Chapter 1

Life is widely—and plastically—dissipative

When alive, cells spend biochemical energy. How much? What are its mathematical consequences? What are its biological consequences? In this salvo chapter, we briskly pose these questions in turn.

1.1 A lively census of dissipation

Across contexts, cell states, and organisms, the rates that cells and organisms consume energy vary dramatically. The useful free energy processed and dissipated by cells adopts many forms—as ATP gradients, nutrient transformations, ion gradients, fundamental losses as heat, and beyond. This diversity gives way to useful insight when cellular energy expenditures are compared in common thermodynamic units of energetic equivalents expended per time. To think about these scales, consider the power spent per volume across living matter, a spectrum glimpsed in Figure 1.1. From energy-limited microbes to the flight muscle cells of the honey bee, the power densities of cells vary more than an

incredible seven orders of magnitude, as we compute and show in Figure 1.1(B-C). How do such lively metabolisms compare to their inorganic counterparts? Figure 1.1(A) gives a comparative visual answer to that question by familiar examples from the world of everyday macroscopic experience. Some of the examples shown there are drawn from our astronomical neighbors, including the Sun, joined by technologies developed by humans.[1] The most powerful states of living matter, including exponentially growing microbes like *E. coli*, are so dissipative as to be within a factor of a few hundred of the power density inside nuclear reactor cores. On the slow end, environmental microbes in ocean sediments consume power over a million times slower by volume, representing an extraordinary dissipative plasticity.

In the spirit of the Bionumbers database [1], we aim to take stock of the facts and data that have been amassed to describe the energies and powers associated with living matter. We give an even fuller census in the plot of Figure 1.2, which also illustrates how widely certain cell types tend to change their own dissipative status by exertion.[2] There, we also highlight a new experimental measurement, a basal metabolic rate of the opportunistic pathogen *P. aeuriginosa* reported by Ciemniecki *et al.* [2], whose value and attendant questions so intrigue us that we engage with it across many chapters of this thesis.

Understanding why the fastest-growing cells spend the amount of energy they do rewards us with useful rules of thumb that organizes our appreciation of the whole ladder of dissipations visible in both Figs. 1.1 and 1.2. Consider exponentially growing *E. coli*, with a typical doubling time of $\tau \sim 20 - 30$ min (or few $\times 10^3$ s). If each cell to be duplicated has order $N \sim$ few $\times 10^6$ proteins; each protein is composed of $\ell \sim$ few $\times 10^2$ amino acids; and the cost of each such amino acid incorporation is $\eta \sim 4$ ATP equivalents,

[1] Note that although for millennia humans have had control of fire, we are of the mind that the capacity to make things cold is even more impressive; both heating and cooling technologies are shown in Fig. 1.1(A).

[2] Note that in Fig. 1.2, we omitted the slowest metabolizing ocean microbes earlier depicted in Fig. 1.1.

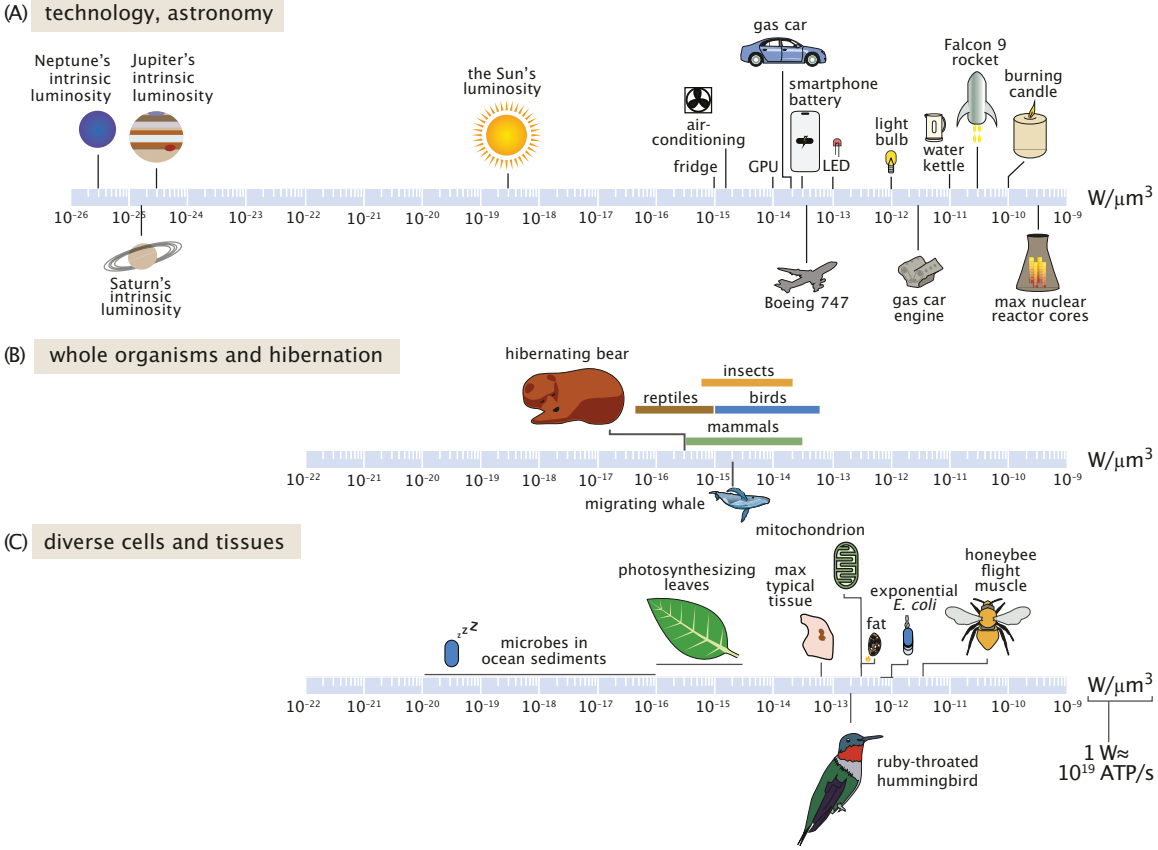

Figure 1.1: Comparison of the power densities of inanimate matter and living matter. The power density of organisms, and their plasticity, is impressive even on astronomical terms. (A) Astronomical, technological, and abiotic sources with familiar dissipations. We converted stellar data from NASA to power densities; other quantities were computed from first principles or manufacturer information. (B) Power densities at the organismal scale, including the fascinating process of hibernation. (C) Power densities at the cellular and tissue scales.

then just the cost of duplicating the proteome into a new daughter cell is

$$\text{power to duplicate proteome: } \frac{N\ell\eta}{\tau} \tag{1.1}$$

$$\approx \frac{(\text{few} \times 10^6 \text{ proteins}) \left(\frac{\text{few}\times 10^2 \text{ aa}}{\text{protein}}\right) \left(\frac{\text{few ATP}}{\text{aa}}\right)}{\text{few} \times 10^3 \ s} \tag{1.2}$$

$$\sim 10^6 \text{ ATP/s/cell.} \tag{1.3}$$

This simple calculation also invites us to recall the extremely useful scale that one ATP

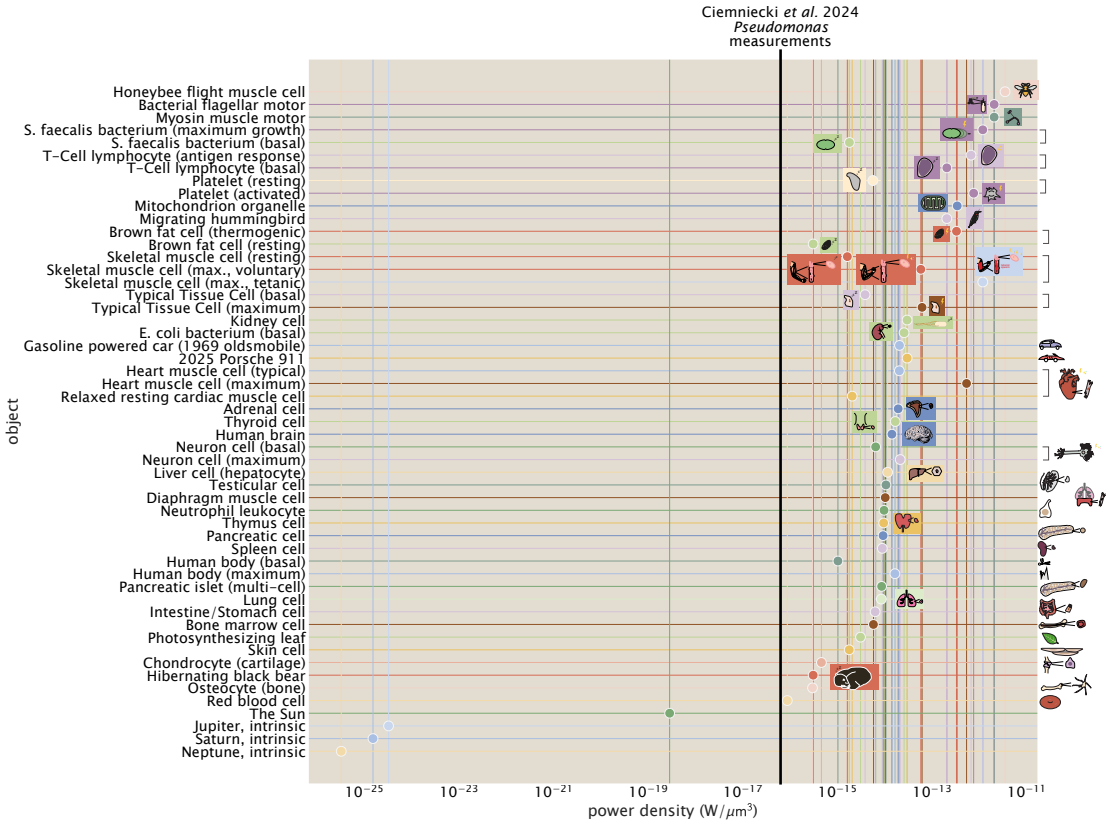

Figure 1.2: Power densities across contexts, cell states, and different organisms. This representation given in units of power/volume (on a log scale) shows that the power densities of organisms and cells vary over many orders of magnitude.

hydrolysis's worth of energy typically liberates approximately $20\ k_BT$ worth of energy at temperature T, where k_B is Boltzmann's constant (Bionumbers ID (BNID) 103385 [1]); at typical temperatures $T \sim 300\ K$ found comfortable by an array of life, this implies that 1 Watt $\approx 10^{19}$ ATP/s. Accordingly, that power *E. coli* pays to duplicate its proteome is also expressible as $\sim 10^{-13}\ W/\mu m^3$ in its one cubic micron volume. Consulting Fig. 1.1(C), we recall that the total metabolic rate of exponential *E. coli* is measured as $\sim$ few $\times\ 10^6$ ATP/s/μm^3 to 10^7 ATP/s/μm^3 (or, equivalently, few $\times\ 10^{-13}$ W/μm^3 to 10^{-12} W/μm^3). In other words, duplicating proteins is so expensive that it is within a factor of few to ten of the total power budget of rapidly growing *E. coli* cells. Protein duplication appears to explain much, though not all, of the scale for why fast *E. coli* cells

consume the energy that they do [3].

What about the remaining, non-protein-duplicating, dissipation measured in fast growing cells? Does the remainder of the power budget have obvious (and precisely accounted for) biophysical destinies? Intriguingly, carefully tracking known expenditures (towards other accumulation of biomass and known biochemistry) does not always fully account for the measured total metabolism. Such outstanding energy expenditures were suggested back in the 1980s [4], but more fastidious modern accounting still ratifies "missing" fractions of the energy budget in common examples of aerobic metabolism. Writing in their 2023 study deploying detailed flux balance analysis modeling, and fully invoking sophisticated modern insights about physiology and constraints, Mori, Hwa, and coworkers write, "Surprisingly, the ATP generated during the biosynthesis of building blocks from glucose almost balances the demand from protein synthesis, the largest energy expenditure known for growing cells. This leaves the bulk of the energy generated by fermentation and respiration unaccounted for" [5]. One practical way these unexplained energy expenditures plague our field's collective quantitative understanding is the need for flux balance analysis to invoke *ad hoc* multiplicative inefficiency or mechanistically-unmotivated "maintenance" factors in order to match total metabolic rates [6, 7].

Not only do fast-growing cells spend large (sometimes only partly explained) fluxes of ATP, they also sustain large standing pools of ATP. Interestingly, across many cell types from bacteria to human muscle, brain, and retina, the steady-state concentration of ATP is consistently on the scale of a few millimolar in concentration, varying by only a factor of a few in typical homeostatic conditions [8]. This is a big standing pool, namely a few million molecules per cubic micron. While it is of course ultimately the instantaneous ratio of ATP to ADP and phosphate levels (not merely the total ATP concentration) that sets the energetic value of an ATP hydrolysis event, this interesting abundance further hints at a

broad standing capacity for nonequilibrium operations in many cells.

Regarded collectively, these numbers loudly announce some irresistible and important mystery stories. Towards what rich purposes could these plausible, hidden biological dissipations be directed? Which of those purposes can *only* be accomplished out of equilibrium, and which are reproducible without spending energy? What empirically accessible signatures would help us distinguish between equilibrium and nonequilibrium operations in biological networks? If, mathematically and physiologically, a function *could* be performed without spending as much energy as biology might do, what abstractions better replace a brittle analogy of "efficiency," and more expansively accommodate the high-dimensional priorities that really shape an organism? What are the limits below and above which functions underlying life is impossible? These are questions ripe for human wonder.

While we built initial intuition and met questions by appealing to the high dissipation end of the biological dissipative ladders, we can—and rather, must—learn from the fullness of the dissipative spectrum. The high end of the spectrum invites marveling at the regimes and mechanisms of extraordinary biomass accumulation, high-flux component turnovers, and the buzz of honeybee wings. The low end of the spectrum, in the great "struggle for existence," promises to teach us about the true boundaries between living and nonliving matter. In between lies principles of physiological adaptation that beckon discovery.

1.2 Mathematical notions of (non)equilibrium

Having confronted just how out of equilibrium biology often is, we now recall what it means mathematically for a system to be in or out of equilibrium. To set the stage for this discussion, imagine a system of interest like a cell, or a small subset of molecules inside the cell, that can be found in one of a finite number of states, as shown in Fig. 1.3(A). If the

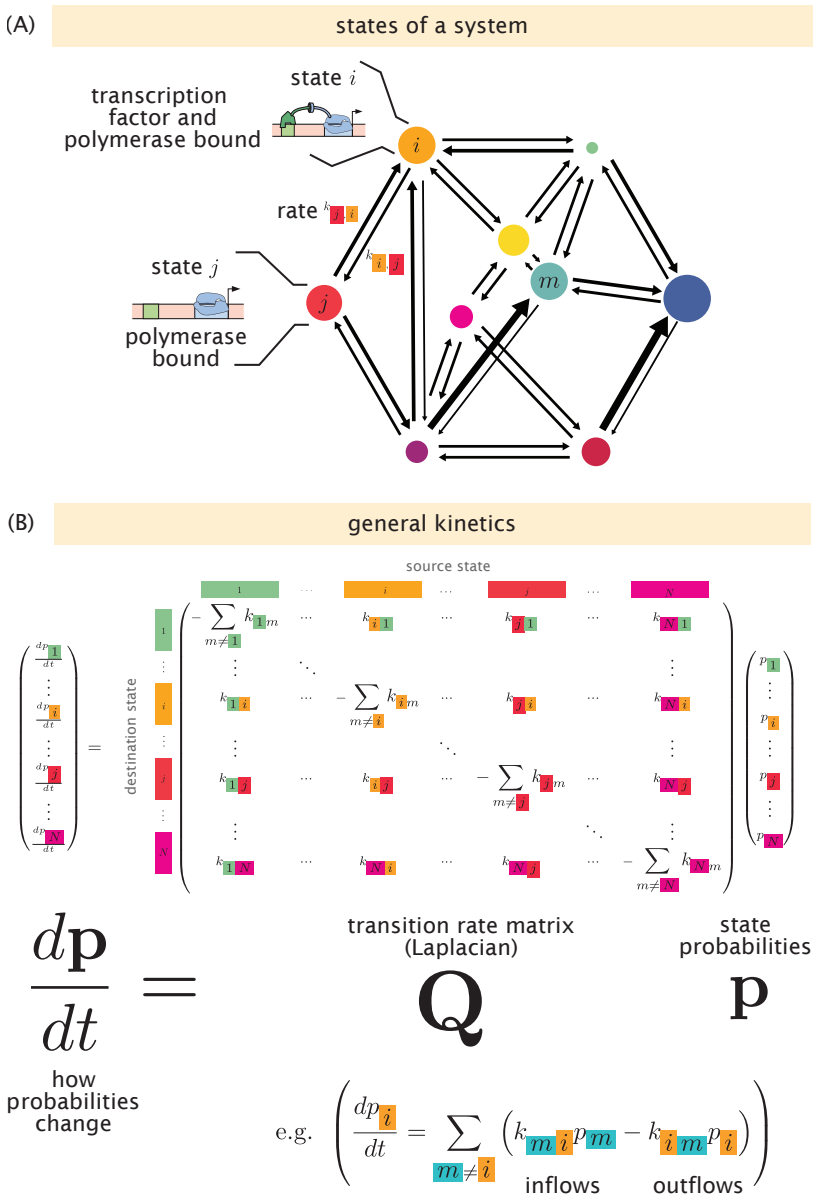

Figure 1.3: States of a system represented as a network (a finite state, continuous time Markov chain). (A) States of a cell, genome, or set of molecules schematized as a network or graph. Transitions between any pair of eligible connected states i and j occur with characteristic rate k_{ij}. (B) The probabilities $\mathbf{p}(t)$ of being found in any of the states change depending on the balance of inflows and outflows to each state, which can be mathematically summarized by the matrix equation $\frac{d\mathbf{p}}{dt} = \mathbf{Q}\mathbf{p}$, where the transition rates k_{ij} populate the entries of the transition rate (Laplacian) matrix $\mathbf{Q}$.

system transitions from state i to j at a characteristic rate k_{ij}, then the probability $p_i(t)$ of being in state i at time t changes depending on the balance of inflows of probability into

state i from other states $m \neq i$, and vice versa. That is,

$$\frac{dp_i}{dt} = \sum_{m \neq i} \left(k_{mi} p_m - k_{im} p_i \right), \tag{1.4}$$

as shown in Fig. 1.3(B). Equivalently, these rules for how the probabilities of all the states change can be tidily written as a matrix equation by squeezing the probabilities of every state into a vector $\mathbf{p} \equiv [p_1, \ldots p_N]^\top$ and the transition rates into a matrix $\mathbf{Q}$ whose i, jth entry represents the transition rate from state j to state i (and the diagonal entries of $\mathbf{Q}$ are set to ensure that probability is conserved to sum to one).[3] Or, all together (as shown in Fig. 1.3(B)),

$$\frac{d\mathbf{p}}{dt} = \mathbf{Q}\mathbf{p}. \tag{1.5}$$

This expression is also known as the Master Equation for the stochastic trajectory of the system. If the system's set of states is sufficiently well behaved (including, for instance, the requirement that every state i is reachable by some path from any other state j), then then there is a probability vector $\mathbf{p}_{\text{ss}}$ over states for which the the probability will not change, $\left.\frac{d\mathbf{p}}{dt}\right|_{\mathbf{p}_{\text{ss}}} = \mathbf{0}$. In general, all that the steady-state condition enforces is that, for all states $i \in [1, \ldots, N]$, the net inflow and net outflow of probability mass for every state balances each other (as schematized by Fig. 1.4(A)). This condition imposes N algebraic constraints on the behavior of the system.

[3] This convention, where the indices of k_{ij} are transposed to be stored in Q_{ji}, may seem artificial, but this is just cosmetic. When considering the transition from state i to j, the columns of $\mathbf{Q}$ (containing i) are regarded as the source states, and the rows of $\mathbf{Q}$ (containing j) are the destination states.

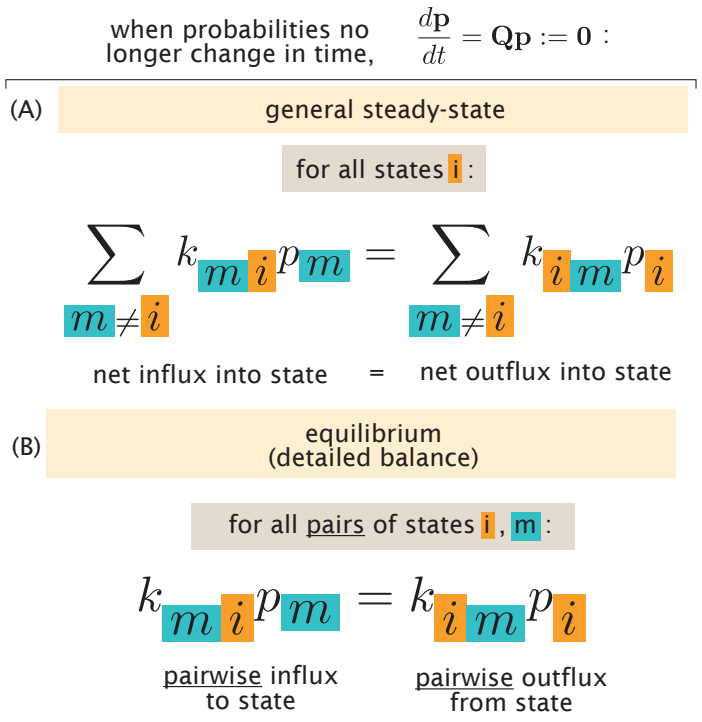

Figure 1.4: Steady-states in and out of equilibrium. (A) Conditions on the elements of the transition matrix and probabilities for a system that has reached a nonequilibrium steady state. (B) The special steady state known as equilibrium demands specific conditions on the transition rates and probabilities.

However, one very special way of achieving this steady-state condition is by a more granular constraint, considered on every pair of states. Specifically, instead of merely requiring that the *net* inflow and net outflow balance each other for each state, as in Fig. 1.4(A), consider instead the scenario where each and every *pair* of states enjoys a mutual balance of probability flow. That is, if,

$$\text{detailed balance (equilibrium): } \text{flux}_{i\to j} = k_{ij}p_i = k_{ji}p_j = \text{flux}_{j\to i}, \tag{1.6}$$

then the system is referred to as being in detailed balance. Shown in Fig. 1.4(B), this condition is exactly the condition of thermodynamic equilibrium (when the system is otherwise closed to material inputs and Markovian). Crucially, for many state network topologies, steady-state distributions can be reached without satisfying this much more restrictive constraint of local detailed balance enforcing equilibrium. Notice that there are a greater number, $N(N-1)/2$, of constraints imposed on the system compared to merely the steady-state condition. This mathematical framing concretely grounds our more empirical and philosophical allusions to dissipation in the preceding discussion to a specific, underlying gulf between defined mathematical universes.

An incredibly powerful consequence of the constraint of detailed balance is given by Boltzmann's equation, which assures that at equilibrium, the probability of a state i with an energy parameter E_i is proportional to $\exp[-E_i/k_BT]$. This relationship, whose form facilitates a huge fraction of successful statistical mechanical calculations, is so formidable and useful in calculations because it gives a *local* rule for predicting the relative probability of a system being in a state. This idea is depicted in Fig. 1.5. However, Boltzmann's equation only applies to systems in equilibrium. One theme we value and pursue aggressively with this thesis is to explore candidate "cognitive technologies" that could adopt the guiding navigational role that Boltzmann's equation gives, but for

nonequilibrium domains.

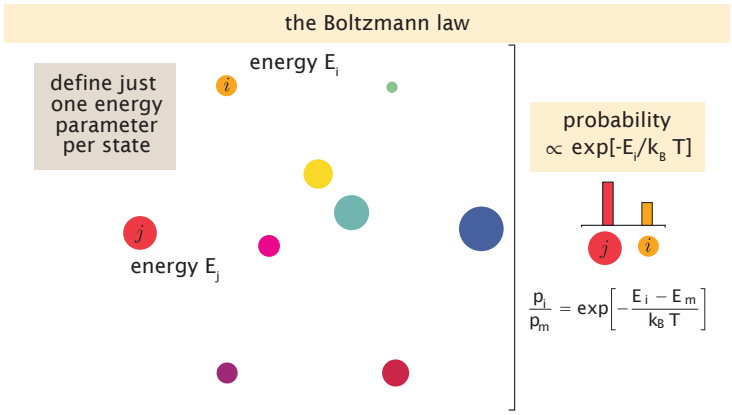

Figure 1.5: The Boltzmann law for computing the equilibrium probabilities of different microstates.

1.2.1 Implications of this dissipation for theory

Although the primordial mathematical differences between equilibrium and nonequilibrium behaviors are quickly stated (viz., the statements of Fig. 1.4), their actual consequences for the attainable complexity of biological functions of interest is a much richer, more interesting, and more subtle question. This is due to at least two surprising facts.

First, purely equilibrium systems are themselves capable of surprisingly complex and interesting behaviors, especially if they are large enough (and given enough time). This capability essentially underlies the success of sufficiently large Boltzmann machines to reproduce arbitrary probability distributions (probabilistic models by definition instantiated to be at equilibrium). Further, recently a body of stimulating work illustrates just how capable equilibrium self-assembly can be, in a formal computational sense [9].

Second, nonequilibrium systems are regularly capable of looking nearly equilibrium. This is a theme we now explore in two evocative vignettes.

A thermodynamic irony: the curious success of equilibrium theory.

In a famous article in the early 1960s, Eugene Wigner talked about the "unreasonable effectiveness of mathematics in the natural sciences." That justly famous essay highlights the repeated successes, both big and small, of recasting our understanding of the natural world in mathematical form. There are similarly curious "effectivenesses" to be celebrated in the natural sciences. One of the most remarkable aspects of many quantitative dissections of biological phenomena is the power of equilibrium thinking, one of the most unreasonably effective theories in all of science — with "equilibrium" even serving as the primary tool used to understand the degree of ionization of atoms in stars —- a manifestly out of equilibrium situation. A compelling biological example of this effectiveness is given in Figure 1.6 in the context of gene expression, as we now describe.

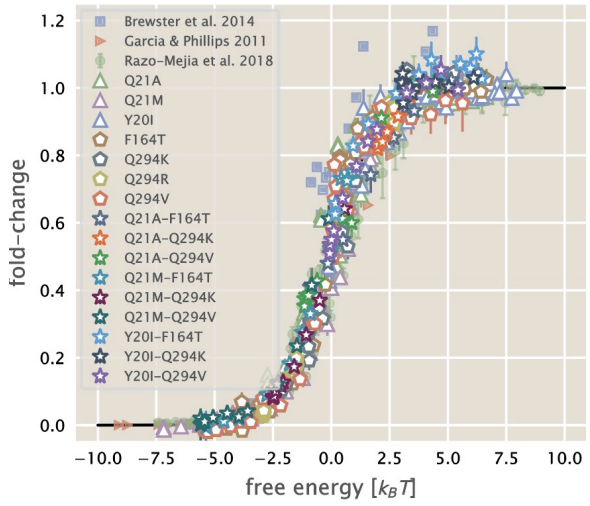

Figure 1.6: Gene expression for a great diversity of different situations all fall onto one master curve. The data collapse shown here highlights two key points: i) the effectiveness of equilibrium models of the process of transcription and ii) the emergence of "natural variables" that reveal what the cells "cares about" as opposed to the tuning variables used in the laboratory such as how tightly transcription factors bind to DNA, their copy number and the concentration of some inducer. (This figure is reproduced from Chure *et al.* [10].)

There are many ways to tune the level of gene expression of a bacterium. For example, we can artificially alter the sequence of the binding site for some transcription factor, thus

changing its K_d, and hence, how tightly the transcription factor binds the DNA. Alternatively, by tuning the ribosomal binding site for the mRNA that produces those transcription factors, we can control the copy number of the transcription factor. Or, we might tune the amount of some inducer, which essentially changes the number of *active* transcription factors. How do we account for all of these different ways of perturbing the level of gene expression?

As already seen in Figure 1.6, one route to computing the different states of occupancy of a promoter, and thereby, it's level of gene expression is the use of the Boltzmann law. In the context of transcription, the Boltzmann protocol asks us to enumerate the microscopic states of occupancy of some promoter (i.e. bound by polymerase, empty, bound by repressor, etc.) and then to compute their respective probabilities by invoking the Boltzmann distribution $p_i \propto \exp\left[-E_i/k_BT\right]$. For all of the different experiments reflected in Figure 1.6, the Boltzmann distribution was used to compute a single quantity that we will call p_{bound}, the probability that the promoter is occupied by RNA polymerase and thereby, producing mRNA. Without entering into any of the important technical details on the experimental or theoretical fronts, the reason for showing Figure 1.6 is to highlight how subtle the question of equilibrium and nonequilibrium really is in the biological setting. Let's now turn to a larger scale example in the context of embryonic development that illustrates the same subtlety.

Subtlety in pattern formation

The topic of how inorganic matter is turned into living organisms is one whose importance is felt in subjects other than the study of the growth of microbes. Indeed, its borders reach far beyond the traditional confines of the subject we call biology. In the 19th century, the subject of thermodynamics arose in large measure because of attempts to understand the workings of heat engines and how energy could be converted between different currencies.

Even discovering that there are different energetic currencies, especially the nature of the "kind of motion we call heat" was a great achievement [11, 12]. Unfortunately, the subject of thermodynamics felt short of its original name since almost always, the classical use of that term refers to thermostatics, the study of thermal systems in their terminal privileged state of equilibrium in which they are no longer changing in time. One of the deepest questions in modern science is how to generalize the ideas of traditional equilibrium systems to handle the full complexity of nonequilibrium steady states of living systems. Living organisms are *thermodynamic* systems in the true sense of the word in that they are both thermodynamic and involve the interconversion of matter and energy at every single moment that a chunk of matter is "alive."

The extraordinary numbers we are staring at about the scales of dissipation might give the impression that the physical consequences of nonequilibrium are always biologically obvious or conspicuous. To remember to disabuse ourselves of falling into this impression, here we offer an evocative thought experiment that clarifies that the accomplishments of energy expenditure are often subtle or barely perceptible at the scale of mean free paths of molecules—but of course emergently and importantly functional at larger scales or on statistical average.

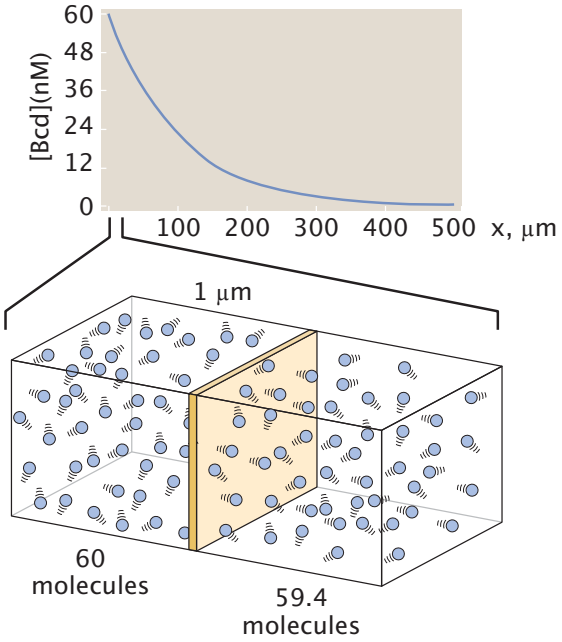

Figure 1.7: The superficially artificial nature of the distinction between equilibrium and nonequilibrium. The long axis of the 500 μm long fly embryo is discretized the into a collection of 1 μm boxes, the difference in the number of Bicoid molecules from one box to the next is less than a single molecule.

As shown in Figure 1.7, the subject is far more subtle than to simply make pronouncements that living organisms are out of equilibrium. In particular, there we see that given the measured Bicoid gradient in the developing fly embryo, if we discretize the 500 μm long embryo into a collection of 1 μm boxes, the difference in the number of Bicoid molecules from one box to the next is less than a single molecule. Clearly these individual molecules "know" nothing about a gradient: the distinction between equilibrium and nonequilibrium requires a macroscopic concept of fluxes.

In our view, these vignettes showing that equilibrium and nonequilibrium systems can surprisingly resemble each other makes even more interesting the effort to precisely define the boundaries of reachable behavior of each class of system, especially in biology. This question finds echoes in all of the chapters of this thesis.

In our view, the studies considered here help point the way to finding more general thermodynamic principles that describe active matter in all of its generality.

Bibliography

[1] Ron Milo, Paul Jorgensen, Uri Moran, Griffin Weber, and Michael Springer. Bionumbers—the database of key numbers in molecular and cell biology. *Nucleic acids research*, 38(suppl_1): D750–D753, 2010.

[2] John A Ciemniecki, Chia-Lun Ho, Richard D Horak, Akihiro Okamoto, and Dianne K Newman. Mechanistic study of a low-power bacterial maintenance state using high-throughput electrochemistry. *Cell*, 187(24):6882–6895, 2024.

[3] Ron Milo and Rob Phillips. *Cell biology by the numbers*. Garland Science, 2015.

[4] Ian S Farmer and Colin W Jones. The energetics of escherichia coli during aerobic growth in continuous culture. *European journal of biochemistry*, 67(1):115–122, 1976.

[5] Matteo Mori, Chuankai Cheng, Brian R Taylor, Hiroyuki Okano, and Terence Hwa. Functional decomposition of metabolism allows a system-level quantification of fluxes and protein allocation towards specific metabolic functions. *Nature Communications*, 14(1):4161, 2023.

[6] Amit Varma and Bernhard O Palsson. Metabolic capabilities of escherichia coli: I. synthesis of biosynthetic precursors and cofactors. *Journal of theoretical biology*, 165(4):477–502, 1993.

[7] Hong Zeng and Aidong Yang. Modelling overflow metabolism in escherichia coli with flux balance analysis incorporating differential proteomic efficiencies of energy pathways. *BMC systems biology*, 13:1–18, 2019.

[8] Jack V Greiner and Thomas Glonek. Intracellular atp concentration and implication for cellular evolution. *Biology*, 10(11):1166, 2021.

[9] Cameron Chalk, Salvador Buse, Krishna Shrinivas, Arvind Murugan, and Erik Winfree. Learning and Inference in a Lattice Model of Multicomponent Condensates. In Shinnosuke Seki and Jaimie Marie Stewart, editors, *30th International Conference on DNA Computing and Molecular Programming (DNA 30)*, volume 314 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 5:1–5:24, Dagstuhl, Germany, 2024. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. ISBN 978-3-95977-344-7. doi: 10.4230/LIPIcs.DNA.30.5. URL https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.DNA.30.5.

[10] Griffin Chure, Manuel Razo-Mejia, Nathan M Belliveau, Tal Einav, Zofii A Kaczmarek, Stephanie L Barnes, Mitchell Lewis, and Rob Phillips. The energetics of molecular adaptation in transcriptional regulation. *bioRxiv*, page 638270, 2019.

[11] A. Einstein and L. Infeld. *The Evolution of Physics*. Cambridge University Press, Cambridge, England, 1938.

[12] SG Brush. *The Kind of Motion We Call Heat*. North Holland, Amsterdam, 1986.