5 Paradoxes

> According to the second law of thermodynamics, disorder, once created, is almost impossible to destroy.[1]

5a Maxwell's demon

> The first thought experiment in which mention is made of information as a parameter with physical significance and linked with entropy is Maxwell's demon concept ... introduced ... in a letter ... on 1867.[2]

A question of evident practical and theoretical importance: is it possible to convert thermal energy into mechanical energy, having only one (hot) reservoir?

> Maxwell's demon ... passes "hot" molecules into one half of the cylinder, and "cold" molecules in the other half by the opening and closing of a microscopic door in the wall dividing a gas cylinder into two halves.[3]

Is it necessarily a supernatural demon able to violate physical laws? Or can it be material? We do not try to make it practical, just possible in principle.

The question becomes easier if we replace a 'serial' demon with a 'parallel' one, and a gas with a spin system. A demon measures all spins simultaneously, and sorts them: all (-1) spins to the left, all $(+1)$ spins to the right. Two nearly equal spin systems at (extremely) different temperatures appear,

[1] S. Lloyd and W.H. Zurek, "Algorithmic treatment of the spin-echo effect", Journal of Statistical Physics **62**:3/4, 819–839 (1991); see page 819.

[2] A. Moue, K. Masavetas, H. Karayianni, "Maxwell's demon: thermodynamics of information gaining and information processing", in: Proc. HERMCA-2005 (7th Hellenic European Conference on Computer Mathematics and its Applications); see page 1.

[3] Moue et al., page 1.

and so, mechanical energy can be extracted (as in Sect. 4e). No cold reservoir is needed. Really? No; there is a catch.

First, a digression. It may seem that at every instant we observe the whole world simultaneously. This illusion appears because the speed of light is very large (in our everyday units).[1]

Similarly, it may seem that our mental decisions are not thermodynamic. Of course, our brains consume energy and produce heat, but our decisions do not. This illusion appears because the Boltzmann constant is very small (in our everyday units).[2]

If the demon is material then its phase space must carry a measure, and its dynamics must be invertible and measure preserving, unless the demon interacts with the environment, in which case the joint dynamics must be invertible and measure preserving. Also the energy must be preserved.

It cannot happen that the system 'the demon and the spins' passes from an arbitrary spin configuration to a sorted one unless a trace remains inside the demon or outside it. Accordingly, something like a cold reservoir must exist inside or outside the demon.

Let us assume that the demon is discrete (like a spin system), deterministic (like a computer), and its energy is negligible.[3] It was believed ('the von Neumann–Landauer limit') that any computer must dissipate at least $k_{\mathrm{B}}T \cdot \ln 2$ of heat per any irreversible bit operation (such as 'and'), but fortunately this is not the case. The so-called reversible computing[4] is a way to design any computer in such a way that it does not dissipate energy.[5] Thus we may treat any one-to-one transformation of the demon's phase state as a feasible dynamics.[6] The same holds for the product phase state of the 'demon and spins' system and energy preserving one-to-one transformations. Moreover, a transformation not preserving energy can be used if the demon can transfer mechanical energy from/to its environment as needed. Thus, the demon need not sort the spins, it can just turn all the spins to (-1) and release their energy as work. The only problem is, to be one-to-one.

For every state $X = (x_1, \dots, x_n) \in \{-1, 1\}^n$ of the spin system there exists a one-to-one transformation $f_X : \{-1, 1\}^n \to \{-1, 1\}^n$ such that $f_X(X) = -\mathbb{1} = (-1, \dots, -1)$, the ground state (of least energy). Such a demon is an oracle: it knows the spin configuration beforehand! A material

[1] 299 792 458 m/s (exactly).

[2] $1.38 \cdot 10^{-23}$ J/K.

[3] The general case leads to similar conclusions.

[4] See 'Reversible computing' in Wikipedia.

[5] Or rather, the dissipation can be made arbitrarily small. The result of the computation leaves the computer, and the computer returns to its initial state.

[6] Computability is not an issue.

demon is assumed to have no such abilities.

It may happen that $f(X) = -\mathbb{1}$ just by chance, but such a lucky demon is of no use. With a small probability the given spin configuration is already sorted, and so, mechanical energy can be extracted, with no demon. Such miracles of exponentially small probability are of no interest.

Dealing with a system of n spins we may imagine a demon with a memory of n bits. What is the distinction between the spin system and the memory? Two distinctions: energy and relaxation. For the spin system, the energy is proportional to the sum of the spins; for the memory, the energy is zero, and relaxation does not happen. Here is a very simple model of relaxation for the spin system: just a random perturbation of the spins, applied every short time, independently. The spins are volatile (only their sum is constant); the memory is reliable.

Denoting the spins (at a given instant) by $X = (x_1, \dots, x_n) \in \{-1, 1\}^n$ and the demon's memory by $Y = (y_1, \dots, y_n) \in \{-1, 1\}^n$ we introduce for example the transformation

$$(X, Y) \mapsto (X, XY)\,, \tag{5a1}$$

that is, $y_k := x_k y_k$. It is one-to-one (think, why). Let initially $Y_0 = \mathbb{1}$, then after the transformation we get $Y_1 = X_0$. The demon measures the microscopic state of the spin system and remembers it. The spin system is not affected. However, it is quickly randomized by relaxation.

From demon's viewpoint, immediately after the measurement the microstate of the spins is known, no more random. Thus, its entropy is zero, and all its energy is mechanical. Work can be extracted, but this should be made quickly, before the relaxation. Well, demons can be very swift.

Here is even simpler one-to-one transformation:

$$(X, Y) \mapsto (-Y, X)\,.$$

Let initially $Y_0 = \mathbb{1}$, then after the transformation we get $X_1 = -\mathbb{1}$ and $Y_1 = X_0$; all the heat is converted into work, and the initial spin array is remembered by the demon. The entropy of the spin system is decreased dramatically. Not for nothing, however: doing so, the initially powerful demon becomes powerless! Some resource is spent.

Is the array $Y_0 = \mathbb{1}$ special? No, it is not. For every $y \in \{-1, 1\}^n$ we may take $f_y : (X, Y) \mapsto (-yY, X)$, then $f_y(X, y) = (-\mathbb{1}, X)$. Then, is the demon with Y_1 in its memory really powerless? Yes, it is. Not because Y_1 is itself worse than $\mathbb{1}$, but because Y_1 depends on X_0. In order to apply f_{Y_1} the demon must be an oracle.

We may imagine a programmable demon whose transformation is not hard-wired. But then a second demon, after reprogramming the first demon, becomes powerless.

If a cold reservoir is available, the demon can use it. Let $Z_2 = -\mathbb{1}$ be the state of a spin system at zero (that is, $+0$) temperature, and $Y_2 = Y_1$ the state of demon's memory (the obsolete information). The one-to-one transformation

$$(Z, Y) \mapsto (Y, -Z)$$

gives $Y_3 = \mathbb{1}$ and $Z_3 = Y_2$. The demon is powerful again; the obsolete information is dumped into the cold reservoir;[1] the latter is thus heated; some mechanical energy is taken by the demon and wasted into heat.

After all, the demonic cycle is similar to the thermodynamic cycle treated in Sect. 4d, 4e. Some thermal energy together with some entropy is received from the hot reservoir; the entropy together with a part of the energy is sent to the cold reservoir; the rest of the energy is converted into work.

Everyone knows the important role of observers in quantum theory: a measurement influences the object. The role of observers in statistical physics is less evident, but still important. Here, a measurement need not influence the object, but there is another problem: a measurement increases the entropy of the observer! A human observer gets no more than several bits of information; the corresponding entropy, several times k_B, is practically negligible. However, a demon observer can drain the object of all entropy.

For us humans the distinction between mechanical and thermal energy is important. Thus, we should not take demon's viewpoint. Even if a demon knows everything, still, we do not. For us the state of demon's memory is random, and has a (quite large) informational entropy.

The informational entropy of the combined system 'demon and spins' is invariant under all one-to-one transformations.[2] What about thermodynamic entropy? It is controversial, whether this notion is applicable to demon's memory, or not. Let us agree that it is not. Then we may say that a demon can convert the thermodynamic entropy of spins (or something else) to the informational entropy of itself, and conversely. Still, it cannot change the total entropy.[3]

---

[1]There, it will be quickly destroyed by relaxation, which is of no importance.

[2]Clearly, $\sum_{x,y} p(x,y) \ln p(x,y) = \sum_{x,y} p(f(x,y)) \ln p(f(x,y))$ for every one-to-one f. And more generally, if $f : \Omega \to \Omega$ is an invertible transformation preserving a measure μ, and ν is a probability measure on Ω absolutely continuous w.r.t. μ, then f sends ν into a measure ν_1 of the same differential entropy: $H_\mu(\nu_1) = H_\mu(\nu)$, since $\frac{\mathrm{d}\nu_1}{\mathrm{d}\mu}(f(\cdot)) = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}(\cdot)$. A non-invertible measure preserving transformation can increase entropy, which cannot happen when μ is a counting measure.

[3]Relaxation can increase the total entropy; this point is somewhat moot, like the very

Note that the informational entropy of the 'demon and spins' system is not at all the sum of entropies, unless the two are independent (recall 3e5). In contrast, thermodynamics treats entropy as additive, since relaxation makes (interacting) subsystems nearly independent.

Conversions between thermodynamic and informational entropy are treated by the so-called thermodynamics of computation.

5b Gibbs' paradox

> But the increase of entropy due to the mixing of given volumes of the gases at a given temperature and pressure would be independent of the degree of similarity or dissimilarity between them.[1]

> It has always been believed that the Gibbs' Paradox embodied profound thought. That it was intimately linked up with something so important and completely new could hardly have been foreseen.[2]

The importance of one-to-one transformations (Sect. 5a) implies importance of one-to-one correspondence between physical microstates and the mathematical objects that describe them. Till now, Ω^n was used as the phase space of an n-particle system. This is correct for a system of n pairwise distinguishable ('numbered') particles. Anyway, each function of the form $f^{(n)}$ ('macroscopic observable'), being invariant under the permutation group S_n (that acts on Ω^n by measure preserving transformations), may be treated as a function on the quotient space $\Omega^n_{\text{sym}} = \Omega^n/S_n$ of orbits. The quotient measure μ^n_{sym} on Ω^n_{sym} makes the natural projection $\Omega^n \to \Omega^n_{\text{sym}}$ measure preserving. The distribution of $f^{(n)}$ on (Ω^n, μ^n) is equal to the distribution of the corresponding function on $(\Omega^n_{\text{sym}}, \mu^n_{\text{sym}})$. This is why till now we were able to use (Ω^n, μ^n) successfully. However, troubles will appear soon.

When μ is nonatomic, Ω^n_{sym} may be treated as the set of all n-point subsets of Ω. Or equivalently, of measures on Ω consisting of n atoms of mass 1 each. When μ is atomic, elements of Ω^n_{sym} cannot be treated as subsets of Ω, but still can be treated as measures on Ω consisting of atoms of integral masses, with the total mass n.

As usual, a discrete model is easier to understand. Thus, imagine first a container V of just two points, and n particles in V (many particles may

idea of relaxation.

[1] W. Gibbs, 'On the equilibrium of heterogeneous substances', 1878.

[2] E. Schrödinger, 'Statistical thermodynamics', Cambridge, 1948.

occupy the same place). Or equivalently we may imagine n spins. If the particles are distinguishable then we have $|\Omega^n| = 2^n$ possible states. A demon can use this system (if at zero temperature) for dumping n bits of obsolete information. Now imagine that the particles are indistinguishable; then we have only $|\Omega^n_{\text{sym}}| = n + 1$ possible states;[1] a demon cannot dump n bits, nor even $n/100$ bits to this system. Indistinguishability has dramatic thermodynamic consequences!

According to Sections 3 and 4, the ideal gas (with zero potential) has the entropy $n(\ln V - \frac{3}{2} \ln \beta + \text{const})$.[2] Two independent portions of the gas, each of n particles, in disjoint containers of volume V each, at equal temperatures, have the total entropy $2n(\ln V - \frac{3}{2} \ln \beta + \text{const})$. However, a single portion of $2n$ particles in a volume $2V$ has another entropy $2n(\ln 2V - \frac{3}{2} \ln \beta + \text{const})$ larger by $2n \ln 2$. Why? What is the distinction between "n in V plus n in V" and "$2n$ in $2V$"?

The momentum subsystem, contributing $-\frac{3}{2} \ln \beta + \text{const}$ (per particle), is not guilty. The coordinate subsystem, contributing $\ln V$ (per particle), is guilty. Thus we concentrate on coordinates.

We may embed $\mathcal{V}_1^n \times \mathcal{V}_2^n$ into $(\mathcal{V}_1 \cup \mathcal{V}_2)^{2n}$ by $\big((q'_1, \dots, q'_n), (q''_1, \dots, q''_n)\big) \mapsto (q'_1, \dots, q'_n, q''_1, \dots, q''_n)$. The image is less than $(\mathcal{V}_1 \cup \mathcal{V}_2)^{2n}$ for two reasons. First, it contains only states with exactly n particles in $\mathcal{V}_1$. This is an event of probability $2^{-2n}\binom{2n}{n} \sim 1/\sqrt{\pi n} = \exp o(n)$. Second, it contains only states such that the n particles situated in $\mathcal{V}_1$ have the numbers $1, \dots, n$. Here we have an event of probability $1/\binom{2n}{n} \sim \sqrt{\pi n}\, 2^{-2n} = \exp\big(-2n \ln 2 + o(n)\big)$ responsible for the extra summand $2n \ln 2$ in the entropy.

In other words: given a microstate of $2n$ numbered particles in $\mathcal{V} = \mathcal{V}_1 \uplus \mathcal{V}_2$, we may consider $2n$ bits of information; the k-th bit shows whether the k-th particle is in $\mathcal{V}_1$ or $\mathcal{V}_2$. This bit array is quickly randomized by relaxation. Let us insert a thin partition between $\mathcal{V}_1$ and $\mathcal{V}_2$ (that is, increase the potential near the boundary between them to a high value). Now the bit array is frozen, no more subject to relaxation. The seamingly innocent partition is in some sense demonic! It effectively creates a memory[3] of $2n$ bits, measures some aspect of the microstate of the gas, and stores the result in the memory. Thus, it converts the entropy $2n \ln 2$ from thermodynamic to informational.

You may wonder, how does the partition influence 'practical' thermodynamics of the gas, related to heat engines, pumps etc. The answer is, no influence. What matters is the derivative of the entropy in the temperature.

---

[1] The state $k \in \{0, 1, \dots, n\}$ is of μ^n_{sym}-measure $\binom{n}{k}$.

[2] In fact, const $= \ln(2\pi e m)$.

[3] Read-only memory, if you like.

The extra summand $2n \ln 2$ is rather harmless, since it does not depend on the temperature.

Can particles be distinguishable? Well, a not-so-large organic molecule can easily carry, say, 100 bits of constant information (like a short DNA segment), which is more than enough for having 10^{23} pairwise distinct molecules. However, atoms and small molecules are indistinguishable[1] according to quantum theory. Thus, statistical physics usually deals with indistinguishable particles.

A coordinate microstate of n indistinguishable particles is a point of $\mathcal{V}^n_{\text{sym}}$; the relevant measure on $\mathcal{V}^n_{\text{sym}}$ is not μ^n_{sym} but $(1/n!)\mu^n_{\text{sym}}$ (which is again a classical approximation to something quantal; however, you can get the same idea very simply by discretizing $\mathcal{V}$). Thus, the entropy is not $n \ln V = \ln V^n$ but

$$\ln(V^n/n!) = n \ln V - n \ln n + n + o(n) = n\Big(\ln \frac{V}{n} + 1 \Big) + o(n)\,.$$

For $2n$ particles in the volume $2V$ we get now

$$2n\Big(\ln \frac{2V}{2n} + 1 \Big) + o(n)\,;$$

dividing the container in two (and putting n particles into each part) we get

$$2 \cdot n\Big(\ln \frac{V}{n} + 1 \Big) + o(n)\,;$$

just the same! The partition is no more demonic.

Assume now that $\mathcal{V}_1$ contains n particles of one type, and $\mathcal{V}_2$ contains n particles of another type (say, atoms of argon and krypton). When the partition is removed, the gases mix, and the entropy increases by $2n \ln 2$, from $2 \cdot n\big(\ln \frac{V}{n} + 1\big)$ to $2 \cdot n\big(\ln \frac{2V}{n} + 1\big)$ (ignoring $o(n)$), since the phase state changes from $(\mathcal{V}_1)^n_{\text{sym}} \times (\mathcal{V}_2)^n_{\text{sym}}$ to $\mathcal{V}^n_{\text{sym}} \times \mathcal{V}^n_{\text{sym}}$, $\mathcal{V} = \mathcal{V}_1 \uplus \mathcal{V}_2$.

The extra entropy is equal to 0 if the gases are identical, and $2n \ln 2$ if the gases are different, no matter how much (and in what aspect) different! This is Gibbs' paradox.

Can we use this effect as a universal, infinitely sensitive distinction detector? No, we cannot, since we have no universal device for measuring entropy. (Recall, the extra summand not dependent on the temperature does not manifest itself in thermodynamic cycles.)

The quantum theory reveals something new about the notions 'identical' and 'different'. Any classical distinction means orthogonal state vectors and

[1] Or have a small choice of states.

can be amplified without bound by a measuring device. But a very small distinction is no more classical; it means a nontrivial angle between state vectors, and cannot be amplified (not even a little).[1] It leads to an extra summand in entropy that is less than $2n \ln 2$ and depends continuously on the distinction (the angle).[2]

Still, some questions remain. Can we check whether two given particles are of the same type or not? Several quantum numbers, such as 'flavour', 'strangeness', 'charm' are well-known. Maybe another one, call it 'stupidity', is still unknown. Then the entropy of a gas may differ substantially from what we believe it is. But on the other hand, should we add the extra entropy even if 'stupidity' plays absolutely no role in interactions of these particles with each other or anything else?

Imagine that 'stupidity' takes on two values. Then it is another array of bits. If it is subject to relaxation then its entropy is thermodynamic, otherwise informational. But anyway, if 'stupidity' interacts with other degrees of freedom then it belongs to what we call the gas; otherwise it is rather a closed physical system disconnected from the gas and our apparata, and so, its entropy is irrelevant.

It may happen that 'stupidity' interacts with other degrees of freedom, but slowly. Then it is effectively a separate system on short times, but a part of the gas on long times. Compare it to the (possibly negative) spin temperature mentioned in Sect. 4c.

Amazingly, quantum theory proposes a universal way to check whether two given particles are of the same type or not; this test reveals 'stupidity' even if we have no idea of it.[3], [4] However, this argument is hardly relevant to entropy.

---

[1] "Some authors have asserted, without feeling further comment to be necessary or useful, that two particles are either the same or they are different. Others, following von Neumann (1955) and Schrödinger (1950, 1952), have used the overlap of wavefunctions representing two quantum states as a *continuous* measure of indistinguishability." A.M. Lesk, "On the Gibbs paradox: what does indistinguishability really mean?", J. Phys. A: Math. Gen. **13** (1980), L111–L114; see pp. L111–L112.

[2] "He showed explicitly that the entropy change varies *continuously* from 0 to $2R \ln 2$ as the overlap decreases from unity to 0 (Klein 1948, 1959)." Lesk, p. L112.

[3] See "Hanbury Brown and Twiss effect" in Wikipedia.

[4] There is an element of convention in regarding particles as distinguishable or indistinguishable". Lesk, page L112.

5c Spin echo

> In the spin-echo effect, the disorder of the spins first increases, then decreases dramatically.[1]

A macroscopic portion of paraffin, solid at room temperature, contains (among other microscopic degrees of freedom) nuclear spins. Initially most spins are (-1), which manifests itself by a macroscopic magnetic field. The field decays after $\sim 10^{-5}$ sec because the spins interact with the environment. After 0.001 sec an apparatus influences the spins by a carefully adjusted radio frequency pulse. Still, no magnetic field. But after another 0.001 sec, amazingly, the magnetic field reappears suddenly, lasts for $\sim 10^{-5}$ sec and decays again. This phenomenon is called the spin echo.[2]

Spin echo seems to refute irreversibility of relaxation and increase of entropy. However, a closer look reveals the following.

In a liquid paraffin, molecules walk at random (which is called self-diffusion). If the paraffin cools and solidifies, a random configuration of molecules freezes. Molecules oscillate but do not walk.[3] Their configuration is no more subject to relaxation. Thus, a part of the thermodynamic entropy of the liquid paraffin turns into informational entropy of the solid paraffin.[4]

On the first 0.001 sec of the experiment, the spins measure some of the frozen degrees of freedom,[5] which is similar to (5a1):

$$(X,Y) \mapsto (X,XY)\,; \quad (X,1\!\!1) \mapsto (X,X)\,,$$

where X is the bit array of the frozen information (read-only memory, if you like), and Y is the array of nuclear spins; we may treat Y as a (read-write) memory, since its relaxation time is ~ 0.01 sec. The informational entropy of X is $n \ln 2$; the informational entropy of (X,X) is still $n \ln 2$, not $2n \ln 2$.

On the last 0.001 sec of the experiment the measurement is undone:

$$(X,Y) \mapsto (X,XY)\,; \quad (X,X) \mapsto (X,1\!\!1)\,.$$

The informational entropy is $n \ln 2$, still. The total entropy never decreases, if treated appropriately.

---

[1] Lloyd and Zurek, page 838.

[2] This description is rather a caricature. In fact, the spins are quantal; their state is a point of a three-dimensional ball (rather than $[-1,1]$); the apparatus generates a static magnetic field and more than one radio frequency pulse; etc. See "Spin echo" in Wikipedia, and Sect. III(H) in the famous article: E.L. Hahn, "Spin echoes", Phys. Rev. **80**:4 (1950), 580–594.

[3] Sometimes they jump, but rarely.

[4] So-called frozen (or quenched) disorder; see "Quenched disorder" in Wikipedia.

[5] Namely, magnetic field at the nuclei, due to neighbour molecules.

Index