The demons haunting thermodynamics

Katie Robertson

ARTICLES YOU MAY BE INTERESTED IN

(); https://doi.org/10.1063/PT.6.3.20220113a

The three physicists
Physics Today **74**, 42 (2021); https://doi.org/10.1063/PT.3.4700

Einstein would be doubly amazed
Physics Today **75**, 46 (2022); https://doi.org/10.1063/PT.3.5063

How a fake Kepler portrait became iconic
Physics Today **74**, 10 (2021); https://doi.org/10.1063/PT.3.4825

Ten billion years ago, galaxies were already running out of gas
Physics Today **74**, 20 (2021); https://doi.org/10.1063/PT.3.4897

(); https://doi.org/10.1063/PT.6.1.20211005a
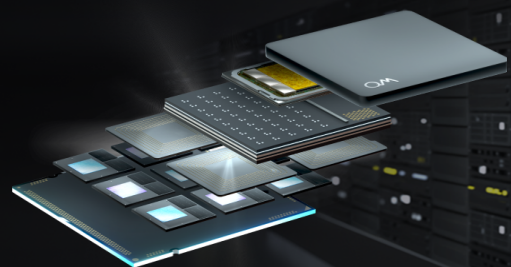
THE DEMONS HAUNTING

Katie Robertson

The contradictory implications of statistical mechanics have worried physicists and philosophers for centuries. Does our present-day knowledge allow us to exorcize those philosophical demons?

Thermodynamics is a strange theory. Although it is fundamental to our understanding of the world, it differs dramatically from other physical theories. For that reason, it has been termed the "village witch" of physics.[1] Some of the many oddities of thermodynamics are the bizarre philosophical implications of classical statistical mechanics. Well before relativity theory and quantum mechanics brought the paradoxes of modern physics into the public eye, Ludwig Boltzmann, James Clerk Maxwell, and other pioneers of statistical mechanics wrestled with several thought experiments, or demons, that threatened to undermine thermodynamics.

Despite valiant efforts, Maxwell and Boltzmann were unable to completely vanquish the demons besetting the village witch of physics—largely because they were limited to the classical perspective. Today, experimental and theoretical developments in quantum foundations have granted present-day researchers and philosophers greater insights into thermodynamics and statistical mechanics. They allow us to perform a "quantum exorcism" on the demons haunting thermodynamics and banish them once and for all.

Loschmidt's demon and time reversibility

Boltzmann, a founder of statistical mechanics and thermodynamics, was fascinated by one of the lat- ter field's seeming paradoxes: How does the irreversible behavior demonstrated by a system reaching thermodynamic equilibrium, such as a cup of coffee cooling down or a gas spreading out, arise from the underlying time-reversible classical mechanics?[2] That equilibrating behavior only happens in one direction of time: If you watch a video of a wine glass smashing, you know immediately whether the video was in rewind or not. In contrast, the underlying classical or quantum mechanics are time reversible: If you were to see a video of lots of billiard balls colliding, you wouldn't necessarily know whether the video was in rewind or not. Throughout his career, Boltzmann pursued a range of strategies to explain irreversible equilibrating behavior from the underlying reversible dynamics.

Boltzmann's friend Josef Loschmidt famously objected to those attempts. He argued that the underlying classical mechanics allow for the possibility that the momenta are reversed, which would
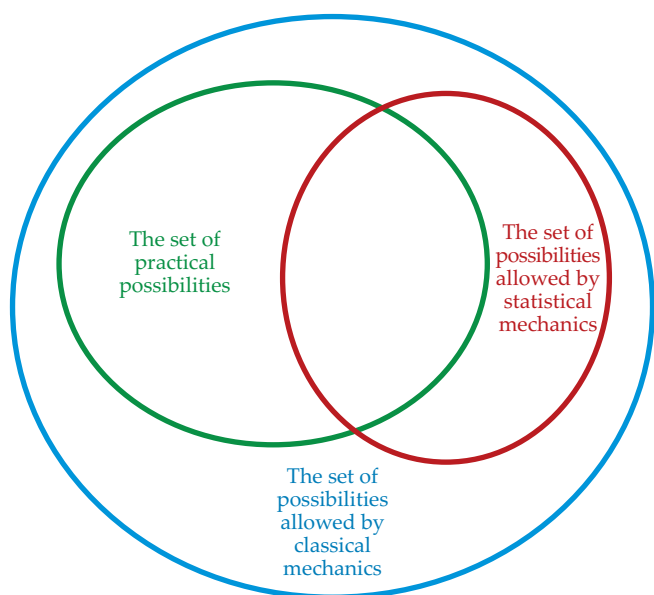
Katie Robertson is a Leverhulme Early Career Fellow at the University of Birmingham in the UK. Her research focuses on the philosophical implications of thermodynamics.

THERMODYNAMICS
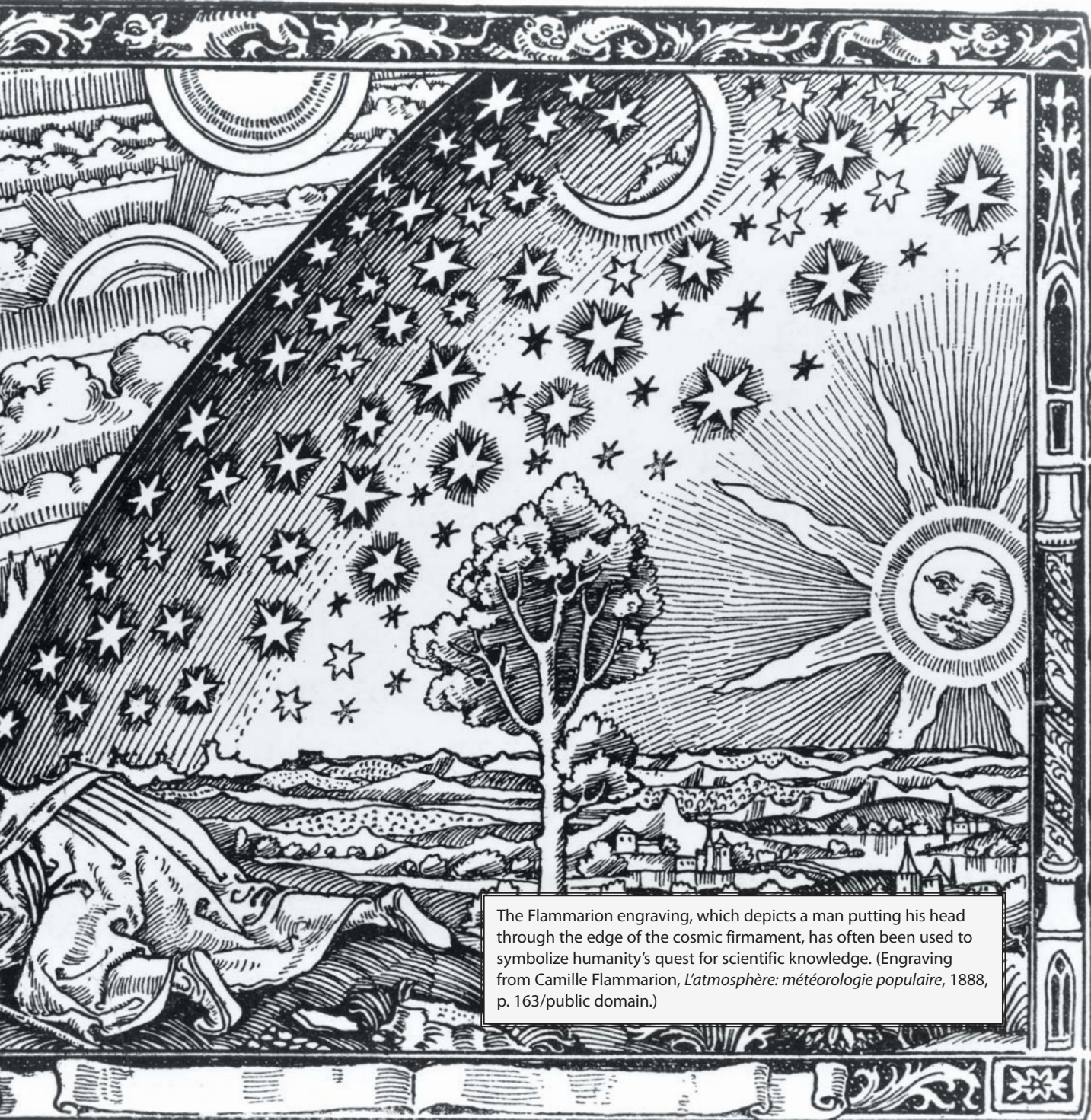
The Flammarion engraving, which depicts a man putting his head through the edge of the cosmic firmament, has often been used to symbolize humanity's quest for scientific knowledge. (Engraving from Camille Flammarion, *L'atmosphère: météorologie populaire*, 1888, p. 163/public domain.)

FIGURE 1. THE VARYING SETS of possibilities allowed by various physical theories can be depicted by intersecting ovals. The set of possibilities allowed by classical mechanics (the blue oval) contains both the set of possibilities allowed by statistical mechanics (the red oval) and the set of possibilities that can be practically realized in real life (the green oval). Loschmidt's demon is possible under the classical regime because that paradigm allows for the possibility of a system's momenta being reversed. Statistical mechanics, on the other hand, says that systems can't decrease their entropy on average, which rules out any demonic activities. Although Ludwig Boltzmann believed that a demon that reverses the momenta is practically impossible, the spin-echo experiment has proven that it can be done in rare cases. Loschmidt's demon is thus inside the blue and green ovals but not the red oval.

lead to the gas retracing its steps and "anti-equilibrating" to the earlier, lower-entropy state. Boltzmann challenged Loschmidt to try to reverse the momenta, but Loschmidt was unable to do so. Nevertheless, we can envision a demon that could. After all, it is just a matter of practical impossibility—not physical impossibility—that we can't reach into a box of gas and reverse each molecule's trajectory.

Technological developments since Loschmidt's death in 1895 have expanded the horizons of what is practically possible (see figure 1). Although it seemed impossible during his lifetime, Loschmidt's vision of reversing the momenta was realized by Erwin Hahn in 1950 in the spin-echo experiment, in which atomic spins that have dephased and become disordered are taken back to their earlier state by an RF pulse. If it is practically possible to reverse the momenta, what does that imply about equilibration? Is Loschmidt's demon triumphant?

Unlike the other two demons we will encounter, we can make our peace with Loschmidt's. It turns out that the spin-echo experiment is a special case; most systems approach equilibrium instead of retracing their steps back to nonequilibrium states. But Loschmidt's demon vividly reminds us that the underlying laws of mechanics allow for a system to retrace its steps. Why don't we see that possibility? Why doesn't a gas compress back into a smaller volume? Why don't eggs unsmash or cups of coffee spontaneously warm up?

The answer lies in the distinction between laws and initial conditions. Consider a stone thrown into a pond. The initial condition—the stone hitting the pond—explains why we see ripples going outward. In contrast, we never see ripples converge inward and propel a stone from the pond's depths because the required initial condition would be fiendishly difficult to set up. Similarly, the initial conditions typical in systems involving gases explain why they approach equilibrium. But special initial conditions with finely tuned correlations could lead to instances of anti-equilibration, such as your coffee spontaneously getting hotter or stones being propelled out of ponds. In other words, anti-equilibration is possible according to the microdynamic laws of physics, but only if systems have highly atypical initial conditions.

Maxwell's nimble-fingered demon

By far the most famous hypothetical demon in physics is the one conjured up by Maxwell in 1867 (see figure 2). He envisioned a being that observes individual molecules in a gas-filled box with a partition in the middle. If the demon sees a fast-moving gas molecule, it opens a trapdoor in the partition that allows fast-moving molecules through while leaving slow-moving ones behind. Repeatedly doing that would allow the
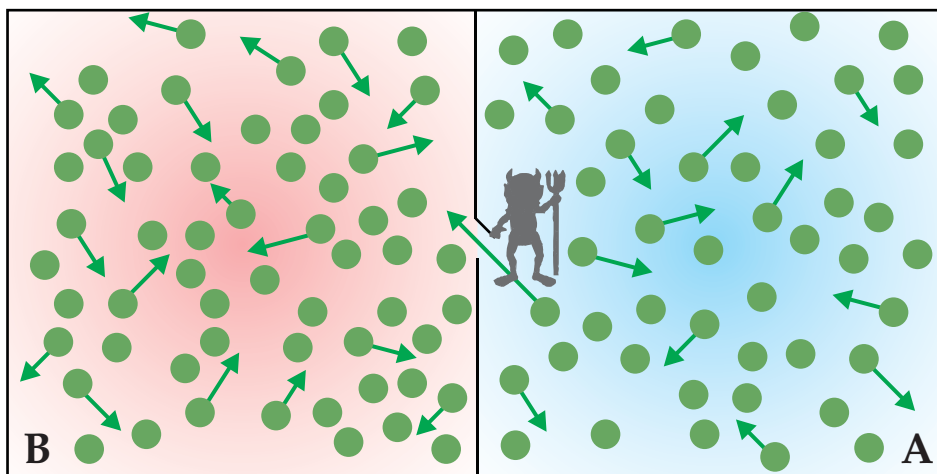

FIGURE 2. MAXWELL'S DEMON is a hypothetical being that can observe individual molecules in a gas-filled box with a partition in the middle separating chambers A and B. If the demon sees a fast-moving gas molecule, it opens a trapdoor in the partition to let fast-moving molecules into chamber B while leaving slow-moving ones behind. Repeating that action would allow the buildup of a temperature difference between the two sides of the partition. A heat engine could use that temperature difference to perform work, which would contradict the second law of thermodynamics.

buildup of a temperature difference between the two sides of the partition. A heat engine could use that temperature difference to perform work, which would contradict the second law of thermodynamics.

Is Maxwell's demon in the same category as Loschmidt's demon—namely, a mere matter of practical difficulty rather than physical impossibility? Maxwell thought so. According to philosopher of physics Wayne Myrvold, Maxwell believed that "only our current, but perhaps temporary, inability to manipulate molecules individually . . . prevents us from doing what the demon would be able to do."[3]

When Maxwell was writing over 150 years ago, the possibility of manipulating individual molecules might have seemed far-fetched, but that's no longer the case today. From purpose-built experimental apparatuses to molecular machines found in nature, devices similar to Maxwell's demon abound. For example, biomolecular machines use ratchet-style mechanisms[4] akin to the version of Maxwell's demon devised by Richard Feynman in a 1962 lecture.

Moreover, researchers have seemingly been able to realize Maxwell's demon in experiments. One group in Tokyo led by Masaki Sano devised a demon-style experiment in 2010 (see the article by Eric Lutz and Sergio Ciliberto, PHYSICS TODAY, September 2015, page 30). Using a tilted optical lattice to manipulate a particle, Sano's team created a "spiral staircase" that, on average, the particle tended to descend. By using a CCD camera, the experimenters monitored the fluctuations in the particle's position in real time. When the particle fluctuated upward, they altered the voltage and trapped it in a higher position, much like the demon shutting the trapdoor (see figure 3). By repeating that process, Sano was able to gradually move the particle upward and do work.

Are such ingenious devices genuine Maxwellian demons? Do they invalidate the second law? Although their mechanisms seem to be demonic, some careful entropic accounting is in order. A process contravenes the second law only if the entropy of the total system decreases. In a familiar example, the entropy of an ideal gas decreases during isothermal compression, but the compensating increase in the heat bath means that the system's total entropy increases. Is there a compensating entropic increase in the environment that thwarts any attempt to violate the second law?

That question has been vigorously debated since Maxwell's demon was first posited.[5] Although some philosophers of physics disagree,[6] many physicists now believe that there is an entropy cost associated with the demon's activities. Because those ingenious devices lead to entropy increases else-where in the greater system, none of them truly violate the second law. The entropic costs stem from the demon's operation. To run, it must perform a feedback operation: If the molecule is fast moving, the demon opens the door, but if the molecule is slow, the demon closes the door.

That requires the demon to have a memory, which must be reset at the end of the cyclic process. But resetting the memory has an entropic cost, which can be quantified by a principle proposed by Rolf Landauer in 1961. It states that entropy increases by $k_B \ln 2$ per bit of information that is reset, where k_B is Boltzmann's constant. In other words, erasing information will cost you. Landauer's principle thus forges a connection between thermodynamics and information theory—although the precise nature of their relationship remains controversial.

Nonetheless, to my eyes, Landauer's principle explains why no matter how ingenious or nimble fingered today's experimenters may be, they can't build engines that reliably violate the second law of thermodynamics and solve the global energy crisis. Once we peek behind the scenes and account for the environment, we see that today's alleged Maxwellian demons are deft illusionists rather than true magicians.

Much of the activity in contemporary thermal physics arises from the melding of quantum information theory with thermodynamics. Can going quantum release the demon from the shackles imposed by Landauer's principle? Sadly, it cannot. The principle holds for all forms of dynamics that preserve phase-space volume, and both classical and quantum mechanics fulfill that criterion. Moreover, there may even be extra entropic costs associated with quantum operations: The Landauer limit cannot be reached by quantum computation.[7]

Quantum steampunk

Maxwell's philosophical speculations about the nature of thermodynamics and statistical mechanics extended beyond his demon. To reconcile those probabilistic theories with his classically trained worldview, Maxwell made two philosophical claims: First, thermodynamics applies only to systems with many degrees of freedom, and second,
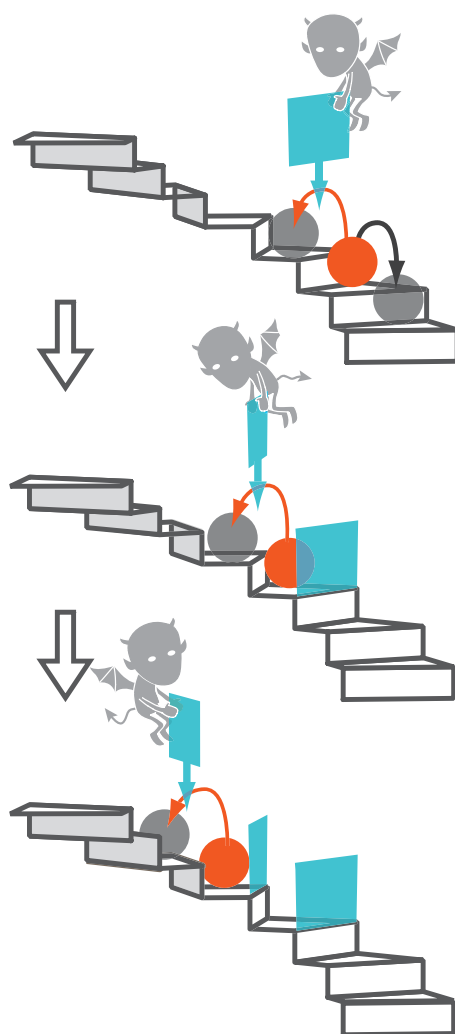
FIGURE 3. A DEMON-STYLE EXPERIMENT was devised by Masaki Sano's group in 2010. It used a tilted optical lattice to create a "spiral staircase" that the particle tended to descend over time. The experimenters monitored the fluctuations in the particle's position in real time, and when the particle fluctuated upward, they altered the voltage and trapped it in a higher position. As illustrated here, the experiment can be analogized to Maxwell's demon shutting the trapdoor. (Adapted from S. Toyabe et al., *Nat. Phys.* **6**, 988, 2010.)

it is anthropocentric and contingent on our human viewpoint. Do those philosophical postulates hold up today?

Experimental and theoretical developments in thermodynamics since the mid 20th century have demonstrated that Maxwell's first claim was incorrect. In Maxwell's time, thermodynamics was characterized by the steam engines that powered the Industrial Revolution, but today the thermodynamic revolution—the subfield that Nicole Yunger Halpern has termed "quantum steampunk"—is on the atomic scale.[8] Quantum heat engines, for example, were first proposed in 1959 by Derrick Scovil and Erich Schulz-DuBois, who demonstrated how a three-level maser could function as a heat engine. With the advent of quantum information theory, those tiny thermodynamic systems now provide fodder for an entire subfield.[9] Other types of quantum thermal machines use microscopic systems such as multilevel atoms, qubits, and quantum dots as the working substance in a heat engine.

How do quantum and classical heat engines differ? Additional resources are available in the quantum regime: Entanglement and coherence can be used as "fuel." Still, no one has found a way to cheat the second law.[10] Perhaps that is to be expected. After all, Seth Lloyd held that "nothing in life is certain except death, taxes, and the second law of thermodynamics."

Nonetheless, the two types of engines differ in fascinating ways. In traditional thermodynamics, maximum Carnot efficiency is achieved only when processes are carried out quasistatically, or infinitely slowly, which means that the power generated tends toward zero. That constraint led to the development of finite-time thermodynamics, and within that framework there are other limits to efficiency. Quantum machines can be more efficient than their classical counterparts in that finite-time regime—but both remain bounded by the Carnot limit.[11]

If thermodynamics is not limited to macroscopically large systems, is it universal? Many physicists believe that it is. Albert Einstein once said that "it is the only physical theory of universal content concerning which I am convinced that, within the framework of applicability of its basic concepts, it will never be overthrown."[12] Nowadays thermodynamics is used to understand topics as varied as quantum thermal engines, globular clusters of stars, black holes, bacterial colonies, and—more controversially—the brain.[13]

Is thermodynamics anthropocentric?

What about Maxwell's second philosophical claim, that thermodynamics is a feature of our perspective on reality? As he wrote in an 1877 *Encyclopædia Britannica* article, the distinction between ordered and disordered motion that is fundamental to thermodynamics "is not a property of material things in

FIGURE 4. LAPLACE'S DEMON is a fictional entity endowed with the knowledge of the position and momentum of every molecule in the universe. (*Laplace's Demon*, by Ele Willoughby, linocut, 2011.)

themselves, but only in relation to the mind which perceives them."[14] Maxwell's viewpoint has proven influential over the years. Percy Bridgman, for example, echoed Maxwell when he asserted in 1941 that "thermodynamics smells more of its human origin than other branches of physics—the manipulator is usually present in the argument."[15]

Why is that? Consider the honeybee as an example. The insect sees a garden very differently than we do because its eyes are sensitive to a different part of the electromagnetic spectrum than ours are. The claim that thermodynamics is anthropocentric, or observer dependent, implies that thermodynamic features such as entropy might look different—or not exist at all—if we were a different type of creature. In that view, thermodynamics would be analogous to a pair of rose-tinted glasses through which we understand and perceive the world but do not see how it actually looks.

In that way, Maxwell's ideas tie thermodynamics to beings like us. Because quantum mechanics has made many people comfortable with the observer being seemingly ineliminable from physics, that might not seem unusual. But Maxwell wasn't appealing to some generic observer like the honeybee.

He believed that thermodynamics is specifically anthropocentric. As he wrote in the same *Encyclopædia Britannica* article, "It is only to a being in the intermediate stage, who can lay hold of some forms of energy while others elude his grasp, that energy appears to be passing inevitably from the available to the dissipated state."[16] Understanding that anthropocentrism is fraught with challenges. For example, it seems undeniable that cups of coffee cool down regardless of what we know about them or our perspective on reality.

How worrying we find the possibility of anthropocentrism is determined by nothing less than our stance on the scientific enterprise itself. Are scientists actually learning about the deep nature of reality in a manner independent of our perspective? Or is science just a mere tool or instrument that we should use to "shut up and calculate"? The debate over scientific realism, as it is termed, has raged unresolved for centuries. But recent developments in quantum thermodynamics offer some hope for those who would like to rid thermodynamics of its human smell.

Leaving ignorance out of it

In classical statistical mechanics, the key postulate—often called the fundamental assumption—is that each accessible microstate of a system must be equally likely. But how should we understand probabilities in statistical mechanics? That question has received considerable attention over the years from trailblazers such as Boltzmann, Paul Ehrenfest, and Tatiana Ehrenfest-Afanasyeva. Here we will narrow our attention to one dominant view, popularized by physicist Edwin Jaynes, which argues that the fundamental assumption of statistical mechanics stems from our ignorance of the microscopic details. Because the Jaynesian view emphasizes our own ignorance, it implicitly reinforces the idea that thermal physics is anthropocentric. We must assume each state is equally likely because we don't know which exact microstate the system is in.

Here we are confronted by our third and final philosophical specter: the demon first articulated by Pierre Simon Laplace in 1814 (see figure 4). Laplace's demon is a hypothetical observer that knows the position and momentum of every molecule in the universe. In other words, it knows the exact microstate of every system in the universe.

In statistical mechanics, the entropy of a system is commonly expressed by the Gibbs formula, $S_G = \int \rho \ln \rho \, d^N\mathbf{q} \, d^N\mathbf{p}$, where $\rho(\mathbf{q}, \mathbf{p})$ represents a probability distribution, such as the microcanonical distribution, over the phase space of positions and momenta, $\{\mathbf{q}_1, \ldots, \mathbf{q}_N; \mathbf{p}_1, \ldots, \mathbf{p}_N\}$, that the system of N particles could occupy. But for Laplace's demon, $\rho = 1$, because it knows the system's exact microstate with certainty. That omniscience means the demon would calculate the system's Gibbs entropy to be zero! The Jaynesian view of statistical mechanical probabilities thus has a radical consequence: It means that the value assigned to the Gibbs entropy depends on our knowledge of the world.

Does Laplace's demon threaten the Jaynesian view of statistical mechanics? Not quite. Fortunately, it, too, can be exorcised by shifting to a quantum perspective on statistical mechanics. In classical statistical mechanics, probabilities are an additional ingredient added to the system's microdynamics. According to the Jaynesian view, they are a necessary step because of our ignorance. But in the quantum case, probabilities are already an inherent part of the theory, so there is no need to add ignorance to the picture. In other words, the probabilities from statistical mechanics and quantum mechanics turn out to be one and the same.

But in quantum mechanics, the Born rule implies that a quantum state encodes probabilities of different measurement outcomes. How can those probabilities give rise to the familiar probability distributions from statistical mechanics? That question is especially tricky because quantum mechanics assigns an isolated system a definite state known as a pure state. In contrast, statistical mechanics assigns such a system an inherently uncertain state known as a maximally mixed state, in which each possibility is equally likely. On the face of it, statistical mechanics and quantum mechanics appear to clash.

The distinctively quantum nature of entanglement holds the key to resolving that seeming conflict[17] (see figure 5). Consider a qubit that is entangled with a surrounding heat bath.
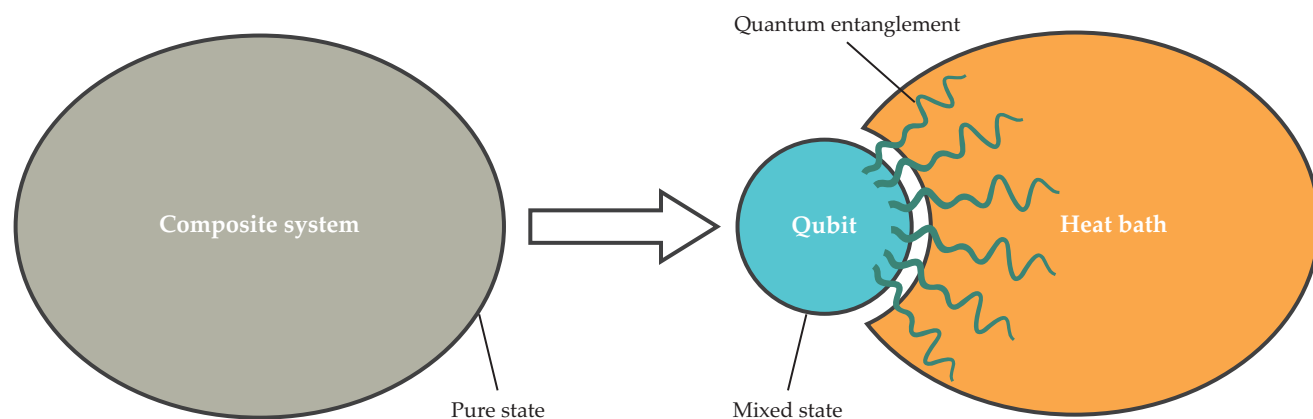

FIGURE 5. QUANTUM ENTANGLEMENT explains why Laplace's demon is vanquished. Consider a qubit and a heat bath that are entangled (at right). If the qubit is taken on its own, it will be in a mixed state, as will the heat bath. But the composite system of the qubit and the heat bath (at left) is in a pure state because when taken as a whole, it is isolated. Assuming the environment is sufficiently large, for almost any pure state that the composite system is in, the qubit will be in a state very close to the state it would be assigned by traditional statistical mechanics. Thus the system acts as if the fundamental assumption of statistical mechanics were true. The probability distribution assigned by statistical mechanics is indistinguishable from the quantum state, which implies that statistical mechanics does not require Jaynes's introduction of "ignorance." Laplace's demon is therefore defanged. (Adapted from S. Deffner, *Nat. Phys.* **11**, 383, 2015.)

Because they are entangled, if one of the two systems is taken on its own, it will be in an intrinsically uncertain state known as a mixed state. Nonetheless, the composite system of the qubit taken together with the heat bath is in a pure state because when taken as a whole, it is isolated. Assuming that the surrounding environment—namely, the heat bath—is sufficiently large, then for almost any pure state that the composite system is in, the qubit will be in a state very, very close to the state it would be assigned by traditional statistical mechanics.

In other words, the system under study—the qubit—behaves as if the composite system were in a maximally mixed state, namely, as if each microstate of the composite system is equally likely. The nature of the probabilities is ultimately quantum, but the system acts as if the fundamental assumption of statistical mechanics were true. The quantum description thus leads to a probability distribution indistinguishable from that of statistical mechanics.

How does that conclusion vanquish Laplace's demon? Quantum mechanics assigns probabilities to events not because we don't know their exact value but because both we and the demon cannot know that value. Probabilities are an intrinsic and inescapable part of quantum mechanics. When it describes the entangled system taken on its own, Laplace's demon cannot know any more than us.

Arthur Eddington proclaimed in 1928 that the second law of thermodynamics held "the supreme position among the laws of Nature." Any theory that argued against it, he wrote, would "collapse in deepest humiliation."[18] Nearly 100 years later, Eddington has yet to be proven wrong.

REFERENCES

1. J. Goold et al., *J. Phys. A: Math. Theor.* **49**, 143001 (2016).
2. C. Cercignani, *Ludwig Boltzmann: The Man Who Trusted Atoms*, Oxford U. Press (1998).
3. W. C. Myrvold, *Stud. Hist. Philos. Sci. B* **42**, 237 (2011).
4. V. Serreli et al., *Nature* **445**, 523 (2007).
5. H. Leff, A. F. Rex, eds., *Maxwell's Demon 2: Entropy, Classical and Quantum Information, Computing*, CRC Press (2002).
6. O. Shenker, M. Hemmo, *Entropy* **22**, 269 (2020); J. Norton, *Stud. Hist. Phil. Sci. B* **42**, 184 (2011).
7. D. J. Bedingham, O. J. E. Maroney, *New J. Phys.* **18**, 113050 (2016).
8. N. Y. Halpern, *Sci. Am.*, May 2020, p. 60.
9. H. E. D. Scovil, E. O. Schulz-DuBois, *Phys. Rev. Lett.* **2**, 262 (1959).
10. A. Levy, D. Gelbwaser-Klimovsky, in *Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions*, F. Binder et al., eds., Springer (2018), p. 87; B. Gardas, S. Deffner, *Phys. Rev. E* **92**, 042126 (2015).
11. J. Klatzow et al., *Phys. Rev. Lett.* **122**, 110601 (2019).
12. A. Einstein, *Autobiographical Notes*, P. A. Schilpp, trans. and ed., Open Court (1979), p. 31.
13. K. Friston, *Nat. Rev. Neurosci.* **11**, 127 (2010).
14. J. C. Maxwell, "Diffusion," in *The Encyclopædia Britannica: A Dictionary of Arts, Sciences, and General Literature*, 9th ed., vol. 7, Adam and Charles Black (1877), p. 220.
15. P. W. Bridgman, *The Nature of Thermodynamics*, Harvard U. Press (1941), p. 214.
16. Ref. 14, p. 220.
17. S. Lloyd, *Nat. Phys.* **2**, 727 (2006); S. Popescu, A. Short, A. Winter, *Nat. Phys.* **2**, 754 (2006); S. Goldstein et al., *Phys. Rev. Lett.* **96**, 050403 (2006).
18. A. S. Eddington, *The Nature of the Physical World*, Cambridge U. Press (1928), p. 74. **PT**