Stephen Klein
FQXi Essay Contest Submission

Laplace's Demon – Thwarted by Modern Physics, or Does He Know Something We Don't?

The question of fate has perplexed mankind throughout all of recorded history. In the Western world, the question has been interpreted through many lenses – Reformist theologists claim that the Christian God preordains every event in the universe; metaphysical libertarians not only embrace "free will," but also reject the notion of physical determinism; solipsists throw their hands up and avoid the question altogether, denying any sort of external, objective reality. In my opinion, the most intellectually honest approach is the agnostic, empirical approach taken by science.

When removing spirituality or any other metaphysical quantities from the equation, the question becomes purely physical. More specifically, the question becomes, "when analyzing systems of physical objects, can one predict how they will evolve over time, given enough information at the beginning?" Whether it was his intention or not, Newton seemed to provide an answer when formalizing his physical laws. Newton's second law of motion says

$$\Sigma F = \dot{p},$$

where the left-hand side represents the sum of the forces acting on an object, and the right-hand side is the change in the object's momentum with respect to time. This can be rewritten as

$$\Sigma F = m\ddot{\mathrm{x}},$$

which is a second order differential equation in x, the object's position. From this, it can be seen that one only needs two initial conditions – its position and velocity at the start (wherever one decides to set time to zero) – to know exactly how the object with evolve over time. Obviously, a system involving an enormous number of objects would be impossible for any single man or woman to track; however, there is no *theoretical* limit for a being or machine that *could* keep track of all of the particles, given their initial states. In 1814, the famous French scholar Pierre-Simon Laplace imagined a demon who was capable of such a feat. This demon would be able to know every particle in the universe's state at a given time, and thus, be capable of extrapolating the entire past and future of the universe. This thought experiment concludes that Newtonian physics implies a rigid, deterministic universe; every event in the universe was set into motion according to the conditions of the Big Bang.

Modern developments in physics seem to indicate that there is indeed a limit to how much can be known about a system. Quantum mechanics defies the existence of Laplace's demon. The formal mathematical language of quantum theory is linear algebra, involving vectors, which represent states of particles, and operators, who act on these vectors to produce a new state. In linear algebra, operators, say $\hat{A}$ and $\hat{B}$, are said to either commute, or not commute, according to the commutator of the two operators:

$$\left[\hat{A},\hat{B}\right] = \widehat{\mathrm{A}}\widehat{\mathrm{B}} - \widehat{\mathrm{B}}\widehat{\mathrm{A}} = 0 \qquad \textbf{\textit{Commuting operators}}$$

$$\left[\hat{C},\widehat{D}\right] = \hat{\mathrm{C}}\widehat{\mathrm{D}} - \widehat{D}\hat{\mathrm{C}} \neq 0. \qquad \textbf{\textit{Non-commuting operators}}$$

For any physical measurement in QM, these operators are compositions of position and momentum operators. They act on a complex vector $|\Psi\rangle$ (representing the state of a particle or an entangled system) to produce the same vector times a real-valued constant. That is,

$$\widehat{\mathrm{A}}\,|\Psi\rangle = \mathrm{a}\,|\Psi\rangle,$$

where the constant a is an eigenvalue, and physically, represents the measured value of the physical operator on the eigenfunctions (i.e., the name of a vector which produces an eigenvalue). Two operators which commute can share eigenfunctions, meaning you could apply both of them, in any order, to an eigenfunction and extract measurement values for both. One of the consequences of this formalism is that the operator for position, $\hat{x}$, and the operator for momentum, $\hat{p}$, do not commute. This means that one can never find simultaneous eigenvectors for both position and momentum; that is, not even Laplace's demon could actually know both conditions at any given time. This is summarized famously in Heisenberg's Uncertainty Principle:

$$\sigma_x \sigma_p = \frac{\hbar}{2},$$

in which σ_x is the uncertainty in position, σ_p is the uncertainty in momentum, and $\hbar$ is Planck's constant. Quantum formalism sets a definite limit on the deterministic universe predicted by Newtonian mechanics.

Entropy is another physical quantity at odds with predictability. In thermodynamic statistics, entropy is related to the number of unique configurations a given system can be arranged – i.e. the microstates of the system:

$$S = -k_B \Sigma p_i ln(p_i).$$

In the above formula, S is entropy, k_B is Boltzmann's constant, and p_i is the probability of a given microstate. Given a sufficient amount of time, it's been experimentally verified that physical systems tend to evolve towards macrostates with the largest multiplicity of microstates, Ω. This is the Second Law of Thermodynamics. In other words, the entropy of the universe is always increasing with respect to time. This physical law has some deep implications with regard to information. Specifically, the thermodynamic entropy S is analogous to Shannon entropy H:

$$H = -\Sigma p_i log_b p_i,$$

where the logarithmic base b is related to the types of questions one would be asking to gain information about the system. For example, a yes or no question is a binary one, and so one would use $b = 2$ if analyzing the system this way. The two entropies aren't merely mathematically symmetric; they are directly related. For example, any system's microscopic configuration, say an Einstein solid with N atoms with q quantized energy units, can be described with some discrete number of yes or no questions – “Does the Nth atom contain 1 or more quanta?,” “Does it contain 10?,” etc. An incremental increase of the solid’s thermodynamic entropy, dS, requires *more* probing, dH, in the form of yes or no questions in order to be fully described. In other words, when a system’s thermodynamic entropy increases, quantifiable information is lost. The Second Law of Thermodynamics implies that one needs to play a deranged, never-ending game of “21 Questions” to keep up to date with a system. With a system as large as the universe in question, Laplace’s demon would need to ask and answer these questions as fast as they arrive, ad infinitum.

Information isn’t actually *lost*. Despite needing to play catch up, the demon still can find it encoded somewhere in the universe, according to unitarity of quantum mechanics. The notion that information could be permanently erased would give a lot of physicists nightmares, as information conservation is fundamental to the theories they build. For example, in 1974, Stephen Hawking made a such a proposal with his theoretical prediction that black holes emit radiation. This Hawking radiation, which Hawking himself thought to be completely thermal in nature, would seem to imply that a black hole that doesn’t absorb mass is doomed to eventually evaporate away. Prior to Hawking’s discovery, physicists were content to say that the information that crosses a black hole’s event horizon stays encoded in the hole itself. Hawking contended that this radiation doesn’t contain any extractable information about the contents of the black hole, and so eventually, when the black hole completely dissipates, and information is lost. Is this the final nail in the demon’s coffin? Not quite. Physicists, notably Hawking’s colleagues Leonard Susskind and John Preskill, were appalled at this suggestion. Many solutions have been put forth, from the intuitive suggestion that the information indeed *does* seep out with the radiation, to the mind-bending idea that the information is stored in a “baby universe” separate than our own universe. The most popular belief is rooted in string theory, named the anti-de Sitter/conformal field theory (AdS/CFT) correspondence, which claims that the information lies on the event horizon itself.

This is far from a foregone conclusion however, and in an effort to propose an actual mechanism for which the information leaks from the black hole, physicists are faced with a dilemma. Physicist Joe Polchinski and his colleagues rigorously argued that the black hole information paradox necessarily pits quantum mechanics against general relativity. Namely, one of the two must give: either quantum theory's unitarity, which calls for the conservation of information, or general relativity's principle of equivalence, which says that an observer can't tell the difference between uniform acceleration and gravity. If these theories, on which most of our modern understanding of the universe is based, are incompatible, then it's fair to question whether or not our conclusions about the demon hold up.

How do these inconsistencies in our physical theories arise? For one, it's important to notice that many of the conclusions we reach in physics are based in the mathematical formalism we adopt. Personally, in my relatively short experience with physics as an undergraduate student, the distinction between physics and math is much clearer to me than it was as a first semester freshman. When others would remark, "Oh, you must be good at math," when learning I was going to study physics, I would try to reply humbly in agreement. Two years later, I can now safely say that I am *not* good at math. Mathematics attempts to expand the framework set up by axioms with no external guidance; wherever the axioms lead, so the math goes. Physics takes whatever it needs from mathematical formalism in order to accurately model nature, sometimes excluding parts of the formalism deemed unnecessary or incompatible with physics. This can be seen, for example, when finding the compatible eigenvalues of the z-component, L_z, and total angular momentum, L^2 of a hydrogen orbital using ladder operators. Letting the L_z top rung eigenvalue be hl, and the bottom rung eigenvalue be $h\bar{l}$, we come across two solutions relating the two:

$$\bar{l} = l + 1; \qquad \bar{l} = -l.$$

Mathematically, there is no preference; the two statements are equally valid. But physically, the first statement claims that the bottom rung eigenvalue is greater than the eigenvalue at the top. Knowing this is wrong, physicists throw out the first statement as nonsensical. While the formalism of quantum mechanics is incredibly successful, it remains one of the most open-ended theories regarding physical interpretation, including the reality of the wavefunction and the randomness of measurement.

So long as the mathematics are successful at modeling the universe, it sticks around in physics. But this doesn't mean that the mathematics are an implicit reflection of *reality*. For centuries, Newton's inverse square law describing gravity,

$$\overrightarrow{F_g} = -\frac{GM_1M_2}{r^2}\hat{r},$$

(where G is the gravitational constant, M_1 and M_2 are the two masses, and r is their relative distance), was generally accepted to be true. However, for over 200 years, there was an unsolved discrepancy between Mercury's observed orbit and the orbit predicted by Newtonian gravity. Einstein was able to solve the matter with his theory of general relativity, but this begs the question of what this new theory fails to explain – which we have already seen regarding information and black holes, and quantum gravity at large. It is paramount to understand that mathematics is a necessary tool in analyzing physical systems, but not equivalent to the physics, and more importantly, nature itself.

Furthermore, our mathematical formalism and analysis of nature are unavoidably human. That might sound trivial, but it's a necessary distinction to make if we are to contrast our experience with Laplace's demon. Our cognition, which is physically manifest in our prefrontal cortex, is inextricably linked to biology thanks to evolution. Given that we have no other intelligent life around with which to compare notes, it's an open question whether the logic we employ in our mathematics is actually absolute, or whether it's painted by our subjective experience and/or derivative of our animal ancestors'. Our observations, which form the basis of our physical theory, can be called into question as well. Time, for example, plays an intimate role in all of our physical theories, whether it exists as an independent variable, such as in Newtonian mechanics, or as linked to space itself, as in general relativity. The time symmetry associated with the latter implores some physicists to claim that there *is* no flow of time as we experience it – the four-dimensional spacetime that makes up the universe exists as a static "block," and any interpretation otherwise must be a mental construct.

A more familiar and grounded example of how our humanity paints our answer to Laplace's demon is the existence of our intrinsic bias. Despite scientist's best efforts to remain impartial, their prior beliefs undeniably guide their positions when answering questions. When Stephen Hawking implied the non-conservation of information in black holes, Leonard Susskind and others were so appalled that they (in good fun) declared "war" on Hawking, which Susskind detailed in his popular science book "The Black Hole War." The interpretation of quantum mechanics is hotly debated, with more than a dozen popular and influential theories. The great Einstein himself, a hard determinist, fundamentally opposed the most orthodox position: the Copenhagen interpretation. This interpretation regards particles to be in actual *physical* superpositions prior to measurement, leading to fundamentally randomly measurement values. He famously said in a letter to fellow quantum giant Max Born, "The theory produces a good deal but hardly brings us closer to the secret of the Old One. I am at all events convinced that He does not play dice."

Despite the holes in our physical theories and the valid questions of our impartiality as human beings, can we regard our physics as predictive? Within the context of human experience, absolutely. Quantum mechanics, in spite of its open-ended interpretations, has revolutionized our computing thanks to band theory and the development of semiconductors. Modern navigation and the GPS owe themselves to the precise prediction of photon red-shifting according to general relativity. Thermodynamic and Shannon entropy are the foundations of chemical engineering and information technology, respectively. But do these theories, which imply there is a fundamental limit one can know, defeat Laplace’s demon? The answer is as unknowable as whether or not these shackles of unpredictability are projections of our humanity. Not satisfied? Ask Laplace’s demon – he might know something I don’t.

References

1. Leonard Susskind (7 July 2008). "Introduction". *The Black Hole War: My Battle with Stephen Hawking to Make the World Safe for Quantum Mechanics*. Hachette Inc. ISBN 978-0-316-01640-7.

2. Preskill, John. "Black hole information bet". *California Institute of Technology*.

3. Almheiri, Ahmed; Marolf, Donald; Polchinski, Joseph; Sully, James (11 February 2013). "Black holes: complementarity or firewalls?". *Journal of High Energy Physics*. 2013 (2): 62. arXiv:1207.3123