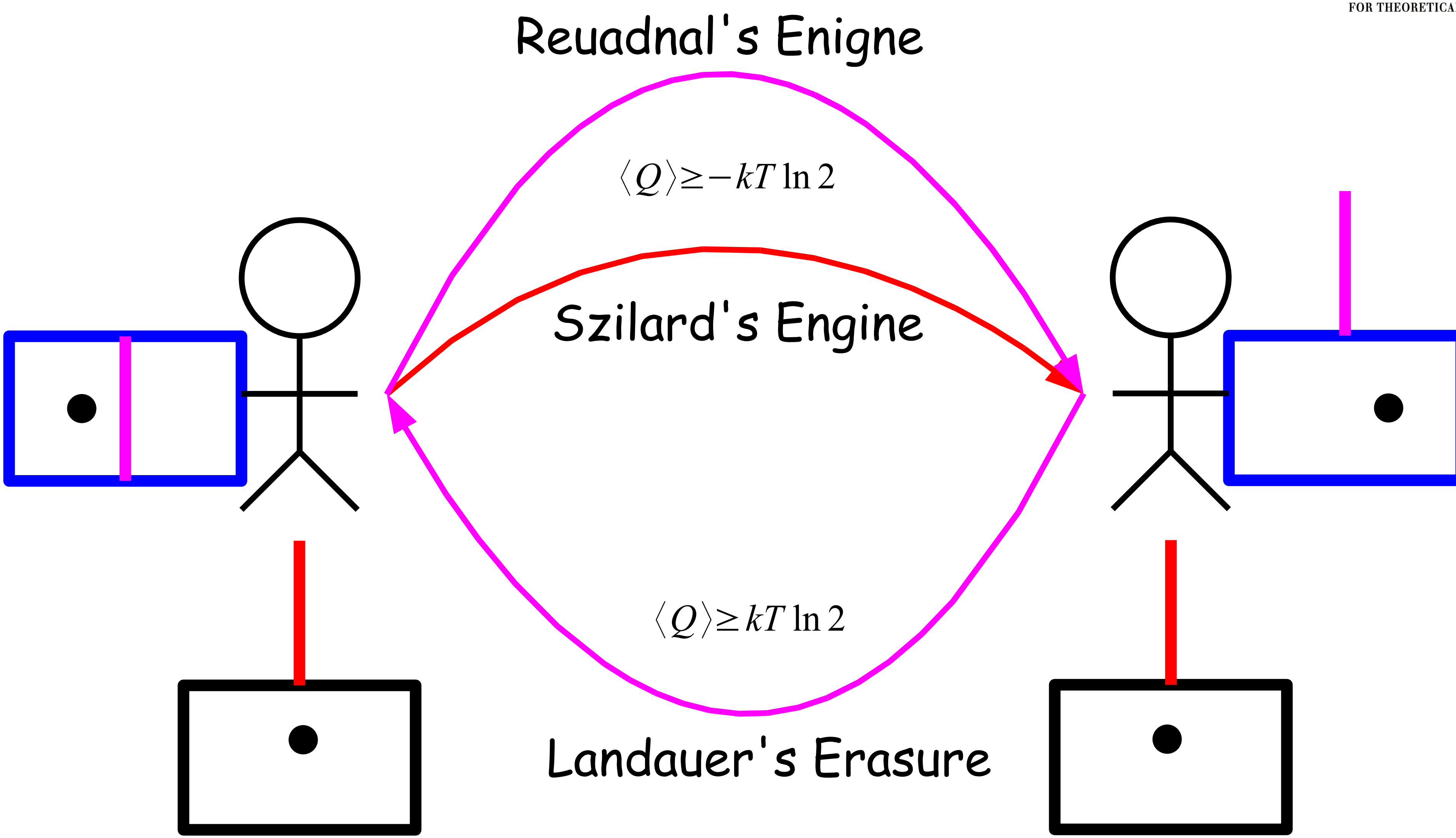


Landauer vs Dralizs

- Either Dralizs or Landauer are valid means of performing erasure
  - Both have the same costs
- Landauer's *Principle* seems valid
  - But not as a principle, “only” as a theorem.
- What establishes the *minimum* cost?
  - Not any particular example of *erasure*!
  - Only the *opposite* path, can establish a minimum!

PERIMETER INSTITUTE
FOR THEORETICAL PHYSICS

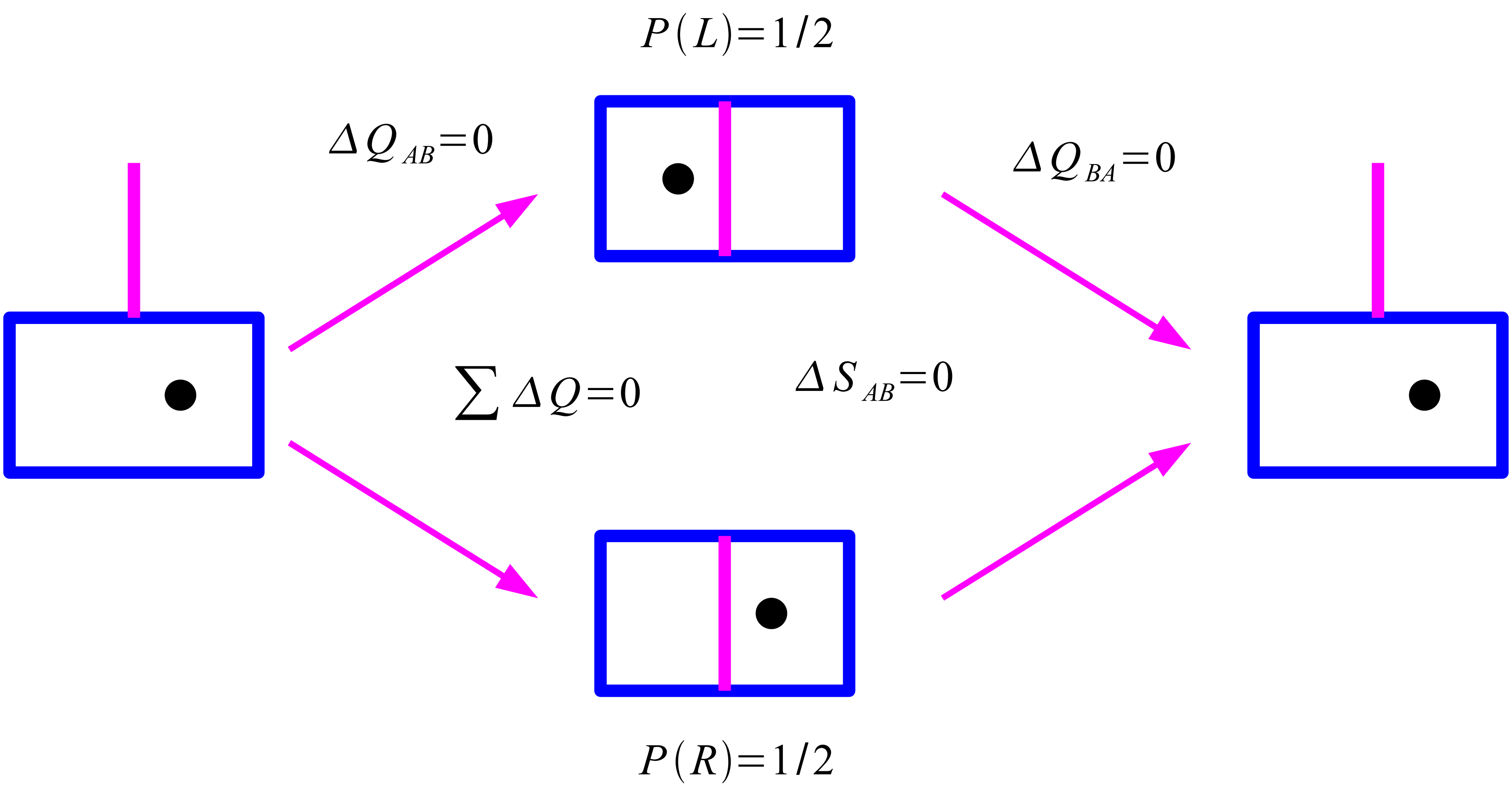
Cyclic paths



Landauer's "Principle"

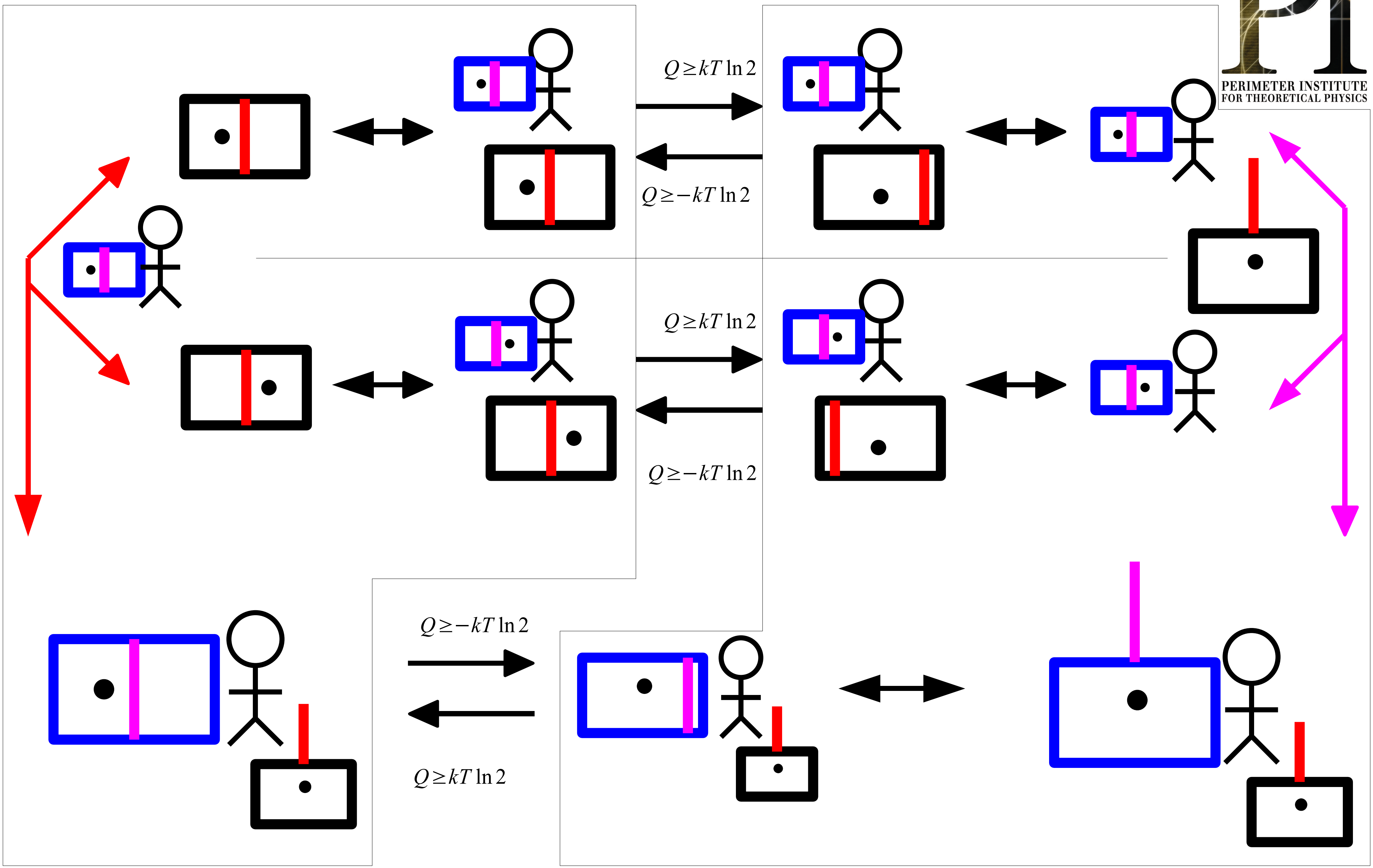
- It is the existence of the Szilard Engine process that guarantees (by the statistical mechanical entropy theorem) that one cannot do better than the Landauer Erasure process.
  - (Equivalently the Reuadnal Enigne path may guarantee this also)
- The combined cycle can have a net cost of zero.
  - So all processes involved, including Landauer Erasure, are statistically mechanically *reversible*

Cyclic paths



$$\Delta S = k \ln 2$$

PERIMETER INSTITUTE
FOR THEORETICAL PHYSICS



Maxwell through the Looking Glass:
From Szilard to Landauer and back again.

PIAF Sydney, February 2008

Conclusions

PERIMETER INSTITUTE FOR THEORETICAL PHYSICS

- The focus should not be upon particular paths or processes connecting pairs of states
  - This causes excessive attention to the details of a particular process (correlations, information, measurement, fluctuations).
  - Resulting explanations are tied to an understanding of a particular process and lack generality.
- The focus should be upon the existence of optimal cycles incorporating such states
  - But the explanation must not be tied to the specifics of a *particular* optimal cycle.
- Presupposed ideas of what statistical mechanical entropy is, ought to be, or is *for*, are not *necessarily* helpful
  - Objective, subjective, microscopic, macroscopic, index of irreversibility...
  - A statistical mechanical theorem may be derived that does not presuppose any notion of entropy. It may, if one so chooses, be used to *define* a statistical mechanical entropy.

Conclusions

- Recent criticisms seem misguided.
  - Statistical mechanics provides a theorem about the intraconvertibility of heat and work.
    - This produces the Gibbs-von Neuman entropy measure (but does *not* presuppose it)
  - Probability distributions over macroscopic states reduce the convertibility of heat into work.
    - A consequence of unitarity
    - True *regardless* of the origin or understanding of the "probability".
- Landauer's Principle is valid, but has been badly formulated.
  - It is not a *principle*, but it *is* a valid theorem.
  - Erasure requires heat generation, but is not necessarily irreversible.
- Information theory is neither necessary nor sufficient to understand the operation of the Szilard Engine.
- Maxwell's Demons do not exist! *A General Proof*!
  - Subject to four assumptions.
    - Which may be challenged! But are only sufficient, not necessary.

PERIMETER INSTITUTE FOR THEORETICAL PHYSICS