Quantropy

John Baez

There is an extensive analogy between statistical mechanics and quantum mechanics. In statistical mechanics, a system in thermal equilibrium occupies all possible states, each with a probability depending on the energy of that state. In quantum mechanics, a system takes all possible paths with the passage of time, each with an amplitude depending on the action of that path.

This analogy is famous and widely used in physics. But what does it really mean? Is it just a mathematical 'trick', or something deeper? I believe it is trying to tell us something about information.

Occam's razor says that the best model of a system is the simplest one that fits the data we have. As noted by Jaynes [8] and Solomonoff [11], the concept of 'simplicity' can be made quantitative using information theory. In these terms, Occam's razor says the best model specifies the least amount of information needed to fit the data. But since entropy is another name for unspecified information, we can say this another way: our model should maximize entropy subject to the constraints given by the data. This is the **principle of maximum entropy**.

Statistical mechanics is governed by the principle of maximum entropy. A system in thermal equilibrium occupies states with precisely the probabilities that maximize entropy subject to the constraints given by what we know about the system.

In preliminary calculations, I have found that the analogy between statistical mechanics and quantum mechanics extends to include a concept analogous to entropy, which I call 'quantropy'. In quantum mechanics, the amplitudes for a system to take different paths are precisely those that extremize quantropy.

Quantropy is not the same as the entropy of a quantum system; it is something new. For example, it is a *complex* number instead of a *real* number. It governs *dynamical* rather than *static* systems. Nonetheless, we can ask: does the tight relation between information and entropy extend to include quantropy? Can we use quantropy to understand dynamics in quantum mechanics using a new generalization of Occam's razor? To answer these questions, the first step is to carefully work out the role quantropy plays in quantum mechanics.

In what follows, I start by reviewing some variational principles in physics, and noting that the principles governing statics all follow from the principle of maximum entropy. Then I argue that the principles governing dynamics follow from the principle of stationary quantropy. Finally, I sketch a research program, to be funded by this grant, that could study this principle and explore its implications.

1 Variational principles in physics

1.1 Statics

Static systems at temperature zero obey the **principle of minimum energy**: they seek to minimize their energy. Energy is typically the sum of kinetic and potential energy:

$$E = K + V$$

where the potential energy V depends only on the system's position, while the kinetic energy K also depends on its velocity. The kinetic energy usually has a minimum at velocity zero. In classical physics, this lets our system minimize energy in a two-step way. It minimizes kinetic energy by staying still. It minimizes potential energy by finding the right place to stay still. So classically, statics at zero temperature is usually described using the **principle of minimum potential energy**.

In quantum physics, a tradeoff is required: thanks to the uncertainty principle, we can't simultaneously minimize potential and kinetic energy. This makes minimizing their sum more tricky. However, the principle of minimum energy still holds.

What about static systems at nonzero temperature? These are studied in the subject called 'equilibrium thermodynamics', which is governed by the **principle of minimum free energy**. In equilibrium at any fixed temperature, a closed system will minimize its free energy

$$F = E - TS$$

where T is the temperature and S is the entropy. Note that this principle reduces to the principle of minimum energy when $T = 0$.

But where does the principle of minimum free energy come from? We can understand it using probability theory. Suppose for simplicity that our system has a finite set of states, say X, and the energy of the state $x \in X$ is E_x. Instead of our system occupying a single definite state, let us suppose it can be in *any* state, with a probability p_x of being in the state x. Then its **entropy** is, by definition:

$$S = -\sum_x p_x \ln p_x .$$

The expected value of the energy is

$$E = \sum_x p_x E_x.$$

Now suppose our system maximizes entropy subject to a constraint on the expected value of energy. Using the method of Lagrange multipliers, this is the same as maximizing

$$S - \beta E$$

where β is a Lagrange multiplier. When we go ahead and maximize this, we find the system chooses a 'Boltzmann distribution':

$$p_x = \frac{\exp(-\beta E_x)}{\sum_x \exp(-\beta E_x)}.$$

But what does this mean? The quantity β is the reciprocal of the temperature T, at least in units where Boltzmann's constant is set to 1. So, when the temperature is positive, maximizing $S - \beta E$ is the same as minimizing the free energy $F = E - TS$.

In short, all the variational principles in statics, in order of increasing generality:

- the principle of least potential energy
- the principle of least energy
- the principle of least free energy

are special or limiting case of the principle of maximum entropy, as long as we maximize entropy subject to the relevant constraints.

In 1948, Shannon [10] had an insight that cast these facts in a brand new light: he realized that entropy is the same as *unspecified information*. This led Jaynes [8] to realize that the principle of maximum entropy can be seen as a precise, quantitative version of Occam's razor. In other words: our probabilistic description of a system in equilibrium should contain as little information as possible, beyond that needed to fit the data we observe.

1.2 Dynamics

Now suppose things are changing as time passes, so we are doing 'dynamics' instead of mere statics. In classical mechanics we can imagine a system tracing out a path $\gamma(t)$ as time passes from one time to another, for example from $t = t_0$ to $t = t_1$. The **action** of this path is typically the integral of the kinetic minus potential energy:

$$I(\gamma) = \int_{t_0}^{t_1} (K(t) - V(t))\, dt$$

where $K(t)$ and $V(t)$ depend on the path γ. The **principle of least action** says that if we fix the endpoints of this path, the system will follow the path that minimizes the action subject to these constraints.

The situation becomes more interesting in quantum mechanics. Here Feynman proposed that instead of our following a single definite path, it can follow *any* path, with an amplitude $a(\gamma)$ of following the path γ. And he proposed this prescription for the amplitude:

$$a(\gamma) = \frac{\exp(iI(\gamma)/\hbar)}{\displaystyle\int \exp(iI(\gamma)/\hbar)\, D\gamma}$$

where $\hbar$ is Planck's constant. He also gave a heuristic argument showing that as $\hbar \to 0$, this prescription reduces to the principle of least action.

Unfortunately the integral over all paths is hard to make rigorous except in certain special cases. Since this issue is a bit of a distraction here, let us talk more abstractly about 'histories' instead of paths with fixed endpoints, and consider a system whose possible histories form a finite set, say X. Suppose the action of the history $x \in X$ is I_x. Then Feynman's 'sum over histories' prescription says the amplitude of the history x is

$$a_x = \frac{\exp(iI_x/\hbar)}{\displaystyle\sum_x \exp(iI_x/\hbar)}$$

The formula here looks very much like the Boltzmann distribution:

$$p_x = \frac{\exp(-E_x/T)}{\displaystyle\sum_x \exp(-E_x/T)}$$

This gives rise to the following famous analogy:

Statics	**Dynamics**
statistical mechanics	quantum mechanics
probabilities	amplitudes
Boltzmann distribution	Feynman sum over histories
energy, E	action, I
temperature, T	i times Plancks constant, $i\hbar$
entropy	???
free energy	???

However, this analogy is missing two items, which would be good to fill in. Most importantly: *what is the analogue of entropy?* The principle of minimum entropy is our way of making Occam's razor precise using information theory. It guides our understanding of statistical mechanics. Surely we should seek its analogue in quantum mechanics!

1.3 Quantropy

Since the Boltzmann distribution

$$p_x = \frac{\exp(-E_x/T)}{\sum\limits_x \exp(-E_x/T)}$$

comes from the principle of maximum entropy, we might hope Feynman's sum over histories formulation of quantum mechanics:

$$a_x = \frac{\exp(iA_x/\hbar)}{\sum\limits_x \exp(iA_x/\hbar)}$$

comes from a maximum principle too.

Unfortunately Feynman's sum over histories involves complex numbers, and it doesn't make sense to maximize a complex function. However, when we say nature minimizes or maximizes something, it sometimes behaves like a bad student who applies the first derivative test and quits there: it just finds a 'stationary point', where the first derivative is zero. For example, paths in classical mechanics need not minimize the action: they may merely be stationary points. This is good for us, because stationary points still make sense for complex functions.

So, let us try to derive Feynman's prescription from a **principle of stationary quantropy**.

Suppose we have a finite set of histories X and each history $x \in X$ has a complex amplitude $a_x \in \mathbb{C}$. Assume these amplitudes are normalized so that

$$\sum_x a_x = 1$$

since that is what Feynman's normalization actually achieves. We can try to define the **quantropy** of a by:

$$Q = -\sum_x a_x \ln a_x.$$

One might fear this is ill-defined when $a_x = 0$, but that is not the main problem; in the study of entropy we typically set $0 \ln 0 = 0$, and everything works fine. The more serious problem is that the logarithm has different branches: we can add any multiple of $2\pi i$ to our logarithm and get another equally good logarithm. To deal with this, let us assume we have chosen a specific logarithm for each number a_x, and suppose that when we vary them they don't go through zero, so we can smoothly change their logarithm as they move. (Clearly this issue deserves further study.)

Next, suppose each history x has an action $I_x \in \mathbb{R}$. Recall that in statistical mechanics, the probabilities give a stationary point of entropy subject to a constraint on the expected energy. So, following our analogy chart, let us seek amplitudes a_x that give a stationary point of the quantropy Q subject to a constraint on the **expected action**:

$$A = \sum_x a_x I_x.$$

The term 'expected action' is a bit odd, since the numbers a_x are amplitudes rather than probabilities. While we could try to justify it from how expected values are computed in Feynman's formalism, I am mainly using this term because A is analogous to the expected value of the energy, which we saw earlier.

To find a stationary point of Q subject to a constraint on A, we can use the method of Lagrange multipliers, which reduces this task to finding a stationary point of

$$Q - \lambda A$$

where λ is a Lagrange multiplier. This quantity λ is analogous to the inverse temperature $\beta = 1/T$ in statistical mechanics, so our analogy chart suggests we should take

$$\lambda = 1/i\hbar.$$

Finding stationary points of $Q - \lambda A$ is not hard; a calculation [1] shows they occur when

$$a_x = \frac{\exp(-\lambda I_x)}{\sum_x \exp(-\lambda I_x)}$$

If we choose λ as above, *this is precisely Feynman's sum over histories formulation of quantum mechanics!*

So, the principle of stationary quantropy indeed gives the correct amplitudes for time evolution in quantum mechanics. In the process, it also gives a stationary point of the **free action** $I - i\hbar Q$. This 'free action' is the quantum analogue of the free energy, and it completes our analogy chart:

Statics	**Dynamics**
statistical mechanics	quantum mechanics
probabilities	amplitudes
Boltzmann distribution	Feynman sum over histories
energy, E	action, I
temperature, T	i times Planck's constant, $i\hbar$
entropy, S	quantropy, Q
free energy, $E - TS$	free action, $I - i\hbar Q$

When $\hbar \to 0$, free action reduces to action, so we recover the principle of stationary action in classical mechanics.

2 Research Proposal

The challenge now is to understand quantropy in more detail, and especially its relation to information theory. My plan is as follows:

1. Write and publish a paper explaining the basic idea of quantropy, how it arises naturally from the analogy between statistical mechanics and quantum mechanics, and how it can be seen as a novel extension of the concept of information.

2. Show that quantropy and free action obey a host of identities analogous to those obeyed by entropy and free energy in thermodynamics. I have already begun this [2], but there is more to do, and all this work needs to be published.

3. Explicitly compute the quantropy and free action in some examples, and investigate their properties. I have already begun this for the n-dimensional quantum harmonic oscillator [3]. Since the space of histories for this system is not finite, the sums in my expository treatment above become path integrals that require some work to deal with in a rigorous way. Here my experience with mathematically rigorous quantum field theory will come in handy [5]. The assistance of a graduate student, to be funded by this grant, will also be very helpful here.

4. Study the sense in which quantropy generalizes information. In physics the technique of Wick rotation lets us think of time as 'imaginary inverse temperature': the role of $\exp(-\beta E)$ in statistical mechanics is

sometimes taken over by $\exp(itE/\hbar)$ in quantum mechanics. Similarly, while there is a well-known way to compute probabilities from amplitudes in quantum mechanics, it can also be useful to treat amplitudes as 'complex probabilities'. I am almost done writing a book with Jacob Biamonte on this topic [4]. Extending this line of thought, we can try to see quantropy as 'complex information'. However, the meaning of this remains obscure, and demands further thought! This part of the project is more risky than the rest, but it holds the promise of bigger payoffs, since it is deeply foundational in nature.

5. Study the extent to which Occam's razor can be formalized using information theory and the principle of maximum entropy, and then *generalized* using the principle of stationary quantropy. The aforementioned work of Jaynes [8] went a long way toward clarifying the former question, but there is a somewhat separate body of work going back to Solomonoff [11], which uses 'algorithmic information' to measure the complexity of a model. Various forms of algorithmic information has been extensively studied, along with their connection to entropy [7, 9]. However, it is only rather recently that Mike Stay and I have generalized all the usual formulas of thermodynamics to 'algorithmic thermodynamics', using the fact that algorithmic information is actually a *special case* of the entropy of a probability distribution [6]. So, there is more to do when it comes to clarifying the conceptual foundations of the principle of maximum entropy. When it comes to quantropy, the field is wide open.

To pursue these goals, I am proposing a 2-course reduction in my teaching load during the period of the grant. (Math department faculty at U. C. Riverside teach 4 quarter-long courses per year, so this would be a 25% teaching load reduction.) I am also asking for support for one graduate student at 49% for 6 academic months per year for two years. This would reduce their usual teaching assistant duties and enable them to put time into this project, helping me with calculations and writing papers. All this would enable me to pursue work on quantropy, and more generally the foundations of information theory and its interaction with quantum theory, which I would otherwise be hard pressed to do.

References

[1] J. Baez, Quantropy (Part 1), *Azimuth*, 22 December 2011.

[2] J. Baez, Quantropy (Part 2), *Azimuth*, 10 February 2012.

[3] J. Baez, Quantropy (Part 3), *Azimuth*, 18 February 2012.

[4] J. Baez and J. Biamonte *Quantum Techniques for Stochastic Processes*, with Jacob Biamonte, in preparation.

[5] J. Baez, I. E. Segal and Z. Zhou, *An Introduction to Algebraic and Constructive Quantum Field Theory*, Princeton University Press, 1992.

[6] J. Baez and M. Stay, Algorithmic thermodynamics, *Math. Struct. Comp. Sci.* **22** (2012), 771–787.

[7] C. S. Calude, *Information and Randomness: An Algorithmic Perspective*, Springer, Berlin, 2002.

[8] E. T. Jaynes, *Probability Theory: The Logic of Science*, Cambridge U. Press, Cambridge, 2003.

[9] M. Li and P. Vitányi, *An Introduction to Kolmogorov Complexity Theory and Its Applications*, Springer, Berlin, 2008.

[10] C. E. Shannon, A mathematical theory of communication, *Bell System Technical Journal* **27** (1948) 379–423, 623–656.

[11] R. J. Solomonoff, A formal theory of inductive inference, Part I, *Inform. Control* **7** (1964), 1–22.