Seminars in Cell and Developmental Biology 175 (2025) 103646

Contents lists available at ScienceDirect

Seminars in Cell and Developmental Biology

journal homepage: www.elsevier.com/locate/semcdb

Review

Approximating the living

Rob Phillips

Division of Biology and Biological Engineering and Department of Physics, California Institute of Technology, Pasadena, CA, USA

ARTICLE INFO

Keywords:
Mathematics in biology
Representations
Models in biology

ABSTRACT

Is a herd of wildebeest better thought of as a series of individual animals, each with its own glorious and unmanageable volition, or as a field of moving arrows? Are the morphogen gradients that set up the coordinate systems for embryonic anterior–posterior patterning a smooth and continuous concentration field or instead a chaotic collection of protein molecules each jiggling about in the haphazard way first described by Robert Brown in his microscopical observations of pollen? Is water, the great liquid ether of the living world, a collection of discrete molecules or instead a perfectly continuous medium with a density of ≈ 1000 kg/m^3? In this article, I will argue that these questions pose a false dichotomy since there are many different and powerful representations of the world around us. Different representations suit us differently at different times and it is often useful to be able to hold these seemingly contradictory notions in our heads simultaneously. Indeed, mathematics is not only the language of representation, but often is also the engine of reconciliation of such disparate views. In a letter to Alfred Russel Wallace on 14 April 1869, Charles Darwin noted that Lord Kelvin's "views on the recent age of the world have been for some time one of my sorest troubles". Here, I will argue that one of the highest attainments of the scientific enterprise is a coherent picture of the world, a picture in which our stories about the geological age of the Earth are coherent with our stories of how whales populated the oceans, our understanding of the living jibes with our understanding of the inanimate, our insights into the dynamics of genes and molecular structures are consonant with our physical understanding of the laws of statistical physics. The underpinnings of such coherency are often best revealed when viewed through the lens of mathematics.

Contents

> "Do I contradict myself?
>
> Very well then I contradict myself,
>
> (I am large, I contain multitudes.)"
>
> - Walt Whitman, Song of Myself, 51

1. Darwin and Kelvin reconciled

The natural sciences encounter the mystery of the world around us with an apparently limitless collection of "I wonder" questions. The age of the Earth is one such question that has tantalized and troubled people for millennia [1–3]. In parallel, others have wondered how the largest animals to have ever graced our planet moved from their ancestral life on land to "people" the world's oceans [4–6]. The tools of science allow us to little by little pull back the shroud which hides the

---

E-mail address: phillips@pboc.caltech.edu.

https://doi.org/10.1016/j.semcdb.2025.103646
Received 9 December 2024; Accepted 8 August 2025

scientific understanding of these mysteries. This special volume asks a deep question about the modern conception of the living world: "when should mathematical models be used in biology?" My plan is to extend the scope of that question by thinking more broadly about the role of mathematics in the natural sciences, illustrating that as in cases such as the Kelvin–Darwin debate, it is inevitable that seemingly quite distinct areas of science can and must be reconciled. Often, that reconciliation is a quantitative one based upon careful measurements and equally careful mathematical reasoning. The basis of such reconciliations in part derives from the philosophical requirement that quantitative data such as the age of the Earth or the timing of the loss of genes for enamel synthesis in whales demands quantitative models. The dictum of "quantitative data demands quantitative models", applies as much to the theory of the dynamics of the conversion of uranium to lead as to the theory of the molecular clock that ticks off the divergence of molecular sequences such as the pseudogenes found in the whale [7].

Let us dig a little more deeply into this question of quantitative representations and reconciliations, starting with more of the details on how Kelvin produced one of Darwin's "sorest troubles". I do not believe it is controversial to argue that our understanding of biological evolution and geology need to be self consistent. In the 19th century, Lord Kelvin's article "On the Age of the Sun's Heat" made several attempts to work out the age of the Sun, exploring multiple hypotheses including imagining the Sun as a giant chemical furnace. Kelvin concluded on that faulty hypothesis that by burning fuel with an energy density of $\approx 10^7$ J/kg (typical numbers for materials from foods to fuels): "The chemical theory is quite insufficient, because the most energetic chemical action we know, taking place between substances amounting to the whole Sun's mass, would only generate about 3000 years' heat". He brought these same quantitative sensibilities to bear on the question of the age of the Earth. For the science of his time, his hypothesis making and corresponding mathematics were on the mark. However, because of missing out on both radioactivity and the fact that there is internal heat convection within the Earth, his estimates were way off — thus the "sore troubles" of Darwin [1,2,8]. More recently, these same kinds of questions played out in fascinating ways in the context of the Galapagos Islands of Ecuador which dates the oldest islands as roughly 4.5 million years old [9], but with the divergence time of the famed tortoises from their South American relatives even farther in the past [10].

My first thesis in this paper is that mathematics is often a powerful engine of reconciliation. One such reconciliation that we all take so for granted that we do not even notice it is analytic geometry, the marriage of space and number we carry out every time we make a graph. Often, these graphs are made in response to some measurement such as the speed of a molecular motor as a function of ATP concentration [11–13] or the level of gene expression as a function of inducer concentration [14]. Hence, a framing of the question that is the basis of this present issue is to note that one of the paths to making scientific knowledge is by making measurements. I am definitely opposed to a narrow and one-dimensional view of the scientific process, sometimes dogmatically canonized in high-school caricatures as the scientific method. For example, there are times when simple unfettered observation, neither with an ambition for measurement or hypothesis making can still engender deep scientific insight [15]. Here, by way of contrast, I argue that the moment we make a measurement, we have invited mathematics in the front door. When we make graphs relating some variable x to another variable y, we have made a full mathematical embrace of the unification of geometry and number canonized by Déscartes now many centuries ago, and the theories used to describe such data must respond in the language of space and number as well.

We can construct a long list of examples where in my opinion, quantitative biological data requires us to respond mathematically, not verbally. Already decades ago now, single-molecule biophysicists began to characterize both the load and ATP dependence of motor velocities (for examples, see Refs. [11–13,16]), properties that are at the core of how these motors conspire to build structures such as the spindle that segregates chromosomes. In a different corner of biology, beautiful amputation experiments have been performed in which one flagellum of *Chlamydomonas* is removed with the result that the amputated flagellum starts to grow back while the uncut flagellum shrinks [17]! Amazing advances in fluorescence microscopy now make it possible to count the number of mRNAs in individual cells and from it to determine the mRNA distribution [18,19]. Here too such histograms are no longer the province of verbal descriptions and instead call for a theory of the transcriptome. Measurements change the conversation. The graphs, spreadsheets and databases of modern biology, whether of transcriptomes [20,21] and proteomes [22,23] or migratory patterns of marine mammals (see the Movebank Database and Happywhale for several examples), require us to respond in kind by developing quantitative hypotheses and using mathematics to work out their experimental implications.

There are many facets to the question of how to think about the use of mathematics in the study of the living. Here I will focus on the twin pillars of representation and reconciliation as the basis of thinking about the place of mathematics in the study of the living. In this paper I am going to focus on the "mathematics" side of the question. My sense is that there is so much we can learn by trying to better understand what mathematics is and does. As noted by mathematician Alain Connes in his article "Advice to the Beginner" in the *Princeton Companion to Mathematics*, "Mathematics is the backbone of modern science and a remarkably efficient source of new concepts and tools to understand the 'reality' in which we participate". As a tool for understanding the natural world, mathematics provides a way to sharpen our hypotheses and their implications [24]. In the case of Kelvin and Darwin, it took nearly a century for the different threads to come together in such a way that our quantitative understanding of the ages of the Earth and the Sun could be said to align with our quantitative understanding of the evolution of life on Earth.

Throughout this article I will make reference to many distinct examples, providing a quantitative catalog of sorts, with the hope that they will reveal the unexpected ways that ideas within mathematics are linked in inspiring ways. In turn, perhaps even more surprising, is when those ideas reach beyond their home in traditional mathematics to form the basis of insights in the natural sciences.

2. A language whose characters are triangles

Before fully embracing the animating question of "when should mathematical models be used in biology?" we first explore a vision of how we might think of the role of mathematics in the study of the natural world more broadly. Whenever we can, we will let mathematicians speak for themselves. More than four centuries ago, Galileo was struck by the quantitative character of his discoveries such as that the incremental distance traveled by a falling body in successive instants goes as the odd integers [25–28]. That is, after one second, the object has fallen a distance $\Delta s \approx 5$ m, in the next second, the object falls a distance $3\Delta s$, in the next second the object falls a distance $5\Delta s$. With the allied insight that the sum of the first n odd integers is n^2, Galileo articulated the idea that we now know in the formula

$$s = \frac{1}{2} g t^2, \tag{1}$$

a formula that is in modern pedagogy tied to the plug-and-chug philosophy of elementary physics exams rather than the confluence of experimental acumen and surprising hypothesis making which characterized its origins. Indeed, so moved was he by his various insights into the mechanics of the world, both celestial and terrestrial, Galileo asserted: "[The universe] cannot be read until we have learnt the language and become familiar with the characters in which it is written. It is written in mathematical language, and the letters are triangles,

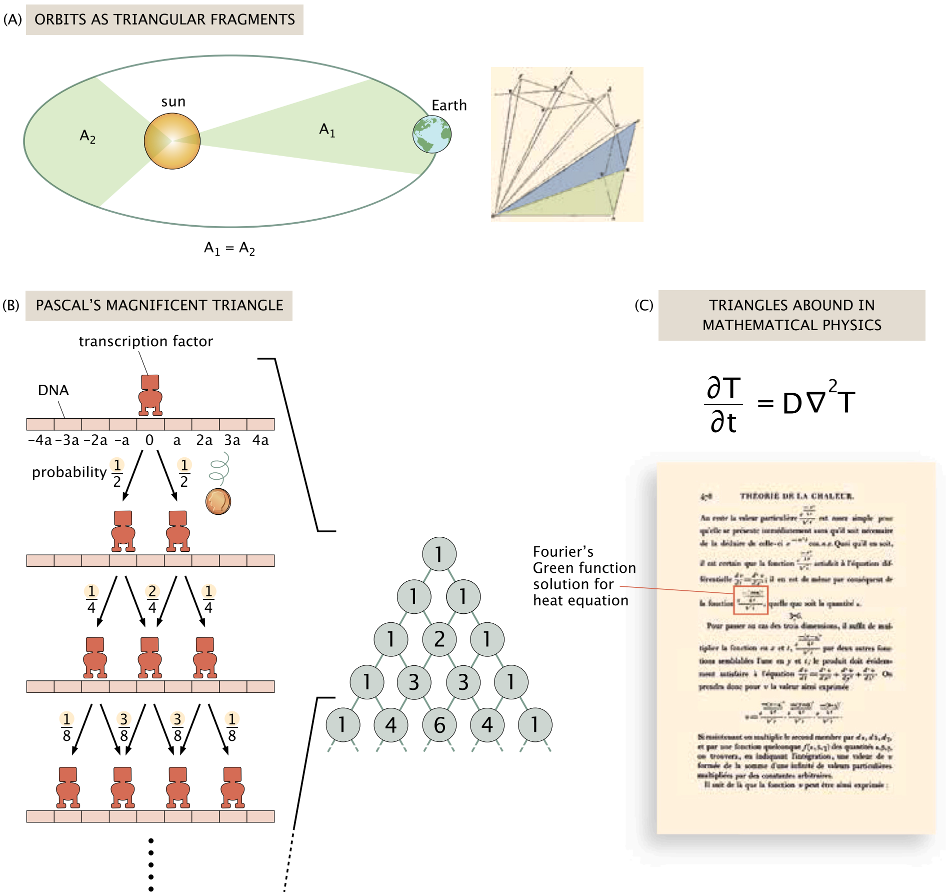


Fig. 1. A language whose characters are triangles. (A) Newtonian orbits as a succession of triangles. In his *Principia Mathematica* (figure on the right), Newton proved Kepler's law on equal areas being swept in equal times (i.e. $A_1 = A_2$) by showing the equality of triangles, two of which are shown in green and blue here. (B) Pascal's triangle and the dynamics of binary choices. Diffusion of transcription factors along DNA can be thought of as a random walk with the probabilities of different positions after different times dictated by the entries in Pascal's triangle. (C) The Laplacian denoted by an upside down triangle is used in many of the partial differential equations of mathematical physics including the heat equation shown here. The page from Joseph Fourier's "Analytical Theory of Heat" offers his solution to the heat equation shown at the top using what we now call a Green function.

circles and other geometrical figures, without which means it is humanly impossible to comprehend a single word". (Galileo Galileo Opere Il Saggiatore p. 171) With tongue in cheek, Fig. 1 shows just how right Galileo was with three disparate examples of the way in which triangles have figured centrally in the mathematical description of nature.

Similar sentiments to those articulated by Galileo were expressed centuries later in a now famous article by Eugene Wigner [29] in his piece on "The Unreasonable Effectiveness of Mathematics in the Natural Sciences" which highlights the many unexpected ways that mathematics insinuates itself into the way we think and talk about the world around us. Wigner tells a humorous story of the presence of π in the famed Gaussian distribution

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \tag{2}$$

Wigner admits that it is indeed remarkable that a number that quantifies the ratio of the circumference and the diameter of a circle could make such an unexpected announcement in the context of a problem such as the distribution of some variable of interest such as human height.

Most of the time when using mathematics, we are fully aware that our treatment of the question of interest is approximate. In his wonderful book "Mathematics: A Very Short Introduction", famed mathematician Timothy Gowers takes stock of the difference between the perception of mathematics and its reality [30]. "Most people think of mathematics as a very clean, exact subject. One learns at school to expect that if a mathematical problem can be stated succinctly, then it will probably have a short answer, often given by a simple formula. Those who continue with mathematics ⋯ soon discover that nothing could be further from the truth... most of the time one must settle for a rough estimate instead. Until one is used to estimates, they seem ugly and unsatisfying. However, it is worth acquiring a taste for them, because not to do so is to miss out on many of the greatest theorems and most interesting unsolved problems of mathematics". The second main thesis of my article is that mathematical approximations are the lifeblood of the physical modeling of the natural world too. Fig. 2 provides a glimpse into the power of mathematical representations and approximations. There we see the master craftsman Isaac Newton at work during one of the multiple times he computed ln $(1 + x)$ for $x = 1/10$. In particular, we see Newton approximate the logarithm (also the area under the curve $y = 1/x$ we can see in the top left of the figure) using the series

$$\ln(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} + \cdots. \tag{3}$$

For the special case of $x = 1/10$, this takes the beautiful form

$$\ln\left(1 + \frac{1}{10}\right) = \frac{1}{10} - \frac{1}{200} + \frac{1}{3000} - \frac{1}{40,000} + \frac{1}{500,000} + \cdots. \tag{4}$$

The two downward pointing triangular pyramids of numbers seen in the figure reveal Newton separately computing the positive and negative terms in his approximate answer out to more than 50 decimal places.

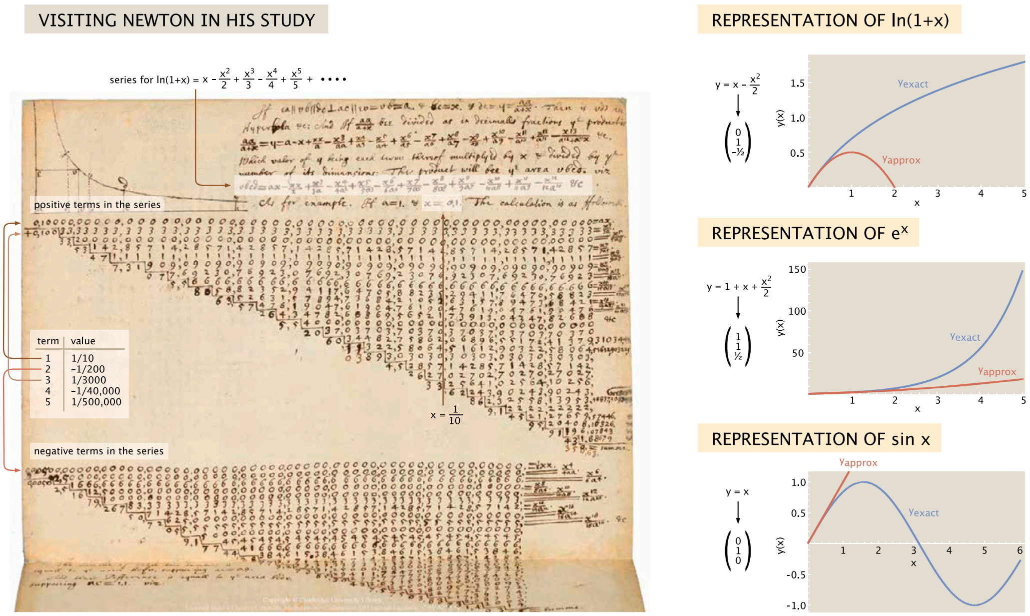


Fig. 2. Representing functions by polynomials and series. (A) Newton's "College Notebook" features a huge variety of examples of the explorer in action. Here he relates the area of the hyperbola and the logarithm to a series. The series for ln $(1+x)$ is highlighted and Newton chooses $x = 0.1$ and proceeds to calculate an approximate value of ln $(1+x)$ to more than 50 decimal places. There are several examples of these calculations in the Newton papers including MS Add. 4000 - 20r and MS Add. 3958 - 78v. (B) Three examples of replacing functions (blue) with corresponding second-order polynomials (red), meaning that each such function is entirely defined by three numbers!

3. Mapmaking: Representations and approximations

Timothy Gowers is not the only one who had deep insights into the always approximate way in which we represent the world around us. Virtuoso master craftsman of the imagination and the written word, the Argentinian Jorge Luis Borges, wrote a short story masterpiece that is only a paragraph long entitled "On Exactitude in Science" which reports on an empire in which the cartographer's guild made maps that were the size of the empire. The occupants of Borges' empire "saw that that vast map was useless". Like Borges, Picasso understood the importance of representation and approximation as evidenced by his eleven 1945 lithographs entitled "The Bull". These inspiring drawings celebrate abstraction and force us to wonder what are the minimalistic ways we can find of representing the super complex world we all wander around in. Mathematics is one of the most important ways we have of constructing maps that are not the size of the empire itself. Fig. 2 shows how using only three numbers in the form of the coefficients a_0, a_1 and a_2 in the representation $y = a_0 + a_1 x + a_2 x^2$, we can represent the functions $\ln(1+x)$, e^x and $\sin x$, about which we will have more to say below.

One of our most compelling witnesses for the transformative power of mathematics for our thinking is William Thurston. In a series of inspiring essays and online posts (see Thurston on math*overflow*), he gave a beautiful rendering of his views on the nature of mathematics [31]. One way in which he crystallized his thinking might surprise many people — he took issue with the idea that mathematics is only about proving theorems [32] — rather, holding that mathematics is a way of thinking that formalizes the search for formal patterns. But perhaps more importantly, Thurston argues that the necessity of the mathematical mindset derives from the fact that: "The world does not suffer from an oversupply of clarity and understanding", and mathematics offers a path to both. One of the central arguments of this essay is the idea that how we represent biological systems can alter how we think about them. To see how different representations can alter our perceptions both conceptually and psychologically, we begin with a nonbiological childhood example of "long division". On this topic, Thurston notes: "I remember as a child, in fifth grade, coming to the amazing (to me) realization that the answer to 134 divided by 29 is 134/29 (and so forth). What a tremendous labor saving device. To me, '134 divided by 29' meant a certain tedious chore while 134/29 was an object with no implicit work". Let us really take Thurston seriously. I invite the reader to first stare at the representation

$$29\overline{)134} \tag{5}$$

and to think about how that feels and what it means to you. Like Thurston, to me, this expression immediately conjures labor and hard thinking and mental gymnastics and trial and error. For most of us this symbolic representation conveys an algorithm, a specific prescription for creating a decimal representation of a particular fraction, namely, 4.620689655172413793..., . Alternatively, let us stare at the completely different

$$\frac{134}{29}. \tag{6}$$

The latter representation is a fraction and either a position somewhere along the real number line, or taking the fraction 1/29 and stacking it up against itself 134 times.

As we muse on the relationship between mathematics and biology it is important that we embrace the approximate and at first blush, apparently contradictory ways in which we represent the world around us. The importance of our point of view is especially evident when we use maps such as that of Manhattan shown in Fig. 3. Few cities are so noted for such a spectacularly orderly Cartesian coordinate system with its orthogonal grid of avenues and streets. In that figure, we choose an origin of coordinates at the intersection of Avenue of the Americas and Central Park South. The blue vector points to an address somewhere on the East Side not far from Rockefeller University. We can defiantly turn our backs on the convenient Cartesian coordinates of the end point of that vector, instead deciding to specify all New York City addresses with polar coordinates (r, θ) as a bird might. There is nothing wrong with such a description, but it is decidedly inconvenient for human navigation and travel. A more serious example of the price we pay for approximate representations in our maps comes from a careful study of a world map. The geographical reality, despite the

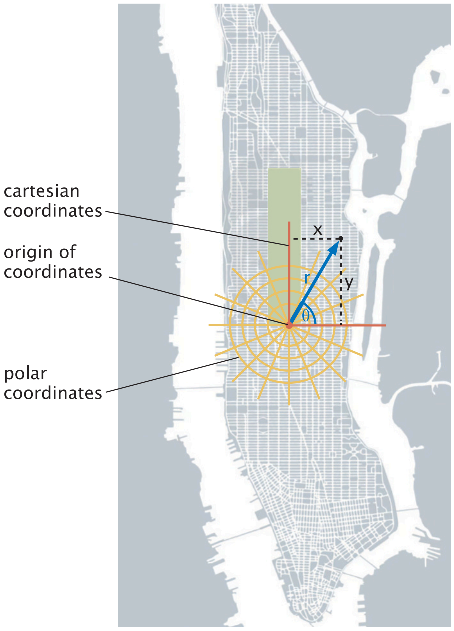


Fig. 3. A map of Manhattan and a coordinate system gone terribly wrong. The Cartesian coordinate system is more natural in terms of the rectangular grid of streets. However, in either Cartesian or polar coordinates, the blue vector points to a destination and that vector is indifferent to the coordinates used to describe it.

evidence of what we see on this distorted planar representation, is that the area of the United States and Canada are both approximately 10 million km^2, while that of Greenland is roughly 5 times smaller at ≈ 2 million km^2. That said, it is not a particularly nuanced take on such maps to call them "wrong" since for generations mathematicians have formulated theorems that they understand very well that tell us about the distortions that must be present if we adopt a planar representation of objects (such as continents) on a sphere. Holding a more refined idea of representation and approximation can guard us against falling into the knee-jerk and dismissive trap of referring to models of biological phenomena as "wrong", which is often neither helpful nor even true. Many of these models are "wrong" in a similar way that a Mercator Projection is wrong.

As an introduction to the simultaneous ambiguity and precision of different representations, Fig. 4 provides a gallery of examples. On a light note, we begin by thinking about the way we and others represent ourselves. It is intriguing to imagine that certain personality tests such as the Enneagram attempt to embody the entirety of a human's personality in a single scalar parameter. The famed Briggs–Myers test used in some job interviews increases that complexity by considering a person as a four component vector. A much more nuanced, but perhaps biased view of ourselves might be provided either in the form of an autobiography or a biography, characterized by a roughly 10^6 dimensional vector of letters, punctuation marks and spaces.

An equally surprising and perhaps more successful representation shown in the second panel of the figure characterizes the distribution of prime numbers. This example is meant to both inspire our readers and to clarify the perhaps surprising realization of how often mathematics concerns itself with approximations. We can represent the primes by a function which increases by one every time we hit the next prime as we walk along the x-axis, a function sometimes denoted the staircase of primes and shown in red in the figure [33]. To put the problem in perspective, we do best to quote from a beautiful article by Don Zagier [34] who says "There are two facts about the distribution of prime numbers of which I hope to convince you so overwhelmingly that they will be permanently engraved in your hearts. The first is that, despite their simple definition and role as the building blocks of the natural numbers, the prime numbers grow like weeds among the natural numbers, seeming to obey no other law than that of chance, and nobody can predict where the next one will sprout. The second fact is even more astonishing, for it states just the opposite: that the primes exhibit stunning regularity, that there are laws governing their behavior, and that they obey these laws with almost military precision". Some of that military precision is revealed statistically in the figure where we see how the mathematicians Legendre, Gauss and Riemann provided beautiful and approximate representations of the staircase of primes [33,35,36].

In the early 1800s, Joseph Fourier set himself the problem of doing nothing less for the study of heat than Newton before him had done for the study of motion. His resoundingly successful approach to these problems appeared in his "Théorie analytique de la chaleur" (Analytical Theory of Heat), a page of which was already shown in Fig. 1(C). As part of that enterprise, Fourier introduced ways of representing mathematical functions that we now refer to as Fourier series and Fourier integrals (or Fourier transforms). The representation shown in the figure both in equation and graphical form is that we can represent any sufficiently well behaved periodic function as a superposition of sines and cosines.

The next example that teaches us how to hold many representations of the same thing in our mind at one time is offered by the symmetry of an equilateral triangle. Here we opt for the minimal symmetry of rotations about the center, noting that the three operations of leaving the triangle alone (I), rotating by 120 degrees (R_{120}) and rotating by 240 degrees (R_{240}) leaves the triangle "unchanged". Further, as seen in the "multiplication table" we can take two operations in succession and this will leave us with an equivalent representation of the triangle. Such a closed set of operations is known as a group in mathematics, and the matrices in the figure show different "representations" of the group of rotations describing the symmetry of an equilateral triangle.

Our next example is the fascinating way in which we can arbitrarily deform a solid material and represent the geometry of that deformation by imagining a 3×3 symmetric matrix known as the strain tensor which lives at every point within the material. As seen in the figure, such a material can undergo a uniform volume expansion, a shear deformation or even a twist, and in all of these cases, the elements of the strain tensor tell us about the relative motions of material points within the materials.

Finally, in thinking about the motility of a parasite such as *Toxoplasma gondii* several different representations can be used simultaneously. First, the shape of the cell as measured using X-ray tomography can be written in terms of the so-called spherical harmonics. But then given that approximate representation of the shape, a second representation using the finite element method can be used to describe the spatial distribution and motion of the actin filaments that drive that motility [37]. What all of these examples share in common is that some part of the world is *represented* mathematically, and there are multiple distinct representations of the same thing.

The list shown in Fig. 4 only scratches the surface of the great diversity of mathematical representations that have been invented and discovered for representing the world around us. As seen above, the shapes of *Toxoplasma* cells can be represented using spherical harmonics. Alternatively, the vibrations of the drums used by a drummer such as Chad Smith of the Red Hot Chili Peppers can be represented using Bessel functions. There are many legendary and intimidating names such as Fourier, Bessel, Legendre, Airy, Laguerre, Hermite, Chebyshev and many others each of which is attached to the special functions of mathematical physics [38]. Before the era of computers, these special functions were the subject of enormous tables [39], calculated laboriously using the kinds of tricks we saw in the hands of Newton in Fig. 2. But another way of thinking of these special functions with their intimidating names is as the "right" basis vectors for representing

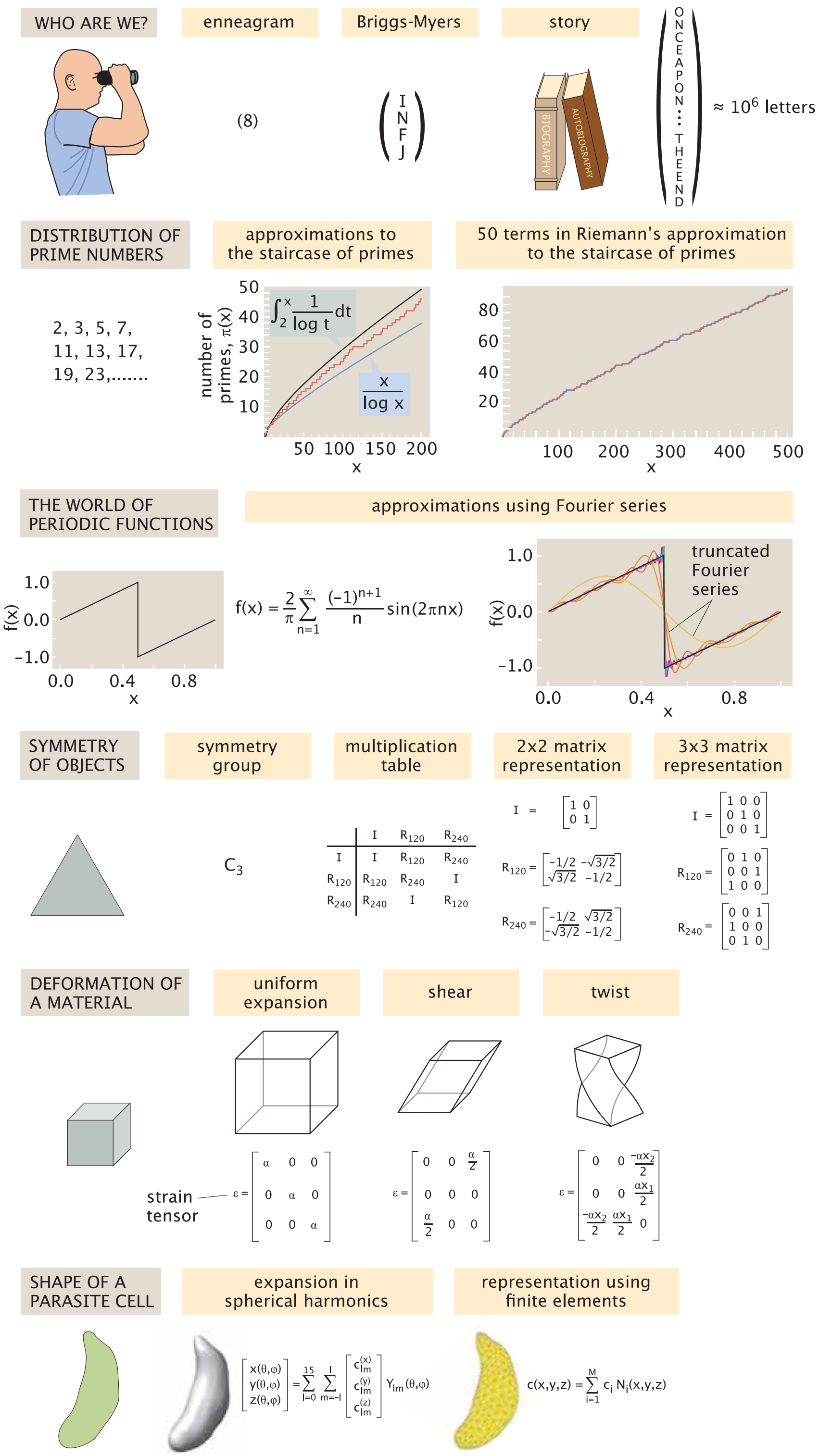


Fig. 4. A variety of examples of representations that formalize patterns and compress topics that might otherwise be described with words. The first row shows different attempts to represent a human being in terms of vectors of different dimensions. The second row shows different treatments of the distribution of prime numbers. The third row shows the all-important representation of a sawtooth function as an example of much broader "Fourier decomposition" of functions. The next row captures the symmetries of an equilateral triangle by three symmetry operations that are closed under multiplication. The next-to-last row shows how we can use the strain tensor to describe the deformations of a solid material. Finally, we show two distinct representations of the shape and concentration field of actin in a parasite cell.

functions in situations with high symmetry such as cubes (Fourier), spheres (Legendre) and cylinders (Bessel).

The advent of the computer has been a complete game changer redefining what we consider a useful or viable representation, with methods such as the finite element method and the composition of functions represented in neural networks making it possible to reimagine what we mean by a representation. The special functions of classical mathematics and mathematical physics are beautiful and powerful, but

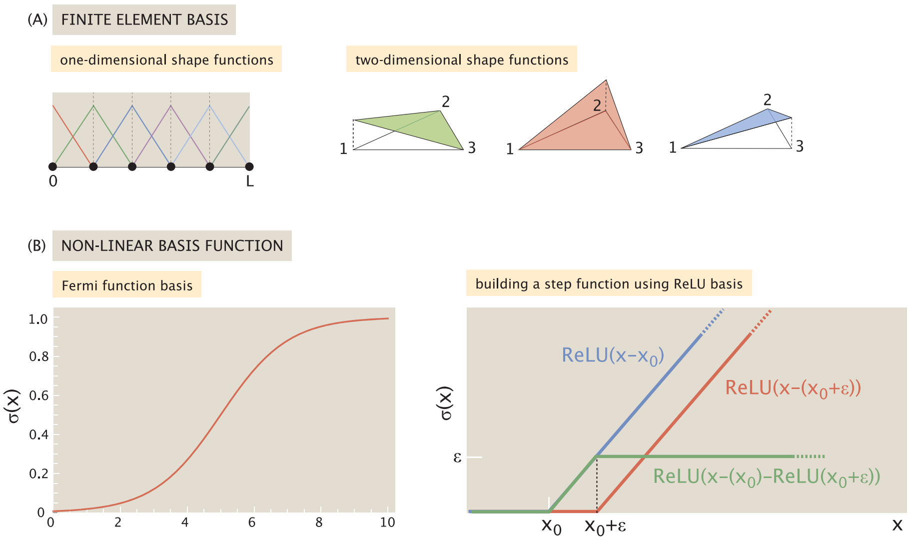


Fig. 5. A new generation of basis functions that are convenient to use on the computer. These basis functions fit squarely within the long history of mathematical approximation and interpolation [40] (A) One dimensional shape functions for representing a function $f(x)$ in the form $f(x) = \sum_i f_i N_i(x)$. The two-dimensional example shows the generalization of tent functions for two-dimensional interpolation. (B) Representation of functions using composition of functions. Several approaches to constructing a "step function" using simple Fermi functions and ReLU functions are illustrated schematically. Given a "step function" we can construct a "delta function" in turn [41].

painfully restrictive in the geometries they consider. By way of contrast, with the advent of computational "shape functions" we now live in a world of very simple basis functions that permit a huge variety in the situations that can be treated. Writ large, these shape functions tell us that if we want to represent some vector field $\mathbf{u}(\mathbf{x})$ we can do so by defining the field at "nodes" with value $\mathbf{u}_i$ on the ith node and between the nodes, we represent the field strictly by interpolation of the form

$$\mathbf{u}(\mathbf{x}) = \sum_{i=1}^{N} \mathbf{u}_i N_i(\mathbf{x}), \tag{7}$$

where $N_i(\mathbf{x})$ is the shape function associated with the ith node. Fig. 5(A) makes these ideas concrete by showing one- and two-dimensional examples of finite element shape functions. In the one-dimensional example, the shape functions $N_i(x)$ adopt a tent like shape centered on each node. If we wanted to represent a concentration field $c(x)$, we would do so as $c(x) = \sum_{i=1}^{N} c_i N_i(x)$, where c_i is the discretized value of the concentration at the ith node. To find the concentration elsewhere, the shape functions imply nothing more than linear interpolation between the nodes. Simple but highly effective!

So far we have seen that we can represent functions in many different ways. We can add up sines and cosines or add up shape functions that look like tents, and in many cases, both of them work! But now in keeping with this theme of the power unlocked by the computer in totally novel kinds of representations, we explore a second class of "simple" basis functions with an unprecedented power of "representation" as shown in Fig. 5(B) [41–43]. In this discussion, we arrive with openness to the unexpected ways in which mathematics that appears to have nothing to do with questions about the natural world end up mattering very much to those endeavors. Machine learning is all the range, leaving no corner of the study of life untouched. And yet, at its core, one of its principal features is that of mathematical representation. In many cases when we use these modern algorithms, just beneath the surface are a series of "weights" that tell us how much we are going to "mix" basis functions like those shown in Fig. 5(B). Those weights are numbers inside of those descriptions and those numbers describe mappings in exactly the same way that the tiny three-dimensional vectors $(0, 1, -1/2)$, $(1, 1, 1/2)$ and $(0, 1, 0)$ were introduced to capture the essences of $ln\,(1+x)$, e^x and $sin\;x$, respectively, as shown in Fig. 2 for small x.

How does it work? As shown in Fig. 5(B), the functions that we use as the basis of our representations have traditionally assumed several forms [43]. One of them is known in physics as the Fermi function in honor of the way that Enrico Fermi showed how to construct a quantum mechanical theory of a monatomic gas. However, that functional form is of much wider significance and is given by

$$\sigma(x) = \frac{1}{1+e^{-x}} \tag{8}$$

with the characteristic sigmoidal shape seen in the figure. A second even simpler functional form is given by

$$\mathrm{ReLU}(x) = x_+ = \max(x, 0) \tag{9}$$

which as shown in the figure is a function that rises linearly as long as x is greater than zero. Though our language here is overly casual, for our present purposes the key insight is to recognize that we can construct steps and spikes out of either the Fermi or ReLU elements. For example, we can make a "spike" from $w_1\sigma(x) + w_2\sigma(x-\epsilon)$ where the "weights" are given by $w_1 = -w_2 = 1$. In the context of some "hidden layer" between an input vector $\mathbf{x}_k$ and the output $\mathbf{v}_k$ we have a nonlinear transformation built using the basis functions shown above an described by the equation

$$\boldsymbol{v}_k = F_k\left(\boldsymbol{v}_{k-1}\right) = \mathrm{ReLU}\left(A_k \boldsymbol{v}_{k-1} + \boldsymbol{b}_k\right). \tag{10}$$

This input–output function takes in the vector $\boldsymbol{v}_{k-1}$, performs the linear transformation $A_k\boldsymbol{v}_{k-1} + \boldsymbol{b}_k$ and finishes by imposing the nonlinear transformation made by the ReLU function.

Nowhere that I know of is the magic of machine learning better explained than by Gilbert Strang who argues: "Here is the indispensable ingredient in the learning function F. The best way to create complex functions from simple functions is by **composition** [43]". Just as the finite element representation of functions provides a powerful way of representing functions on beautifully complex geometries by using simple interpolation, a second kind of numerical representation based upon composition of functions is equally impressive in its simplicity and its reach. In particular, the composition of which Strang speaks is given by

$$F(\boldsymbol{v}) = F_L\left(F_{L-1}\left(\ldots F_2\left(F_1(\boldsymbol{v})\right)\right)\right). \tag{11}$$

One of the most fascinating aspects of these representations is that we can use this composition of functions to create a spike (more formally known as a delta function). The reason this is so important was long ago demonstrated in the context of the so-called Green function shown in Fig. 1(C) which tells us how to build up solutions by adding up the contributions due to individual spikes.

Here we have only scratched the surface of the myriad of different ways that we can represent our understanding of the world around us using mathematics. Always in the back of my mind, I hold dear both the Borges story and the Picasso lithographs, both focused on finding abstract, approximate and reduced descriptions of the world we see around us. When writing approximate mathematical descriptions of the natural world this is not naivete or ignorance or laziness, it is the height of sophistication. Maps the size of the empire are not helpful.

4. Success stories in biological representation

One of the delights of exploring the world from a mathematical perspective is both the unexpected origins of some of our most fundamental mathematical ideas and the almost bizarre way in which those ideas end up in places we would not expect. The great mathematicians Blaise Pascal and Pierre de Fermat carried on a prolonged correspondence about seemingly trivial games of chance, only for us to realize centuries later that as noted by ET Jaynes, probability is the language of science [44]. The utter seriousness of the mathematics of these games of chance really came into its own in the hands of Pierre Simon de Laplace who was so confident in his estimates of the mass of Jupiter and Saturn he was willing to make a bet [45]. In particular, speaking of the estimate of a certain M. Bouvard on the mass of Saturn, Laplace notes: "His calculations give him the mass of Saturn equal to the 3512th part of that of the Sun. Applying to them my formulae of probability, I find that it is a bet of 11,000 against one that the error of this result is not 1/100 of its value, or that which amounts to almost the same - that after a century of new observations added to the preceding ones, and examined in the same manner, the new result will not differ by 1/100 from that of M. Bouvard". Amazing (and he was right). In modern times, this mathematics has become the basis of much of the enterprise of exploring, comparing and understanding genomes.

Another apparently idle amusement in the early 1700s was a game one might have contemplated on a Sunday afternoon. As told by Euler in his paper that launched modern graph theory [46]: "The problem which I am told is widely known, is as follows: in Königsberg in Prussia, there is an island A, called *the Kneiphof*; the river which surrounds it is divided into two branches, as can be seen in Fig. 6(A), and these branches are crossed by seven bridges, a, b, c, d, e, f and g. Concerning these bridges, it was asked whether anyone could arrange a route in such a way that he would cross each bridge once and only once". In a daring act of graph theory abstraction, Euler showed that try as they might, no one out on a Sunday afternoon stroll would ever succeed in meeting this challenge.

Roughly one hundred years later with the advent of an ever improving phenomenological understanding of the input–output properties of electrical circuits such as the schematic shown in Fig. 6(B), Gustav Kirchhoff articulated the laws of such circuits that now bear his name and with it extended the subject of graph theory to include the matrix tree theorem. Little could he know that a century and a half later those ideas would become what I think of as a candidate for one of the most profound unifying insights for the mathematicization of much of modern biology [47–49]. In particular, one of the ways we can frame many problems in biology is as questions of input–output functions of the form

$$\mathbf{v} = \mathbf{g}(\mathbf{x}), \tag{12}$$

where $\mathbf{x}$ is some vector of "inputs", $\mathbf{v}$ is some vector of outputs and $\mathbf{g}(\mathbf{x})$ is the input–output function itself. Interestingly, by linking ideas from statistical physics and graph theory, one can write down mathematical descriptions of response functions for situations ranging from enzyme action to the response of membrane proteins such as ligand-gated ion channels and chemotaxis receptors to the genetic circuits that preside over much of cellular physiology such as the one shown in Fig. 6(C) [50]. For the cases of interest here, these response functions can be written as rational functions of the form

$$\text{Response}(x) = \frac{F(x)}{G(x)} = \frac{a_0 + a_1 x + a_2 x^2 + \cdots a_N x^N}{b_0 + b_1 x + b_2 x^2 + \cdots b_M x^M} \tag{13}$$

requiring a theory of the coefficients a_0, $a_1 \cdots$ and b_0, $b_1 \cdots$ like those we saw in the very low-dimensional context in Fig. 2 where we represented quadratic polynomials as three-vectors of the form (a_0, a_1, a_2). Interestingly, graph theory gives a very clean and transparent way of determining the probabilities of each of the states (i.e. the nodes on the graph), and hence of the response function in Eq. (13) [48,51].

More than a century of brilliant efforts on input–output functions have resulted in a variety of different models for the coefficients a_0, $a_1 \cdots$ and b_0, $b_1 \cdots$. One of the classic case studies concerns the oxygen binding proteins myoglobin and hemoglobin. We know that the degree of saturation of the myoglobin and hemoglobin molecules in a sperm whale is a function of the oxygen partial pressure as it takes its last breath before a dive. More than 100 years ago, Archibald Hill wrote down a description of this binding that we now know as the Hill function which for hemoglobin he wrote as

$$p_{bound}(x) = \frac{\left(\frac{x}{K}\right)^n}{1 + \left(\frac{x}{K}\right)^n}, \tag{14}$$

where x is the concentration of O_2 and K is its allied dissociation constant. As Hill himself tells us, this functional form provides a summary of the occupancy of hemoglobin (the example he used, though it has been applied much more broadly). If we think of the huge topic of input–output functions in biology, then the kind of characteristics embodied in the Hill approach include a representation of leakiness (the amount of output even in the absence of input, $p_{bound}(0)$), dynamic range, EC_{50} (the concentration at which the output reaches half its maximum, $EC_{50} = K$) and the sensitivity as measured by the slope of the input–output curve (usually in logarithmic variables) at the midpoint. It is instructive to hear Hill himself commenting on his thinking: "My object was rather to see whether an equation *of this type* can satisfy all the observations, than to base any direct physical meaning on n and K [52]". He goes further in his 1913 paper noting [53] "In point of fact n does not turn out to be a whole number, but this is due simply to the fact that aggregation is not into one particular type of molecule, but rather into a whole series of different molecules: so that Eq. (1) (our Eq. (14)) is a rough mathematical expression for the sum of several similar quantities with n equal to 1, 2, 3, 4 and possibly higher integers". We think it important to remember that the Hill function is a phenomenological description of equilibrium binding. In that sense, the measured binding curves of some macromolecule such as hemoglobin are an abstract representation of our understanding of the function of that molecule.

Hill's work was just the beginning. One of the great advances of 20th century biophysics was the Monod–Wyman–Changeux model of allostery [50,54–58] which gave its own description of the binding of oxygen to hemoglobin [59–61]. In this case, the average number of bound O_2 molecules per hemoglobin can be written precisely in the form of Eq. (12) as

$$\langle N([O_2])\rangle = 4\,\frac{e^{-\beta\varepsilon}\frac{[O_2]}{K_R}(1+\frac{[O_2]}{K_R})^3 + \frac{[O_2]}{K_T}(1+\frac{[O_2]}{K_T})^3}{\mathrm{e}^{-\beta\varepsilon}(1+\frac{[O_2]}{K_R})^4 + (1+\frac{[O_2]}{K_T})^4}. \tag{15}$$

Here we use angular brackets ⟨thing⟩ to represent the average of the thing within the brackets, and hence, $\mathbf{v} = \langle N([O_2])\rangle$, is the average number of O_2 molecules per hemoglobin and $\mathbf{x} = [O_2]$. The allosteric model of hemoglobin recognizes two conformational states of the protein known as tense (T) and relaxed (R) and those two states each have their own K_ds labeled by K_T and K_R, respectively.

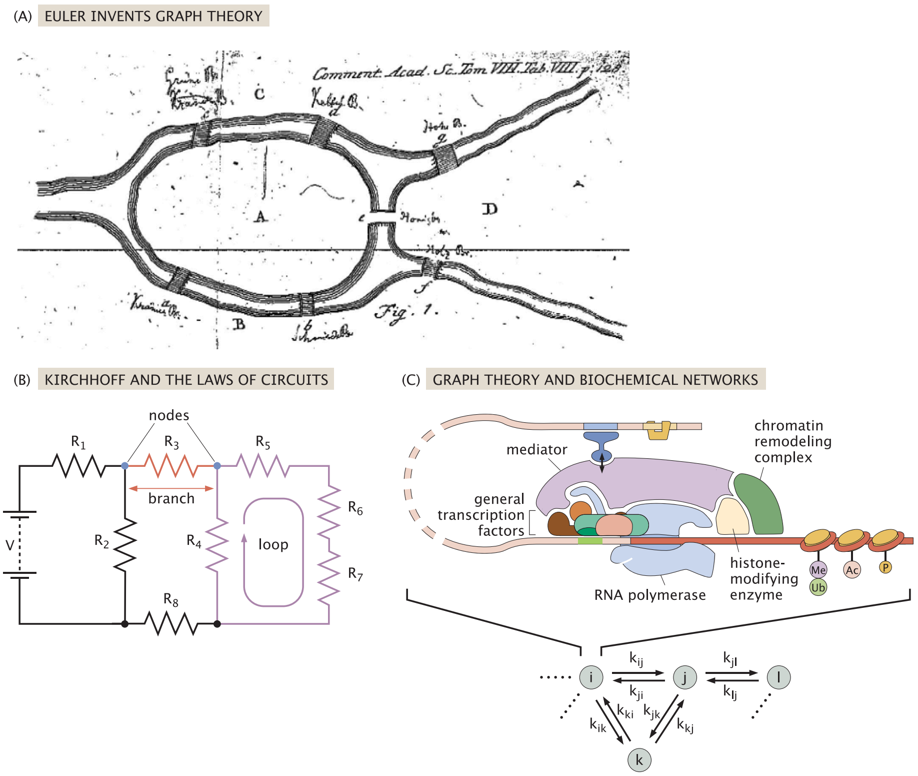
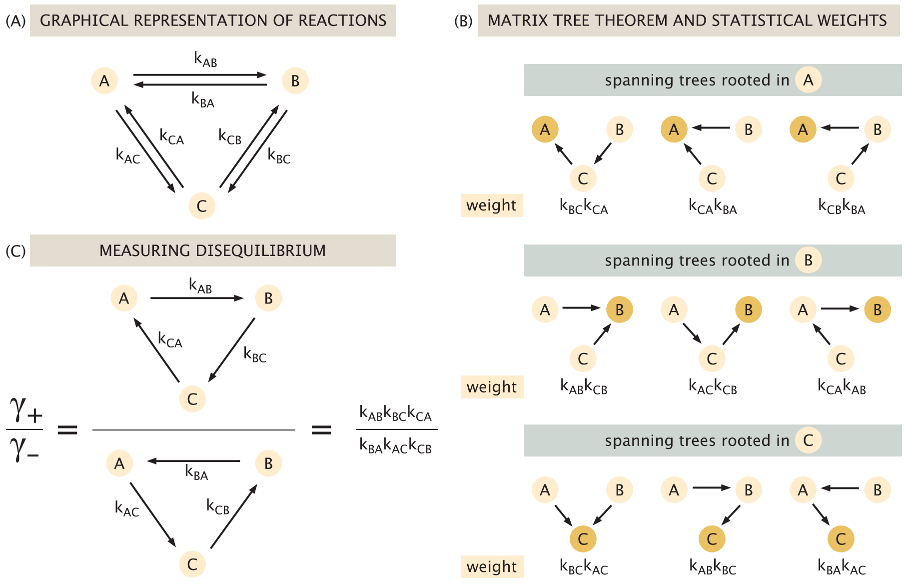


Fig. 6. The rise of graph theory and its biological implications. (A) Figure from Euler's original paper on the problem of the bridges of Königsberg. (B) Current flow in circuits can be computed on the basis of Kirchhoff's laws. (C) The decoration of a promoter by polymerase and attendant proteins such as transcription factors can be thought of as a graph with each node in the graph corresponding to one particular state of occupancy. Transitions between these states are the edges of the graph and associated with rate constants as shown in the figure. This part of the figure is inspired by many articles by J. Gunawardena and collaborators.



$$\frac{\gamma_+}{\gamma_-} = \frac{k_{AB}k_{BC}k_{CA}}{k_{BA}k_{AC}k_{CB}}$$

Fig. 7. Graph theory treatment of a triangular reaction network. (A) The states of the system and the rate constants describing the transitions between those states are defined. (B) The rooted spanning trees for all three nodes of the graph. Each row shows all the rooted spanning trees at that node as well as their statistical weights. (C) Fundamental ratio that reveals whether detailed balance is broken around the cycle.

The graph theory ideas of Kirchhoff's matrix tree theorem, like the MWC model, are of great generality. All the states of a biochemical network are nodes on the graph as shown in Fig. 6(C). For the special case of a triangular network, those nodes are shown in Fig. 7(A). The edges on the graph correspond to biochemical transitions between states. Often, the goal of studying such a biochemical network is to obtain the steady-state probabilities. This is especially pertinent for biology since the conception of equilibrium states so prevalent in physics need here to be replaced by the more general notion of nonequilibrium steady states. The matrix tree theorem tells us how to write these probabilities in terms of the so-called rooted spanning trees as highlighted in Fig. 7(B) and resulting in state probabilities. By using the statistical weights

$$p_A = \frac{k_{BC}k_{CA} + k_{CA}k_{BA} + k_{CB}k_{BA}}{k_{BC}k_{CA} + k_{CA}k_{BA} + k_{CB}k_{BA} + k_{AB}k_{CB} + k_{AC}k_{CB} + k_{CA}k_{AB} + k_{BC}k_{AC} + k_{AB}k_{BC} + k_{BA}k_{AC}}. \tag{16}$$

Box I.

that are obtained by summing the weights from each spanning tree as shown in Fig. 7(B), we can write the steady-state probability as in Box I. A more transparent rendering of this result is gotten by dividing top and bottom by the numerator resulting in

$$p_A = \frac{1}{1 + \frac{k_{AB}\left(k_{CB}+\frac{k_{AC}k_{CB}k_{BA}}{k_{AB}k_{BA}}+k_{CA}\right)}{k_{BA}\left(k_{CB}+\frac{k_{BC}k_{CA}k_{AB}}{k_{BA}k_{AB}}+k_{CA}\right)} + \frac{k_{AC}\left(k_{BC}+\frac{k_{AB}k_{BC}k_{CA}}{k_{AC}k_{CA}}+k_{BA}\right)}{k_{CA}\left(k_{BA}+\frac{k_{CB}k_{BA}k_{AC}}{k_{CA}k_{AC}}+k_{BC}\right)}}. \tag{17}$$

This result has several important features. First, as noted by examining Fig. 7(C), there is a very special ratio that measures the extent to which the biochemical circuit is driven out of equilibrium. Further, for the case in which $\gamma_+/\gamma_- = 1$, the terms with parentheses all cancel out and we recover the very simple functional form found when the network is in equilibrium.

In the remainder of this section, we explore several biological phenomena, each of which can be thought of as requiring theories of the coefficients in Eq. (13). For our present purposes, we will specialize to the case in which the output is a scalar that describes the magnitude of some output function such as the level of gene expression or the amplitude of some signaling response. In the case of the activity of chemotaxis receptors, using the tools of statistical mechanics, we can write the activity of the molecular circuit of interest as

$$p_{active}(c) = \frac{(1+\frac{c}{K_A})}{(1+\frac{c}{K_A}) + e^{-\beta\Delta\epsilon}(1+\frac{c}{K_I})}, \tag{18}$$

where K_A is the dissociation constant for binding in the active state and K_I is the dissociation constant for binding in the inactive state. As it stands, as shown in the top left panel of Fig. 8, a more sophisticated variant of this equation [62] can be used to characterize a variety of chemotaxis mutants. But the description really reveals its power when we rewrite the equation as

$$p_{active} = \frac{1}{1 + \mathrm{e}^{-\beta(\epsilon_I-\epsilon_A)}\frac{1+\left(c/K_I\right)}{1+\left(c/K_A\right)}}. \tag{19}$$

which invites the interpretation

$$p_{active} = \frac{1}{1+e^{-\beta F_{eff}}}, \tag{20}$$

where F_{eff} is an effective free energy that measures the relative free energy of the microscopic states of the system and is defined as

$$F_{eff} = \Delta\epsilon + k_BT\ \ln\frac{\left(1+\frac{c}{K_I}\right)}{\left(1+\frac{c}{K_A}\right)}. \tag{21}$$

Though we will not enter into all of the details here, a similar analysis can be performed in the context of inducible transcription factors as shown in the right panel of the figure.

Note that in both of the cases shown in Fig. 8, these are not fits! They are a reflection of real understanding [62,63]. This understanding implies not only that we think we know how to compute the input–output properties of some molecular pathway, but further, that we know how to figure out which variables the cell cares about, as opposed to the variables manipulated by the person doing the experiment. If the reader gets nothing more out of this article than coming away with an appreciation for the deep insights that come from performing data collapses such as those shown in Fig. 8, the article will have been a success.

5. Embracing the magic

The central argument of this paper is that mathematics is one of the most powerful languages we have for describing the living world. It allows us to vividly see things that in many cases would remain unseen. The fact that mathematics has been a huge part of the success of our description of the physical world is well known. But mathematics has already proven its unreasonable effectiveness in the study of the living world too [64–68]. Some of my own favorite Hall of Fame examples include: the use of the binomial theorem as a tool to uncover the mechanism of carboxysome segregation in cyanobacteria [69], the development of differential equation analyses of the antenna model of microtubule depolymerization by molecular motors [70], the use of dynamical systems theory to provide deep insights into the stability of genetic circuits and oscillators [71–73], the use of reaction–diffusion models to characterize spatial patterning [74] or the use of the mathematics of coin flips to describe the neutral evolution of eye color in flies [75]. One deeply troubling reaction to these kinds of models is a strange obsession with the times in which they are "wrong", with the even more bizarre urge to declare that if one does *not* use mathematics that confers a kind of immunity against making faulty hypotheses [76], a statement so at odds with history it barely deserves being countenanced by a response.

The mathematical frontiers of biology require buying into and embracing the magic that comes on the heels of forcing ourselves to represent our thinking in mathematical form. What is this magic? It is the revelation of surprising and unexpected connections [33,77]. Though this *is* an essay about biology, just consider this equation

$$\sum_{n=1}^{\infty}\frac{1}{n^s} = \prod_{p\ \text{prime}}\frac{1}{1-p^{-s}}. \tag{22}$$

To make sure there is no confusion about notation, let us expand both the sum and the product as

$$\frac{1}{1^s}+\frac{1}{2^s}+\frac{1}{3^s}+\frac{1}{4^s}+\cdots$$
$$=\left(\frac{1}{1-\frac{1}{2^s}}\right)\left(\frac{1}{1-\frac{1}{3^s}}\right)\left(\frac{1}{1-\frac{1}{5^s}}\right)\left(\frac{1}{1-\frac{1}{7^s}}\right)\left(\frac{1}{1-\frac{1}{11^s}}\right)\cdots \tag{23}$$

On the left side, we sum the reciprocals of all of the integers raised to the power s. On the right side, a fascinating product is formed involving the reciprocals of the *prime* numbers raised to the power s. Using a shockingly simple argument, Euler showed the equivalence of these two expressions [78,79] (see "Various observations about infinite series" by Euler as article 32 of the book by Sandifer). Why bring this up in a paper that is ostensibly about the study of living organisms? My reason is that the act of mathematicizing our thinking can have the very surprising side effect of revealing completely unexpected, even freakish connections such as that exhibited in the formula above. As mathematician David Mumford notes (see his post Math & Beauty & Brain Areas on his excellent Archive for Reprints, Notes, Talks, and Blog): "Such links suggest that the world has a hidden unity, previously concealed from our mortal eyes but blindingly beautiful if we stumble upon it".

None of us are surprised anymore by Descartes' brilliant linkage between algebra and geometry. But every time we say something about the biological world using an x–y plot for example, we are tapping into the magic of a mathematical representation of our topic. However, in other cases, these hidden connections remain but a topic of conjecture.

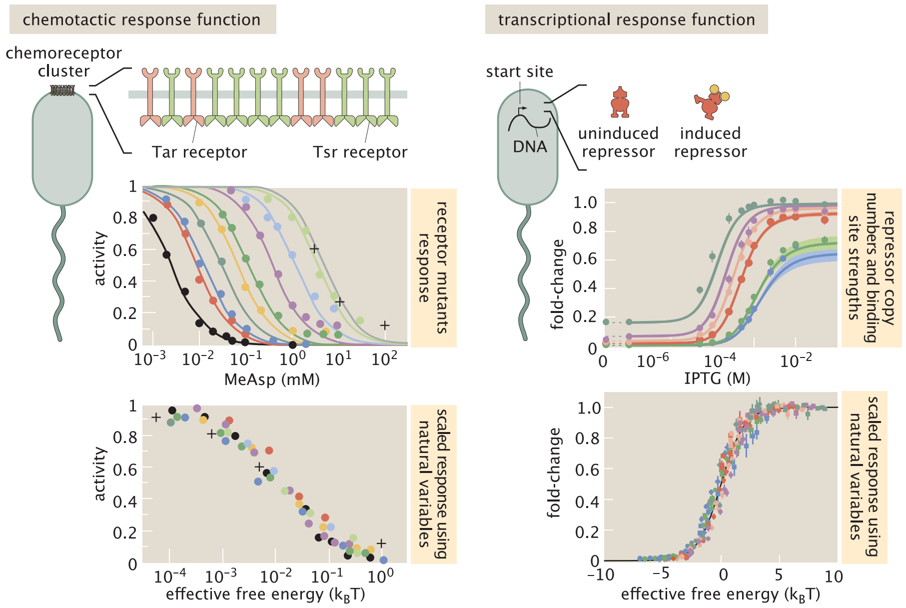


Fig. 8. A gallery of natural variables. Chemotaxis receptors respond to chemoattractants such as methyl aspartate (MeAsp). Different mutants of the receptors have different activity curves, though all such curves fall onto one master curve when plotted with respect to their natural variable as explained in the text. Similarly, the induction of a gene by the inducer IPTG depends upon control variables such as the copy number of the inducible transcription factor or the strength of binding of that transcription factor to the DNA. However, when plotted with respect to their natural variables as explained in the text, there is a universal response.

Mathematics is at peace with conjectures and some of the most famous problems for centuries have been precisely about such conjectures. In our own lifetimes some of the most famous ones such as Fermat' last theorem or the Poincare conjecture have been settled. Others such as the Goldbach conjecture, the twin-prime conjecture or the Riemann hypothesis remain unresolved. When applied to the natural world our mathematical conjectures become a call to experimental action, telling us not only about useful new representations of our problems, but also suggesting new directions in the laboratory.

To my mind, one of the biggest blights on modern science is the supposed dichotomy between different fields and the constant attempts to defend one field over another, or worse yet, to convince us that some particular approach is the most important or most useful in the study of the living. Far more important than the dichotomy between different fields is their coherency and their capacity to demand coherency. This special issue is focused on the ways in which mathematical models will help us better understand the nature of the living. My answer to the question behind all the articles in this special issue is: we should use all the tools at our disposal all the time to try to answer interesting questions about the natural world. However, I also hold true to the idea that our models and theories need to offer a proportional response to our measurements. If people are going to go to all the trouble of providing us with 20,000 dimensional vectors (i.e. transcriptomes), then we better find a way to make a quantitative theory of those 20,000 dimensional vectors. But in the spirit of the famed quote from John F Kennedy, let us not only consider what mathematics and physics can do for biology, but also what biology can do for them. History has shown again and again that whenever those fields turn themselves to new domains of enquiry (such as Fourier's incredibly productive analysis of the phenomena of heat which fell outside of the purview of Newtonian mechanics), they are always enriched not only with fascinating new problems, but invariably with new principles. My hope and sense is that the study of the nature of life using the language of mathematics will catapult both biology and physics forward in the best of ways revealing both new phenomena and new principles.

Declaration of competing interest

I was invited by Wallace Marshall to submit an article for the special volume in which the question of when should mathematical models be used in biology is addressed. I declare no competing interests.

Acknowledgments

I am so grateful to the NIH for support through award numbers DP1OD000217 (Director's Pioneer Award) and NIH MIRA 1R35 GM118043-01. The trust and financial support of this great institution make it possible for today's scientists to grapple with the endless fascination of trying to understand the nature of life. Though my Pioneer Award ended long ago, the foundations laid down by that support will never go away. Participating in the CZI Theory Institute Without Walls which has as one its central themes the question of what is the "right" representation of biological systems has been a huge privilege and deeply inspiring with special thanks to Pankaj Mehta, Vincenzo Vitelli, Madhav Mani, Evelyn Tang and Jane Kondev for really pushing me to broaden my perspective. I am grateful to Gabe Salmon, Jane Kondev, Madhav Mani and Hernan Garcia for comments on the manuscript and related discussions. Willem Kegel many important discussions on input–output functions and data collapse as well as the history of the Hill function. Marc Kirschner continues to inspire me with his expansive and insightful vision of what the study of life can look like.

References

As usual in these kind of subjective essays on various aspects of science, my referencing is idiosyncratic. Many of my references are offered as entry points into a vast literature which I have found to be trustworthy guides into areas that I know little about. For example, the two books of Brent Dalrymple have been an inspiring way to learn about the hard work and long debate surrounding the "I wonder" question of the age of the Earth. The "scholarly" references are for sure an endorsement of those works as a way to learn more about the technical details of some topic, but the absence of many of the other important references on each such topic are in no way a negative judgment on those works.

[1] F.M. Richter, Kelvin and the age of the earth, J. Geol. 94 (3) (1986) 395–401.

[2] G. Brent Dalrymple, The Age of the Earth, Stanford University Press, Stanford, California, 1991.

[3] G. Brent Dalrymple, Ancient Earth, Ancient Skies: The Age of Earth and its Cosmic Surroundings, Stanford University Press, Stanford, California, 2004.

[4] Carl Zimmer, At the Water's Edge : Macroevolution and the Transformation of Life, Free Press, New York, 1998.

[5] J.G.M. Thewissen, S. Bajpai, Whale origins as a poster child for macroevolution, BioScience 51 (12) (2001) 1037–1049.

[6] J.G.M. Thewissen, Jacqueline Dillard, The Walking Whales : From Land to Water in Eight Million Years, University of California Press, Oakland, California, 2015.

[7] R.W. Meredith, J. Gatesy, W.J. Murphy, O.A. Ryder, M.S. Springer, Molecular decay of the tooth gene enamelin (ENAM) mirrors the loss of enamel in the fossil record of placental mammals, Plos Genet. 5 (9) (2009).

[8] P.C. England, P. Molnar, F.M. Richter, Kelvin, Perry and the age of the earth, Am. Sci. 95 (4) (2007) 342–349.

[9] K. Bailey, Potassium-argon ages from Galapagos islands, Science 192 (4238) (1976) 465–467.

[10] A. Caccone, J.P. Gibbs, V. Ketmaier, E. Suatoni, J.R. Powell, Origin and evolutionary relationships of giant Galapagos tortoises, Proc. Natl. Acad. Sci. USA 96 (23) (1999) 13223–13228.

[11] M.J. Schnitzer, S.M. Block, Kinesin hydrolyses one ATP per 8-nm step, Nature 388 (6640) (1997) 386–390.

[12] M.J. Schnitzer, K. Visscher, S.M. Block, Force production by single kinesin motors, Nat. Cell Biol. 2 (10) (2000) 718–723.

[13] William R. Schief, Rutilio H. Clark, Alvaro H. Crevenna, Jonathon Howard, Inhibition of kinesin motility by ADP and phosphate supports a hand-over-hand mechanism, Proc. Natl. Acad. Sci. 101 (5) (2004) 1183–1188.

[14] M. Razo-Mejia, S.L. Barnes, N.M. Belliveau, G. Chure, T. Einav, M. Lewis, R. Phillips, Tuning transcriptional regulation through signaling: A predictive theory of allosteric induction, Cell Syst. 6 (4) (2018) 456–+.

[15] Iain McCalman, Darwin's Armada, Viking, Camberwell, Vic., 2009.

[16] Jonathon Howard, Mechanics of Motor Proteins and the Cytoskeleton, Sinauer Associates, Publishers, Sunderland, Mass., 2001.

[17] W.B. Ludington, L.Z. Shi, Q. Zhu, M.W. Berns, W.F. Marshall, Organelle size equalization by a constitutive process, Curr. Biol. 22 (2012) 2173–2179.

[18] A. Raj, C.S. Peskin, D. Tranchina, D.Y. Vargas, S. Tyagi, Stochastic mRNA synthesis in mammalian cells, PLoS Biol 4 (10) (2006) e309.

[19] D. Zenklusen, D.R. Larson, R.H. Singer, Single-RNA counting reveals alternative modes of gene expression in yeast, Nat. Struct. Mol. Biol. 15 (12) (2008) 1263–1271.

[20] H.J. Li, et al., Fly cell atlas: A single-nucleus transcriptomic atlas of the adult fruit fly, Science 375 (6584) (2022) 991–+.

[21] R.C. Jones, et al., The tabula sapiens: A multiple-organ, single-cell transcriptomic atlas of humans, Science 376 (6594) (2022) 711–+.

[22] A. Schmidt, K. Kochanowski, S. Vedelaar, E. Ahrné, B. Volkmer, L. Callipo, K. Knoops, M. Bauer, R. Aebersold, M. Heinemann, The quantitative and condition-dependent *Escherichia coli* proteome, Nature Biotechnol. 34 (1) (2015) 104–110.

[23] R. Balakrishnan, M. Mori, I. Segota, Z.G. Zhang, R. Aebersold, C. Ludwig, T. Hwa, Principles of gene regulation quantitatively connect DNA to RNA and proteins in bacteria, Science 378 (6624) (2022) 1066–+.

[24] J. Gunawardena, Models in biology: 'accurate descriptions of our pathetic thinking', BMC Biol. 12 (2014) 29.

[25] S. Drake, Galileo's discovery of law of free fall, Sci. Am. 228 (5) (1973) 85–92.

[26] S. Drake, J. Maclachlan, Galileo's discovery of parabolic trajectory, Sci. Am. 232 (3) (1975) 102–110.

[27] Galileo Galilei, Stillman Drake, Discoveries and Opinions of Galileo : Including the Starry Messenger (1610), Letter to the Grand Duchess Christina (1615), and Excerpts from Letters on Sunspots (1613), the Assayer (1623), Anchor Books, New York, 1990.

[28] I. Bernard Cohen, The Birth of a New Physics, Greenwood Press, Westport, Conn., 1981.

[29] E.P. Wigner, The unreasonable effectiveness of mathematics in the natural sciences, Comm. Pure Appl. Math. 13 (1) (1960) 1–14.

[30] Timothy Gowers, Mathematics : A Very Short Introduction, Oxford University Press, Oxford ; New York, 2002.

[31] William Thurston, Mathematical education, 2005, ArXiv:Math/0503081.

[32] W.P. Thurston, On proof and progress in mathematics, Bull. Am. Math. Soc. 30 (2) (1994) 161–177.

[33] Barry Mazur, William A. Stein, Prime Numbers and the Riemann Hypothesis, Cambridge University Press, New York, NY, 2015.

[34] D. Zagier, The first 50 million prime numbers, Math. Intelligencer 1 (1977) 7–19.

[35] Daniel N. Rockmore, Stalking the Riemann Hypothesis : The Quest to Find the Hidden Law of Prime Numbers, first ed., Pantheon Books, New York, 2005.

[36] Roland van der Veen, Jan van de Craats, The Riemann Hypothesis : A Million Dollar Problem, Mathematical Association of America, Washington, 2015.

[37] C.L. Hueschen, L. Segev-Zarko, J.H. Chen, M.A. Legros, C.A. Larabell, J.C. Boothroyd, R. Phillips, A.R. Dunn, Emergent actin flows explain distinct modes of gliding motility, Nat. Phys. (2024).

[38] James D. Talman, Special Functions, A Group Theoretic Approach, W. A. Benjamin, Inc., New York, 1968.

[39] M. Abramowitz, I.A. Stegun, Handbook of Mathematical Functions, NBS, Washington, DC, 1964.

[40] E. Hairer, Gerhard Wanner, Analysis by its history, Springer, New York, 2008.

[41] Michael Nielsen, Neural Networks and Deep Learning, Online Book, 2024.

[42] G. Cybenko, Approximation by superpositions of a sigmoidal function, Math. Control. Signal Syst. 2 (1989) 303–314.

[43] Gilbert Strang, Linear Algebra and Learning from Data, Wellesely-Cambridge Press, Cambridge, Massachusetts, 2019.

[44] E.T. Jaynes, G. Larry Bretthorst, Probability Theory : The Logic of Science, Cambridge University Press, Cambridge, UK ; New York, NY, 2003.

[45] Pierre Simon Laplace, A Philosophical Essay on Probabilities, Dover Publications, New York, 1995.

[46] Norman Biggs, E. Keith Lloyd, Robin J. Wilson, Graph Theory, 1736-1936, Clarendon Press, Oxford Oxfordshire ; New York, 1986.

[47] J. Gunawardena, A linear framework for time-scale separation in nonlinear biochemical systems, PLoS One 7 (5) (2012).

[48] I. Mirzaev, J. Gunawardena, Laplacian dynamics on general graphs, Bull. Math. Biol. 75 (11) (2013) 2118–2149.

[49] T. Ahsendorf, F. Wong, R. Eils, J. Gunawardena, A framework for modelling gene regulation which accommodates non-equilibrium mechanisms, BMC Biol. 12 (1) (2014) 102.

[50] R. Phillips, The Molecular Switch : Signaling and Allostery, Princeton University Press, 2020.

[51] T. Ahsendorf, F. Wong, R. Eils, J. Gunawardena, A framework for modelling gene regulation which accommodates non-equilibrium mechanisms, Bmc Biology 12 (2014).

[52] A.V. Hill, The possible effects of the aggregations of the molecules of haemoglobin on its dissociation curves, Proc. Physiol. Soc. (1910).

[53] A.V. Hill, The combinations of haemoglobin with oxygen and with carbon monoxide. I., Biochem. J. 7 (5) (1913) 471–480.

[54] J.C. Gerhart, A.B. Pardee, Enzymology of control by feedback inhibition, J. Biol. Chem. 237 (3) (1962) 891–+.

[55] J. Monod, J.P. Changeux, F. Jacob, Allosteric proteins and cellular control systems, J. Mol. Biol. 6 (1963) 306–329.

[56] J. Monod, J. Wyman, J.P. Changeux, On the nature of allosteric transitions: A plausible model, J. Mol. Biol. 12 (1965) 88–118.

[57] J.P. Changeux, 50 years of allosteric interactions: the twists and turns of the models, Nature Rev. Mol. Cell Biol. 14 (12) (2013) 819–829.

[58] J. Gerhart, From feedback inhibition to allostery: the enduring example of aspartate transcarbamoylase, Febs J. 281 (2) (2014) 612–620.

[59] R. Milo, J.H. Hou, M. Springer, M.P. Brenner, M.W. Kirschner, The relationship between evolutionary and physiological variation in hemoglobin, Proc. Natl. Acad. Sci. USA 104 (43) (2007) 16998–17003.

[60] O. Rapp, O. Yifrach, Using the MWC model to describe heterotropic interactions in hemoglobin, PLoS One 12 (8) (2017).

[61] O. Rapp, O. Yifrach, Evolutionary and functional insights into the mechanism underlying body-sizerelated adaptation of mammalian hemoglobin, Elife 8 (2019).

[62] J.E. Keymer, R.G. Endres, M. Skoge, Y. Meir, N.S. Wingreen, Chemosensing in *Escherichia coli:* two regimes of two-state receptors, Proc. Natl. Acad. Sci. USA 103 (6) (2006) 1786–1791.

[63] L.R. Swem, D.L. Swem, N.S. Wingreen, B.L. Bassler, Deducing receptor signaling parameters from in vivo analysis: LuxN/AI-1 quorum sensing in *Vibrio harveyi*, Cell 134 (3) (2008) 461–473.

[64] P.C. Nelson, Biological Physics : Energy, Information, Life, W.H. Freeman and Co., New York, 2004.

[65] W.S. Bialek, Biophysics : Searching for Principles, Princeton University Press, Princeton, NJ, 2012, p. xii, 640.

[66] P.C. Nelson, Physical Models of Living Systems, W.H. Freeman & Company, New York, NY, 2015.

[67] W. Bialek, Perspectives on theory at the interface of physics and biology, Rep. Progr. Phys. 81 (1) (2018).

[68] R.E. Goldstein, Are theoretical results 'results'? Elife 7 (2018).

[69] D.F. Savage, B. Afonso, A.H. Chen, P.A. Silver, Spatially ordered dynamics of the bacterial carbon fixation machinery, Science 327 (5970) (2010) 1258–1261.

[70] V. Varga, J. Helenius, K. Tanaka, A.A. Hyman, T.U. Tanaka, J. Howard, Yeast kinesin-8 depolymerizes microtubules in a length-dependent manner, Nat. Cell Biol. 8 (9) (2006) 957–962.

[71] J.L. Cherry, F.R. Adler, How to make a biological switch, J. Theoret. Biol. 203 (2) (2000) 117–133.

[72] T.S. Gardner, C.R. Cantor, J.J. Collins, Construction of a genetic toggle switch in *Escherichia coli*, Nature 403 (6767) (2000) 339–342.

[73] M.B. Elowitz, S. Leibler, A synthetic oscillatory network of transcriptional regulators, Nature 403 (6767) (2000) 335–338.

[74] L. Wolpert, Positional information and the spatial pattern of cellular differentiation, J. Theoret. Biol. 25 (1) (1969) 1–47.

[75] P. Buri, Gene-frequency in small populations of mutant *Drosophila*, Evolution 10 (4) (1956) 367–402.

[76] S.F. Gilbert, Achilles and the tortoise: Some caveats to mathematical modeling in biology, Prog. Biophys. Mol. Biol. 137 (2018) 37–45.

[77] Harold M. Edwards, Riemann’s Zeta Function, Dover, Dover Publications, Mineola, NY, 2001.

[78] L.J. Goldstein, History of prime number theorem, Amer. Math. Monthly 80 (6) (1973) 599–615.

[79] Charles Edward Sandifer, The Early Mathematics of Leonhard Euler, Mathematical Association of America, Washington, DC, 2007.