Entropy growth and information gain in operating organized systems

Cite as: AIP Advances 9, 125041 (2019); https://doi.org/10.1063/1.5128315
Submitted: 18 September 2019 . Accepted: 05 December 2019 . Published Online: 26 December 2019

Guibert U. Crevecoeur

© 2019 Author(s).
Entropy growth and information gain in operating organized systems

Cite as: AIP Advances 9, 125041 (2019); doi: 10.1063/1.5128315
Submitted: 18 September 2019 • Accepted: 5 December 2019 • Published Online: 26 December 2019

Guibert U. Crevecoeur

AFFILIATIONS
FPS Economy–DG Energy, Brussels B1000, Belgium

E-mail: guibert.crevecoeur@hotmail.com

ABSTRACT
In the framework of thermodynamics of irreversible processes, patterns of macroscopic evolution of operating organized systems from various fields (engineering, biology, cosmology) are coupled to the increase in their entropy. An extension of Boltzmann’s equation is proposed to characterize the entropy evolution. It is shown that such a “top-down” approach allows us to merge empirical data in a single inclusive model. A method is proposed to quantitatively assess the minimum semantic information gained during the life of the systems. This allows us to compare systems with different types of organization and lifespans. An example of calculation is given for the universe. The method also offers a challenging view to “bottom-up” approaches in progress.

I. INTRODUCTION
The thermodynamic concept of entropy was introduced by Clausius. It is defined for isolated systems at equilibrium. According to the second law of thermodynamics, spontaneous changes in entropy can never be negative. Boltzmann showed that the entropy of the macroscopic configuration (macrostate) of an isolated system at equilibrium was proportional to the logarithm of the number of microscopic configurations (microstates) consistent with it. Gibbs developed a relationship between entropy and the logarithm of the probability distribution of the microstates. Shannon linked the concept to information. Information entropy is related to unknown quantities that can be described by probability distributions. It was further noted that an entropy increase could theoretically occur at no energy cost.

Nowadays, there is still no definition of the entropy of open systems (in nonequilibrium states). In addition, thermodynamics gives no attention to the time evolution of entropy. Operating organized systems are open and evolve during their operation. Attempts have been made to extend the concept of entropy to operating organized systems. Schrödinger postulated a local decrease in entropy in living organisms, thanks to their nonrandom distribution of microstates. He first called this “negative entropy” (renamed “negentropy” by Brillouin and followers) but admitted later in a footnote that the wording was not appropriate. Prigogine showed that the rate of production of entropy was decreasing toward a minimum in irreversible processes. More recent research refers to fluctuation theorems (an approach from microscopic to macroscopic). Jarzynski derived a manner to obtain equilibrium free-energy differences from nonequilibrium measurements in a “bottom-up” approach: “As we turn our attention to smaller systems, however, statistical fluctuations become more prominent . . . a proper accounting of fluctuations allows us to rewrite familiar inequalities of macroscopic thermodynamics as equalities.” Other fields of development include stochastic thermodynamics (Markovian processes), ergodic theory, Kolmogorov entropy, and black-hole entropy.

Thermodynamics and cybernetics are at the roots of this article. Cybernetics deals with “all forms of behavior in so far as they are regular, or determinate, or reproducible. The materiality is irrelevant.” It will therefore often be referred to as feedback loops.

This article analyzes how to quantitatively assess the information gain in operating organized systems.

II. TOWARD A DEFINITION OF ENTROPY GROWTH IN OPERATING ORGANIZED SYSTEMS
An operating organized system is a group of interlocking parts operating together with a common goal (based on Ref. 26). It is
open and shows internal organization with feedback loops inside it and with its environment. Various examples of feedback loops in operating organized systems are given in Refs. 26–39. Before starting to operate, the systems considered have reached a given state of organization. This is the result of the previous assembly due to self-organized criticality (SOC), 52,53 or natural evolution, or human design, etc. A car, an engine, a computer, a biological cell, a star, a galaxy, and the universe itself are operating organized systems.

Phenomena far from equilibrium such as Bénard cells, 43 Rayleigh–Bénard rolls, 44 or Belousov-Zhabotinsky structures 45 show ordered behavior under physical constraints but are not operating organized systems. Their order vanishes with the constraints, and they do not show operating life.

An operating organized system may involve intelligent agents, but not necessarily—as sometimes required for complex adaptive systems (CASs). For instance, expressions like the following are used for CASs: the systems are characterized “by nodes (agents) that make decisions based on past actions and on new input” 16 or in the cases of social science and ecosystems, “partly independent operators work together in a network and give rise to an emergent complex behavior.” 20 An operating organized system anyway adapts to external constraints and/or internal stressors, thanks to its organization. Its history has an influence on its further evolution. Operating organized systems are usually nonergodic, 36,46 they operate (“live”) in a changing environment up to “collapse” (meaning that the system is no longer in an organized state allowing it to operate, thus equivalent to being “dead”); examples are given in Refs. 50 and 51.

Here, we focus on the following operating organized systems: creeping metals, 26–28 thermal equipment of a power plant, cars, computers, living organisms (where favorable mutations can happen 7 and the law of requisite variety applies 27,47,48 ), intermediate organisms (viruses, bacteriophages, 59 tardigrades, 60 robots), control systems in living organisms (immune, nervous, circulatory systems, . . . , general adaptation syndrome, 57,58 evolution of species, and the universe. For the universe, the theory of force is the cosmological “Λ-CDM model” (“Lambda–Cold Dark Matter” model) based on general relativity. It assumes that the universe is a closed system governed by internal forces with gravitation binding all masses together. 60–62 However, hierarchical structure formation is observed, 60 and gravitation may play an important role in increasing its entropy as gravitation makes sparse matter condense into stars which would end as black holes. 62

More details on operating organized systems can be found in Ref. 25.

In the following, information can be any kind of signal: physical, chemical, biological, structural, and procedural. It is not Shannon’s computational information. In contrast with Shannon’s information which “reflects the amount of statistical correlation between systems” (called “syntactic information”), several authors draw attention to another kind of information called “semantic information” that refers to “those correlations which carry significance or “meaning” for a given system.” 60” They recall that semantic information plays an important role in many fields, “including biology, cognitive science, and philosophy,” and that there has been a long-standing interest in formulating a broadly applicable and formal theory of semantic information. It is this kind of information we are focusing on here.

In their recent article, Kolchinsky and Wolpert introduce such a theory of semantic information. They propose a “bottom-up” framework “grounded in the intrinsic dynamics of a system coupled to an environment” and “applicable to any physical system, living or otherwise.” They therefore define semantic information as “the syntactic information that a physical system has about its environment which is causally necessary for the system to maintain its own existence.” 60

Here, we extend this definition to the whole information needed for the system to “maintain its own existence” (i.e., to operate) and also to the “meaningful” information created by the system during its own existence (which may be useful otherwise, e.g., for other systems, like information for devising structures, objects, tools, or new types of operating organized systems, for teaching, for inventing theories, etc.). We propose this extension because at the present stage, tools are lacking to distinguish between the two. A full theory of semantic information is still in its early stages. Notwithstanding this, we observe that meaningful information is gained for the operating organized system to operate—as compared to a heap of sparse parts—and is created by the system during its life. The total semantic information for a system would then be the sum of all information. We therefore take any increase in this information to be “information gain.”

Now as mentioned in the Introduction, there is still no definition of the entropy of open systems, and thermodynamics pays no attention to the evolution of entropy. As the following text is devoted to operating organized systems (which are usually open) and to their entropy evolution, we make the following assumption: the use of Boltzmann’s equation—Eq. (1) hereafter—can be extended to the description of the evolution of operating organized systems.

This allows then to provide background to a “top-down” theory based on plenty of observational data from different fields which have already been analyzed. It is shown how integrating the data in a general framework leads to the assessment of the semantic information gain.

A modern way of presenting entropy is to say that it measures the quantity of microscopic details that are lost when moving from a description at the atomic scale to a macroscopic description. Entropy allows us to measure the uncertainty about what occurs microscopically or at scales many orders of magnitude smaller than the observation scale. It is a coarse grain property.

In statistical thermodynamics, the entropy is given by the natural logarithm of the number “W” of microstates of a system consistent with its macrostate (at Boltzmann’s constant),

\[ S = k_B \ln W. \]  

This exactly squares with the thermodynamic entropy for the special case of a closed system with an ideal gas at equilibrium (with an equally probable occupation of the microstates). However, for anything else, including a closed system with a real gas at equilibrium, it leads to wrong predictions as the interactions between subparts and molecules are not taken into account.

To assess “W” for an operating organized system, one would divide the phase space of microstates into boxes of different sizes, each box corresponding to a macrostate. The more consistent the microstates are with the macrostate, the bigger the box. 63 It is like for a photograph: one accepts more (fine grain) or less (coarse grain) accuracy of details of the macrostate.
Now the system is not at equilibrium. It operates in time. When the system starts operating for the first time, it is in a microstate corresponding to its macrostate. Immediately afterward, during the first increment of time \( dt \), the system operates. It goes through a sequence of new microstates in the phase space. The number "\( W \)" increases.

One has "\( W(t) \)" entailing an entropy evolution "\( S(t) \)" that in a coarse grain view can be assumed to be continuous,

\[
S(t) = \int_0^t \dot{S}dt + \text{const.} \tag{2}
\]

And Eq. (1) becomes \([S_E(t) \text{ stands for "evolution entropy"}]

\[
S_E(t) = k_B \ln(W(t)) \tag{3}
\]

with \( W(t) \) giving the number of microstates consistent with a given macrostate (a given box) at time "\( t \)." The system successively explores boxes of different dimensions, each box corresponding to a macrostate. With high coarse graining, one could just have 4 boxes, e.g., for a living organism: youth, adulthood, senescence, and death. With more fine graining, there would be more boxes to explore. The entropy increase in the different boxes is added by virtue of the logarithm (entropy is an extensive property): the evolution is cumulative. In line with the second law of thermodynamics, it proceeds toward less organization (more uniformity and more "disorder").

We will now examine how operating organized systems evolve.

III. EVOLUTION OF OPERATING ORGANIZED SYSTEMS

The approach followed here is "top-down." Such an approach is in line with the way thermodynamics was created.

It is an established empirical feature that the pattern of evolution of operating organized systems can be expressed by an equation of the kind \( t^\beta e^{\alpha t} \), i.e., combining a power law with an exponential cutoff.\(^{67-70}\)

\[
E(t) = k' e^{\alpha t}(t + 1)^\beta \tag{4}
\]

with the following:

- \( \alpha = 1/t_i \) (where \( t_i \) is the "critical time"\(^{71}\) of the system, i.e., the time from which the system could become liable to collapse)
- \( 0 < \beta < 1 \)
- \( k' = \text{constant} \)
- \( (t + 1)^\beta \approx t^\beta \) to avoid infinites in some computations with logarithms.\(^{72}\)

The shape of the curve given by Eq. (4) is shown in Fig. 1 (with the following parameters taken: \( k' = 0.00065, \alpha = 0.0002, \beta = 0.15, \) and time \( t \) in arbitrary units, for the sake of visualizing in Figs. 1–6).

This is a coarse grain view of what occurs microscopically. The curve shows the following three steps for the aging (or evolution in time) of the system:

1. "Youth": The cumulative number of changes (adaptations/defects) increases first quickly and then at a decreasing rate (the first part of the curve: concave from below).
2. "Active life": The cumulative number of changes increases at a minimum rate (the second part of the curve: quasilinear).
3. "Senescence": The cumulative number of changes increases at a quasixponential rate (the third part of the curve: convex from below) up to collapse.
Equation (4) also gives the cumulative number of failures (or the cumulative hazard rate) of a modified 2-parameter Weibull probability distribution with parameter \( \beta < 1 \) and \( k' \) containing the scale-parameter. This modified Weibull distribution has its cumulative distribution function given by

\[
F(t) = 1 - \exp\left(-e^{\alpha t}\left(t \gamma / \beta\right)\right)
\]  
(5)

instead of

\[
F(t) = 1 - \exp\left(-\left(t / \gamma \right)^{\beta}\right)
\]

(6)

for the standard 2-parameter Weibull distribution. This means that the exponent “\(-\left(t / \gamma \right)^{\beta}\)” has been replaced by “\(-e^{\alpha t}\left(t \gamma / \beta\right)\),” which is the same as “\(-E(t)\)” in Eq. (4). The MLE (“maximum likelihood estimation”) for such modified Weibull distribution has exactly been computed for both uncensored and right-censored data. The Weibull distribution is found back by making \( \alpha = 0 \) (\( e^{\alpha t} = 1 \)).

This modified 2-parameter Weibull distribution also allows us to fit bathtub shaped curves of the kind shown in Fig. 2 and reliability curves of the kind shown in Fig. 3.\(^{74-75}\)

As examples of cases where Eq. (4) and the curve in Fig. 1 are found, one may quote the following: creep curves in metals (see textbooks and standards on creep),\(^{52,76}\) laboratory creep tests,\(^{55}\) electronic computers,\(^{78}\) diesel engines,\(^{79}\) power plants,\(^{80}\) chemicals,\(^{81,82}\) biological organisms,\(^{83-87}\) phylogenetic depth,\(^{88}\) evolution of species over the last \(600 \times 10^6\) years,\(^{89,51}\) and scale parameter of the universe.\(^{70,90}\) It must be emphasized that Eq. (4) and the corresponding curve are a coarse grain view of how the universe evolves. A fine grain view would include the “microscopic”—at the universe’s scale—organization of the universe with its planets, stars, constellations, clusters, superclusters, etc. as reflected by the cosmic microwave background radiation.\(^{62}\) See also Ref. 63.

The derivative of Eq. (4) gives the rate of production of changes in function of time,

\[
E(t) = \frac{dE(t)}{dt} = E(t)\left(\alpha + \frac{\beta}{t + \alpha}\right).
\]

(7)

The shape of the curve is shown in Fig. 2. Curves with such a shape are called “bathtub curves” in reliability analysis. In line with the curve in Fig. 1, the curve in Fig. 2 shows the following three steps for the aging (or evolution in time) of the system: (1) “youth” (decreasing rate of production of changes); (2) “active life” (minimum rate of production of changes); and (3) “senescence” (increasing rate of production of changes).

Bathtub curves are of current use in the reliability analysis of mechanical and electrical devices. One speaks of the “failure (or hazard) rate.” For living organisms, one would rather speak of the “mortality curve.”

As examples, one may quote the following: all kinds of “mortality curves” giving the rate of deaths in a population of living organisms of given categories in function of age, e.g., humans,\(^{97}\) cells such as lymphocytes,\(^{97}\) and AIDS,\(^{94-97}\) and the bathtub curve in reliability analysis of mechanical and electrical devices,\(^{98}\) foam glass under compressive stress,\(^{99,100}\) and nuclear power plants.\(^{101}\)

The exponential of minus Eq. (4) gives the “reliability” in function of time (as usual in reliability applications),

\[
R(t) = e^{-E(t)}.
\]

(8)
The shape of the curve is given in Fig. 3. In line with the curves in Figs. 1 and 2, the curve shows the following three steps for the aging (or evolution in time) of the system: (1) "youth" (high and then slowing down the decrease in survival expectation corresponding to "infant illnesses": convex curve from below); (2) "active life" (sustained quasilinear decrease in survival expectation); and (3) "senescence" (accelerated decrease in survival expectation: concave curve from below).

Examples are reliability of mechanical, electrical, and electronic systems, 1 reliability of lighting systems, 2,3,4 and all kinds of "survival" or "survivorbility curves (type 1)" in a population of living organisms of a given category in function of age, e.g., humans, 5,6 nematodes, 7 and cells. 8

Finally, the exponential of minus Eq. (7) also gives often observed curves,

$$D(t) = e^{-CE(t)}.$$  (9)

The shape of the curve is shown in Fig. 4 (with C taken as 100 000). In line with the curves in Figs. 1–3, the curve shows the following three steps for the aging (or evolution in time) of the system: (1) "youth" (quickly increasing resistance toward a maximum); (2) "active life" (sustained high resistance near the maximum); and (3) "senescence" (slower and then accelerated the decrease in resistance). Note that the shape of the "hill" can be different in function of the value of $C$, but there is always a sequence: an increase—plateau (maximum)—and a decrease.

Examples are the proliferative capacity of cells, 9 constant strain rate tensile tests, 10 maximal oxygen uptake in men, 11 memory scores, 12,13 bones, 14 muscles, 15 human serum, 112,113 brain in men (when started from time zero), 115 and the general adaptation syndrome. 12,57,58

IV. INFORMATION GAIN IN OPERATING ORGANIZED SYSTEMS

One may assume that operating organized systems will be in a state of lower entropy at the start of operation because of their internal organization. Due to their complexity, they have a relatively small number of microstates consistent with the observed macrostate. For instance, it is admitted that an operating car has low entropy ($S = k_B \ln 10^{10}$) compared to a disassembled car with sparse parts ($S = k_B \ln 10^{81}$). 16

As both $W(t)$ and $E(t)$ reflect the global evolution of the system, one may assume proportionality between these parameters and the following relationship: $W(t) = kE(t)$ to be true in the first approximation. Then, Eq. (3) becomes

$$S_t(t) = k_B \ln (kE(t)).$$  (10)

This makes sense as we have seen that $E(t)$ is the cumulative hazard rate of a modified Weibull distribution with $\beta < 1$ (bathtub shaped failure rate model). 7,9 note that the failure probability density $f(t)$ is given by $f(t) = E(t) \cdot R(t)$. And $W(t)$ is the cumulative number of microstates passed through during the evolution of the entropy of the system toward a small structure (second law of thermodynamics).

Using Eq. (4), one finds

$$S_t(t) = k_B \ln (kE(t)) = k_B [\ln (kx^\prime) + at + \beta \ln (t + 1)],$$  (11)

which corresponds to the plain curve in blue in Fig. 5.

In Fig. 5, the curve of $S_t(t)$ has been started from a value of "0" at time "0," thus concentrating on "at + $\beta \ln (t + 1)$." This makes sense as one is interested in entropy differences rather than in absolute entropy values. Similarly, it is not necessary at this stage to know the values of the constants.

The plain straight line in red would then give the corresponding increase in entropy of the system if it were in a nonorganized state, i.e., no longer operating, just the set of sparse parts without interactions (no functional feedback loops). There is just a random statistical increase in the number of microstates consistent with the set of atoms (due to decomposition, degradation, chemical attack, corrosion, erosion, etc.). This makes sense as we know that (1) the entropy must increase and (2) it must join the curve for the operating organized system at $t_c$ with a common tangent to allow for smooth transition. It could also join the curve later if the system would continue to follow Eq. (11) for a while after $t_c$. However, the information gain would be higher (no longer minimum) as the slope of the tangent would be smaller. Nobody actually knows what kind of curve could express the evolution of the entropy of a set of various sparse subparts from a fully damaged operating organized system in permanent contact with the environment while being sufficiently general to cover most types of such systems. Such a curve probably does not exist. The simplest way to meet conditions (1) and (2) above is then to prolong the tangent at time $t_c$ to time zero. Indeed, (a) a slower increase than linear would mean that the sparse parts are still partly organized or even better organized than the operating organized system itself which is nonsense, while (b) a quicker increase than linear, e.g., growing exponential, is always possible but the start of it could always be approximated by a straight line at least for the period from time zero to $t_c$ (looking "old" before $t_c$ would be the exception). Also in case of an oddly shaped entropy growth, it should anyway be averaged by a straight line in the first approximation. Therefore, the straight line with a positive slope and tangent (at time $t_c$) to the curve for the operating organized system would be the best assumption to allow for assessment of the minimum information gain.

The operating organized system pursues its evolution following the blue plain curve until the time $t_c$ which we have called the "critical time" because it is the moment from which the system is liable to collapse. This moment is different in function of the class of operating organized systems under scope, but in the first approximation before a better assessment is made, it can be taken as $t_c = 1/\alpha$ [see Eq. (4) and Ref. 71]. After $t_c$, the risk of collapsing increases, and given the large number of possible microstates of disorganization consistent with a macrostate reflecting the global level of senescence $E(t)$ reached by the non fully organized system, one would expect that the probability distribution of the corresponding ages in a class of systems of the same type would progressively approach a normal distribution. 6,25

The blue plain curve and the red straight line would join at the critical time. As mentioned before, the straight line is obtained from the tangent to the plain curve at $t_c$,

$$S_{\text{non coas}}(t) = mt + b,$$  (12)

with $S_{\text{non coas}}(t)$ being the entropy of the set of sparse subparts of the no longer organized system ("non coas" stands for the "nonoperating
organized system”), “m” being the slope of the tangent (as computed), and “b” being the value obtained by extrapolation to time zero.

The minimum information gain is given by the shaded area between the red straight line and the blue plain curve up to \( t_i \). The evolution of the rate of production of entropy is given for the operating organized system by

\[
\frac{dS}{dt} = k_B \frac{1}{E(t)} \frac{dE}{dt} = k_B \left( \alpha + \frac{\beta}{t+1} \right).
\]

The corresponding curve is the plain curve in blue shown in Fig. 6. Interestingly, it decreases toward a minimum (in line with Prigogine’s theorem that the rate of production of entropy goes toward a minimum in near to equilibrium irreversible processes), but it is not negative as the parameters at the right are all positive.

The dotted red curves in Fig. 6 are examples of evolutions of the rate of entropy production after \( t_i \) to fix ideas. This is to be understood in connection with Ref. 71. Because a zone of likelihood to collapse can settle, it might be that Eq. (13) is no longer valid after \( t_i \) for most systems encountered in practice. Entropy then starts to increase quicker than if the system would have still maintained its organization. The curve for the rate of production of entropy then increases in a more or less accelerated way. The two left curves are examples of such a behavior. On the contrary, the right curve shows that the system first continues to follow Eq. (13): it still behaves like an organized system for a while. However, later on, it starts departing from it due to being in a zone of likelihood to collapse.

Kaila and Annila use Eq. (14) expressing the rate of entropy production as a differential equation of motion. The rate of entropy production is made dependent on a propagator \( \beta \), \( \alpha \) \( \dot{S}(t) = \frac{dS(t)}{dt} = k_B \frac{1}{P} \frac{dP}{dt} = k_B L \), with \( P \) being the probability of motion.

Equation (14) is alike Eq. (13) though developed to specifically reflect the evolution of species by natural selection. The authors deduce that \( dS/dt > 0 \) “until the gradients have vanished and a stationary state \( dS/dt = 0 \) has been reached.” Here, one would say that, according to Eq. (13), the stationary state will rather be given by \( dS/dt = k_B a \) that is close to zero but not nil. However, in practical applications, this stage is usually not reached as the chance that the rate of entropy production will increase again after the critical time \( t_i \) is high (see, e.g., the three dotted red curves in Fig. 6).

V. DISCUSSION AND COMMENTS

A. Capacity of entropy and information gain

The developments above allow one to understand why information is gained during the process despite the accumulation of microscopic defects with time. This is roughly said because “the rate of production of defects is going toward a minimum.”

The minimum gained information is given by the area between the plain straight line in red and the plain curve in blue in Fig. 5. This may be put in the relationship with Gibbs’s “capacity for entropy,” i.e., the amount of entropy that may be increased without changing the internal energy or increasing the volume. It is the difference between the maximum possible entropy of the system and its actual entropy. This is equivalent to information gain. That information gain is the difference between maximum entropy and actual entropy is also a conclusion of Layzer. In a similar way, Landsberg has shown that the disorder (which he defines as the entropy divided by the system’s maximum possible entropy) of the universe, seen as a system, decreases as the system increases the number of its microstates as long as the entropy increases less quickly than the maximal entropy.

B. Information gain and creation of complexity

Now the information gained by an operating organized system during its operation can be at the origin of the creation of new operating organized systems. We humans use the information we have acquired during our life to build sophisticated machines, computers, robots, etc. It is a whole process of production of information allowing among other things to create new operating organized systems. These will by their operation produce additional information allowing us to create new operating organized systems and so on (we learn from their working). Also the first operating organized system producing gain of information must then be no less than the universe itself.

One may question some theoretical reasons (anthropic principle, engine of complexity, etc.) that have been proposed to explain why complexity seems to increase for living organisms over evolutionary periods. Our view is that there is no general trend toward complexity. Only more complex solutions may sometimes arise from local adaptations to challenges with time going (time is available at leisure on evolutionary scales). This is alike a “left censored random walk” as described by Gould. Adaptations are at the roots of what is referred to as “punctuated equilibria” in the evolution of species. The punctuated equilibria occur at smaller scales (fine grain view). At higher scales, they give the feeling of an “uninterrupted evolution,” but it is just a coarse grain view.

C. Information gain and negentropy

When an operating organized system starts operating, it is in a state of low entropy: it shows internal organization. This organization allows it to adapt to environmental challenges inherent in the fact of operating. It trades upon high quality (low entropy) sources of energy and matter to perform its operational works and gives back to the environment low quality (high entropy) energy (heat, etc.) and waste (gases, waste matter, etc.). For living organisms too, we can say with Lehninger that “Living organisms preserve their internal order by taking from their surroundings free energy, in the form of nutrients or sunlight, and returning to their surroundings an equal amount of energy as heat and entropy.”

The global entropy of the operating organized system is always positive and increases when the system operates. There is no “negative entropy” nor “negentropy,” including for living organisms. There is no decrease in entropy during the life of the system as assumed, e.g., in Ref. 126 (“...it is postulated that all living systems are characterized by a continuously, decreasing total entropy level”). One would rather speak of the decrease in the rate of entropy production. Similarly, when it is postulated that “biological death occurs
at some minimum total entropy value, one should rather correct this statement by saying that death occurs after some minimum rate of entropy production has been reached (i.e., at least after the time \( t \)). Indeed, death at minimum total entropy is equivalent to death when the organization is maximum.

Concerning “information entropy,” what Shannon’s entropy measures is the lack of information (important for reliable communication). When this information is “known,” the information entropy diminishes, while the thermodynamic entropy increases accordingly (because of more predictable behavior, the formerly lacking information is now “known,” e.g., statistical results about mutation rates in biological organisms, statistical data such as the mean time between failure for mechanical or electrical devices, etc.). The decrease in information entropy may be called “negentropy,” but it is information gain. The observed increase in information and complexity seemingly accompanying entropy growth is due to the decrease in the rate of production of entropy during the operation of the system which offers an opportunity for information gain.

D. Entropy evolution of an operating organized system

The question may be raised whether the actual evolution of the entropy of an open operating organized system during its operation could have in its broadest generality another shape than the one given by Eq. (10).

Let us assume that we would have \( \Omega(t) \neq \kappa \cdot E(t) \) under the logarithm

\[
S_\Omega(t) = k_B \ln(\Omega(t)),
\]

with \( \Omega(t) \) being the number of microscopic configurations consistent with the macroscopic configuration.

Then, the curve for the rate of production of entropy given by

\[
\dot{S}_\Omega(t) = \frac{dS_\Omega(t)}{dt} = k_B \frac{\Omega(t)}{\Omega(t)}
\]

would have a different shape than the one shown in Fig. 6, but anyway decreasing toward a minimum as required by Prigogine. The evolution of \( \Omega(t) \) would then be

\[
\Omega(t) = e^{\alpha t},
\]

with \( \alpha(t) = \int_0^t \left[ S_\Omega(t) \right] dt \) and be reflected macroscopically by empirical curves different from \( E(t) \) and its derived parameters as given in Figs. 1–4. This would be the general occurrence, and the curves in Figs. 1–4 only are the exception.

But available observations show that it is not the case. The empirical curves which are regularly found with operating organized systems are of the type shown in Figs. 1–4.

The equality \( \dot{S}_\Omega(t) = \kappa E(t) \) should thus be taken in the first approximation as the best one as long as none more general one (covering more observations) is found and a better explanation is given to the curves in Figs. 1–4.

Another way to understand this is to reverse the statement and say that \( E(t) \) and its derived equations reflect the entropy evolution \( S_\Omega(t) \) of the operating organized system during its operation,

\[
E(t) = \frac{1}{\kappa} e^{S_\Omega(t)/k_B}.
\]

It is because Eqs. (3) and (10) exist that the curves in Figs. 1–4 are observed for operating organized systems of various kinds as found in different fields.

E. Thermodynamic entropy of closed systems as the possible limiting case

It is known that the entropy of closed systems with ideal gases subject to experimental constraints goes toward a maximum. This is, e.g., seen by measuring the entropy growth—or simulating with a computer the increase in the number of indistinguishable macrostates—when two ideal gases are progressively mixed: the entropy or the number of indistinguishable macrostates stabilizes after a while showing a global logarithmic curve. One then obtains something very alike the beginning of the plain curve in Fig. 5: it is a logarithmic growth as given by the dotted curve.

In operating organized systems, each change in molecule (or atom, etc.) at any increment of time can become the source of further changes in other neighboring or related molecules. There is a cumulative (“snowball”) effect. This explains the factor “\( e^{\alpha t} \)” in Eq. (4) and the plain curve in Fig. 5.

Now, one can have only the first part of Fig. 1 (start of life) for operating organized systems too, under specific conditions of limited constraints or when the constraints stop acting sufficiently soon.

Then, \( \alpha \) can be made 0 and Eq. (4) becomes

\[
E(t) = k'(t + 1)^\beta.
\]

Similarly, Eq. (11) becomes

\[
S_\Omega(t) = k_B \ln(kE(t)) = k_B \left[ \ln(k' = k + \beta \ln(t + 1) \right].
\]

This is a pure logarithmic evolution alike the one obtained, e.g., when simulating the increase in the number of indistinguishable macrostates when two ideal gases are progressively mixed, as mentioned above.

The entropy evolution of operating organized systems could then be seen as approaching the behavior of more simple laboratory systems when \( \alpha = 0 \) in Eq. (11). The relevance and generality of such similarities must be further studied.

Examples of evolution only in the first stage can be found in various fields, e.g., for the creep of steels when creep takes place under a threshold temperature around 0.4 \( T_m \) (\( T_m \) = melting temperature in K), for learning in general (the “learning curve”), learning as reliability growth of devices and machines or punted equilibrium in biology. One may also include the fact that, at the second stage of resistance of Selye’s general adaptation syndrome, the body’s reaction tends to oppose a resistance to the stress to maintain physiological equilibrium: if the resistance is successful, the stress may disappear and the resistance stage stops.

In the same order of ideas, it is interesting to note that oncologist Lawenda wrote the following as a comment in 2013: “Theoretically (although I am convinced based on my clinical experience), if your stress hormone levels remain elevated for extended periods of time (years-to-decades), the ability of your adrenals to make cortisol and DHEA (dehydroepiandrosterone) can be compromised. If this stress continues, the high levels of cortisol and
DHEA begin to drop as the adrenals eventually "burn out." At this point, the adrenals become exhausted/fatigued and can no longer sustain an adequate response to stress.\textsuperscript{47,53}\textsuperscript{47} If the stress stops sufficiently soon (before extended periods of time would have been reached), the high levels of cortisol and DHEA would not begin to drop.

F. What kind of information is gained?

One must distinguish between the information already existing (previously gained) in the operating organized system before it starts operating and the gain of information during operation.

The first one is all the information needed for good (safe, reliable, a sufficiently long time, etc.) operation.

Then, there is the information gained by the operating organized system during its operation. The sum of all information, that acquired to operate (compared to sparse subparts without interactions) and that acquired during operation, is the total information given in Fig. 5 and computed in Subsection V G. It would be semantic, but most gained information would first be used by the system itself.

An example is given by plants and trees. They are operating organized systems too. In case of attack, they react, adapt, and warn their neighboring plants and trees. This takes place via air and roots using a subtle mix of chemical (and sometimes electrical) information, see, e.g., the sensitivity of willows to pheromones of other willows emitted as response to attack by tent caterpillars and webworms.\textsuperscript{35} Plants communicate with other species of plants, or with animals. Giant sequoias can live up to 3500 years. They have adapted to fire making it an advantage.\textsuperscript{136} This is information used for the systems to thrive.

G. Example of quantitative evaluation of the information gain: The universe

One would wish to compute how much information is at least gained for an operating organized system during its life (in average). This can be made using the method given in Sec. IV. Here, we present the case of the universe.

The standard cosmological model for the evolution of the universe is the $\Lambda$-CDM model.\textsuperscript{62,63} According to this model, the scale-factor of the universe $R$ evolves, as shown by the plain curve in Fig. 7 (it is generally thought that the universe started at the "Big Bang" 13.8 Gys ago). The dotted line corresponds to the present approach using Eq. (4). We see that, from the Big Bang to the present time, the curves of both models are close to each other.

In order to get the dotted curve, the following values of the parameters in Eq. (4) have been determined: $\alpha \ (\text{Myrs}^{-1}) = 0.000 \ 021 \ 029 \ 7$, $\beta = 2/3$, and $\kappa' = 0.000 \ 65$.

This allows us to compute the (semantic) information gain as follows:

- **Step 1:** Evaluation of the critical time $t_c$.
  
  One finds $t_c = 1/\alpha = 47 \ 550 \ \text{Myrs}$.

  This critical time lies beyond the upper and right ends of the figure at a time of about 33 750 Myrs. At that time, the scale-factor will be higher following the $\Lambda$-CDM model than following the present approach.

- **Step 2:** Evolution of $S_H(t)$ as given by Eq. (11).
  
  As only entropy differences are important here, we put the plain curve (in blue) of $S_H(t)$ to start from a value of "0" at time "0," as shown in Fig. 5. Therefore, $\ln(\kappa')$ is not taken into account for both calculations of $S_H(0)$ and $S_H(t_c)$ (and all intermediate times). The resulting curve has the shape of the plain curve (in blue) in Fig. 5. One gets
  
  \[ S_H(t_c) = 8.18 \ k_B J/K. \]

- **Step 3:** Tangent to the plain curve at $t_c$.
  
  The following values are found for the parameters "$m$" and "$b$":
  
  \[ m = 0.000 \ 035 \ 05 \]
  
  \[ b = 6.513. \]

- **Step 4:** Assessment of the information gain (shaded area in Fig. 5).

The following two methods have been used:

1. By solving the integral $\int_0^{t_c} [S_H(t) - S_H(t_c)] dt$

2. Graphically by subtracting the area under the curve for $S_H(t)$ (after approximating the area by dividing it into small rectangles of width 50 Myrs and adding the areas of the rectangles) from the area under the straight line for $S_{\text{non-coal}}(t)$.

One finally obtains a minimum information gain $I_{\text{gain min}}$ of (1) as 15 826 $k_B \ \text{Jysts/K}$ or nats Myrs/$k_B$ (with the integral—time in Myrs) and (2) as 15 783 $k_B \ \text{Jysts/K}$ or nats Myrs (graphically—time in Myrs).

Because of the logarithm of the time in Eq. (11), the computed value of the information gain is timescale dependent (here Myrs: if the time would have been measured in seconds or units of Planck’s time, the figure would have been different). This must be taken into account when comparing systems from different fields: the same timescale must be used.

The same kind of computation as above can indeed be performed for other operating organized systems such as those quoted as examples in the text under Figs. 1, 2, and 4: power plants ($t_i \sim 40$ years), humans ($t_i \sim 70$ years), creeping steel at 813 K ($t_i \sim 200 \ 000 \ \text{h}$).
creeping steel at 1473 K \((t_I \sim 10 \text{ s})\), etc. It shows the importance of time in the process of gaining information.

H. Comparison with the network approach

A whole field of research has been started recently based on mathematical descriptions of big networks. Operating organized systems (some of which being recorded as complex adaptive systems) are characterized by complex networks of links between subparts. This allows for a “bottom-up” approach where mathematical models are built making hypotheses at different stages to simulate the actual behavior of such systems and predict their future. This approach is mainly useful for big networks such as railway systems, road traffic, Internet, etc.\(^{137}\) However, research has also started using this approach to describe the behavior of industrial installations considered as networks with their subparts (pumps, heating devices, engines, turbines, etc.) as nodes. On the base of the big amounts of times series measurements recorded on the subparts, some emergent behavior can be computed and compared to the global behavior of the installation.\(^{138}\) The same approach can be imagined for other operating organized systems such as living organisms, the universe, etc.\(^{139,140}\) This is of peculiar interest when there is no global behavior which is reflected by the evolution of some parameters as shown in the present article. The network approach then opens promising paths notwithstanding the fact that it also means lengthy measurements, huge data collections to be stocked, and progressive adjustments in function of the feedback from the actual behavior. Both approaches, “top-down” with observed patterns of evolution and “bottom-up” with mathematical models applied to networks, should therefore be considered complementary. The “top-down” approach should be privileged when the patterns of evolution are available to form first ideas at low cost. The “bottom-up” approach would then be introduced further for peculiar cases corresponding to very big networks where no patterns of evolution can be found or for operating organized systems when it is worth investing in continuous monitoring for safety or reliability reasons. For instance, in the case of electricity generating units (see the example of Electrosul in Brazil\(^{9}\)), patterns of evolution have been observed, but because of the need for safe and reliable operation, some continuous monitoring might be necessary. This could then be applied to other similar installations.

In another example of complex industrial installation, the “bottom-up” approach with networks has been proposed.\(^{141,142}\) It shows the appearance of a bathtub curve analogous to the one in Fig. 2 and of an entropy growth curve analogous to the blue plain curve in Fig. 5 as emergent features, see, respectively, Figs. 7 and 6 in Ref. 138. It is the evolution of the information entropy which is shown in Fig. 6 in function of the number of symbols as synthesized by the used model. The curve shows the information loss in function of the symbol number. In the example, the chosen installation was a compressor unit in the chemical industry. It involved four subsystems: a compressor, a reducer, a steam turbine, and a supercharger, and some main air passages and storage tanks.

VI. CONCLUSION

Similar empirical curves reflecting the evolution of operating organized systems from different fields (engineering, biology, cosmology, etc.) are put into a relationship with their entropy growth in a “top-down” approach using Boltzmann’s equation. This allows us to assess the minimum gain of semantic information during the life of these systems. As an example, a quantitative computation is presented for the universe. This allows us to compare systems with different types of organizations and lifespans.

REFERENCES

1. R. Clausius, Poggendorf’s Ann. Phys. Chem. 79, 368 (1850).
2. L. Boltzmann, The Second Law of Thermodynamics (Theoretical Physics and Philosophical Problems) (Springer-Verlag New York, LLC, 1974).
3. E. T. Jaynes, Am. J. Phys. 33, 391 (1965).
4. C. E. Shannon, Bell Syst. Tech. J. 27, 379, 623 (1948).
5. C. E. Shannon and W. Weaver, The Mathematical Theory of Communication, 10th ed. (The University of Illinois Press, Urbana, 1964).
6. R. Landauer, IBM J. Res. Dev. 5(3), 183 (1961).
7. J. Vaccaro and S. Barnett, Proc. R. Soc. A 467(2130), 1770 (2011).
8. B. Derida, Lessons on physical statistics at Collège de France (Paris, http://www.college-de-france.fr/site/bernard-derrida/course-2015-2016.htm).
9. A. Domínguez Alvarez, Información y Entropía. La Medida de lo no Observado (RBA, España, 2017).
10. E. Schrödinger, What Is Life?: The Physical Aspect of the Living Cell (Cambridge University Press, 1944).
11. L. Brillouin, J. Appl. Phys. 24, 1132 (1953).
12. In his note to Chap. 6 of What Is Life?, Schrödinger remarks on his usage of the term “negative entropy”: “Let me say first, that if I had been catering for them [physicists] alone I should have let the discussion turn on free energy instead. It is the more familiar notion in this context. But this highly technical term seemed linguistically too near to energy for making the average reader alive to the contrast between the two things.”
13. L. Prigogine, Introduction to Thermodynamics of Irreversible Processes (Charles C. Thomas Publisher, Springfield, Illinois, 1955).
14. C. Jarzynski, Phys. Rev. Lett. 78, 2690 (1997).
15. C. Jarzynski, Phys. Rev. E 56, 5018 (1997).
16. C. Jarzynski, Annu. Rev. Condens. Matter Phys. 2, 329 (2011).
17. C. Van den Broeck, in Physics of Complex Colloids, Proceedings of the International School of Physics “Enrico Fermi” Course 184, edited by C. Bechinger, F. Sciortino, and P. Ziferbo (IOS, Amsterdam, SIF, Bologna, 2013), p. 155.
18. V. Holubec, Non-Equilibrium Energy Transformation Processes, Springer Theses (Springer, Cham, 2014).
19. P. Billingsley, Ergodic Theory and Information (John Wiley & Sons, New York, 1965).
20. F. L. Comen, S. F. Fomin, and Y. G. Sinai, Ergodic Theory (Springer-Verlag, 1981).
21. A. N. Kolmogorov, Dokl. Russ. Acad. Sci. 124, 754 (1959).
22. J. D. Bekenstein, Phys. Rev. D 7, 2333 (1973).
23. S. W. Hawking, Nature 248, 30 (1974).
24. N. Wiener, Cybernetics or Control and Communication in the Animal and the Machine (John Wiley & Sons, New York, 1948).
25. W. R. Ashby, An Introduction to Cybernetics (Chapman & Hall, London, 1956).
26. J. W. Forrester, Principles of Systems (Wright-Allen Press, 1968).
27. F. Jacob, D. Perrin, C. Sanchez, and J. Monod, C. R. Acad. Sci. 250, 1727 (1960).
28. F. E. Hardin, J. C. Hall, and M. Rosbash, Nature 343, 536 (1990).
29. D. S. Robertson, J. Theor. Biol. 152, 469 (1991).
30. E. M. Jazwinski, Science 273, 54 (1996).
31. C. Ahrens and C. Donald, Meteorology Today, 6th ed. (Brooks/Cole Publishing Co., 1999).
32. S. Camazine, J.-L. Deneubourg, N. R. Franks, J. Sneyd, G. Theraulaz, and E. Bonabeau, Self-Organization in Biological Systems, Princeton Studies in Complexity (Princeton University Press, Princeton; Oxford, 2001).
33. K. V. Venkatesh, S. Bhatiya, and A. Ruhela, FEBS Lett. 563(1-3), 234 (2004).
time is exactly given by the metal leading to quickly reduced cross-sectional area and rupture. And this we include the universe although we do not know whether it has an "environment" (as we are inside). We do it because its global evolution (the evolution of the scale factor) is alike that of operating organized systems.

Chandrasekhar, Hydrodynamic and Hydromagnetic Stability (Clarendon Press, Oxford, UK, 1961).

J. T. Lir and T. F. Lin, Int. J. Heat Mass Transfer 44(15), 2889 (2001).

A. N. Zaikin and A. M. Zhrobatinsky, Nature 225, 535 (1970).

D. L. Abel and J. T. Trevor, Phys. Life Rev. 3(4), 211 (2006).

J. E. Mayfield, The Engine of Complexity: Evolution as Computation (Columbia University Press, 2013).

S. E. Phelan, Syst. Pract. Action Res. 12(3), 237 (1999).

W. R. Ashby, Cybernetica 1(2) (1958).

G. U. Crevecoeur, Quantitative General Adaptation (JePublic, Paris, 2008).

G. U. Crevecoeur, J. Big Hist. 2(2), 79 (2018).

G. Smith, Properties of Metals at Elevated Temperatures, 3rd ed. (Mc Graw Hill Book Co., Inc., New York, 1950).

K. Yagi, G. Merckling, T.-U. Kern, H. Irie, and H. Warlimont, Creep Properties of Heat Resistant Steels and Superalloys (Springer-Verlag Berlin and Heidelberg GmbH & Co, 2004).

B. G. Hall, Genetics 126, 5 (1990).

E. Lederberg, Microb. Genet. Bull. 1, 5 (1950).

Wikipedia, https://en.wikipedia.org/wiki/Tardigrade, viewed August 5, 2019.

H. Selye, Nature 138, 32 (1936).

H. Selye, Stress Without Distress (J. B. Lippincott Co., New York, 1974).

M. F. Bear, B. W. Conners, and M. A. Paradiso, Stress Without Distress (Mc Graw Hill, New York, 1966).

G. U. Crevecoeur, Eur. J. Mech. Engrg. 39(4), 219 (1994).

M. Caubo and J. Mathonnet, Rapports du CNRM RA 711, 1969.

L. O. Bueno and J. R. Sobrinho, Rev. Mater. 17(3), 1098 (2012).

D. R. Cox and P. A. Lewis, The Statistical Analysis of Series of Events (Methuen, London, 1966).

H. E. Ascher and H. Feingold, in Repairable Systems Reliability, Lecture Notes in Statistics Vol. 7, edited by D. B. Owen (Marcel Dekker, Inc., New York and Basel, 1984).

M. Th. Schilling, J. C. G. Praca, J. F. de Queiroz, C. Singh, and H. Ascher, IEEE Trans. Power Syst. 3(2), 490 (1988).

L. De Schepper, L. Tielmans, J. Roggen, and L. M. Stals, European Conference on Advanced Materials and Processes (EUROMAT, Aachen, 1989).

U. S. Agarwal and D. V. Khakhari, Nature 360, 53 (1992).

T. Fotis, Y. Zhang, M. S. Pepper, H. Adlercreutz, R. Montesano, P. P. Nawroth, and L. Schweigerer, Nature 368, 237 (1994).

P. C. Caron and B. R. Unsworth, Mech. Ageing Dev. 8, 181 (1978).

K. Kaji and M. Matsuo, Mech. Ageing Dev. 8, 233 (1978).

H. Vlassara, Age 11, 74 (1988).

P. Kannouche, A. R. Fernandez de Henestrosa, B. Coull, A. E. Vidal, C. Gray, D. Zicha, R. Woodgate, and A. R. Lehm, EMBO J. 12(22), 6246 (2002).

R. E. Lenksi, C. Ofria, R. T. Pennock, and C. Adami, Nature 423, 139 (2003).

Wikipedia, https://en.wikipedia.org/wiki/Timeline_of_the_evolutionary_history_of_life#Phanerozoe_Eon, viewed August 1, 2018.

G. U. Crevecoeur, Phys. Essays 30(3), 255 (2017).

As usual, a point above a time varying parameter means its derivative in function of time.

Z. Shao, W. Gao, Y. Yao, Y. Zhuo, and J. E. Riggs, Mech. Ageing Dev. 67, 239 (1993).

H. U. Luder, Mech. Ageing Dev. 67, 247 (1993).

D. H. Ho, A. U. Neumann, A. S. Perelshon, W. Chen, J. M. Leonard, and M. Markowtiz, Nature 373, 123 (1995).

J. M. Coffin, Science 267, 483 (1995).

X. Wei, S. K. Ghosh, M. E. Taylor, V. A. Johnson, E. A. Emini, P. Deutsch, J. D. Lilson, S. Bonhoeffer, M. A. Nowak, B. H. Hahn, M. S. Saag, and G. M. Shaw, Nature 373, 117 (1995).

J. A. Levy, Microbiol. Rev. 57(1), 183 (1993).

L. Lee, Technometries 22(2), 193 (1980).

H. Straussen, G. Claeys, G. Heylen, G. Crevecoeur, and C. Mæs, in 22nd European Conference on Acoustic Emission Testing, Aberdeen, UK (Robert Gordon University, 1996).

C. Mæs, A. Van Moffaert, H. Frederix, and H. Straussen, Phys. Rev. B 57(9), 4987 (1998).

N. Meyer, D. Rieck, and I. Tweer, Alterung in Kernkraftwerken, 2nd ed. (Greenpeace, Hamburg, 1998).

V. Wang, S. Gong, and S. Grzybowski, Energies 4, 1362 (2011).
corresponding to the shaded zone before a given system, but we know for sure that it will be at minimum than the one be omitted. We cannot predict how much information will actually be gained in diminishing slopes of the tangent and thus an increasing shaded zone and more liable to collapse (but not necessarily immediately—see Ref. 71). In actual cases, this is an actual minimum as ”m” is ”a” as seen from Eq. (11) when βi becomes negligibly small compared to at and can be omitted. We cannot predict how much information will actually be gained in a given system, but we know for sure that it will be at minimum than the one corresponding to the shaded zone before ti.

This is a minimum as ”m” is found from the slope of the tangent to the curve at time ti, the time from which the operating organized system becomes liable to collapse (but not necessarily immediately—see Ref. 71). In actual cases, most systems will continue to follow the curve given by Eq. (11) which will result in diminishing slopes of the tangent and thus an increasing shaded zone and more information gain. The slopes diminish as the limit value of “β” as seen from Eq. (11) when βln(t + 1) becomes negligibly small compared to at and can be omitted. We cannot predict how much information will actually be gained in a given system, but we know for sure that it will be at minimum than the one corresponding to the shaded zone before ti.