

Physical foundations of evolutionary theory

Arto Annala and Stanley Salthe

Communicated by B. Weber, Fullerton, USA

Keywords. Evolution theory, non-Hamiltonian systems, second law of thermodynamics, natural selection.

Abstract

The theory of evolution by natural selection is herein subsumed by the 2nd law of thermodynamics. The mathematical form of evolutionary theory is based on a re-examination of the probability concept that underlies statistical physics. Probability regarded as physical must include, in addition to isoenergetic combinatorial configurations, also energy in conditional circumstances. Consequently, entropy as an additive logarithmic probability measure is found to be a function of the free energy, and the process toward the maximum entropy state is found equivalent to evolution toward the free energy minimum in accordance with the basic maxim of chemical thermodynamics. The principle of increasing entropy when given as an equation of motion reveals that expansion, proliferation, differentiation, diversification, and catalysis are all ways for a system to evolve toward the stationary state in its respective surroundings. Intriguingly, the equation of evolution cannot be solved when there remain degrees of freedom to consume the free energy, and hence evolutionary trajectories of a non-Hamiltonian system remain intractable. Finally, when to-and-from flows of energy are balanced between a system and its surroundings, the system is at the Lyapunov-stable stationary state. The principle of maximal energy dispersal, equivalent to the maximal rate of entropy production, gives rise to the ubiquitous characteristics, conventions, and regularities found in nature, where thermodynamics makes no demarcation line between animate and inanimate.

1 Introduction

The theory of evolution by natural selection [12] is often regarded as the most general description of living nature. Despite its broad scope and de-

scriptive power, Darwin's theory, when without a physical basis provided in a mathematical form, is not amenable for rigorous analyses to draw unambiguous conclusions. In quest of the fundamental form of evolutionary theory, Fisher produced a mathematical model of population fitness increase as dependent upon variation in the fitness of various types in a population [16]. In this formulation, which was inspired by the principle of increasing entropy, fitness must increase in any surviving population, as entropy must everywhere. However, no connection is established to actual physical quantities and relationships. Of course, some may question why should evolutionary theory have a definite physical foundation given in a mathematical form at all? But such doubts would be equivalent to thoughts that evolution would entail some unobservable elements. Namely, every observation, as any other form of interaction, is an energy transduction process that complies with the conservation of energy, which in turn is secured when expressed in mathematical form.

It is no new thought to search for the physical foundations of the evolutionary theory. This task was taken up already by Boltzmann soon after the publication *On the Origin of Species* 150 years ago. Boltzmann understood evolution as a probable process, a likely sequence of events that is far from being miraculous. Therefore Boltzmann placed probability P as the cornerstone of his statistical theory for many-body systems [9]. It is a powerful theory. Irrespective of the system's complexity, the stationary state condition $dP/dt = 0$ at the end point of evolution yields effortlessly, e.g., using the method of Lagrange multipliers, the most probable partition of entities in the system [2]. Despite the ease in obtaining the equilibrium properties, the statistical theory of Boltzmann was limited to conserved systems, thereby encompassing only a part of the classical thermodynamics formulated by Carnot, Clausius, and Clayperon. Since the limits of contemporary statistical physics ultimately stem from the definition of probability, the concept of probability is re-examined in this study. When probability is understood as physical, then statistical physics does also include open systems and allows us to discern that the theory of evolution by natural selection can be viewed as an expression of the 2nd law of thermodynamics.

2 Physical probability

During the course of history the probability concept has been associated on the one hand with combinatorics and on the other hand with conditional

circumstances. The former dates back to Descartes, Fermat, Pascal, and others in the context of gambling, and the latter originates from Bayes in the context of acquiring information [25]. In terms of physics, the Cartesian notion inheres in energetically invariant, stationary systems, whereas the Bayesian view relates to energetically varying, evolutionary systems. Since Boltzmann based his statistical physics on the enumeration of isoenergetic configurations that are commonly referred to as microstates, his theory was limited to conserved systems. However, when circumstantial conditions are included in the probability measure, statistical physics is also applicable to non-conserved systems, such as biological systems, which are open to flows of energy from and to their respective surroundings.

A chemical reaction mixture is an example of an open system that will naturally progress toward the most probable, i.e., the maximum entropy, state. That stationary state is equivalent to the free energy minimum as that depends on the surrounding conditions, as stated by Le Châtelier's principle [8]. A chemical system such as a mixture of molecules in a reaction vessel at an organic chemistry lab or in a living cell may be comprised of a myriad of molecules. Such an arbitrary system is pictured in statistical physics as an energy level diagram (Figure 1) where distinguishable molecules occupy distinct energy levels. Reactions bring about changes in

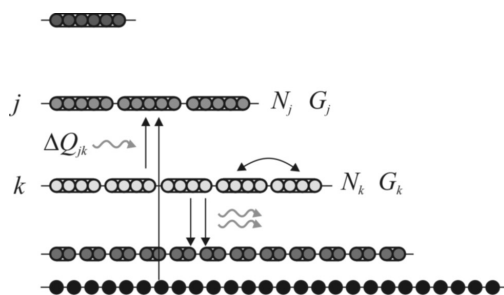


Figure 1. Energy level diagram of a system with a diversity of entities. A pool of j entities (shown in grayscale), in numbers N_j , is associated with an energy density $\phi_j = N_j \exp(G_j/k_B T)$. According to the 2nd law the system evolves from one state to another, more probable one, by diverse dissipative jk transformations (shown as vertical arrows and horizontal wavy arrows) that diminish the energy density differences, i.e., the free energy $\Sigma k_B T \ln \phi_k + \Delta Q_{jk} - k_B T \ln \phi_j$, until the stationary state in the respective surroundings has been attained. At the dynamic steady state, absorption and emission balance each other resulting in only an exchange of entities (shown as a double-ended arrow).

the populations at diverse levels. These transformations from one molecular state to another are depicted as endergonic and exergonic transitions, whereas conformational exchange is isergonic.

The probability P_j for a particular molecule (indexed with j) to exist depends on its substrates, as well as on energy in the surroundings that couples to the transformation from the substrates to the product. When any one vital substrate k is missing entirely ($N_k = 0$), there is no jk -synthesis that could yield the j -product. Likewise, no product will be obtained when the surroundings cannot provide quanta for the endergonic jk -reaction. Conversely, the process will halt when the surroundings cannot accept any quanta from the exergonic jk -reaction. Furthermore, the yield, i.e., the probability of the product will be low when the product is expensive in energy, G_j , in comparison to energies G_k of its substrates. Therefore P_j depends on the difference $\Delta G_{jk} = G_j - G_k$ between the j -products and k -substrates given in the familiar exponential form. That difference can be bridged by the energy influx ΔQ_{jk} from the surroundings that couples to the jk -transformation. These circumstantial considerations for the j -molecule to exist yield the conditional probability [46, 54]:

$$P_j = \prod_k N_k e^{-(\Delta G_{jk} - i \Delta Q_{jk})/k_B T}, \quad (1)$$

where the total energy difference $-\Delta G_{jk} + i \Delta Q_{jk}$ is normalized by the average energy per particle $k_B T$ (or RT per mole). The flux of quanta (bosons) is explicitly denoted by i as orthogonal to the density difference between energy repositories (fermions) that exclude each other in space. In the continuum these two forms of energy are customarily referred to as scalar and vector potentials. In the thermodynamics context their orthogonality is seldom explicitly marked but is familiar, for example, from a situation where a charge that is accelerating in an electromagnetic potential energy gradient is radiating perpendicular to its trajectory. Also in mechanics, the scalar and vector potential energy gradients are referred to as the irrotational and divergence-free components of the force.

In addition to the circumstantial conditions given by eq. (1) the probability depends on the isergonic configurations. The k -substrates that are incorporated in the j -product as indistinguishable (symmetric) copies are numbered by the (stoichiometric) degeneracy g_{jk} . Likewise, identical copies of the j -product are numbered by N_j . These combinations are, as usual, taken into account by factorials [46, 54]:

$$P_j = \left(\prod_k (N_k e^{-(\Delta G_{jk} - i \Delta Q_{jk})/k_B T})^{g_{jk}} / g_{jk}! \right)^{N_j} / N_j!. \quad (2)$$

This nested, self-similar formula means that each k -substrate is considered as a product of some earlier evolutionary process [43]. For example, elements that make molecules are regarded themselves as yields of nuclear reactions in stellar syntheses. In turn, the molecules are considered substrates for cellular assembly processes, and cells are ingredients for individual development, and so on.

Since thermodynamics pictures everything in terms of energy, the scale-independent formula (eq. (2)) that was derived exemplifying chemical reactions specifically, is generally applicable to diverse natural processes. Admittedly, it may appear to some unthinkable to regard abiotic processes as evolutionary, or conversely, to look upon evolving biotic systems as undergoing merely time-dependent physical processes. However, it remains impossible to identify any unambiguous difference between the animate and the inanimate other than the binding of history by the animate in the genetic (DNA) information storage, whereas the past inanimate processes are recorded, e.g., in rocks and sediments. Certainly the biological information allows organisms to assemble mechanisms to access and consume free energy rapidly, but irrespective of kinetic pathways the direction of overall thermodynamic process is toward the free energy minimum.

The total probability P of the entire reaction mixture is obtained by considering reactions over all levels as statistically independent [46, 54] because of the scale differences between the levels [43].

$$P = \prod_j P_j = \prod_j \left(\prod_k (N_k e^{-(\Delta G_{jk} - i \Delta Q_{jk})/k_B T})^{g_{jk}} / g_{jk}! \right)^{N_j} / N_j!. \quad (3)$$

Admittedly, the obtained formula is somewhat cumbersome for computations due to the factorials. To this end, Stirling's approximation $\ln N_j! \approx N_j \ln N_j - N_j$ is applied. When $\ln P$ is multiplied by k_B , the additive measure of the system, known as entropy,

$$S = k_B \ln P = k_B \sum_j \ln P_j \approx k_B \sum_j N_j \left(1 - \sum_k A_{jk} / k_B T \right), \quad (4)$$

is obtained. It is a function of populations N_j and the free energy $A_{jk} = \Delta \mu_{jk} - \Delta Q_{jk}$, referred to also as affinity or exergy. The free energy is the motive force that directs the transforming flow $\partial_t N_j$ from N_k to N_j . The logarithmic density difference is customarily given as the chemical potential difference

$$\Delta \mu_{jk} = \mu_j - \sum_k \mu_k = k_B T \left(\ln \phi_j - \sum_k g_{jk} \ln \phi_k / g_{jk}! \right).$$

In the continuum approximation the two components of the motive force are referred to as the gradients of scalar and vector potentials. For example, in electrodynamics the 2nd law as an equation for flows of energy is Poynting's theorem where a change in scalar potential gives rise to current in an electric field and dissipation as light radiated into the surroundings [21].

The formula of entropy obtained via Stirling's approximation is accurate when N_j is large. Thus, the form of eq. (4) implicitly assumes that S is a sufficient statistic for the average energy density $k_B T$ [31]. In other words, the system defined by a common $k_B T$ would not change substantially in its energy content due to a change in free energy. This assumption may fail for a given $\ln P_j$ when N_j is small, e.g., when a new j -species of particles emerges or one is extinguished. The statistical assumption may also fail when a new route for energy transduction opens up or an old one closes. However, these critical events do not present challenges to the scale-independent formalism as such, because then one inspects the assembly and disassembly processes at a lower level of hierarchy where the subsystems do remain sufficiently statistical. For example, the emergence of a novel mechanism is considered as an evolutionary process where sufficiently statistical constituent systems integrate by mutual interactions toward a common average energy density $k_B T$. When the system has matured to the stationary state, it can be regarded as an invariant mechanism of energy transduction. It is then, in turn, a sufficiently statistical constituent in the evolutionary process that is going on at a higher level of the hierarchy. In other words, when nature evolves and organizes itself as systems within systems, the above formalism bridges levels of the nested dynamical hierarchy [6, 42].

Eventually, at the very lowest levels of hierarchical organization, it may not be possible to dissect the system to yet finer sufficiently statistical subsystems. When even a single quantum brings about substantial changes in the system, entropy is not the best measure because the probability (eq. (3)) will increase in discrete jumps during quantized evolution [54].

3 The equation of evolution

At the thermodynamic stationary state all energy density differences between substrate and product populations have leveled off in the respective surroundings. It is easy to see from the equation of S (eq. (4)) that when the free energy has been all exhausted ($A_{jk} = 0$), the entropy maximum

is simply the sum over all populations, as usual. Although the stationary-state partition can be calculated for fixed energy and entity numbers using the Lagrange multipliers, it is instructive to take the time derivative of entropy to obtain the equation of motion, known as the principle of increasing entropy [46, 54]:

$$\partial_t S = k_B \partial_t \ln P = k_B \partial_t P / P = k_B L = - \sum_{j,k} \partial_t N_j A_{jk} / T \geq 0. \quad (5)$$

The 2nd law of thermodynamics in this form, as an equation of motion, is conceptually simple. It says: energy flows from heights to lows as soon as possible. This imperative for the extremum is also known as the maximum entropy production principle [14, 25, 52, 59], and the principle of least time [41]. The image of energy flowing and naturally selecting the fastest ways to level off energy density differences is valuable in its mathematical form because it can be analyzed to see whether it is in fact the governing law that accounts for the diversity of phenomena. The self-consistent formulation requires that the flow [30, 46],

$$\partial_t N_j = - \sum_{j,k} \sigma_{jk} A_{jk} / k_B T, \quad (6)$$

is proportional to the free energy A_{jk} via a conductance σ_{jk} . The linear form, familiar from Onsager's reciprocal relations [33], is consistent with the notion that the system is sufficiently statistical. Otherwise, a high flow between N_k and N_j would force the conducting system, parameterized by the coefficient σ_{jk} , itself to evolve. Such a case is by no means unusual but is actually characteristic of critical and emergent phenomena, e.g., when a new species appears. That assembly process of sufficiently statistical constituents can then be analyzed at a lower level of hierarchy using the same self-similar formalism. Thus, for example, population genetics does not study changes in organism morphology, but instead studies the frequencies of genes in the gene pool of a population.

The simple equation of evolution (eq. (5)), which can also be concisely written as $\partial_t P = LP$, is surprisingly insightful. To begin with, it cannot be solved. In other words, it cannot be integrated to a closed form because there is no conserved quantity. The open system is either gaining energy from its surroundings or losing energy to the surroundings and hence it does not have a norm. Therefore, there is no unitary transformation to determine eigenvalues (energies) and eigenvectors (trajectories). The equation of motion reveals that the driving forces, i.e., the free energy (A_{jk}) and

the flows ($\partial_t N_j$), depend on each other. Consequently, when there are three or more degrees of freedom, the evolutionary courses are intractable in detail. At best the time-dependent processes can be simulated step-by-step so that potentials and $k_B T$ are updated quanta-by-quanta due to the net energy flux between the system and its surroundings. Since forces and flows are inseparable, the evolving system is non-Hamiltonian. When the potential difference is diminishing, the flow is decreasing, which in turn is reducing the driving force further and so on. Thus, a flock of one species is a stock for another. It is impossible to know the time and place when a particular substrate or any other entity will engage in a reaction or a process because when one does react, forces on others are affected, and evolution redirects its course.

Since nature is in motion not only by way of chemical reactions but also via various transport phenomena such as diffusion, it is of interest to rewrite the equation of evolution also for the continuum (Figure 2). To that end an equation for flows of energy is obtained by multiplying both sides of eq. (5) with T , and replacing the chemical potential difference by the scalar potential gradient and denoting the quantized fluxes to or from the surroundings by the vector potential gradient in the orthogonal direction. The left-hand side $T \partial_t S$ identifies with a change in the kinetic energy [54],

$$\partial_t 2K = - \sum_{x,y,z} (v_x \partial_x U - i \partial_t Q_x). \quad (7)$$

In the continuum form the 2nd law states the conservation of energy. Changes in the potential energy U and dissipation Q are balanced by changes in the kinetic energy $2K$.

Equation (7) can also be derived directly from Newton's 2nd law of motion that relates the force $\mathbf{F} = d\mathbf{p}/dt$ with the change in momentum, $\mathbf{p} = m\mathbf{v}$, by taking the time derivate with respect to both \mathbf{v} and m , then multiplying both sides with \mathbf{v} and finally denoting $m = E/c^2$ and $E = c^2 Q/v^2 = n^2 Q$ where the index $n = c/v$ is the ratio of energy densities in the system and its surroundings. The change in mass is often omitted as insignificant, but then Newton's 2nd law of motion is limited only to conserved systems. Therefore the $\partial_t m$ term is conceptually very important in understanding evolution as an energy transduction process between an open system and its surroundings. The mass change denotes the changes in interaction energy when the system transforms from one state to another. The release of interaction energy as a loss of mass is impressive in nuclear reactions, but loss of mass can also be measured in chemical reactions. Admittedly, in the transport processes dissipation is very small.

Nevertheless, one coordinate differs from another just as one molecule differs from another when a reaction is dissipative. Newton's 2nd law as the equation for the continuity is also known as the Cauchy momentum equation, which is the most general form of the Navier–Stokes equation [1, 4].

The directional derivatives $v_x \partial_x$ and ∂_t (eq. (7)) prompts one to picture the evolving system as a landscape that is leveling off by flows of energy (Figure 2). The flows channel along the paths of least action. That principle [11] is the integrated form of the 2nd law [26]. The evolutionary path, pictured as a directed arc, is given by the change in the kinetic energy, the Rayleigh–Onsager relation [33]. In a locality where the landscape does not curve much, the arc can be approximated by a straight line so that the Pythagoras theorem holds. Then the Lorentz transformation applies and the path can be determined.

The free energy landscape can be given concisely by a field equation. Then each point on the manifold is represented by a vector field which is the gradient of the 4-potential [53]. Generally, a curved but flattening landscape defines an evolving system that is stepping from one stationary state to another by breaking one symmetry for another. Due to the fluxes from the system to its surroundings, or vice versa, the vector field gradients, that is, Lie's derivatives, do not vanish because, for the non-conserved system, the flow (\mathbf{v}) is not collinear along the irrotational force component ($m\mathbf{a}$) but departs by the divergence-free component ($\mathbf{v}\partial_t m$) [49]. In plain language the

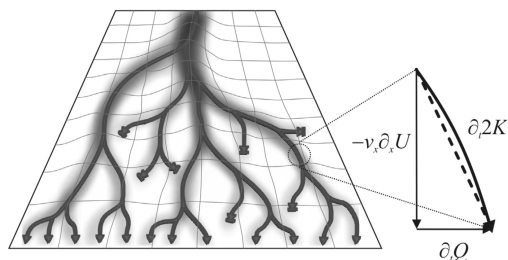


Figure 2. The curved energy landscape represents an evolving system of energy densities. The non-Euclidian manifold is leveling by flows of energy that diminish the energy density differences. The flows direct along the paths of least action to produce the maximal transduction. However the process is intractable when there are alternative paths, as exemplified by branching points, because each flow affects others' driving forces. The expansion of a local region shows how the energy flow, which is approximated by a straight line, represents the change in kinetic energy $2K$, that results from the changes in the scalar U and vector Q potentials.

flow, like a river, is eroding the landscape and thereby redirecting its own course. For a quantized and coherent dissipative process this directionality, which is inherent in time-ordered transformations, is given in the form of a non-vanishing commutation relation $[p, x] = px - xp = -i\nabla$. Its counterpart for a number of decoherent quantized processes is the conservation of energy in action, $2Kt + Ut = Qt$, [54]. Specifically, when evolution has arrived at the thermodynamic stationary state, the landscape is even with conserved currents [39] and the corresponding field equation complies with gauge symmetry [36].

The self-similar balance equation for the flows of energy (eq. (7)), either in its differential or integral forms for the quantized and continuum mechanics, as well as for electromagnetism, shows that the principle of increasing entropy and its equivalent, the principle of least action [11], has been expressed in many mathematical forms of physics [26] – as it should be, since the 2nd law of thermodynamics may perhaps be viewed as the supreme law among all laws of nature. We anticipate that subsequent studies of various evolutionary phenomena as manifestations of the 2nd law will be found to rest on this universal physical principle of maximal energy dispersal extending across scales from microscopic to macroscopic.

4 Exemplary examination of evolutionary phenomena

The physical portrayal of biological evolution by natural selection as a manifestation of the 2nd law of thermodynamics is a most natural tenet because no additional ad hoc elements need to be attributed to the living systems. Whatever the details of macroevolution on a given planet, this process will be one that involves the deployment of species throughout a landscape in such a way as to increase the entropy production from that landscape. Hitherto, it has been difficult to appreciate the explanatory power of thermodynamics because the probability concept has not been fully spelled out to contain both conditional and combinatorial factors. Therefore, the entropy concept has remained poorly understood and even erroneously mixed up with the concept of decoherence, i.e., disorder [45]. Consequently, although life has been seen to be compatible with the laws of physics, the laws of physics have not been seen as conceivable as a determining basis for life [38]. Especially emergent, critical, and non-deterministic phenomena, e.g., the appearance of new species and of chaotic behaviors, have been considered by some to be beyond the conceptual realm of physics. True enough, physics in its reductionist and conserved form is limited to

stationary, deterministic systems, but physics in its full holistic form includes evolutionary, non-deterministic processes. When probability is understood as physical, entropy resolves as a natural statistical measure of a system's free energy status. When the energy fluxes to and from the surroundings are included, the evolutionary processes are found to be inherently non-deterministic, in accordance with observations. Furthermore, the reductionist tenet is understood to fail because the ingredients of emergence include not only the constituents of a system, but also the quanta from its surroundings that incorporate into assembly processes.

From a physical point of view, life in its entirety can be understood by the 2nd law as a thermodynamic process, with no demarcation line between inanimate and animate. Living systems are internally informed dissipative structures [56]. Consequently, no distinct moment or place can be identified as the origin of life. Likewise, no particular reaction or molecule can be pinpointed as fundamental to the initiation of life. Therefore, the principle question how and where life emerged becomes not particularly meaningful in a physical context [5]. It would be like asking, what was the first chemical reaction? Or, where did the first wind stir a warm pond? The processes of life are thermodynamic processes, just as are any others, admittedly highly evolved toward efficacy in energy transduction, and declining large transformations into many smaller ones, and so reducing the explosive possibilities of energy dissipation [44], but in principle no different from those found among inanimate systems.

The revised probability concept comprises not only static combinatorial configurations but also changing conditional circumstances. Most notably the probability depends on the particular ingredients, specifically those found on Earth, as well as the energy density difference between the high-energy radiation from the Sun and the lower-energy thermal radiation of cold Space. The quest to reduce that energy density difference as soon as possible is the driving force that has resulted in the transduction processes, known collectively as the biosphere. During the eons, the chemical potentials of substances on Earth have increased to match the energy density of insolation. In other words, all systems are regarded by thermodynamics as mechanisms to channel energy. The concept of fitness is understood by thermodynamics as involving the ability to access and decrease free energy. In Fisher's formulation [16], the genotypes that reproduce the fastest in given conditions come to prevail in the gene pool. Thus, other things being equal, inasmuch as fast reproduction refers to striving, we can deduce from the Carnot principle that those kinds that produce the most entropy in a population are in line to become the fittest. Naturally, other things

are seldom equal, and so details of adaptive efficacy are required to discover which kind actually wins for the moment. Proliferation will increase the mechanistic capacity to diminish free energy but it is no end in itself. Catalysis will speed up the transduction process but it is no end in itself, either. Adaptation will improve a system's consumption of available free energy, or maintain it in the face of competitors. The most voluminous flows of energy naturally select to channel along the steepest directional gradients; in other words, through those mechanisms that will level off energy density differences in least time. The mechanisms of life that work to raise the chemical potential of matter on Earth toward the energy density in insolation are not destructive explosions. In order for life to exist, explosive and fiery environments will have to have been replaced by more benign conditions. As well, energy flows invoked in striving cannot be so great as to injure the dissipative structure mediating them. When that happens the dissipative structure dissipates. In any case, survival means access to free energy; extinction is inevitable when the vital flow dries up.

Often the theory of evolution by natural selection is seen to be problematic as emphasizing competition while collaboration is apparent in nature. According to thermodynamics, neither rivalry nor co-operation is an end in itself, but whichever mechanism that will best contribute to the increase of entropy production will be naturally selected by the flows of energy. Competing systems will turn to collaboration when their integration will yield an increase in entropy production. Conversely, organizations will break apart when their constituent systems' independent consumption of free energy is more effective in this. Likewise, evolution into species richness is no objective as such but is common because generally more diverse sources of free energy are accessed by diverse mechanisms. No single species is equipped with all interaction mechanisms to draw from all sources. This is because history has continually reworked surviving energy gradients into increasing numbers of different configurations and conformations. Admittedly, the energy transduction mechanisms of living systems are complex and diverse but the underlying thermodynamic principle is simple and distinct. The various ways of energy intake for a system that is initially below the energy density of its surroundings to move toward an overall stationary state are precisely those that we attribute to the living system, i.e., expansion, proliferation, differentiation, catalysis, and adaptation. Boltzmann admired Darwin, which is apparent from the quote that "The animate struggle for entropy, not against it" [29], which has been commonly misconceived.

The universal characteristics of nature are consequences of maximal energy dispersal, not outcomes of particular fortuitous events or rare mecha-

nisms. For example, evolution on a global, just as on a local, scale, bursts by punctuations and settles to stasis [18], something that is apparent from ubiquitous sigmoid growth curves [48]. Initially, when a new supply of free energy is accessed, the driving force is large but the flows are curtailed by rudimentary mechanisms. Subsequently, when new mechanisms will appear, e.g., via random mutations, those that will facilitate the flow further become entrained. However, it will not be too long before the increasing mechanistic capacity will draw so rapidly from the repository of renewable free energy that growth begins to level off toward stasis, i.e., toward a thermodynamic steady state. A series of punctuations and stases results from a series of sigmoid growth processes following each other. Likewise, when the system exploits a non-renewable source, recession is inevitable unless other sources of free energy are found. Each sigmoid, just as a series of sigmoids, when plotted on a log-log plot, follows mostly a power law. A sigmoid is a cumulative curve of a skewed distribution. These nearly log-normal distributions are ubiquitous not only among plant and animal species but also among inanimate structures on all scales [22, 35]. We suggest that these universal characteristics are consequences of the law for maximal energy dispersal.

Along with the general question “what is life”, a few specific questions, such as the ubiquitous choice of L-amino acids in natural proteins, have puzzled biologists, chemists, physicists, and philosophers alike. The ubiquitous chirality consensus and the common genetic code are often associated with a postulated singular emergence of life, but thermodynamics provides another view [24]. These standards, like any other standards, have emerged to facilitate the dispersal of energy. Consider a system containing initially achiral basic constituents that may assemble to right- or left-handed compounds that may further polymerize to pure or mix-handed polymers. The pure-handed constituents polymerize via a single type of reaction because the substrates are identical, whereas mixed-handed substrates are non-standard and thus require more than one type of reaction. If a chemical system happens to emerge with one effective enzyme, it is enough for a pure-handed system to advance in its energy transduction. Conversely, for the mixed-handed system to advance in energy dispersal several catalysts are required. Therefore, a single faster reaction will draw more basic compounds into the pure chiral systems, which may therefore emerge with even more powerful catalysts and evolve even faster. There is a positive feed-back for the pure-handed systems that would soon be competing only with each other for the basic compounds. If either of them happens to advance a little further than the other, it may emerge with even

better mechanisms. This competitive exclusion principle [40] works not only among molecular species but also between individuals and also between economic rivalries. The instability between systems can be demonstrated using the Lyapunov stability criterion [30, 51] now that the formula for entropy (eq. (4)), and also its time derivative (eq. (5)), are known. Accordingly, when a system has attained its free energy minimum state, it remains stable against fluctuations in its internal composition. The stability of a thermodynamic steady state ranges from the familiar one of chemical equilibrium to the homeostatic behavior of a global system [27, 37]. Increasingly larger systems benefit more from standards. Therefore, the high degree of chiral consensus does not mean that it would have been settled at some unique event but may instead suggest that the global ecosystem has evolved into a highly integrated global system, the biosphere. Similarly, international standards such as those of the International Telecommunications Union are currently emerging from the ongoing global economic integration [7].

The vast amount of non-coding DNA in eukaryotes has been considered a mystery [20]. After all, genes make only 1.5% of our genome. From the perspective of the 2nd law, a genome can be viewed as an ecosystem as any other – for example, a forest – where we are not puzzled to find insects eating trees that are the main means of acquiring energy to the forest system [23]. In the same way it is not puzzling that transposomes, short repeats, introns, intergenic regions, and other non-coding segments in our genomes are thriving when using the resources brought into the genomic system through the agency of genes, the “trees” of a cell nucleus. This conclusion about the role of energy dispersal is backed up by inspecting the distribution of genomic sequences. Their distribution in lengths is skewed, just as any other population, e.g., of plants and animals [58]. The principle behind the skewed distribution is the same as for the skewed velocity distribution of Maxwell–Boltzmann. It is the maximum entropy partition. The cumulative curve of any natural distribution is a sigmoid – on a log-log plot a straight line that follows mostly a power law. Another example is the species-area relationship, whose origin has also been a puzzling question [57].

To predict protein folding from the amino acid sequence is regarded as one of the hardest tasks in computational biology [34]. According to the 2nd law, protein folding would be an inherently intractable process [47]. The driving forces of folding, just like any other evolutionary process, are inseparable from the flows of energy that produce the dissipation, known here as the heat of folding, which can be measured by calorimetry. Fold-

ing, as with a chemical reaction, directs along the fastest ways toward the free energy minimum. However, since the flows themselves change the free energy landscape, the problem is intractable, like many other natural processes. In computational terms the process does not complete in deterministic manner in polynomial time [17]. Information that guides the process is not simply contained only in the amino acid sequence since the evolutionary course of a polypeptide system, like any other system, depends on its surroundings. The resulting ensemble of protein folds is the maximum entropy partition in the respective surroundings.

There are certainly many more examples of evolutionary phenomena than were addressed here that would deserve to be examined by the law of maximal energy dispersal. It is reasonable to anticipate obtaining insights into overall courses, distributions, and regularities, but unreasonable to expect determining details. Undoubtedly it would be rewarding to unravel deterministic causes but it is also gratifying to learn why such knowledge is not available.

5 Discussion

Boltzmann's idea to regard time-dependent phenomena of complicated many-body systems as probable sequences of state changes was groundbreaking. However, he adopted an impaired formulation of probability which limited applications of statistical physics to stationary-state systems. The herein re-examined probability concept includes not only combinatorial configurations but also conditional circumstances. In this way, Boltzmann's objective to study also evolving systems by statistical physics has now been re-opened. Entropy, the elusive concept that has been troubled with misconceptions, emerges from the re-formulation of probability as the statistical measure of the system's free energy status. The principle of increasing entropy resolves as an equation of motion that is consistent with classical thermodynamics [10, 19, 50]. Intriguingly, evolution is found to be a non-deterministic process whose trajectories are inherently intractable just as biologists have generally suspected, with their focus on history.

The herein-presented revised form of statistical physics undoubtedly prompts many more questions than were attempted or referred to in this study. For example, the physical foundations of evolution, when given independently of the machinery of energy dispersal, may appear to some as incoherent with the modern evolutionary synthesis, which describes biological evolution by giving a major role to genetic information. As well, the thermodynamic imperative for maximal energy dispersal may seem to

some to be incommensurable with the origin of life research that depicts chemical evolution by giving a big role to autocatalytic processes.

Since thermodynamics values everything in terms of energy, information is not viewed as abstract [32] but bound to its material representations, which in turn are subject to the 2nd law: flows of energy naturally select the fastest ways to level off energy density differences. Therefore, it is no surprise that the distributions of letters, words, length of sentences, length of genes, and others, are skewed, and that their cumulative curves are sigmoid, i.e., power laws on log-log scales. In short, information is physical since “bit” is necessarily embodied in “it”. And so for many purposes it would be best described by thermodynamic entropy rather than by Shannon’s entropy [28]. Put otherwise, variety finds its meaning in energy dispersal [14]. Importantly, thermodynamics would not be limited to the description of syntax but could assign value to semantic meaning as well. The meaning of a message would ultimately be valued by the receiver system depending on how much its entropy production would increase when processing the received message. Thus communication, as any other process, could be viewed by thermodynamics merely as a means, often highly effective, of energy dispersal.

The role of kinetics in evolution is inbuilt in the 2nd law because the flows of energy are driven by the differences in energy densities. This functional form (eq. (6)) is distinct, e.g., from the law-of-mass action where kinetics is written as proportional to populations via varying rate “constants”. The early erroneous formula [55] renders kinetics and thermodynamics formally inconsistent with each other [15], so that evolutionary courses are pictured as crossing over high-energy barriers. In contrast, the principle of increasing entropy given in the form of $\partial_t S$ equation (eq. (5)) and its associated flow $\partial_t N_j$ equation (eq. (6)) describes kinetic courses not as flows across barriers, but as an energy manifold, as the representation of the system itself, working its way down along gradients. Due to dissipation, the thermodynamic gradients change during the evolutionary course. During an irreversible process, new paths open up while others close down. Although the non-deterministic equation of evolution (eq. (5)) cannot be solved analytically, kinetic courses can be studied numerically or simulated to give an ensemble of conceivable trajectories.

Darwin and his contemporaries did not know 150 years ago how evolution works at the molecular level. This lack of knowledge of mechanistic details did not prevent him from seeing a universal principle. Darwin, as is apparent from the famous letter to a friend [13], did not find it particularly problematic to imagine that life emerged from natural processes taking

place in a warm pond. Likewise, Darwin did not find it problematic to adopt the natural selection principle from economics. He did not see impossible demarcation lines between the inanimate and the animate. Thermodynamics recognizes none either. Darwin propagated by way of vivid words the universal principle that now deserves to be viewed more generally as the 2nd law in its mathematical form in order for us to understand more deeply who we are, what we do, and, importantly, why we do anything at all.

Bibliography

- [1] Acheson, D. J., *Elementary Fluid Dynamics*, Oxford University Press, Oxford, 1990.
- [2] Alonso, M. and Finn, E. J., *Fundamental University Physics*, Vol. 3, Addison-Wesley, Reading, MA, 1983.
- [3] Annala, A., *Physical portrayal of computational complexity*, in preparation, 2010.
- [4] Annala, A., *Space, time and machines*, in preparation, 2010.
- [5] Annala, A. and Annala, E., Why did life emerge? *Int. J. Astrobiol.* 7 (2008), 293–300.
- [6] Annala, A. and Kuismanen, E., Natural hierarchy emerges from energy dispersal, *Biosystems*, 95 (2008), 227–233.
- [7] Annala, A. and Salthe, S., Economies evolve by energy dispersal, *Entropy*, 11 (2009), 606–633.
- [8] Atkins, P. W. and de Paula, J., *Physical Chemistry*, Oxford University Press, New York, 2006.
- [9] Boltzmann, L., *Populäre Schriften*, Barth, Leipzig, 1905. [Partially transl.: McGuinness, B., *Theoretical physics and philosophical problems*, Reidel, Dordrecht, 1974.]
- [10] Carnot, S., *Reflexions sur la Puissance Motrice du Feu et sur les Machines Propres a Developper cette Puissance*, Bachelier, Paris, 1824.
- [11] De Maupertuis, P. L. M., *Les Loix du mouvement et du repos déduites d'un principe metaphysique*, *Histoire de l'Acad. Roy. Sci. Belleslett.* (1746), 267–294.
- [12] Darwin, C., *On the Origin of Species*, John Murray, London, 1859.

-
- [13] Darwin, C., In a letter to the botanist Joseph Dalton Hooker, 1871.
- [14] Dewar, R., Information theory explanation of the fluctuation theorem, maximum entropy production and self-organized criticality in non-equilibrium stationary states, *J. Phys. A Math. Gen.*, 36 (2003), 631–641.
- [15] Fersht, A. R., *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*, Freeman, New York, 1999.
- [16] Fisher, R. A., *The Genetical Theory of Natural Selection*, Clarendon Press, Oxford, 1930.
- [17] Fraenkel, A. S., Complexity of protein folding, *Bull. Math. Biol.*, 55 (1993), 1199–1210.
- [18] Gould, S. J., *The Structure of Evolutionary Theory*, Harvard University Press, Cambridge, MA, 2002.
- [19] Gouy, L. G., Sur l'énergie utilisable, *J. de Physique*, 8 (1889), 501–518.
- [20] Gregory, T. R. ed., *The Evolution of the Genome*, Elsevier, San Diego, 2005.
- [21] Griffiths, D., *Introduction to Electrodynamics*, Prentice Hall, Englewood Cliffs, NJ, 1999.
- [22] Grönholm, T. and Annala, A., Natural distribution, *Math. Biosci.*, 210 (2007), 659–667.
- [23] Jaakkola, S., El-Showk, S. and Annala, A., The driving force behind genomic diversity, *Biophys. Chem.*, 134 (2008), 232–238.
- [24] Jaakkola, S., Sharma, V. and Annala, A., Cause of chirality consensus, *Curr. Chem. Biol.*, 2 (2008), 53–58.
- [25] Jaynes, E. T., *Probability Theory. The Logic of Science*, Cambridge University Press, Cambridge, 2003.
- [26] Kaila, V. R. I. and Annala, A., Natural selection for least action, *Proc. R. Soc. A*, 464 (2008), 3055–3070.
- [27] Karnani, M. and Annala, A., Gaia again, *Biosystems*, 95 (2009), 82–87.
- [28] Karnani, M., Pääkkönen, K. and Annala, A., The physical character of information, *Proc. R. Soc. A*, 465 (2009), 2155–2175.
- [29] Knight, R. D., *Physics for Scientists and Engineers with Modern Physics: A Strategic Approach*, Addison-Wesley, San Francisco, 2003.

-
- [30] Kondepudi, D. and Prigogine, I., *Modern Thermodynamics*, Wiley, New York, 1998.
- [31] Kullback, S., *Information Theory and Statistics*, Wiley, New York, 1959.
- [32] Landauer, R., *Information is physical*, *Physics Today*, May (1991).
- [33] Lavenda, B. H., *Nonequilibrium Statistical Thermodynamics*, John Wiley & Sons, New York, 1985.
- [34] Levinthal, C., *Are there pathways for protein folding?*, *J. Chim. Phys.*, 65 (1968), 44–45.
- [35] Limpert, E., Stahel, W. A. and Abbt, M., *Log-normal distributions across the sciences: Keys and clues*, *Bioscience*, 51 (2001), 341–352.
- [36] Lorenz, L., *On the identity of the vibrations of light with electrical currents*, *Philos. Mag.*, 34 (1867), 287–301.
- [37] Lovelock, J. E., *The Ages of Gaia*, Oxford University Press, Oxford, 1988.
- [38] Monod, J., *Chance and Necessity: An Essay on the Natural Philosophy of Modern Biology*, Alfred A. Knopf, New York, 1971.
- [39] Noether, E., *Invariante Variationprobleme*. *Nach. v. d. Ges. d. Wiss. zu Goettingen, Mathphys. Klasse* (1918), 235–257; English translation: Tavel, M. A., *Invariant variation problem*, *Transp. Theory Stat. Phys.*, 1 (1971), 183–207.
- [40] Rosenzweig, M. L., *Species Diversity in Space and Time*, Cambridge University Press, Cambridge, 1995.
- [41] Roshdi, R., *Optique et Mathematiques: Recherches sur L'Histoire de la Pensee Scientifique en Arabe*, Variorum, Aldershot, UK, 1992.
- [42] Salthe, S. N., *Evolving Hierarchical Systems: Their Structure and Representation*, Columbia University Press, New York, 1985.
- [43] Salthe, S. N., *Summary of the principles of hierarchy theory*, *General Systems Bulletin* 31 (2002), 13–17.
- [44] Salthe, S. N., *The natural philosophy of work*, *Entropy*, 9 (2007), 83–99.
- [45] Schrödinger, E., *What is Life? The Physical Aspects of the Living Cell*, Cambridge University Press, Cambridge, 1948.
- [46] Sharma, V. and Annala, A., *Natural process – natural selection*, *Biophys. Chem.*, 127 (2007), 123–128.

- [47] Sharma, V., Kaila, V. R. I. and Annala, A., Protein folding as an evolutionary process, *Physica A*, 388 (2009), 851–862.
- [48] Sneppen, K., Bak, P., Flyvbjerg, H. and Jensen, M. H., Evolution as a self-organized critical phenomenon, *Proc. Natl. Acad. Sci. USA*, 92 (1995), 5209–5213.
- [49] Spivak, M., *A Comprehensive Introduction to Differential Geometry*, Publish or Perish, 3rd edition, Berkeley, CA, 1999.
- [50] Stodola, A., *Steam and Gas Turbines*, McGraw-Hill, New York, 1910.
- [51] Strogatz, S. H., *Nonlinear Dynamics and Chaos with Applications to Physics, Biology, Chemistry and Engineering*, Westview, Cambridge, MA, 2000.
- [52] Swenson, R., Emergent attractors and the law of maximum entropy production: Foundations to a theory of general evolution, *Systems Research*, 6 (1989), 187–199.
- [53] Szekeres, P., *A Course in Modern Mathematical Physics*, Cambridge University Press, Cambridge, 2004.
- [54] Tuisku, P., Pernu, T. K. and Annala, A., In the light of time, *Proc. R. Soc. A*. 465 (2009), 1173–1198.
- [55] Waage, P. and Guldberg, C. M., *Studies Concerning Affinity*, Forhandlinger 35 Videnskabs-Selskabet i Christiania, 1864.
- [56] Wicken, J. S., *Evolution, Thermodynamics, and Information: Extending the Darwinian Program*, Oxford University Press, Oxford, 1987.
- [57] Würtz, P. and Annala, A., Roots of diversity relations, *J. Biophys.*, 2008 (2008), <http://www.hindawi.com/journals/jbp/2008/654672.html>.
- [58] Zhang, J., Protein-length distributions for the three domains of life, *Trends Genet.*, 16 (2000), 107–109.
- [59] Ziegler, H., *An Introduction to Thermomechanics*, North-Holland Series in Applied Mathematics and Mechanics, vol. 21, North-Holland, New York, USA, 1983.

Received October 14, 2009; accepted December 17, 2009.

Author information

Arto Annala, Department of Biosciences, FI-00014 University of Helsinki, Finland;

Institute of Biotechnology, FI-00014 University of Helsinki, Finland; and
Department of Physics, FI-00014 University of Helsinki, Finland.

E-mail: arto.annala@helsinki.fi

Stanley Salthe, Biological Sciences, Binghamton University, Binghamton,
NY 13754, USA.

E-mail: ssalthe@binghamton.edu