Chapter 4

FUNDAMENTAL POWER REQUIREMENTS TO BUILD OR MAINTAIN BIOCHEMICAL GRADIENTS

4.1 Abstract

Organisms organize and respond to a plethora of different gradients. Perhaps the most famous such gradient patterns the anterior-posterior axis of cells in the fly embryo, but many further examples abound. Organelles and cell membranes sustain electrochemical gradients; motile cells respond to chemoattractant gradients; tumors and biofilms develop oxygen gradients; and gradients in biodiversity even develop along evironmental axes such as elevation. In our own experimental work with light-controlled microtubule-motor systems, we and others have found how molecular assemblies spontaneously organize into star-shaped "asters" that feature an approximately spherically-symmetric and exponentially-decaying arrangement of motors; these aster structures and gradients evoke biologically-relevant structures such as the mitotic spindle. Our experiments have also measured the distribution of ATP in space and time as well as spatially-resolved measurements of the power consumed. These measurements report that asters often consume power almost ten times faster while they are forming and changing rapidly than at late stages. We hypothesize this discrepancy between early and late dissipation is due to fundamental differences in the energetic costs to build and maintain such gradients. In this paper, we explore these questions and their broad implications using simple ideas from statistical physics.

4.2 The Power of Biological Processes

Living things build and depend on exquisite patterns of chemicals in space and time. Precisely how much energy must these systems pay to incite these patterns, and then to defy their decay towards equilibrium homogeneity? Does building or maintaining structures demand greater biochemical energy and power expenditures? How do these expenditures stack up against the broader cellular economy of metabolic expenditures?

These big questions enjoy rich history. How cells invest energy over a wide set of tasks is a mystery that has gained urgency in many guises, especially in the 1970s. Early researchers asked how much energy growing microbial cells need to duplicate

their contents, finding the surprise that cells spend significantly more energy than apparently required to duplicate biomass [1, 2]. Recent, even more precise, work and modeling validate these mismatches [3] and highlight that microbes have significant expenditures unexplained by material construction costs alone. Cells far beyond microbes, including eukaryotic and plant cells, also show levels of dissipation awaiting full and complete accounting [4, 5]. Clearly, cells perform many energetically-costly functions beyond copying biomass that can participate in such total metabolic demands. These include kinetic proofreading (explaining how apparently futile cycles of GTP hydrolysis accomplish greater accuracy in protein translation and other biological transformations) [6], and GTP hydrolysis regulating the polymerization of microtubules [7, 8]. Hydrolyzing ATP may also facilitate more sensitive [9] or flexible [10] signaling in gene regulation, or additionally reduce noise in signaling networks [11], than permitted at equilibrium. Beyond these charismatic examples, exactly what other capabilities that cells unlock as they spend energy is a frontier that invites huge discoveries, particularly facilitated by new experimental technologies that resolve cellular dissipation in unprecedented regimes of physiology.

One fundamental destiny for cellular energy expenditures surely must be to assemble the extraordinary patterns of biomolecular components in space and time that orchestrate living matter. Writing in 1970, Francis Crick, Mary Munro, and coworkers pursued imaginative calculations complementing the work of Alan Turing that asked how diffusion and cellular production can establish morphogenetic gradients in embryonic development [12, 13]. These formative works assessed the plausibility and constraints of such mechanisms to establish gradients by particularly focusing on the time required to set up gradients on cellular and organismal length scales, identifying feasible regimes where gradients can be established in acceptable developmental times [13]. The time to assemble structure is just one feature affecting how biological gradients develop, however. More contemporary works often investigate the accuracy attainable while forming discrete sets of combinatorial structures [14] under different kinetic and dissipative protocols. Another blooming thread comes from results of modern stochastic thermodynamics that link the work extractable from a nonequilibrium system in a certain state (say matter arranged in space) from the Kullback Leibler divergence between that initial state and a terminal equilibrium distribution (say matter absent a gradient) [15, 16]. How such advancements may be generalized to understand the energetic costs of building and maintaining continuous biological patterns of special interest, not just from static comparisons of initial and final conditions but at local and temporal detail, is a challenge that gains both

urgency and tractability from the development of new precision measurements.

In this paper, we aim to build on these early efforts (and many others) to carefully characterize the power costs of forming gradients in new experimental measurements. We illustrate calculations in the context of how molecular motors and microtubules assemble asters in experiments we have done in which light is used to transiently crosslink motors. In these experiments, a homogeneous mixture of microtubules and motors is induced to form these ordered structures by light-induced crosslinking of the motors. Our measurements revealed that during the early stages of aster formation, the power is nearly 10-fold higher than at the late stages of aster formation. We suspect that this discrepancy is due to the substantial difference in the energetic cost to build an aster vs to maintain it. In this paper, we explore that hypothesis. While our discussion often transacts in the specific language and details of gradients of molecular motors along microtubules, the ingredients of these calculations are highly generic and may apply to make predictions and infer bounds about the costs and strategies to build versus maintaining gradients in myriad guises across biology, agnostic of particular mechanisms.

Conceptually, the structure of the estimates will all be the same, with the generic functional form

$$\text{power of process} = J_{\text{process}} \times \Delta\mu_{\text{process}}. \tag{4.1}$$

Here we have defined the quantity $\Delta\mu_{\text{process}}$ as the free energy cost of a unit process such as a single step of a motor or the movement of a single molecule up a gradient. J_{process} refers to the flux associated with the process of interest, meaning how many unit processes occur per unit time. For example, in the context of constantly pumping ions up a concentration gradient, $\Delta\mu_{\text{process}}$ refers to the free energy cost of taking a single ion from one side of the membrane to the other. Similarly, the flux in that case would be given by a phenomenological linear transport law relating the concentration jump across the membrane to the flux itself. Using this quantitative structure, we carry out a series of estimates for various models of gradient formation and maintenance.

Power to Maintain a Motor Gradient

As noted in our earlier work, the largest power estimate out of the suite of processes that occur during aster formation is associated with the formation of the gradient of motors. Because of the gradient in motor density across the aster, it is of interest to estimate the free energy required to maintain that gradient. We begin by examining

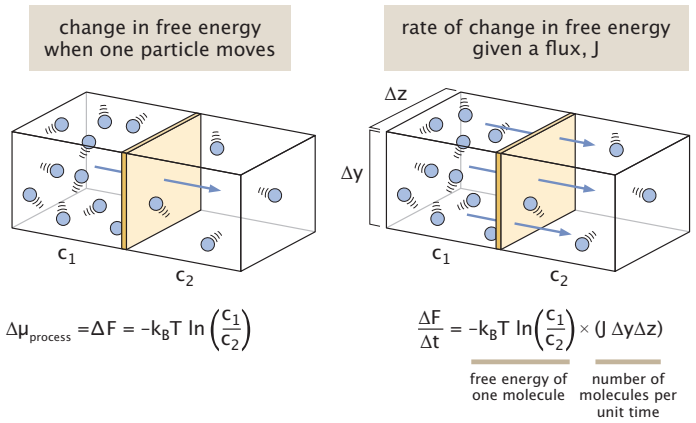

Figure 4.1: **The change in free energy when a particle is transported in a gradient.** (A) The free energy change upon moving a particle from the left reservoir to the right reservoir. (B) The total free energy dissipated as many particles move in the presence of a gradient. Adapted from Hueschen and Phillips, The Restless Cell.

the free energy cost of a one-dimensional gradient to set notation and to explain the concept and follow that discussion by the case of interest involving spherical symmetry. The concept of the estimate is to compute the free energy change when we take one particle from a region with one concentration and put that particle in a nearby region with a slightly different concentration.

Figure 4.1(A) makes this explicit by illustrating the free energy change associated with moving a single molecule from a reservoir at one concentration to a second reservoir at a different concentration. We consider a scenario in which the entirety of the free energy change is entropic (e.g., we do not consider situations involving gravitational or electrostatic potentials), meaning that the free energy is defined as

$$F = -TS, \tag{4.2}$$

where S is the entropy. We interest ourselves in the change in free energy

$$\Delta F = F_{\text{final}} - F_{\text{initial}} = -TS_{\text{final}} + TS_{\text{initial}}. \tag{4.3}$$

To compute the entropy change, we need to compute the entropy of the solutions on both sides of the partition, both before and after we have taken a molecule from the left side and placed it on the right side. The total entropy is given by

$$S_{\text{tot}} = S_1 + S_2 \tag{4.4}$$

where the subscripts refer to the two compartments. Using the Boltzmann definition of entropy, we have

$$S_{\text{tot}}^{(\text{final})} = k_B \ln W_1^{(\text{final})} + k_B \ln W_2^{(\text{final})}, \tag{4.5}$$

with a similar expression for the entropy in the initial state. This can be simplified to the form

$$S_{\text{tot}}^{(\text{final})} = k_B \ln (W_1^{(\text{final})} W_2^{(\text{final})}), \tag{4.6}$$

which makes sense given that the total number of microscopic states is equal to the product of the number of states in the first box and the number of states in the second box.

In a lattice model, we make the abstraction that space is subdivided into tiny lattice sites with a characteristic dimension of 1 nm^3 (i.e., molecular sizes). To compute the number of microstates W, we count the number of ways of arranging our L_i motors among the Ω lattice sites as

$$W_i(L_i) = \frac{\Omega!}{L_i!(\Omega - L_i)!}. \tag{4.7}$$

When $\Omega \gg L_i$ (i.e. the dilute limit), we can make the much simpler approximation

$$W_i(L) = \frac{\Omega^{L_i}}{L_i!}, \tag{4.8}$$

which amounts to the idea that every motor can sit on any of the Ω lattice sites.

We can now write the change in free energy which is strictly entropic as

$$\Delta\mu = -k_B T \left(\ln \frac{\Omega^{L_1+1}}{(L_1+1)!} \frac{\Omega^{L_2-1}}{(L_2-1)!} - \ln \frac{\Omega^{L_1}}{L_1!} \frac{\Omega^{L_2}}{L_2!} \right), \tag{4.9}$$

where we revert to the notation $\Delta\mu$ since this is the free energy change of the unit process of moving one molecule from one side of the partition to the other. This can be simplified to

$$\Delta\mu = -k_B T \ln \frac{L_1!}{(L_1+1)!} \frac{L_2!}{(L_2-1)!} \approx -k_B T \ln \frac{L_2}{L_1}. \tag{4.10}$$

We can rewrite this in a more familiar form using the language of concentrations. If we multiply numerator and denominator within the logarithm by Ωv, where v is the volume of a single lattice site in our lattice model, then $\Omega v = V_{tot}$ and hence $c_1 = L_1/V_{tot}$ and $c_2 = L_2/V_{tot}$, permitting us to write

$$\Delta\mu = k_B T \ln \frac{c_2}{c_1}. \tag{4.11}$$

Now, as seen in the right panel of Figure 4.1 if we want to find the total rate of free energy change, we need to multiply the free energy per particle by the total number

of particles transported between the two adjacent reservoirs per unit time using the flux resulting in

$$\left(\frac{\Delta F}{\Delta t}\right) = -\Delta\mu(J\Delta y\Delta z). \tag{4.12}$$

The factor $J\Delta y\Delta z$ counts the number of molecules carried down the gradient per unit time and the minus sign guarantees that if the left reservoir has more molecules than the left, then $\Delta F/\Delta t < 0$. We now interest ourselves in the case where the concentration is slowly varying, permitting us to write

$$\left(\frac{\Delta F}{\Delta t}\right) = k_B T \ln \frac{c(x+\Delta x)}{c(x)} \times (J\Delta y\Delta z), \tag{4.13}$$

where we have introduced the notation $c_1 = c(x)$ and $c_2 = c(x+\Delta x)$. Note now we switched the sign because we inverted the ratio in the logarithm with the concentration on the right now appearing in the numerator. Note that this expression is valid regardless of whether $c_1 > c_2$ or $c_2 > c_1$ since in those two cases the flux is in opposite directions and our expression reflects that. By invoking Fick's law we now have

$$\left(\frac{\Delta F}{\Delta t}\right) = k_B T \ln \frac{c(x+\Delta x)}{c(x)} \times \left(-D\frac{dc}{dx}\right)\Delta y\Delta z. \tag{4.14}$$

By carrying out the Taylor expansion $c(x+\Delta x) = c(x) + (\partial c/\partial x)\Delta x$, we can rewrite the logarithmic term as

$$\begin{aligned} \ln \frac{c(x+\Delta x)}{c(x)} &= \ln\left[\frac{c(x) + \frac{\partial c}{\partial x}\Delta x}{c(x)}\right] \\ &= \ln\left[1 + \frac{1}{c}\frac{\partial c}{\partial x}\Delta x\right]. \end{aligned} \tag{4.15}$$

Finally, we invoke a second Taylor series in the form $\ln(1+\epsilon) \approx \epsilon$, resulting in the simple and powerful expression

$$\left(\frac{\Delta F}{\Delta t}\right) = -Dk_B T \frac{1}{c(x)}\left(\frac{\partial c(x)}{\partial x}\right)^2 \Delta x\Delta y\Delta z. \tag{4.16}$$

This result tells us the free energy loss for two adjacent planes if the gradient is allowed to dissipate. We can now put this all together to explore the free energy dissipated over a continuous concentration field. In this case, we take the expression seen in eqn. 4.16 and add up the contribution from each set of planes in our discrete representation of the concentration field. We can work out the minimum power to maintain such a gradient where we imagine every time a molecule goes down its gradient, energy is consumed to push it back where it came from. Given this

approach, we can now tackle the question of estimating the power associated with maintaining the radial motor gradient in our asters. However, to do so, we first need to reinterpret the one-dimensional analysis done here to the case of a three-dimensional, but spherically symmetric concentration gradient.

The Power Required to Maintain a Three-Dimensional Gradient

Building on the analysis of the previous section which was performed in one-dimension, we now turn our attention to a three-dimensional structure with a spherically symmetric concentration gradient of motors of the form $c(r)$, where r is the radial distance from the aster center. Figure 4.2(A) shows the amendment that needs to be made to the one-dimensional analysis, where now in the three dimensional case, the "boxes" are spherical shells. Interestingly, the expression we derived earlier goes through essentially unchanged except that now we consider the radial concentration, and instead of integrating over planes along the x-direction, we now integrate over spherical shells in the r direction. Given these adjustments, the power to sustain a gradient is now given as

$$P = k_{\mathrm{B}}TD \int_0^{2\pi} \int_0^{\pi} \int_0^{\infty} \frac{1}{c(r)} \left(\frac{\partial c(r)}{\partial r} \right)^2 r^2 \sin\theta \, dr \, d\theta \, d\phi, \tag{4.17}$$

where D is the diffusion constant and $c(r)$ is the radial concentration profile.

To get a qualitative feeling for the power scales, we begin by considering a fixed, radial distribution of motors described as a decaying exponential of the form

$$c(r) = c_0 e^{-r/\lambda}, \tag{4.18}$$

where c_0 is the motor concentration at $r = 0$ and λ is the decay length. Plugging this profile into eqn. 4.17, we can immediately compute the power that must be expended to prevent the diffusive relaxation of this gradient as

$$\begin{aligned} P &= 4\pi k_{\mathrm{B}}TD \int_0^{\infty} \frac{e^{r/\lambda}}{c_0} \left(-\frac{c_0}{\lambda} e^{-r/\lambda} \right)^2 r^2 \, dr \\ &= \frac{4\pi k_{\mathrm{B}}TDc_0}{\lambda^2} \int_0^{\infty} r^2 \, e^{-r/\lambda} dr. \end{aligned} \tag{4.19}$$

Integrating by parts, we find that the integral evaluates to $\lambda^2 \int_0^{\infty} r^2 \, e^{-r/\lambda} dr = 2\lambda^3$ yielding,

$$P = 8\pi k_{\mathrm{B}}TDc_0\lambda. \tag{4.20}$$

Without the prefactors, this result can also be inferred by dimensional analysis.

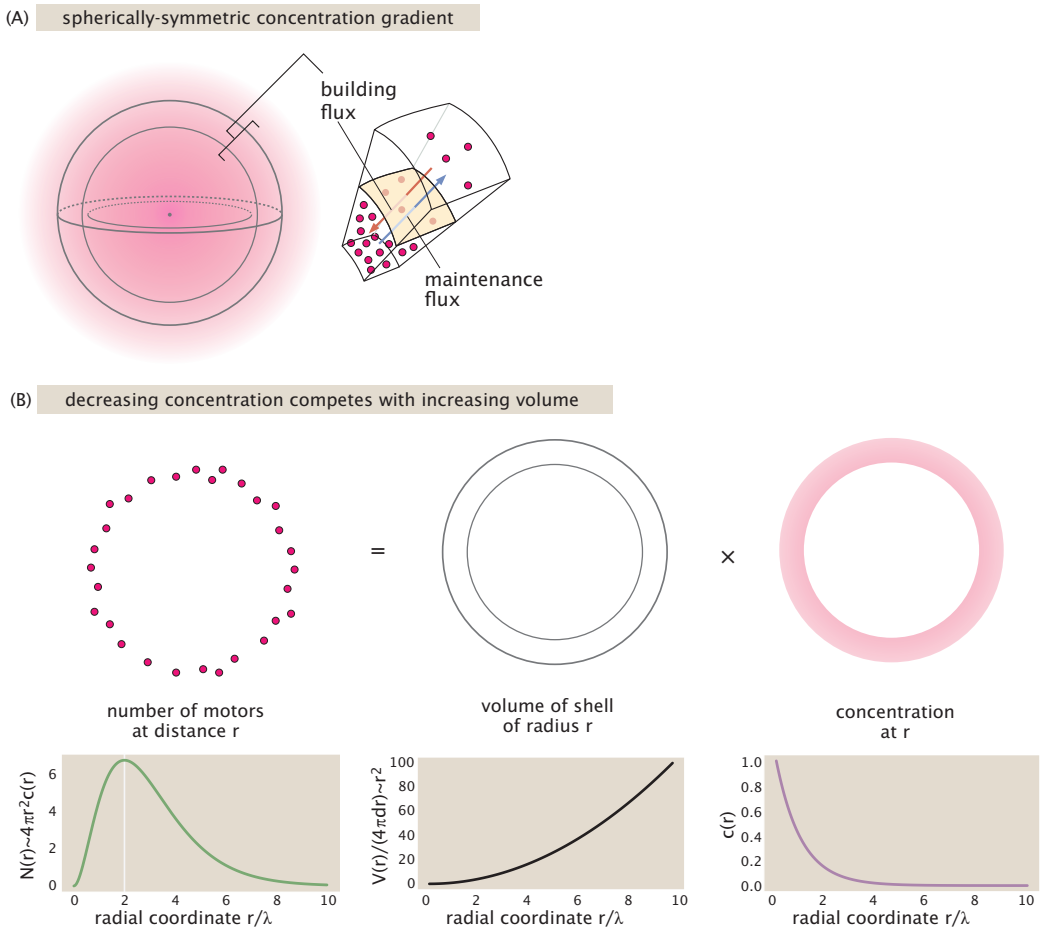

Figure 4.2: **Dissipation in a spherically symmetric concentration gradient.** (A) Maintaining a spherically symmetric concentration gradient results from a competition between the outward diffusive flux and the inward active flux provided by motors moving on microtubules. (B) There is a peak in the distribution of motors as a function of radial distance resulting from the competition between the monotonically decreasing motor concentration and the geometric effect that the spherical shells get larger with increasing r.

Estimate of the magnitude of power to prevent diffusive spreading

To get a feeling for the numbers, we appeal to the measurements described in our experimental paper for which the asters are approximately described by a radial concentration field of the form $c(r) = c_0 e^{-r/\lambda}$, with parameters $c_0 \approx 1\ \mu M$ and $\lambda \approx 30\ \mu m$. Given these values, we first ask roughly how many motors there are in this aster region, found as

$$\#\text{ of motors in aster} = \int_{\text{aster}} dV c_0 e^{-r/\lambda} = 8\pi c_0 \lambda^3. \tag{4.21}$$

This result can be seen by using integration by parts on the radial integral, though intuitively, dimensional analysis recommends this functional form to within numerical factors. Given that $c_0 = 1\mu M \approx 10^3/\mu m^3$ and $\lambda \approx 30\mu m$, we find that the number of motors in the aster is

$$\#\text{ of motors in aster} = 8\pi c_0 \lambda^3 \approx 8\pi \times \frac{10^3}{\mu m^3} \times (f \times 10\ \mu m)^3 \approx 10^9 \text{ motors}. \tag{4.22}$$

If every one of those motors was consuming only one ATP per second, the resulting power is precisely the scale we find in our measurements which is of the order of 10^9 ATP/s.

We can compare this simple order of magnitude guess with the cost to maintain the gradient computed above. In particular, using eqn. 4.20, we find

$$P = 8\pi k_B T D c_0 \lambda \approx (f{\times}10){\times}k_B T{\times}(10\,\frac{\mu m^2}{s}){\times}\left(\frac{10^3}{\mu m^3}\right){\times}(f{\times}10\,\mu m) \approx 5{\times}10^5\,\frac{\text{ATP}}{\text{s}}. \tag{4.23}$$

The numbers we took here for the size of the aster were for " late times" and thus we need to make a more careful analysis for the change in aster size over time to provide a full picture of the dynamics of the power in space and time. The present calculation really serves as a first cut to get a feeling for the numbers.

The Spatial Distribution of Power

The three dimensional dissipation calculated above suggests interesting spatial effects in the measured power. As seen in Figure 4.2, the radially decaying concentration gradient implies that although the concentration may be higher at smaller radii, the number of motors in a given radial shell is not maximum at the origin. This means that there is the possibility of nonmonotonic power as a function of radius. The key idea is illustrated in Figure 4.2(B), where we see that the number

of motors as a function of distance from the origin is peaked. This is the result of a competition between the volume of the shells with increasing r (which increases) and the decrease in motor concentration as a function of distance as shown in the lower right panel.

To get a feeling for this effect, we once again turn to our model of a static, spherically symmetric gradient and show that there is a maximum in the power as a function of radius. Our first task is to compute the power in a spherical shell at radius r. One way to think of the power associated with such a spherical shell at radius r is to evaluate eqn. 4.17 only between r and $r + \Delta r$, resulting in

$$P = 4\pi r^2 k_{\mathrm{B}} T D \frac{1}{c(r)} \left(\frac{\partial c(r)}{\partial r} \right)^2 \Delta r. \tag{4.24}$$

If we now exploit the known concentration profile $c(r) = c_0 \exp(-r/\lambda)$ and substitute it into eqn. 4.24, we find the power in the shell at radius r is given by

$$\frac{\partial F}{\partial t}(r) = -(\text{const.}) \times r^2 e^{-r/\lambda}, \tag{4.25}$$

where we have suppressed all constants such as D, c_0, λ, π, etc. This very simple expression now permits us to show that for our simplified model of a static aster, there is a maximum in the dissipation rate as a function of r given by

$$\frac{d}{dr}\left(\frac{\partial F}{\partial t}(r) \right) = -(\text{const.}) \times \left[2re^{-r/\lambda} - \frac{r^2}{\lambda} e^{-r/\lambda} \right] = 0 \tag{4.26}$$

resulting in a maximum in the power dissipation at radius $r = 2\lambda$

To be thorough, we were also curious about the implications of asymmetries in the aster since three-dimensional imaging demonstrates that our asters are not completely spherical, but have a preferred long axis in the illumination direction. Here we consider the opposite extreme in which the aster is a cylinder instead. In this case, the power as a function of radial distance from the cylinder axis is given by

$$\frac{\partial F}{\partial t}(r) = 2\pi r \Delta z \left(-D \frac{\partial c}{\partial r} \right) \frac{1}{c} \frac{\partial c}{\partial r} \Delta r k_B T. \tag{4.27}$$

This expression has units of power (i.e. J/s) and tells us the power dissipated in the annulus at radius r. This expression raises the question of how the power dissipated depends upon the radius, and specifically, is there a maximum. If we once again exploit the fact that the concentration has cylindrical symmetry and is characterized by $c(r) = c_0 \exp(-r/\lambda)$, then we see that we have

$$\frac{\partial F}{\partial t}(r) = -(\text{const.}) \times r e^{-r/\lambda}, \tag{4.28}$$

where we once again suppress the constants such as D, Δr, etc. This result implies in turn that

$$\frac{d}{dr}\left(\frac{\partial F}{\partial t}(r)\right) = -(\text{const.}) \times \left[e^{-r/\lambda} - \frac{r}{\lambda}e^{-r/\lambda}\right] = 0 \tag{4.29}$$

which implies that the maximum dissipation occurs at $r = \lambda$.

The calculations we have done thus far are both interesting and suggestive for interpreting the measured ATP consumption in our aster experiments. However, the discussion here was based upon a variety of simplifying assumptions, the most important of which is that the aster is fixed in shape over time. However, we know that asters evolve in time from the moment that light is used to cross link the motors leading to a time-dependent aster size and motor concentration. In the next sections, we consider the difference between the power to construct and maintain an aster which explicitly acknowledges this time dependence. Of course, to construct an aster and the motor concentration gradient requires more than diffusion. There are active fluxes and we now examine the power of these fluxes.

The Power Required to Build a Three-Dimensional Gradient

Figure 4.3 shows the power consumption over time during representative aster formation experiments, which suggest that ATP is spent nearly an order of magnitude faster at early times than at late times in the dynamics of aster formation after light-induced motor dimerization occurs. This widens the gap between our largest measured dissipations and our best quantitative estimates for the origins of the power. As a result, in the current section, we graduate our earlier estimates to confront the transient power in the earlier stages of aster formation.

What is the basis of the large discrepancy between power consumption at early and late times in the aster formation process? One provocative possibility is that the formative physics of asters at early times fundamentally demand separate, and plausibly greater, nonequilibrium expenditures than those required to just maintain an aster's structure once it has formed at late times.

To bear on these questions and learn about what they say about the precise destinies of ATP expenditure in pattern formation, here we calculate the free energy cost per time required to both build a concentration gradient in time, and to maintain it at a given state. Combining these calculations with representative phenomenology of motor distributions, we find much support for the latter idea that the power required to assemble a nonuniform profile can be hundreds of times larger than the power cost merely to maintain it (during some stages of aster formation). However, remarkably,

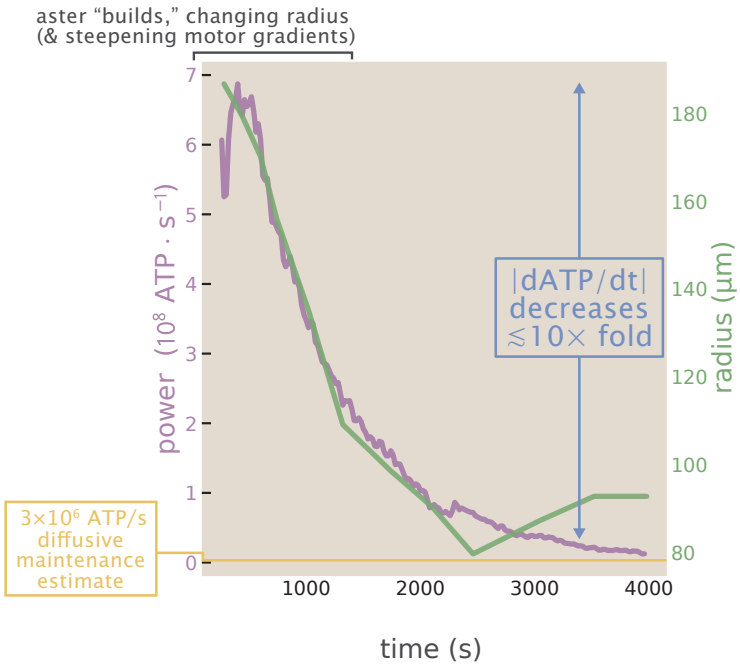

Figure 4.3: **Experimental data motivating the question of how much power it takes to build versus maintain an aster's concentration gradients.** The x-axis is time and the y-axis is the power, making it evident that the power consumption changes over time as the microtubule-motor system evolves towards its nonequilibrium aster steady state.

the relative importance of building power and maintenance power often switches over the course of an aster's trajectory, and more generally can vary over space, time, and values of biophysical parameters. We introduce an amusing and natural "dissipative Peclét number" definable locally or globally in spacetime that neatly organizes and characterizes this competition between dissipative origins.

Building versus maintaining *arbitrary* (spherically-symmetric) gradients

As simple steps towards the enduring question of how much gradients cost energetically, consider some active agent spending energy to maintain—or even grow—a concentration gradient over space. This activity, administering some material flux $\mathbf{J}_A$, must defy the spontaneous diffusive flux $\mathbf{J}_D$ of the substance that tends to relax the concentration gradient to uniformity. For a gradient to persist or build, these diffusive and active fluxes operate in opposing directions, making the concentration evolve according to a net flux $\Delta\mathbf{J} \equiv \mathbf{J}_A + \mathbf{J}_D$.

For concreteness, when the concentration profile is spherically-symmetric (varying

only along the radial coordinate r), and without loss of generality under the convention that more substance is found at small r than at large r (namely $\frac{\partial c}{\partial r} \leq 0$), the agent's flux J_A has a negative sign, $J_A \leq 0$, and the diffusive flux is positive, $J_D = -D\frac{\partial c}{\partial r} \geq 0$. Calling the net flux $J_r = \Delta J$ since it is in the radial direction, these conventions cast the net flux as $J_r = J_D - |J_A|$. (If the gradient is merely maintained, neither decaying nor steepening, then the net flux is zero and $|J_A| = |J_D|$: the active agent moves just enough material thus spending just enough energy to counteract the spontaneous diffusive flux that would result if it were not acting. While building a gradient, the agent's total active flux is thus partitioned into a maintenance flux matching diffusion, plus any attained nonzero net flux $|\Delta J| > 0$ steepening the gradient,

$$|J_A(r,t)| = |J_D(r,t)| + |J_r(r,t)|. \tag{4.30}$$

We can examine the mathematical underpinnings of this question as follows. Let the total free energy density at position $\mathbf{r}$ and time t be $f(\mathbf{r},t)$. Note that the free energy density should depend on time and space only via the concentration field $c(\mathbf{r},t)$, that is $f \equiv f(c(\mathbf{r},t))$. Then the total free energy $F^{\text{tot}}(t)$ of the system is its integral over all space,

$$F^{\text{tot}}(t) = \int d^3\mathbf{r}\; f(c(\mathbf{r},t)), \tag{4.31}$$

and the total power is the time derivative of this total free energy, giving,

$$\begin{aligned} \mathrm{P}^{\text{tot}}(t) &= \frac{dF^{\text{tot}}}{dt} && (4.32)\\ &= \frac{\partial}{\partial t}\int d^3\mathbf{r}\; f(c(\mathbf{r},t)) = \int d^3\mathbf{r}\; \frac{\partial f(c(\mathbf{r},t))}{\partial t} && (4.33)\\ &= \int d^3\mathbf{r}\frac{\partial f(c)}{\partial c}\frac{\partial c(\mathbf{r},t)}{\partial t}. && (4.34) \end{aligned}$$

Note that the chemical potential is by definition the change in free energy with respect to a change in particle number, or (dividing a local volume element by a volume) equivalently the change in free energy density with respect to concentration, $\mu(\mathbf{r},t) \equiv \frac{\partial f}{\partial c}$. In addition, note that conservation of mass gives the standard continuity equation in the concentration field: the local rate at which concentration changes is the divergence of the net flux $\mathbf{J}(\mathbf{r},t)$ without any homogeneous local sources,

$$\frac{\partial c}{\partial t} + \nabla \cdot \mathbf{J}_{\text{net}} = 0, \tag{4.35}$$

so we can express the $\frac{\partial c}{\partial t}$ in the integrand of Eq. 4.34 in terms of the net flux's divergence, $\frac{\partial c}{\partial t} = -\nabla \cdot \mathbf{J}_{\text{net}}$, yielding,

$$\mathrm{P}^{\text{tot}}(t) = \int d^3\mathbf{r} \frac{\partial f(c)}{\partial c} \frac{\partial c(\mathbf{r},t)}{\partial t} \tag{4.36}$$

$$= -\int d^3\mathbf{r}\, \mu(\mathbf{r},t)\, \nabla \cdot \mathbf{J}_{\text{net}}. \tag{4.37}$$

Eq. 4.37 is a volume integral of a scalar field (the chemical potential μ) times the divergence of a vector field (the net flux J_{net}). This permits us to invoke a consequence of the divergence theorem, which (see https://en.wikipedia.org/wiki/Divergence_theorem#Corollaries) delivers the vector identity that,

$$\int dV\, [\mathbf{F} \cdot (\nabla g) + g\, (\nabla \cdot \mathbf{F})] = \oint d\partial V\ \ (g\mathbf{F} \cdot \mathbf{n}) \tag{4.38}$$

$$\rightarrow \int dV\, g\, (\nabla \cdot \mathbf{F}) = \oint d\partial V\ \ (g\mathbf{F} \cdot \mathbf{n}) - \int dV \mathbf{F} \cdot (\nabla g)\,. \tag{4.39}$$

Identifying g as μ and $\mathbf{F}$ as $\mathbf{J}_{\text{net}}$ in this vector identity, we see that the power can be written as a volume integral of a gradient in the chemical potential coupled with the flux, minus a surface integral,

$$\mathrm{P}^{\text{tot}}(t) = \int d^3\mathbf{r}\, \mathbf{J}_{\text{net}} \cdot (\nabla \cdot \mu) - \oint dA\ \mu(\mathbf{J}_{\text{net}} \cdot \mathbf{n}). \tag{4.40}$$

For a system with no explicit dissipative action at infinity, the surface integral should vanish. This gives,

$$\mathrm{P}^{\text{tot}}(t) = \int d^3\mathbf{r}\, \mathbf{J}_{\text{net}} \cdot (\nabla \cdot \mu)\,. \tag{4.41}$$

Given a chemical potential $\mu(r) = \mu_0 + k_B T \ln c(r,t)/c_0$ varying in space (for a standard state concentration c_0 and standard potential μ_0), the local change in chemical potential imposed by moving a particle of substance an infinitesimal distance dr is,

$$d\mu = \frac{\partial \mu}{\partial r} dr \tag{4.42}$$

$$= dr\ k_B T\ \frac{1}{c(r,t)} \frac{\partial c}{\partial r}. \tag{4.43}$$

This free energy cost couples with the earlier fluxes to define the areal power densities σ of maintaining and building the gradient,

$$\underbrace{\sigma_A(r,t)}_{\text{W m}^{-2}\text{ s}^{-1}} \equiv |J_A(r,t)| \frac{\partial \mu}{\partial r} dr = \underbrace{|J_D(r,t)| \frac{\partial \mu}{\partial r}}_{\equiv \rho_D, \text{W m}^{-3}\text{ s}^{-1}} dr + \underbrace{|J_r(r,t)| \frac{\partial \mu}{\partial r}}_{\equiv \rho_r, \text{W m}^{-3}\text{ s}^{-1}} dr. \tag{4.44}$$

Here we also defined the volumetric power densities $\rho_D(r,t)$ and $\rho_r(r,t)$; their values $4\pi r^2\rho(r,t)dr$ give the power attributable to a shell of radial thickness dr around the radial position r, and when integrated radially over space give the total power for these processes.

Integrating this density over all space attacks the question of whether the total power to build a gradient, or that to maintain it against diffusion, dominates. We can express the magnitude of the total power spent by a source as

$$P_{\text{tot}}(t) = \int \left(\frac{\partial \mu}{\partial r} dr\right) \underbrace{4\pi r^2}_{\text{area through which flux passes}} (|J_D| + |J_r|)\,. \tag{4.45}$$

This allows us to explicitly quantify the partitioning of total power into the power P_{build} attributable to building and power P_{maintain} to maintain the gradient, namely,

$$P_{\text{tot}}(t) = \underbrace{\int dr\, 4\pi r^2 \frac{\partial \mu}{\partial r} |J_D(r,t)|}_{\equiv P_{\text{maintain}}(t)} + \underbrace{\int dr\, 4\pi r^2 \frac{\partial \mu}{\partial r} |J_r(r,t)|}_{\equiv P_{\text{build}}(t)}\,. \tag{4.46}$$

To compute the relative magnitudes of these terms, we must specify the governing fluxes.

How does a specified time evolution of a concentration field $c(r,t)$ set the fluxes governing these dissipative expressions? Recall that by continuity, a flux has a divergence that sets the rate of change of concentration; if the flux $\mathbf{J} \equiv [J_r, J_\theta, J_\phi]$ is spherically-symmetric (making all angular derivatives vanish, $\partial_\theta, \partial_\phi \to 0$), then,

$$\frac{\partial c(r,t)}{\partial t} = -\vec{\nabla} \cdot \mathbf{J} \tag{4.47}$$

$$= -\frac{1}{r^2}\frac{\partial}{\partial r}\left[r^2 J_r\right]. \tag{4.48}$$

This means that

$$-r^2 \frac{\partial c}{\partial t} = \frac{\partial}{\partial r}\left[r^2 J_r\right], \tag{4.49}$$

or integrating with respect to space from $r = 0$ to $r = r$,

$$-\int_0^r dr\, r^2 \frac{\partial c}{\partial t} = r^2 J_r\Big|_{r=r} - r^2 J_r\Big|_{r=0}\,. \tag{4.50}$$

Since we expect the flux $J_r(r = 0, t)$ to vanish at the origin (since there is no point source or sink there), the latter term vanishes, establishing that the radial flux is given by the radially-integrated rate-of-change of concentration,

$$J_r(r,t) = -\frac{1}{r^2}\int_{r=0}^{r} dr\, r^2 \frac{\partial c}{\partial t}. \tag{4.51}$$

On the maintenance side, Fick's first law tells us the flux required to compensate for diffusion is given by,

$$|J_D(r,t)| = \left| -D\frac{\partial c}{\partial r} \right| \tag{4.52}$$

$$= D\left| \frac{\partial c}{\partial r} \right|. \tag{4.53}$$

Taken together, Equations 4.43, 4.51, and 4.53 give precise answers to the question asked by Eq. 4.46 of how the powers to maintain or build any prescribed concentration profile $c(r,t)$ compare. To gain further insight, we now examine the consequences and lessons of these expressions by adopting specific concentration profiles $c(r,t)$.

Number-conserving, steepening, exponential concentration gradients

We now turn to the question of a spatially varying concentration gradient, $c(r,t)$. The key difference from the examples of section 4.2 is that now we consider the case in which the spatial extent of the gradient is varying over time (i.e. $\lambda(t)$). This model is physically-plausible and—it transpires—analytically-convenient. This model is motivated by two stylized facts. First, motors do exhibit approximately exponential distributions in space at some time points and over relevant spatial regions. See the conceptual description of such distributions in Figures 4.4(A) and (B) as well as exemplary phenomenology of some empirical motor distributions in Figure 4.5.

Second, physically, we believe there really should be conservation of total motors. Consider a spherically-symmetric time varying concentration gradient given by,

$$c(r,t) = c_0(t)\exp\left[-\frac{r}{\lambda(t)}\right], \tag{4.54}$$

where $c_0(t)$ is the concentration at the radial origin of the gradient ($r = 0$), namely the motor peak, and $\lambda(t)$ is a time-varying length scale of the gradient characterizing the spatial evolution of the field. The total number of motors is given by

$$N(t) = \int_{r=0}^{\infty} dr\, 4\pi r^2 c(r,t) = 8\pi\lambda(t)^3 c_0(t). \tag{4.55}$$

Now, to impose conservation of motors, we demand this integral amount to a constant value N, setting the amplitude prefactor $c_0(t)$ as,

$$c_0(t) \equiv \frac{N}{8\pi}\frac{1}{\lambda(t)^3}. \tag{4.56}$$

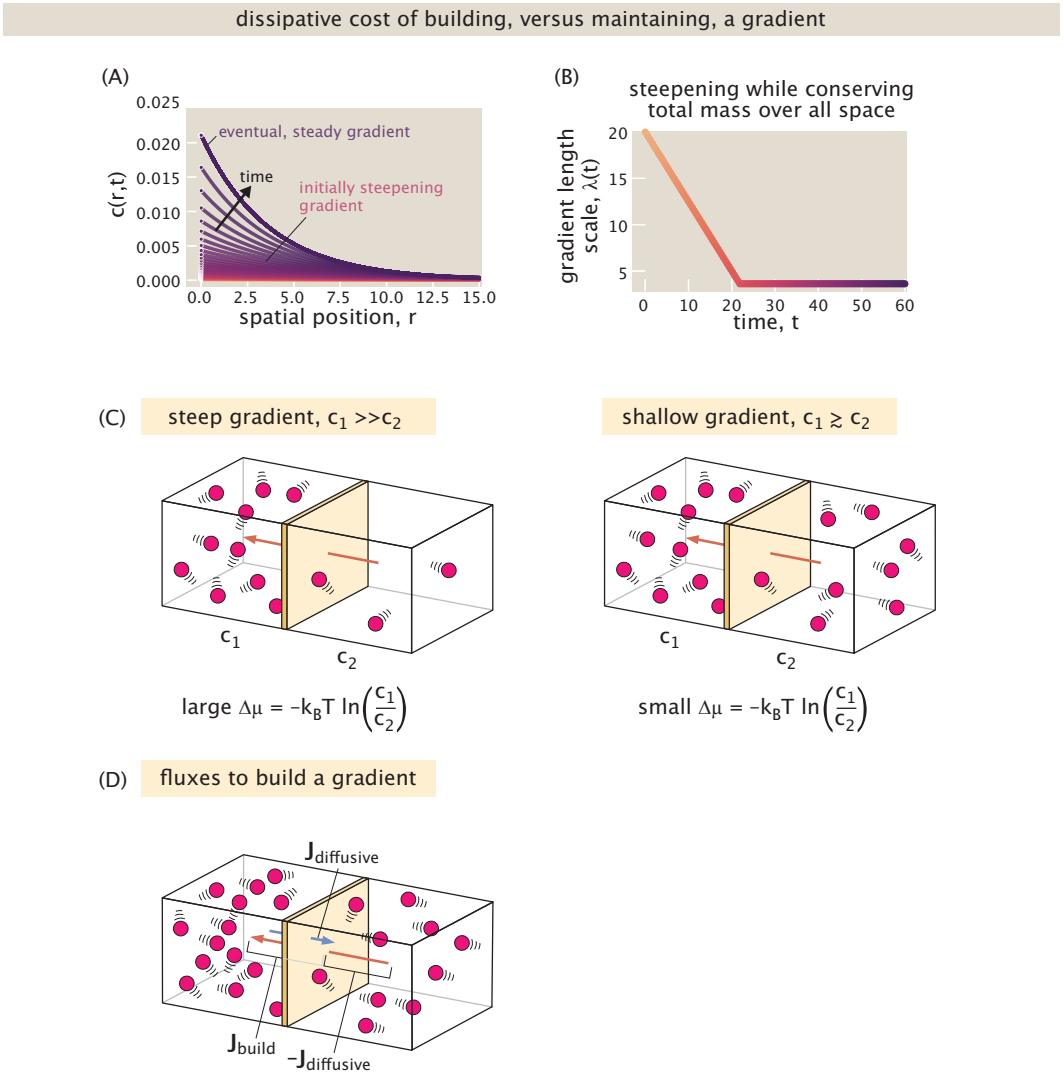

Figure 4.4: **The dynamics of building a gradient.** (A) The concentration profile as a function of time. (B) The length scale of the exponential gradient over time. (C) Schematic of the free energy cost of a steep and a shallow gradient. (D) The fluxes associated with building a gradient.

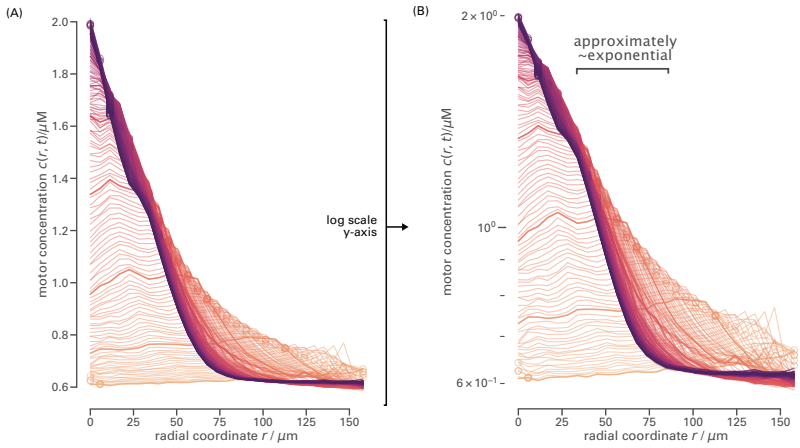

Figure 4.5: **Example motor distributions over space, angularly averaged to show variation of motor concentration over radial coordinate r.** Different lines (in distinct colors) show progression in time. (A) Motors steepen their gradients in time, as visible on linearly-scaled a y-axis. (B) Some regions of the aster are acceptably described as approximately exponentially declining with radial position r away from the center, as demonstrated by acceptably straight line on a semilog plot of the same data in (A).

This sets the spacetime evolution as,

$$c(r,t) = \frac{N}{8\pi}\frac{1}{\lambda(t)^3}\exp\left[-\frac{r}{\lambda(t)}\right]. \tag{4.57}$$

Given such a time-dependent concentration profile, we are poised to examine the power associated with both building and maintaining the gradient.

Resulting (net and diffusive) fluxes of steepening, exponential concentration gradients

Now, we substitute the conservative time evolution profile $c(r,t)$ of Eq. 4.57 into the expression for the net flux derived earlier using the continuity equation. First, we compute the relevant corresponding time derivative $\frac{\partial c}{\partial t}$ as

$$\frac{\partial c}{\partial t} = \frac{N}{8\pi}\left(-3\frac{1}{\lambda(t)^4}\partial_t\lambda\exp\left[-\frac{r}{\lambda(t)}\right] + \frac{1}{\lambda(t)^3}\exp\left[-\frac{r}{\lambda(t)}\right]\frac{r\partial_t\lambda}{\lambda(t)^2}\right) \tag{4.58}$$

$$= \frac{N}{8\pi}\left(\frac{\partial_t\lambda}{\lambda(t)^4}\exp\left[-\frac{r}{\lambda(t)}\right]\left(\frac{r}{\lambda(t)} - 3\right)\right) \tag{4.59}$$

$$= c(r,t)\frac{1}{\lambda(t)}\frac{\partial\lambda}{\partial t}\left(\frac{r}{\lambda(t)} - 3\right). \tag{4.60}$$

Substituting this time derivative into our flux expression Eq. 4.51 gives,

$$J_r(r,t) = -\frac{1}{r^2}\left[\int_{r=0}^{r} dr\, r^2\, \frac{N}{8\pi}\left(\frac{\partial_t \lambda}{\lambda(t)^4}\exp\left[-\frac{r}{\lambda(t)}\right]\left(\frac{r}{\lambda(t)}-3\right)\right)\right] \tag{4.61}$$

$$= -\frac{N}{8\pi}\frac{\partial_t \lambda}{\lambda(t)^4}\frac{1}{r^2}\left[\frac{1}{\lambda(t)}\int_{r=0}^{r} dr\, r^3\, \exp\left[-\frac{r}{\lambda(t)}\right] - 3\int_{r=0}^{r} dr\, r^2 \exp\left[-\frac{r}{\lambda(t)}\right]\right]. \tag{4.62}$$

We evaluate each remaining integral in turn. The first integral $\int_{r=0}^{r} dr\, r^3\, \exp\left[-\frac{r}{\lambda(t)}\right]$ evaluates to $6\lambda(t)^4 - \exp\left[-\frac{r}{\lambda(t)}\right]\lambda(t)\left(r^3 + 3r^2\lambda(t) + 6r\lambda(t)^2 + 6\lambda(t)^3\right)$. The second integral $\int_{r=0}^{r} dr\, r^2 \exp\left[-\frac{r}{\lambda(t)}\right]$ gives $2\lambda(t)^3 - \exp\left[-\frac{r}{\lambda(t)}\right]\lambda(t)\left(r^2 + 2r\lambda(t) + 2\lambda(t)^2\right)$. These give the remarkably-simple weighted-difference

$$\frac{1}{\lambda(t)}\int_{r=0}^{r} dr\, r^3\, \exp\left[-\frac{r}{\lambda(t)}\right] - 3\int_{r=0}^{r} dr\, r^2 \exp\left[-\frac{r}{\lambda(t)}\right] = -\exp\left[-\frac{r}{\lambda(t)}\right] r^3. \tag{4.63}$$

In fact, this simplicity can be anticipated directly by an integration-by-parts result. First, let's start with an integral of the form $\int dr\, r^n e^{-r/\lambda}$. The derivative of the integrand $r^n e^{-r/\lambda}$ is just,

$$\frac{\partial}{\partial r} r^n e^{-r/\lambda} = n r^{n-1} e^{-r/\lambda} - r^n \frac{1}{\lambda} e^{-r/\lambda}. \tag{4.64}$$

Now, integrate each term: this gives $r^n e^{-r/\lambda}$ again on the left side, and two terms involving similar integrals on different integrands on the right side, namely,

$$\int dr\, \left[\frac{\partial}{\partial r} r^n e^{-r/\lambda}\right] = n\int dr\, r^{n-1} e^{-r/\lambda} - \frac{1}{\lambda}\int dr\, r^n e^{-r/\lambda} \tag{4.65}$$

$$\rightarrow r^n e^{-r/\lambda} = n\int dr\, r^{n-1} e^{-r/\lambda} - \frac{1}{\lambda}\int dr\, r^n e^{-r/\lambda}. \tag{4.66}$$

This result directly gives the net flux cancellation above.

Accordingly, the net (building) material flux adopts the charmingly-concise final form,

$$J_r(r,t) = -\frac{N}{8\pi}\frac{\partial_t \lambda}{\lambda(t)^4}\frac{1}{r^2}\left[-\exp\left[-\frac{r}{\lambda(t)}\right] r^3\right] \tag{4.67}$$

$$= r\frac{\partial_t \lambda}{\lambda(t)}\frac{N}{8\pi}\frac{1}{\lambda(t)^3}\exp\left[-\frac{r}{\lambda(t)}\right] \tag{4.68}$$

$$= r\frac{\partial_t \lambda}{\lambda(t)} c(r,t). \tag{4.69}$$

This building flux expression exhibits some crucial properties.

First, we check this expression's units: $\frac{\partial_t \lambda}{\lambda}$ gives units of per time; r gives units of length; and $c(r,t)$ gives units of inverse cubic length. Together these quantities give per length squared per time, as required of any flux. Second, consider the sign of this net flux. A gradient that is steepening has a shrinking characteristic length scale $\lambda(t)$ in time, making $\partial_t \lambda < 0$; this makes $J_r(r,t) < 0$ correspondingly negative (net inwards flux). This sign of the net flux is physically correct: for an extant concentration gradient with greatest concentration in the center at small r and smaller concentration at larger r, we have $\frac{\partial c}{\partial r} < 0$, so a source flux that steepens this imbalance will be indeed pointed inwards, and the opposing diffusive flux $J_D = -D\frac{\partial c}{\partial r} > 0$ will be positive (pointing outwards).

Third, note that the net flux is proportional to temporal rate of change of the length scale λ divided by its current value, which is equivalent to $\frac{\partial_t \lambda}{\lambda(t)} = \frac{\partial \ln \lambda/\lambda_0}{\partial t}$ for some reference length λ_0. That is, the magnitude of the net flux is larger when the governing length scale changes its value faster relative to/*in units of* its current value. If the governing length scale does not change at all in time, $\partial_t \lambda = 0$, then the net flux also vanishes.

Third, note that the flux scales with the radial position r. This can counteract the tendency for concentrations to be smaller at larger r. Specifically, at a fixed time (e.g. setting the value of $\frac{\partial_t \lambda}{\lambda}$ to some fixed constant) the spatial position r where the magnitude of this net building flux J_r is largest is peaked (nonmonotonic) in r at some optimum r_*, since there is a geometric competition between the $\sim r$ term and the exponentially declining gradient in r. The position r_* of this maximum net building flux satisfies

$$0 := \frac{\partial J_r}{\partial r} = \frac{\partial_t \lambda}{\lambda}\left(r \underbrace{\frac{\partial c}{\partial r}}_{\leq 0} + c(r,t) \right) \tag{4.70}$$

$$\rightarrow r_* = \frac{-c(r,t)}{\frac{\partial c}{\partial r}} \tag{4.71}$$

$$\rightarrow r_* = \lambda(t). \tag{4.72}$$

Here, we have learned the following interesting, elegant, nonobvious fact: the maximal material net flux occurs at the radial position which is exactly the current governing exponential length scale. This spatially-maximized flux is $J_r(\lambda, t) =$

$\partial_t\lambda\ c(\lambda,t) = \partial_t\lambda\ \frac{N}{8\pi}\frac{1}{\lambda(t)^3}\frac{1}{e}$. Fourth, note that the flux scales with the absolute local concentration $c(r,t)$: scaling the concentration everywhere by some multiplicative factor α increases the corresponding flux by the same amount. Many of these properties are illustrated for the example of a linear ramp in time of the governing length scale $\lambda(t)$) in Figure 4.6.

For comparison to this net building flux, we also compute the corresponding maintenance (diffusive) flux under this simple conservative model of concentration $c(r,t)$ of Eq. 4.57, finding,

$$J_D(r,t) = -D\frac{\partial c}{\partial r} \tag{4.73}$$

$$= -D\left(-\frac{1}{\lambda(t)}\frac{N}{8\pi}\frac{1}{\lambda(t)^3}\exp\left[-\frac{r}{\lambda(t)}\right]\right) \tag{4.74}$$

$$= D\frac{1}{\lambda(t)}c(r,t). \tag{4.75}$$

Comparing building (net) and maintenance (diffusive) *material* fluxes

Now, we compare the magnitudes of the building and maintenance fluxes given by Eqs. 4.69 and 4.75 by computing their ratio. Importantly, this local ratio of the material transport flux to maintain the gradient, versus that to build the gradient, is also exactly interpretable as the ratio of power densities accomplishing these processes at each point in spacetime. This follows since the volumetric power densities of both processes are each proportional to $\frac{\partial\mu}{\partial r}\Big|_{(r,t)}$ as specified by Eq. 4.44, namely,

$$\frac{\rho_r(r,t)}{\rho_D(r,t)} = \frac{|J_r(r,t)|\frac{\partial\mu}{\partial r}}{|J_D(r,t)|\frac{\partial\mu}{\partial r}} \tag{4.76}$$

$$= \frac{|J_r(r,t)|}{|J_D(r,t)|}. \tag{4.77}$$

Proceeding to compute this ratio of the net flux $J_r(r,t)$ to the maintenance diffusive flux $J_D(r,t)$ (and recalling $|J_r| = -J_r$, we see that

$$\mathrm{Pe}^\circ(r,t) \equiv \frac{|J_r(r,t)|}{J_D(r,t)} = \frac{-r\frac{\partial_t\lambda}{\lambda(t)}c(r,t)}{D\frac{1}{\lambda(t)}c(r,t)} \tag{4.78}$$

$$= \frac{r(-\partial_t\lambda)}{D}. \tag{4.79}$$

$$= \frac{r^2/D}{r/(-\partial_t\lambda)} = \frac{\text{diffusion time}}{\text{“advection” time}}. \tag{4.80}$$

$$\tag{4.81}$$

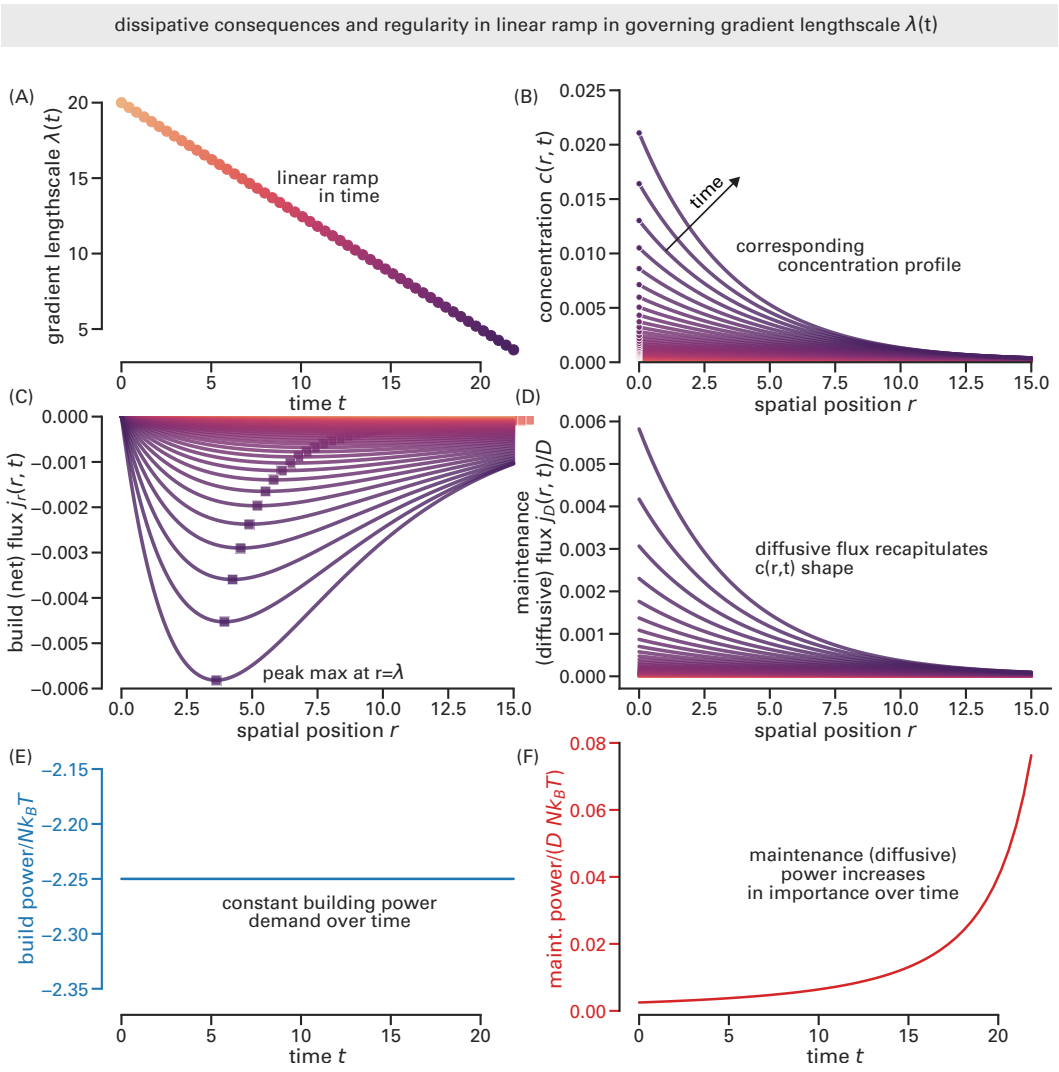

Figure 4.6: **Example of the dissipative behaviors of a steepening (number-preserving) exponential gradient in time, where the characteristic length scale $\lambda(t)$ decreases as a linear ramp in time, $\lambda(t) = \lambda_0 - \gamma t$.** (A) The illustrative linear ramp of the length scale in time, $\lambda(t) = 20 - 0.75t$, for arbitrary units of λ, t. (B) The corresponding concentration profile $c(r,t)$ over space and time (the latter indicated by distinct colors). (C) The resulting net material flux $J_r(r,t)$ required to build the gradient in this prescribed manner. Squares emphasize that the maximum flux occurs at the radial coordinate $r_* = \lambda(t)$. (D) The material flux J_D required to maintain the gradient at its current state. (E) The corresponding power $P_{\text{build}}(t)$ required to enact the steepening trajectory of the concentration gradient. This linear ramp shows a particularly simple (boring) form as $\partial_t \lambda$ is a constant for this example. (F) The maintenance (diffusive) power P_{maintain} increases with the gradient steepness, so also increases in time. (In all these traces, we measured concentration in units of the prefactor $N/8\pi$ and took the diffusion constant to be $D = 1$.)

Thus, under this model, we derive the profoundly surprising result that the relative importance of the net flux J_r needed to *build* the gradient, compared to the flux needed to *maintain* the gradient, depends *only* on whether $r\frac{\partial\lambda}{\partial t} \overset{?}{>} D$. **Interestingly, this ratio can be interpreted as a sort of local material "Péclet number" for each spatial coordinate r and (global) "advective velocity" $\partial_t\lambda$ that compete with diffusion D.** We denote this dimensionless value $\mathrm{Pe}^\circ(r,t)$, where the superscript recalls this is a material (and local) value. When this ratio is smaller than one, the diffusive flux dominates; when the ratio exceeds one, the building flux dominates.

Accordingly, in our exponential concentration gradient setting, for any nonequilibrium control scheme accomplishing a time-varying characteristic length scale $\lambda(t)$ (yielding some value of rate-of-change $\partial_t\lambda \neq 0$), there exists some radial position r past which the magnitude of the flux J_r needed to build the gradient will always exceed the magnitude of the flux J_D needed to just maintain the gradient.

Generality of the local form of this ratio being interpreted as a Péclet number. How special to the details of our chosen illustrative example $c(r,t)$ is the fact that we can precisely interpret the ratio of a net flux to a diffusive flux as some governing effective Péclet number? In fact, not special at all; *any* concentration profile or flux will admit such a description, as we now show. Dimensionally speaking, such a ratio is recognizable as,

$$\left[\frac{J_r}{J_D}\right] = \frac{\frac{1}{\text{area}} \times \frac{1}{\text{time}}}{D\frac{\text{concentration}}{\text{length}}} \tag{4.82}$$

$$= \frac{\frac{1}{\text{length}^2} \times \frac{1}{\text{time}}}{D\frac{1}{\text{length}}\frac{1}{\text{length}^3}} \tag{4.83}$$

$$= \frac{\text{length}^2 \times \frac{1}{\text{time}}}{D} \tag{4.84}$$

$$= \frac{\text{length} \times \frac{\text{length}}{\text{time}}}{D} \tag{4.85}$$

$$= \frac{\text{length} \times \text{velocity}}{D}, \tag{4.86}$$

which is exactly the form of a Péclet number.

Building and maintenance power densities are maximized at *distinct* parts of space

The radial dependence of the local Péclet number in Eq. 4.81 hints that the relative importance of building and maintenance can change over space. To understand this

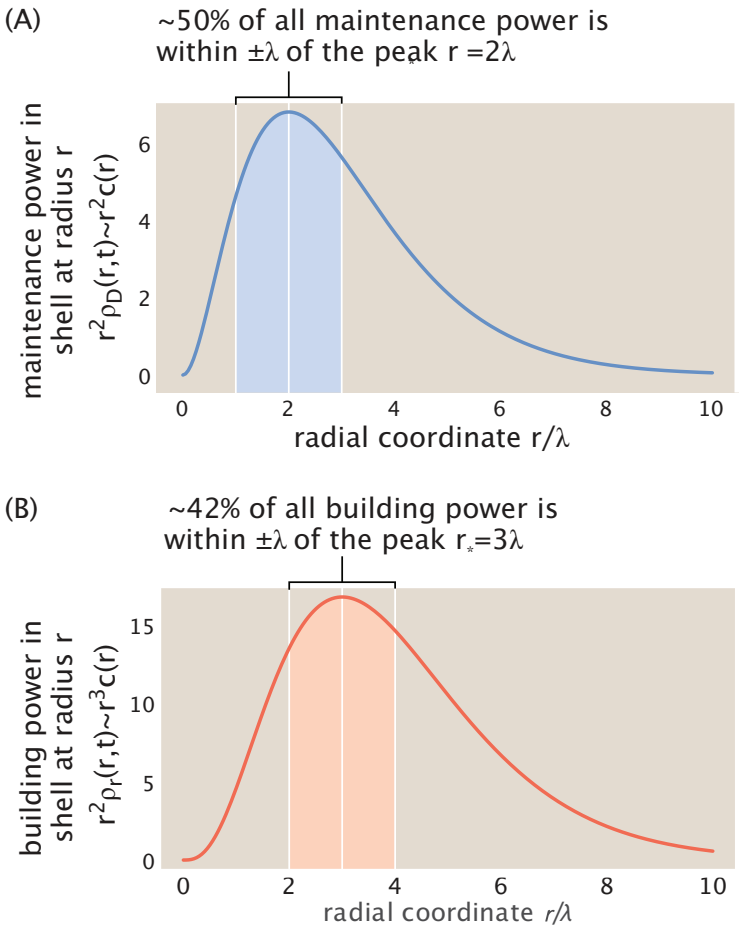

Figure 4.7: **Contrasting the locations in space where distinct dissipative mechanisms are expected to be maximal.** (A) The power attributable to a shell of radius r and unit thickness required just to maintain an exponential gradient, as specified by Eq. 4.93. This maintenance power has a spatial maximum at a radial position of $r = 2\lambda$; this localization reproduces the localization of substance in space: about half of all power over all space (from $r = 0$ to $r = \infty$) is found just within a shell within a decay length $\pm\lambda$ on either side of the dissipatively- maximal radius. (B) The power attributable to a shell of radius r and unit thickness required to further build a gradient, as specified by Eq. 4.99. This building power shows both shifted and more diffuse localization behavior than maintenance power; for instance, only $\sim 42\%$ of the total power is found within a shell at the distinct maximum of $r = 3\lambda$.

spatial dependence, we now explore further how the underlying power in a shell at a radial coordinate r varies in space quite distinctly for these two mechanisms. We preview these contrasting behaviors in Fig. 4.7.

The power attributable to a shell of radius r and thickness dr presenting a volumetric

power density $\rho(r,t)$ is

$$\text{power}(r,t) = \rho(r,t)\, 4\pi r^2\, dr. \tag{4.87}$$

Therefore, the factor $4\pi\rho(r,t)r^2$—a power per radial length increment—controls the spatial distribution of dissipation. Accordingly, we now consider the shape of this radial power density for both maintenance and building power. Substituting the fluxes captured by Eqs. 4.69 and 4.75 and chemical-potential variation (Eq. 4.105) appropriate for our exponential gradient gives the maintenance radial power density as,

$$\begin{aligned}
r^2\rho_D(r,t) &= r^2\left(|J_D(r,t)|\frac{\partial\mu}{\partial r}\right) && (4.88)\\
&= r^2\left(D\frac{1}{\lambda(t)}c(r,t)\right)\left(-k_BT\frac{1}{\lambda(t)}\right) && (4.89)\\
&= r^2c(r,t)\left(-\frac{D}{\lambda(t)^2}\,k_BT\right) && (4.90)\\
&= r^2\left(\frac{N}{8\pi}\frac{1}{\lambda(t)^3}\exp\left[-\frac{r}{\lambda(t)}\right]\right)\left(-\frac{D}{\lambda(t)^2}\,k_BT\right) && (4.91)\\
&= r^2\exp\left[-\frac{r}{\lambda(t)}\right]\left(-\frac{ND}{8\pi\lambda(t)^5}\,k_BT\right) && (4.92)\\
&\sim r^2\exp\left[-\frac{r}{\lambda(t)}\right]. && (4.93)
\end{aligned}$$

Similarly, the building radial power density is given as,

$$\begin{aligned}
r^2\rho_r(r,t) &= r^2\left(|J_r(r,t)|\frac{\partial\mu}{\partial r}\right) && (4.94)\\
&= r^2\left(r\frac{\partial_t\lambda}{\lambda(t)}c(r,t)\right)\left(-k_BT\frac{1}{\lambda(t)}\right) && (4.95)\\
&= r^3c(r,t)\left(-\frac{1}{\lambda(t)}\frac{\partial_t\lambda}{\lambda(t)}\,k_BT\right) && (4.96)\\
&= r^3\left(\frac{N}{8\pi}\frac{1}{\lambda(t)^3}\exp\left[-\frac{r}{\lambda(t)}\right]\right)\left(-\frac{1}{\lambda(t)}\frac{\partial_t\lambda}{\lambda(t)}\,k_BT\right) && (4.97)\\
&= r^3\exp\left[-\frac{r}{\lambda(t)}\right]\left(-\frac{N}{8\pi}\frac{\partial_t\lambda}{\lambda(t)^5}k_BT\right) && (4.98)\\
&\sim r^3\exp\left[-\frac{r}{\lambda(t)}\right]. && (4.99)
\end{aligned}$$

Since the exponentially-decreasing concentration $c(r,t)$ in space competes with the geometric prefactors of r^2 or r^3, respectively, in Eqs. 4.93 and 4.99, both of these

power densities have local maxima in space; these maxima occur when the power laws just start losing to the exponentially decreasing concentration field. However, notably, the spatial positions where these maxima occur are found at distinct radial optima. Specifically, the maximum of the radial maintenance power density occurs at $r_* = 2\lambda$, whereas the maximum of the radial building power density occurs at $r_* = 3\lambda$.

The majority of either type of power occurs within a moderately close vicinity of the respective maxima. To quantify just how much power is concentrated around these positions in space, we compute the fraction of this power found within a tolerance (radial distance) $\pm\delta$ of the maximum r_*, namely what we call the *localization fraction*,

$$f(\delta) \equiv \frac{\int_{r_*-\delta}^{r_*+\delta} dr\, r^2 \rho(r,t)}{\int_0^\infty dr\, r^2 \rho(r,t)}. \tag{4.100}$$

We find that the maintenance power density shows the localization fraction (about its maximum $r_* = 2\lambda$),

$$f_D(\delta) = \frac{\int_{r_*-\delta}^{r_*+\delta} dr\, r^2 \exp\left[-\frac{r}{\lambda(t)}\right]}{\int_0^\infty dr\, r^2 \exp\left[-\frac{r}{\lambda(t)}\right]} \tag{4.101}$$

$$= \frac{\frac{1}{e^2} 2\lambda \left(\left(\delta^2 + 10\lambda^2\right) \sinh\left[\frac{\delta}{\lambda}\right] - 6\delta\lambda \cosh\left[\frac{\delta}{\lambda}\right]\right)}{2\lambda^3}. \tag{4.102}$$

To get an intuitive sense of the scale at which this maintenance power is localized in space, we numerically evaluate this localization fraction with $\delta = \lambda$, which asks for the fraction of power in a shell of total radial thickness 2λ about its maximum $r_* = 2\lambda$. We find $f_D(\lambda) \approx 0.4965$, namely about $\approx 50\%$ of all the maintenance power over all space is localized to this radial region (of within a decay length on either side of the dissipatively-maximal radial position).

Next, we contrast this localization behavior to that of the building power, about its

distinct maximum $r_* = 3\lambda$. We see that,

$$f_r(\delta) = \frac{\displaystyle\int_{r_*-\delta}^{r_*+\delta} dr\, r^3 \exp\left[-\frac{r}{\lambda(t)}\right]}{\displaystyle\int_0^{\infty} dr\, r^2 \exp\left[-\frac{r}{\lambda(t)}\right]} \tag{4.103}$$

$$= \frac{\frac{1}{e^3} 2\lambda \left(6\lambda \left(2\delta^2 + 13\lambda^2\right) \sinh\left[\frac{\delta}{\lambda}\right] - \delta(\delta^2 + 51\lambda^2) \cosh\left[\frac{\delta}{\lambda}\right]\right)}{6\lambda^4}. \tag{4.104}$$

Again numerically evaluating the localization fraction with a tolerance given by the characteristic length scale, $\delta = \lambda$, we see that $f_r(\lambda) = 0.4237$: about $\approx 42\%$ of the power to build the gradient is localized in this region within a decay length in either direction around its distinct spatial maximum.

These power densities make clear a significant and interesting difference. The maintenance power density inherits its localization behavior identically from concentration. That is, the fraction of all dissipation found within a shell centered at the dissipatively-maximal radius is exactly the same as the fraction of all substance found within the same shell. (This fact, reflected by Eq. 4.89, follows since our model of an exponential concentration gradient shows that diffusive flux J_D is simply proportional to the concentration field $c(r,t)$ with no geometric prefactors and the gradient of chemical potential $\frac{\partial \mu}{\partial r}$ is constant in space.) That maintenance power density is in literal thrall to concentration is in substantial contrast to the distinct weighting of the building power, whose integrand enjoys an extra radial $\sim r^{+1}$ weight factor (as reflected by Eq. 4.95 and Eq. 4.69). Figure 4.7 summarizes and illustrates these contrasts in dissipative localization. Last, Figure 4.8 gives another high-level schematic overview of these spatial relationships.

Comparing building (net) and maintenance (diffusive) *powers*

For our specific exponential gradient, we observe the unusual fact that $\frac{\partial \mu}{\partial r}$ is independent of space r, namely,

$$\frac{\partial \mu}{\partial r} = k_B T \frac{\frac{\partial c}{\partial r}}{c(r,t)} = -k_B T \frac{1}{\lambda(t)}. \tag{4.105}$$

Armed with all the ingredients we need, we now compute each power term in turn as

$$P_{\text{build}}(t) = -4\pi \int dr\, r^2 \, \frac{\partial \mu}{\partial r} \left(-\frac{1}{r^2} \int_{r=0}^{r} dr\, r^2 \frac{\partial c}{\partial t}\right) \quad \text{(by Eq. 4.51)}, \tag{4.106}$$

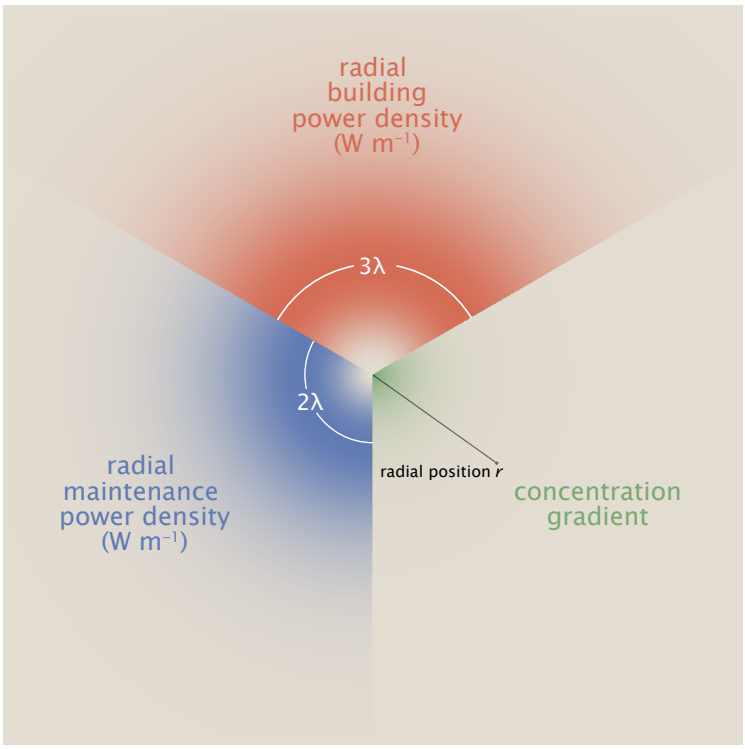

Figure 4.8: **Visualization of the distinct spatial localizations of maintenance power (blue, lower left); building power (red, top); and concentration (green, lower right).**

and

$$P_{\text{maintain}}(t) = 4\pi \int dr\, r^2 \frac{\partial \mu}{\partial r} \left(-D \frac{\partial c}{\partial r}\right) \quad \text{(by Fick)}. \tag{4.107}$$

Substituting this and the remaining ingredients in our particular exponential gradient

scenario, we see,

$$P_{\text{build}}(t) = -4\pi \frac{\partial \mu}{\partial r} \int_0^\infty dr\, r^2 \left(r \frac{\partial_t \lambda}{\lambda(t)} c(r,t) \right) \tag{4.108}$$

$$= -\frac{\partial \mu}{\partial r} \frac{\partial_t \lambda}{\lambda(t)} \int_0^\infty dr\, 4\pi r^2\, r c(r,t) \tag{4.109}$$

$$= -4\pi \frac{\partial \mu}{\partial r} \int dr\, r^2 \left(r \frac{\partial_t \lambda}{\lambda(t)} \frac{N}{8\pi} \frac{1}{\lambda(t)^3} \exp\left[-\frac{r}{\lambda(t)} \right] \right) \tag{4.110}$$

$$= -\frac{\partial \mu}{\partial r} \frac{4\pi}{8\pi} N \frac{\partial_t \lambda}{\lambda(t)} \frac{1}{\lambda(t)^3} \int_0^\infty dr\, r^3 \exp\left[-\frac{r}{\lambda(t)} \right], \quad \text{(Mathematica gives } \int_0^\infty dr\, r^3 e^{-r/\lambda} = 6\lambda^4\text{)}, \tag{4.111}$$

$$= -\frac{\partial \mu}{\partial r} \frac{N}{2} (\partial_t \lambda) \frac{1}{\lambda^4} 6\lambda^4. \tag{4.112}$$

$$= \frac{\partial \mu}{\partial r} 3N(-\partial_t \lambda). \tag{4.113}$$

$$= -3Nk_BT \frac{(-\partial_t \lambda)}{\lambda(t)} \tag{4.114}$$

$$= 3Nk_BT \frac{\partial \ln \lambda(t)/\lambda_0}{\partial t}. \tag{4.115}$$

(Here we could also just recall that for integer n, the value of $\int_0^\infty dr\, r^n e^{-r/\lambda}$ is $n!\lambda^{n+1}$.)

The maintenance power is

$$P_{\text{maintain}}(t) = \frac{\partial \mu}{\partial r} \int dr\, 4\pi\, r^2 \left(D \frac{1}{\lambda(t)} c(r,t) \right) \tag{4.116}$$

$$= \frac{\partial \mu}{\partial r} D \frac{1}{\lambda(t)} \underbrace{\int dr\, 4\pi\, r^2\, c(r,t)}_{=N, \text{ by Eq. 4.55}} \tag{4.117}$$

$$= \frac{\partial \mu}{\partial r} D \frac{1}{\lambda(t)}. \tag{4.118}$$

$$= -Nk_BTD \frac{1}{\lambda(t)^2}. \tag{4.119}$$

Thus the ratios of these two powers is

$$\text{Pe}^{\lightning}(t) \equiv \frac{P_{\text{build}}(t)}{P_{\text{maintain}}(t)} = \frac{\frac{\partial \mu}{\partial r} 3N(-\partial_t \lambda)}{\frac{\partial \mu}{\partial r} D \frac{1}{\lambda(t)} N} \tag{4.120}$$

$$= \frac{3\lambda\, (-\partial_t \lambda)}{D}, \tag{4.121}$$

where we have denoted this ratio as $\text{Pe}^{\lightning}(t)$ to emphasize its nature as a dissipative (and spatially-global) Péclet number. Is this unitless? The numerator is length squared per second, consistent with the denominator, so yes.

(Interestingly, we see that taking a spatial coordinate of $r = 3\lambda$ in the local flux ratio of $\text{Pe}^{\circ}(r,t)$ in Eq. 4.81 makes that local flux ratio there equal to this total radially-integrated power ratio $\text{Pe}^{\lightning}(t)$ of Eq. 4.121.)

How does the relative importance of building versus maintenance powers change over time, as gradients steepen?

Our foregoing analysis was mathematically general, applying to any exponential gradient regardless of the trajectory of the characteristic length scale $\lambda(t)$. To proceed further, however, we must now specify some possible candidate time-evolutions $\lambda(t)$. How will the relative importance of maintenance versus building power be conserved or vary across diffrent $\lambda(t)$ profiles? Consider these three examples, which hint at the diversity of attainable behaviors.

Figure 4.6 illustrates characteristic behaviors of a concentration gradient steepening under a linearly-decreasing gradient length scale $\lambda(t)$ in time.

1. For a simple linear gradient in time, $\lambda(t) = \lambda_0 - \gamma t$ (for $t \leq \frac{\lambda_0}{\gamma}$), we see that $(\partial_t \lambda)\lambda(t) = -\gamma(\lambda_0 - \gamma t)$. This magnitude also decreases in time. In other words, the relative importance of building power to maintenance power must monotonically decrease over time. This is consistent with the fact that the building power is a constant (while maintenance power increases) in this case: see Eq. 4.114. See Figure 4.6.

2. For a power law shrinkage of the exponential gradient length scale in time, $\lambda(t) = Bt^{-\alpha}$ (with $\alpha \geq 0$ to assure that the gradient steepens), observe that

$$\partial_t \lambda = -B\alpha T^{-\alpha-1} = -\alpha \frac{\lambda(t)}{t}.$$

   Therefore, $(\partial_t \lambda)\lambda(t) = -\alpha \frac{\lambda(t)^2}{t}$, whose magnitude decreases with time t (since $\lambda(t)$ declines monotonically). In other words, the relative importance of building power to maintenance power must monotonically decrease over time.

3. For a more unusual (e.g. somewhat contrived) profile like $\lambda(t) = (1-t)^{\eta}$ for $\eta \in [0, 1/2]$, and $t < 1$, we appear to find a profile where—in contrast to the earlier examples—the magnitude of $(\partial_t \lambda)\lambda(t)$ increases in time, making the building power *increasingly* important relative to the maintenance power.

Comparing building (net) and maintenance (diffusive) *energies*

Under our model, during a time τ, the total energy spent on building is the time integral of Eq. 4.114,

$$E_{\text{build}}(\tau) = \int_{t=0}^{\tau} dt\ P_{\text{build}}(t) \tag{4.122}$$

$$= \int_{0}^{\tau} dt\ \frac{\partial \mu}{\partial r} 3N \partial_t \lambda \tag{4.123}$$

$$= 3N \int_{0}^{\tau} dt\ \left(-k_B T \frac{1}{\lambda(t)}\right) \partial_t \lambda \tag{4.124}$$

$$= -3N k_B T \int_{0}^{\tau} dt\ \frac{\partial_t \lambda}{\lambda(t)} \tag{4.125}$$

$$= -3N k_B T \int_{0}^{\tau} dt\ \frac{\partial \ln \lambda/\lambda_0}{\partial t} \tag{4.126}$$

$$= -3N k_B T \ln\left[\frac{\lambda(\tau)}{\lambda(0)}\right]. \tag{4.127}$$

Remarkably, this energy to build is *independent* of the precise timecourse $\lambda(t)$, only depending on the starting and ending length scale of the gradient!

We note with amusement that this expression Eq. 4.127 adopts precisely the same form (including numerical prefactors!) as the calculation for the work needed to contract a (microtubule) gas from an initial volume V_i and a final volume V_f in Eq. 3.146 specifying $W = -N k_B T \ln \frac{V_f}{V_i}$, if we take these volumes to be of order the characteristic length scale cubed, $V_i \sim \lambda(0)^3$ and $V_f \sim \lambda(\tau)^3$, where the cubic relation delivers our factor of three in Eq. 4.127.

In light of our discussion and estimates in this section: the fact that the order-of-magnitude estimate $\langle P_{\text{build}} \rangle = E_{\text{build}}/T = \frac{\int_0^T dt\ P_{\text{build}}(t)}{T}$ (for an observation/construction time T) is much smaller than instantaneous measured power values does *NOT* in fact say that building power is small compared to actual total power expenditures. This discrepancy says much more about the fact that *instantaneous* power values can differ a lot from the *average* of the same values, than that the *instantaneous* fundamental/theoretical building power is small compared to *instantaneous* measured total power values!

In contrast the energy to maintain is the time integral of Eq. 4.118, giving

$$E_{\text{maintain}}(\tau) = \int_{t=0}^{\tau} dt\ P_{\text{maintain}}(t) \tag{4.128}$$

$$= \int_{t=0}^{\tau} dt\ \frac{\partial \mu}{\partial r} D \frac{1}{\lambda(t)} N \tag{4.129}$$

$$= DN \int_{t=0}^{\tau} dt\ \left(-k_B T \frac{1}{\lambda(t)}\right) \frac{1}{\lambda(t)} \tag{4.130}$$

$$= -DNk_B T \int_{0}^{\tau} dt\ \frac{1}{\lambda(t)^2}. \tag{4.131}$$

Notice, as expected, that such a maintenance energy accumulates even when the length scale λ is *not* changing in time ($\partial_t \lambda \neq 0$)—and at any finite constant λ, E_{maintain} will grow with longer τ.

It is hard to bound this maintenance energy without stipulating a precise timecourse for $\lambda(t)$. But, to reach for a rough feel, say that we only know that $\lambda(t)$ decreases monotonically in time on the time interval of interest, such that $\lambda(t = \tau) \leq \lambda(t) \leq \lambda(t = 0)$. Then we can conclude the simple fact that,

$$\int_0^{\tau} dt\ \frac{1}{\lambda(0)^2} \leq \int_0^{\tau} dt\ \frac{1}{\lambda(t)^2} \leq \int_0^{\tau} dt\ \frac{1}{\lambda(\tau)^2} \tag{4.132}$$

$$\rightarrow \frac{\tau}{\lambda(0)^2} \leq \int_0^{\tau} dt\ \frac{1}{\lambda(t)^2} \leq \frac{\tau}{\lambda(\tau)^2}, \tag{4.133}$$

which means that

$$DNk_B T \frac{\tau}{\lambda(0)^2} \leq |E_{\text{maintain}}(\tau)| \leq DNk_B T \frac{\tau}{\lambda(\tau)^2}. \tag{4.134}$$

Accordingly, the ratio of these two energies is only constrained here to lie between,

$$\frac{3Nk_B T \ln\left[\frac{\lambda(\tau)}{\lambda(0)}\right]}{DNk_B T \frac{\tau}{\lambda(\tau)^2}} \leq \frac{|E_{\text{build}}(\tau)|}{|E_{\text{maintain}}(\tau)|} \leq \frac{3Nk_B T \ln\left[\frac{\lambda(\tau)}{\lambda(0)}\right]}{DNk_B T \frac{\tau}{\lambda(0)^2}} \tag{4.135}$$

$$\rightarrow \frac{3}{D} \frac{\ln\left[\frac{\lambda(\tau)}{\lambda(0)}\right]}{\frac{\tau}{\lambda(\tau)^2}} \leq \frac{|E_{\text{build}}(\tau)|}{|E_{\text{maintain}}(\tau)|} \leq \frac{3}{D} \frac{\ln\left[\frac{\lambda(\tau)}{\lambda(0)}\right]}{\frac{\tau}{\lambda(0)^2}}. \tag{4.136}$$

Empirical estimates for an example aster's motor gradients

We now consider how well the theoretical insights described above jibe with our experimental measurements. It appears that the major obvious trends—e.g., that the relative importance of building and maintenance powers can switch dramatically in time across trajectories—will be widely preserved across aster phenomenology.

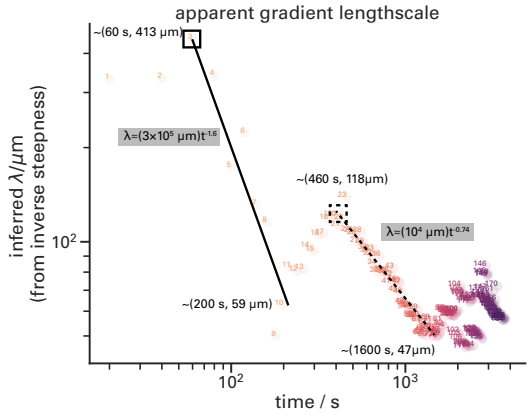

Figure 4.9: **Temporal changes in apparent length scale of motor gradients of the data in Figure 4.5.** At left, highlighted by a solid black line as a guide to the eye, an early-time steepening in the motors gives a steep approximate power law $\hat{\lambda} \approx (3 \times 10^6\ \mu\text{m})(t/1s)^{-1.6}$. The solid black square indicates a reference point $(t, \hat{\lambda}(t)) \approx (60\ s, 413\ \mu m)$ with an approximate rate-of-change of $\partial_t \hat{\lambda} \approx -11\ \mu m/s$ taken from the slope of the line. At right, a slower but still appreciable steepening resumes, marked by the dashed black line, with an approximate power law fit of $\hat{\lambda} \approx (10^4\ \mu m)(t/1s)^{-0.74}$. The dashed black square indicates a reference point at $(t, \hat{\lambda}) \approx (460\ s, 118\ \mu m)$ with an approximate rate of change of $-0.2\ \mu m/s$ taken from the slope of the dashed line.

To get a very rough sense of scale for how motor gradients actually change their characteristic length scales $\lambda(t)$ in time, we considered the same motor distributions shown earlier in the data of Figure 4.5. At each time, we found the radial position r_s at which the motor distribution was steepest in space. Then, assuming (coarsely, but not unreasonably) that these gradients were indeed locally exponentially decaying at this steepest point, we computed the approximate effective exponential gradient length scale as $\hat{\lambda}(t) = -\frac{c(r_s,t)}{\frac{\partial c}{\partial r}}$. Figure 4.9 depicts how this aster changes its apparent length scale.

Next, we—just by eye, not rigorously—extracted an apparent power law fit to each of two distinct dynamical regimes visible in the length scale's trajectory. This allows us to substitute empirically estimated values into the dissipative Péclet number of Eq. 4.121. The first regime, marked by the solid black square in Figure 4.9, implies that the building power $P_{\text{build}}(t)$ exceeds the maintenance power P_{maintain} by approximately a multiplicative factor of order,

$$\frac{P_{\text{build}}(t)}{P_{\text{maintain}}(t)} = \frac{3\lambda\ (-\partial_t \lambda)}{D} \tag{4.137}$$

$$\approx \frac{3(413\ \mu m)(-(-11\ \mu m/s))}{40\ \mu m^2/s} \tag{4.138}$$

$$\approx 340\times . \tag{4.139}$$

The second, more slowly-steepening, regime appears to register a more equal balance between building power and maintenance power, giving about

$$\frac{P_{\text{build}}(t)}{P_{\text{maintain}}(t)} = \frac{3\lambda\ (-\partial_t \lambda)}{D} \tag{4.140}$$

$$\approx \frac{3(118\ \mu m)(-(-0.2\ \mu m/s))}{40\ \mu m^2/s} \tag{4.141}$$

$$\approx 1.8\times . \tag{4.142}$$

These estimates, however coarse, argue quantitatively that **maintenance power could be dwarfed by building power to the tune of order a few hundred in the early times of an aster formation**, but show that the maintenance cost grows much more competitive at late times. □

References

[1] A. H. Stouthamer. "Theoretical study on amount of ATP required for synthesis of microbial cell material." In: *Antonie Van Leeuwenhoek Journal of Microbiology* 39.3 (1973), pp. 545–565.

[2] Ian S. Farmer and Colin W. Jones. "The energetics of Escherichia coli during aerobic growth in continuous culture." In: *European Journal of Biochemistry* 67.1 (1976), pp. 115–122.

[3] Matteo Mori et al. "Functional decomposition of metabolism allows a system-level quantification of fluxes and protein allocation towards specific metabolic functions." In: *Nature Communications* 14.1 (2023), p. 4161.

[4] Ana P. Alonso et al. "Carbon conversion efficiency and central metabolic fluxes in developing sunflower (Helianthus annuus L.) embryos." In: *The Plant Journal* 52.2 (2007), pp. 296–308.

[5] Michael J. Buono and Fred W. Kolkhorst. *Estimating ATP resynthesis during a marathon run: A method to introduce metabolism*. 2001.

[6] John J. Hopfield. "Kinetic proofreading: A new mechanism for reducing errors in biosynthetic processes requiring high specificity." In: *Proceedings of the National Academy of Sciences of the United States of America* 71.10 (1974), pp. 4135–9.

[7] T. L. Hill and Marc W. Kirschner. “Bioenergetics and kinetics of microtubule and actin filament assembly-disassembly.” In: *Int Rev Cytol* 78 (1982), pp. 1–125.

[8] Tim J. Mitchison and Marc W. Kirschner. “Properties of the Kinetochore In vitro. II. Microtubule Capture and ATP-Dependent Translocation.” In: *Journal of Cell Biology* 101.3 (1985), pp. 766–777.

[9] Javier Estrada et al. “Information integration and energy expenditure in gene regulation.” In: *Cell* 166.1 (2016), pp. 234–244.

[10] Sara D. Mahdavi et al. “Flexibility and sensitivity in gene regulation out of equilibrium.” In: *Proceedings of the National Academy of Sciences* 121.46 (2024), e2411395121.

[11] Christopher C. Govern and Pieter Rein ten Wolde. “Energy dissipation and noise correlations in biochemical sensing.” In: *Physical Review Letters* 113.25 (2014), p. 258102.

[12] Francis Crick. “Diffusion in embryogenesis.” In: *Nature* 225.5231 (1970), pp. 420–422.

[13] Mary Munro and Francis H. Crick. “The time needed to set up a gradient: Detailed calculations.” In: *Symposia of the Society for Experimental Biology*. Vol. 25. 1971, pp. 439–453.

[14] Florian M. Gartner and Erwin Frey. “Design principles for fast and efficient self-assembly processes.” In: *Physical Review X* 14.2 (2024), p. 021004.

[15] Massimiliano Esposito and Christian Van den Broeck. “Second law and Landauer principle far from equilibrium.” In: *Europhysics Letters* 95.4 (2011), p. 40004.

[16] Hong Qian. “Relative entropy: Free energy associated with equilibrium fluctuations and nonequilibrium deviations.” In: *Physical Review E* 63.4 (2001), p. 042103.