A Possible Way to Avoid Heat Death

Nitaya Buntao[1], Narunchara Katemee[2], and
Vladik Kreinovich[3]

[1,2]Department of Applied Statistics
Faculty of Applied Science
King Mongkut's University of
Technology North Bangkok
1518 Pibulsongkram Rd.
Bangsue, Bangkok 10800 Thailand
[1]taltanot@hotmail.com
[2]Narun_ess@hotmail.com

[3]Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
vladik@utep.edu

Abstract

A disturbing consequence of the traditional thermodynamics is the possibility of *heat death*, when the Universe arrives at the state with the largest possible value of the entropy and all the processes will stop. In this paper, we show that one possible way to avoid this consequence is to consider situations in which the entropy never attains its maximum – and thus, the heat death state is not possible. We show that such situations can have physical sense – e.g., they naturally appear in boostrap models.

Prediction in physics: general case. In order to formulate the problem, let us briefly recall the main ideas of statistical physics and thermodynamics; for details, see, e.g., [1].

One of the main objectives of physics is to predict what will happen in different systems. To make this prediction, we observe and/or measure the state of the system, and then use the equations to predict how this state will change with time.

Example. In mechanics, we measure the positions and velocities of all the particles. After that, once we know how exactly the particle interact, i.e., how exactly the forces acting on each of these particles depend on their locations and velocities, we can predict the positions and velocities of all the particles.

In particular, in celestial mechanics, Newton's laws enable us to predict the motion of the planets, comets, satellites, etc.

Prediction in physics: case of changing objects. This idea works well if we are only interested in the *motion* of the objects. However, in many practical situations, objects not only move, they also *change*: water evaporates, hot objects become colder, etc. In such situations, we are interested in learning not only how the objects move, but also how they change.

Theoretically, exact prediction is possible for changing objects. Theoretically, this prediction problem is similar to the previous one since, according to modern physics, all observed changes in the macro-bodies can be explained by changing positions and velocities of the molecules that form these bodies. For example:

- when a body becomes colder, it simply means that the average speed of the molecules decreases;
- when a solid crystal melts, it means that the locations of the molecules change from the original regular grid to a a more chaotic random one.

Ideally, if we know the exact position and velocity of each the molecules, i.e., if we know the *micro-state*, then, in theory, we can predict their positions and velocities in the future moments of time as well.

For changing objects, exact prediction is not practically possible. In practice, however, this prediction is not possible. Indeed, each macro-object (gas, liquid, solid, etc.) consists of a large number of molecules $\approx 10^{23}$, and it is not possible

- neither to measure all their positions and velocities
- nor to store and process this information.

What is possible to measure: notion of a macro-state. Since we cannot measure the micro-state, what can we measure? We cannot measure the position of each individual particle, but we can measure density at different points, i.e., the number of particles which are located in the vicinity of this point. We cannot measure the velocity of *each* individual particle, but we can measure the *number* of particles with the given velocity.

The state as described by the measurable characteristics is called a *macro-state.* Each macro-state consists of many different micro-states.

Enter probabilities. As we have just mentioned, a macro-state consists of different micro-states. Different micro-states can lead to different future states. So, if we only know the macro-state, we cannot uniquely determine how an object will change, we can, at best, describe the probabilities of different possible future states. Computing such probabilities is the main objective of *statistical physics.*

Main assumption of statistical physics: all micro-states are equally probable. The main assumption of statistical physics is that all the micro-states are equally probable.

Under this reasonable assumption, the probability of each macro-state s is simply proportional to the total number $N(s)$ of micro-states in this state. If several macro-states s are possible, then the most probable one is the one with the largest possible value of $N(s)$.

So, for making predictions, it is important to be able to compute the value $N(s)$.

How we can compute $N(s)$. As we have mentioned, in the macro-state, we do not know which particle (molecule) is in what state, we only know how many particles are in different states.

Let us assume that out of N particles,

- the p_1-th portion $n_1 = p_1 \cdot N$ are in the state s_1,
- the p_2-th portion $n_2 = p_2 \cdot N$ are in the state s_2,
- $\ldots$,
- and the p_k-th portion $n_k = p_k \cdot N$ are in the state s_k.

To describe a micro-state, we must know which particle is in which state. To describe a micro-state, first, out of N particles, we need to select n_1 particles which are in the state s_1. The number of such selection is equally to the number of combinations, i.e., to

$$\frac{N!}{n_1! \cdot (N - n_1)!} \tag{1}$$

For each such assignment, we must determine which of the remaining $N - n_1$ particle are in the state s_2. There are n_2 particle in the state s_2, so we must select n_2 particles out of $N - n_1$ available ones. The number of such selections is again equal to the number of combinations, i.e., to

$$\frac{(N - n_1)!}{n_2! \cdot (N - n_1 - n_2)!} \tag{2}$$

We need this many selection for each of the original $\dfrac{N!}{n_1! \cdot (N - n_1)!}$ selections, so overall, we have

$$\frac{N!}{n_1! \cdot (N - n_1)!} \cdot \frac{(N - n_1)!}{n_2! \cdot (N - n_1 - n_2)!} = \frac{N!}{n_1! \cdot n_2! \cdot (N - n_1 - n_2)!} \tag{3}$$

ways to allocate particles to states s_1 and s_2. We also need to find n_3 particles in state s_3 among the remaining $N - n_1 - n_2$ one; there are $\dfrac{(N-n_1-n_2)!}{n_3! \cdot (N-n_1-n_2-n_3)!}$ such possibilities, to overall, we have

$$\frac{N!}{n_1! \cdot n_2! \cdot (N-n_1-n_2)!} \cdot \frac{(N-n_1-n_2)!}{n_3! \cdot (N-n_1-n_2-n_3)!} =$$

$$\frac{N!}{n_1! \cdot n_2! \cdot n_3! \cdot (N-n_1-n_2-n_3)!} \tag{4}$$

ways to allocate particles to states s_1, s_2, and s_3. Continuing with the states s_4, s_5, $\ldots$, s_k, we conclude that there are

$$N(s) = \frac{N!}{n_1! \cdot n_2! \cdot n_3! \cdot \ldots \cdot n_k!} \tag{5}$$

ways to allocate particles into k groups.

Towards an easier-to-analyze approximate expression for $N(s)$. To get an easier-to-analyze approximate expression for $N(s)$, let us consider replacing the product with an easier-to-analyze sum. This can be done if we take the logarithm of $N(s)$:

$$\ln(N(s)) = \ln(N!) - \ln(n_1!) - \ln(n_2!) - \ldots - \ln(n_k!). \tag{6}$$

Since $n_i = N \cdot p_i$, we get

$$\ln(N(s)) = \ln(N!) - \ln((N \cdot p_1)!) - \ln((N \cdot p_2)!) - \ldots - \ln((N \cdot p_k)!). \tag{7}$$

Here, for every n, we have $n! = 1 \cdot 2 \cdot \ldots \cdot n$ and thus,

$$\ln(n!) = \ln(1) + \ln(2) + \ldots + \ln(n). \tag{8}$$

The right-hand side of the expression (8) is an integral sum

$$f(x_0) \cdot \Delta x + f(x_1) \cdot \Delta x + \ldots + f(x_m) \cdot \Delta x, \tag{9}$$

with $x_{k+1} = x_k + \Delta x$, corresponding to $f(x) = \ln(x)$ and $\Delta x = 1$. For a large number of points m, this integral sum is asymptotically equal to the corresponding integral $\int_{x_0}^{x_m} f(x)\,dx$. Thus, we conclude that for large n, we have

$$\ln(n!) \approx \int_1^n \ln(x)\,dx. \tag{10}$$

Using the formula of integration by parts

$$\int u\,dv = uv - \int v\,du \tag{11}$$

with $u = \ln(x)$ and $v = x$, we conclude that

$$\int \ln(x)\,dx = x \cdot \ln(x) - \int x \cdot \frac{1}{x}\,dx = x \cdot \ln(x) - \int 1\,dx = x \cdot \ln(x) - x, \quad (12)$$

and thus,

$$\ln(n!) \approx (x \cdot \ln(x) - x)|_1^n = n \cdot \ln(n) - n - 1 \approx n \cdot \ln(n) - n. \quad (13)$$

Substituting this expression into the formula (7), we conclude that

$$\ln(N(s)) \approx N \cdot \ln(N) - N - N \cdot p_1 \cdot \ln(N \cdot p_1) + N \cdot p_1 - \ldots -$$

$$N \cdot p_k \cdot \ln(N \cdot p_k) + N \cdot p_k. \quad (14)$$

Since $\sum\limits_{i=1}^{k} p_i = 1$, we have

$$-N + N \cdot p_1 + \cdot + N \cdot p_k = 0, \quad (15)$$

and thus, the expression (14) can be simplified into

$$\ln(N(s)) \approx N \cdot \ln(N) - N \cdot p_1 \cdot \ln(N \cdot p_1) - \ldots - N \cdot p_k \cdot \ln(N \cdot p_k). \quad (16)$$

Taking into account that $\ln(N \cdot p_i) = \ln(N) + \ln(p_i)$, we get

$$\ln(N(s)) \approx N \cdot \ln(N) - N \cdot p_1 \cdot \ln(N) - N \cdot p_1 \cdot \ln(p_1) - \ldots -$$

$$N \cdot p_k \cdot \ln(N) - N \cdot p_k \cdot \ln(p_k). \quad (17)$$

Here, since $\sum\limits_{i=1}^{k} p_i = 1$, we have

$$N \cdot \ln(N) - N \cdot p_1 \cdot \ln(N) - \ldots - N \cdot p_k \cdot \ln(N) - N \cdot p_k = 0, \quad (18)$$

and thus, the expression (17) can be simplified into

$$\ln(N(s)) \approx -N \cdot p_1 \cdot \ln(N) - \ldots - N \cdot p_k \cdot \ln(p_k), \quad (19)$$

i.e., into

$$\ln(N(s)) \approx N \cdot S, \quad (20)$$

where

$$S \overset{\text{def}}{=} -\sum_{i=1}^{k} p_i \cdot \ln(p_i) \quad (21)$$

is called the *entropy* of the probability distribution $p_1, \ldots, p_k$. So, in terms of entropy, we have

$$N(s) = e^{\ln(N(s))} \approx e^{N \cdot S}. \quad (22)$$

A stable state is the state with the largest value of entropy. We have mentioned that the probability of different macro-states is proportional to $N(s)$. Thus, the higher the entropy, the larger the probability of a macro-state.

Moreover, for large $N \approx 10^{23}$, even a small decrease in S leads to a drastic decrease in $N(s)$ and thus, in the probability. In other words, the probability of a system being in a macro-state with the largest possible value of the entropy is so much larger that the probability of being in any other state that we can simply conclude that the system *is* in the state with the largest entropy.

Of course, this conclusion does not mean that a system is *always* in this state: we can start with an artificially prepared state in which the distribution will be different, but eventually, due to random changes in a micro-state, the system will move into its most probable state – the state with the largest entropy.

In many practical situations, this conclusion makes physical sense. In many situations, the above conclusion makes perfect sense. For example, suppose that we start with a gas that has different density and/or different temperature at different locations. Then,

- warmer parts will heat the cooler ones,
- molecules will spread from the areas with higher density to the areas with lower density,

and eventually, we will reach the maximum entropy state in which both the velocity and the density are the same over all locations.

Heat death: an undesirable consequence of the above conclusion. For the usual macro-systems, the above conclusion – that every system will eventually reach the maximum entropy state (of thermal equilibrium) – is reasonable. However, when applied to the Universe as a whole, it leads to a somewhat unexpected conclusion that the Universe will eventually reach the state with the largest entropy after which there will be no further changes.

This strange final state, in which no further changes are possible, is called a *heat death*.

How physicists approach this problem. The traditional physical approach to the heat death problem is to show that while the heat death state is possible, it is not actually attained in the expanding Universe; see, e.g., [3].

Our main idea. To further avoid the above undesirable conclusion of a heat death of the Universe, we will show that it is possible to have a model of the Universe in which there is no state with the largest entropy – and therefore, heat death is not possible.

This idea further develops arguments described in [4, 5, 6].

To formalize this idea, we need to recall the ideas of statistical physics and thermodynamics related to the notion of temperature.

Statistical physics ideas related to temperature: a brief reminder. In statistical physics, the notion of temperature comes from the fact that energy E is preserved: once we start with a state with a given value of the energy E_0, the resulting states always have the same value of the energy.

Let E_i denote the energy of a state s_i. Then, the overall energy of a macro-state in which we have the p_1-th portion of the particles in the state s_1, the p_2-th portion of the particles in the state s_2, ..., is equal to

$$E = N \cdot p_1 \cdot E_1 + \ldots + N \cdot p_k \cdot E_k. \tag{23}$$

Thus, possible macro-states are limited to states $p_1, \ldots, p_k$ for which

$$p_1 \cdot E_1 + \ldots + p_k \cdot E_k = \frac{E_0}{N}. \tag{24}$$

The maximum entropy state can be thus determined as a state for which the entropy $-\sum\limits_{i=1}^{k} p_i \cdot \ln(p_i)$ attains the largest value under the constraints (24) and

$$p_1 + \ldots + p_k = 1. \tag{25}$$

Applying the Lagrange multiplier method to this constrained optimization problem, we get the unconstrained optimization problem of maximizing the expression

$$-\sum_{i=1} p_i \cdot \ln(p_i) + \lambda_1 \cdot \left(\sum_{i=1} p_i \cdot E_i - \frac{E_0}{N}\right) + \lambda_2 \cdot \left(\sum_{i=1} p_i - 1\right). \tag{26}$$

Differentiating this expression with respect to p_i and equating the derivative to 0, we conclude that

$$-\ln(p_i) - 1 + \lambda_1 \cdot E_i + \lambda_2 = 0, \tag{27}$$

i.e., that $\ln(p_i) = -a - \beta \cdot E_i$ for some constants $a = 1 - \lambda_2$ and $\beta = -\lambda_1$, and that

$$p_i = \frac{1}{Z} \cdot e^{-\beta \cdot E_i}, \tag{28}$$

where we denoted $Z \overset{\text{def}}{=} e^a$.

Once we know β, we can find the normalizing coefficient Z from the condition that $\sum\limits_{i=1}^{k} p_i = 1$, as

$$Z = \sum_{j=1}^{k} e^{-\beta \cdot E_j}. \tag{29}$$

Thus, we have

$$p_i = \frac{e^{-\beta \cdot E_i}}{\sum\limits_{j=1}^{k} e^{-\beta \cdot E_j}}, \tag{30}$$

and the total energy $E(\beta)$ of this macro-state is equal to

$$E(\beta)=\sum_{i=1}^{k}p_i\cdot E_i=\frac{\sum\limits_{i=1}^{k}E_i\cdot e^{-\beta\cdot E_i}}{\sum\limits_{j=1}^{k}e^{-\beta\cdot E_j}}. \tag{31}$$

It is known that when the value β increases, the energy $E(\beta)$ decreases. Indeed, differentiating $E(\beta)$ with respect to β, we conclude that

$$\frac{dE}{d\beta}=\frac{-\left(\sum\limits_{i=1}^{k}E_i^2\cdot e^{-\beta\cdot E_i}\right)\cdot\left(\sum\limits_{j=1}^{k}e^{-\beta\cdot E_j}\right)+\left(\sum\limits_{i=1}^{k}E_i\cdot e^{-\beta\cdot E_i}\right)^2}{\left(\sum\limits_{j=1}^{k}e^{-\beta\cdot E_j}\right)^2}. \tag{32}$$

This expression can be described in terms of p_i (formula (30)) as

$$\frac{dE}{d\beta}=-\left(\sum_{i=1}^{k}p_i\cdot E_i^2\right)+\left(\sum_{i=1}^{k}p_i\cdot E_i\right)^2, \tag{33}$$

i.e., as $\dfrac{dE}{d\beta}=-M[E^2]+(M[E])^2$, where $M[\cdot]$ denotes the mean value. Thus, this derivative is equal to minus variance and is, therefore, always non-negative. Hence, $E(\beta)$ is a decreasing function of β.

The value β must be determined from the condition that the overall energy $E(\beta)$ is equal to $\dfrac{E_0}{N}$.

The larger β, the fewer particles have higher energy. This ties in well with the macro-understanding of the temperature: higher temperatures mean that we have more fast high-energy particles. It turns out that the above-defined parameter β is actually inverse proportional to the temperature T: $\beta=\dfrac{1}{k_0\cdot T}$ for some coefficient k_0.

Continuous approximation. For systems with a large number of particles, there are so many possible states that, in effect, instead of listing all their energies $E_1,\ldots,E_k$, it is more reasonable to consider the *energy density* $d(E)$ – number of different states per unit energy interval. This idea is similar to the usual way of describing a macro-object by its density at different locations, and not by the exact coordinates of all its $N\approx 10^{23}$ particles.

In this approximation, due to the formula (27), the probability density function $\rho(E)$ describing the maximum-entropy energy distribution of states has the form

$$\rho(E)=\frac{1}{Z}\cdot d(E)\cdot e^{-\beta\cdot E}, \tag{34}$$

where the normalizing coefficient Z can be found from the condition that

$$\int_0^\infty \rho(E)\,dE = 1, \tag{35}$$

as

$$Z = \int_0^\infty d(E) \cdot e^{-\beta \cdot E}\,dE. \tag{36}$$

The condition that the average energy is equal to $e_0 \stackrel{\text{def}}{=} \frac{E_0}{N}$ takes the form

$$e_0 = \int_0^\infty E \cdot \rho(E)\,dE = \frac{1}{Z} \cdot \int_0^\infty E \cdot d(E) \cdot e^{-\beta \cdot E}\,dE \tag{37}$$

Case of macro-objects. For usual macro-objects, the state density $d(E)$ grows as a power of E – thus slower than any exponential function $e^{\beta \cdot E}$. As a result, the integral $\int E \cdot d(E) \cdot e^{-\beta \cdot E}\,dE$ always converges.

- For $\beta = \infty$, when all the particles are concentrated in the state with $E = 0$, this integral is equal to 0.
- For $\beta \to 0$, when $e^{-\beta \cdot E} \to 1$, this integral tends to a divergent integral $\int E \cdot d(E)\,dE = +\infty$.

Thus, for each value e_0, we can find some intermediate value β for which the resulting energy will be equal exactly to e_0. In other words, for usual macro-bodies, the maximum entropy state exists for all possible values of energy e_0.

Possibility to avoid heat death: mathematical analysis. What if the state density $d(E)$ grows faster than the power of E? For example, what if $d(E)$ grows exponentially with E, as $d(E) \approx A \cdot E^b \cdot e^{\beta_0 \cdot E}$ for some values A, b, and $\beta_0 > 0$?

In this case, the value $E(\beta)$ decreases with β.

- For $\beta = \infty$, we have $E(\beta) = 0$.
- For $\beta > \beta_0$, the integral

  $$\int E \cdot d(E) \cdot e^{-\beta \cdot E}\,dE = \int E \cdot A \cdot E^b \cdot e^{-(\beta - \beta_0) \cdot E}\,dE \tag{38}$$

  converges and thus, the value $E(\beta)$ is well defined.
- However, for $\beta < \beta_0$, the above integral diverges and thus, such values β are not possible.

If the integral converges for $\beta = \beta_0$, i.e., if

$$\int E \cdot d(E) \cdot e^{-\beta_0 \cdot E}\,dE \approx A \cdot \int E^{b+1}\,dE < +\infty \tag{39}$$

(which happens when $b + 2 < 0$), then the corresponding value $E(\beta_0)$ is the largest value that can be attained for a stable (maximum-entropy) state.

So, if the actual energy e_0 is larger than $E(\beta_0)$, then for this value of energy, there is no maximum-entropy state – and thus, heat death is impossible.

Possible physical meaning of the above energy density. The energy density $d(E) \approx A \cdot E^b \cdot e^{\beta_0 \cdot E}$ with $b < -2$ was actually considered in physics in the so called *bootstrap theory* (see, e.g., [2, 7] and references therein). This theory was based on the assumption of *particle democracy:* crudely speaking, no particle is truly elementary, every particle can be viewed as consisting of others. This theory correctly predicted the empirical distribution of known elementary particles – which indeed has the form $d(E) \approx A \cdot E^b \cdot e^{\beta_0 \cdot E}$.

Researchers developing this theory noticed that in this theory, temperature is bounded by a value corresponding to β_0 – the features that we used in our arguments.

The original bootstrap theory turned out to be too simplified to be correct, but it served as an important first approximation for developing the modern approach of string theory [7].

This physical meaning leaves a possibility that the actual energy spectrum indeed has the form $d(E) \approx A \cdot E^b \cdot e^{\beta_0 \cdot E}$ for which the heat death can be avoided.

Acknowledgments. The authors are thankful to Sa-aat Niwitpong for his help.

References

[1] R. Feynman, R. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Addison Wesley, Boston, Massachusetts, 2005.

[2] S. Frautschi, Physics at the origin of time, *Annals of New York Academy of Sciences*, 1973, Vol. 224, pp. 355–364.

[3] S. Frautschi, Entropy in an expanding Universe, *Science*, 1982, Vol. 217, pp. 593–599.

[4] V. Kreinovich, *A note on limit temperatures*, Technical Report, Informatika Research Center, St. Petersburg, Russia, 1989.

[5] V. Kreinovich, *A model of statistical bootstrap*, Technical Report, Informatika Research Center, St. Petersburg, Russia, 1989 (in Russian).

[6] V. Kreinovich, *Bootstrap and synchrotron radiation*, Technical Report, Informatika Research Center, St. Petersburg, Russia, 1989 (in Russian).

[7] L. Smolin, *The Life of the Cosmos*, Oxford University Press, New York, 1997.