

Chapter 2

Oscillations on Networks

To better understand how to build, design, and operate nanoscale machines, we have to understand more about what makes dynamics on the smallest scales different from dynamics we observe in our everyday lives. By dynamics, we refer to the displacements, oscillations, and momentum transformations that give rise to observable behavior. Although popular science is full of interesting discussions of how the world of quantum mechanics leads to wonderful and nonintuitive dynamics, this is only part of the story of why small is different. Classical dynamics, “the science of the 19th century,” plays a central role. The reason, in short, is “scaling.”

Consider an object in a fluid medium. It is subject to buffeting by molecules of the medium, which deliver kicks that knock the mass off-course, but conservation of momentum also causes the same molecules to sap momentum from any directed motion the body may possess. Newton’s law states that $F = ma$, but the total force on the object will be a combination of endogenous forces (which are those forces still acting on the body even in a vacuum) and the buffeting and damping forces from the viscous medium. If the endogenous forces are extensive, meaning they are proportional to the size of the object, they scale as R^3 , where R is the effective radius

of the object. In contrast, the fluid medium acts on the surface area of the object, so the forces from the bath causing buffeting and damping scale as R^2 . Thus, the ratio of the endogenous forces, which are the forces we would rely on to do useful work, to the bath forces, which impart noise, scale as R . The smaller an object gets, the smaller the endogenous forces become in relation to the bath forces. At a critical length scale, which would depend on the precise forces involved and the nature of the immersion medium, the buffeting and damping by the bath would completely swamp the object's ability to do persistent work. This scaling has been well explored in fluid mechanics via a number of ratios to understand the balance of various factors [2]. Reynolds number, in particular, captures the change in viscous dynamics relative to inertial dynamics as length scales shrink. It is defined as $\text{Re} = VL/\nu$, where V is the velocity of the object, L is a characteristic length scale, and ν is the kinematic viscosity. When Re is small, inertial forces are overwhelmed by viscous forces, and any directed motion is rapidly quenched.

Small objects inherently live in the world of small Re . In the low Reynolds number regime, dynamic motion such as oscillations cannot be sustained. This chapter will explore ways to create oscillatory behavior in the low Reynolds number regime. It has direct application to the design of nanoscale machines and the operation of proteins within the body.

2.1 The Limits of Oscillations in Overdamped Systems

Oscillations are ubiquitous. We celebrate them and attempt to harness them. Naturally, this interest drives us to study them. There has been no shortage of analysis of the simple harmonic oscillator in all of its variations, but much of the periodicity around us is not equivalent to a mass on a spring. For example, the beating heart is driven by molecular motors which exist in the low-Reynolds number regime, where viscous damping overwhelms inertial forces. For these molecular constituents, buffeting by solvent molecules prevents coherent oscillations from persisting on a timescale longer than the mean time between collisions, which is on order picoseconds [3]. Despite this, we observe the coordination of overdamped components to produce periodic behavior [4]. Studying this coordination on a problem-by-problem basis has uncovered some conceptual principles to designing oscillatory behavior in the overdamped regime, but few truly fundamental laws exist [5, 6, 7, 8, 9]. This work bounds the performance of all discrete-state over-damped oscillatory systems, providing a new look at the necessary conditions for creating coherent oscillations in overdamped systems.¹

When the energy landscape of an overdamped system can be divided into distinct basins of attraction with barriers higher than $k_B T$, the system will tend to reach a local equilibrium within a basin of the energy landscape before fluctuations stochas-

¹We all possess an intuitive comfort with oscillations, but we have to formalize this notion for our analysis. To separate coherent oscillations from random fluctuations, we demand that oscillations be predictable and have a characteristic timescale. Predictability implies that the autocorrelation of a signal will have distinct peaks or troughs corresponding to the period of the oscillations.

tically drive it over a barrier into a neighboring basin. Under these conditions, it is common and appropriate to model each basin as a distinct state, with a fixed rate of transitioning from one state to another [10, 11, 12]. These systems are finite state first-order Markov processes² and can be modeled by the master equation:

$$\frac{dp_i(t)}{dt} = \sum_j^N T_{ij} p_j(t) - \sum_{j \neq i}^N T_{ij} p_i(t), \quad (2.1)$$

where T_{ij} is the transition rate from state j to state i . For introductions to the master equation and its numerous physical applications see [13, 14]. We will assume that all rates are time-independent, meaning no external factors change the rates (but does not necessarily mean that the system is closed). We also make the assumption that \mathbf{T} is an irreducible matrix, enforcing the trivial condition that we are not modeling multiple mutually isolated systems. Finally, we assume that the systems conserve probability, which can always be enforced by adding states to the system to represent sinks. The solution to eq. (2.1) is $\mathbf{p}(t) = \exp(\mathbf{T}t)\mathbf{p}(0)$, where \mathbf{T} is matrix notation for T_{ij} , $T_{ii} = -\sum_{j \neq i} T_{ji}$ and $\mathbf{p}(t)$ is vector notation for $p_j(t)$ [15]. Systems represented by the master equation are completely described by the transition rate matrix, \mathbf{T} , and the initial conditions $\mathbf{p}(0)$. The complete solution is

$$\mathbf{p}(t) = \sum_j \mathbf{v}_j e^{\lambda_j t} (\mathbf{V}_j^{-1} \cdot \mathbf{p}(0) + a_j(t)), \quad (2.2)$$

where \mathbf{v}_j is the j^{th} eigenvector and \mathbf{V}^{-1} is the inverse of the matrix of eigenvec-

²This means we exclude systems with high degrees of quantum coherence or those that are underdamped and therefore inertial. The systems in question completely thermalizes before changing states. Without loss of generality, we will only describe the probability of occupying a given state.

tors.³ Hence, characterizing the properties of \mathbf{T} also characterizes the dynamics of the system [16, 17, 13]. Because the time dynamics of individual modes are ultimately determined by the eigenvalues of \mathbf{T} (see eq. (2.2)), we will be concerned with these eigenvalues and how they relate to oscillations. We first explore these eigenvalues and prove how they constrain the possible oscillations to be fewer than the number of states in the system. Second, I provide examples of these limits by exploring the quality of oscillations in a hypothetical stochastic clock, showing how both microscopic oscillations and macrostates are constrained by the number of states in the system. I conclude by proposing some experiments which may cast direct light onto the physical realization of these bounds on oscillations.

2.2 The Limits on Oscillations

To understand the oscillations in the system represented by \mathbf{T} , we consider the relative contributions of different eigenmodes. From the Perron-Frobenius theorem, all eigenvalues of \mathbf{T} have nonpositive real parts, so all but the $\lambda = 0$ equilibrium mode decay away. Eigenvectors with nonzero imaginary eigenvalues oscillate in magnitude as they decay. As we see in eq. (2.2), after a time $(\text{Re } \lambda_i)^{-1}$, mode i 's contribution to $\mathbf{p}(t)$ will have substantially diminished. If there is an imaginary part to λ_i , before decaying mode i will oscillate $|\text{Im } \lambda_i / \text{Re } \lambda_i|$ times. Because each oscillatory mode will have a resonance independent of the other modes, the overall quality of oscillations

³Dennery, P. & Kryzwicki, A. *Mathematics for Physicists*, Dover 1996.

is given by :

$$\mathcal{Q} = \frac{1}{2} \max_i |\text{Im } \lambda_i / \text{Re } \lambda_i|. \quad (2.3)$$

In closed systems, \mathcal{Q} is the upper bound on the number of oscillations. In an open and homogeneously driven system, \mathcal{Q} describes the coherence of those oscillations, in analogy to the quality-factor of harmonic oscillators. This work establishes upper-bounds on \mathcal{Q} by showing the eigenvalues of \mathbf{T} only exist in specific regions of the complex plane.

Karpelevich's Theorem, as clarified by Ito [18, 19], states that all possible eigenvalues of an N -dimensional stochastic matrix with unit spectral radius ($\max_i |\lambda_i| = 1$) are contained in a bounded region, which we call R_N , on the complex plane, shown in figure 2.1. R_N intersects the unit circle at points $\exp(2\pi ia/b)$, where a and b are relatively prime and $0 \leq a < b \leq N$. The curve connecting points $z = e^{2\pi ia_1/b_1}$ and $z = e^{2\pi ia_2/b_2}$ is described by the parametric equation

$$z^{b_2}(z^{b_1} - s)^{\lfloor N/b_1 \rfloor} = z^{b_1 \lfloor N/b_1 \rfloor} (1 - s)^{\lfloor N/b_1 \rfloor}, \quad (2.4)$$

where s runs over the interval $[0, 1]$ and $\lfloor x/y \rfloor$ is the integer floor of x/y . For example, the curve that connects $z = 1$, corresponding to $(a_1 = 0, b_1 = N)$, with $z = e^{2\pi i/N}$ ($a_2 = 1, b_2 = N$) is $z(s) = (e^{2\pi i/N} - 1)s + 1$.

The rate matrix from eq. (2.1), \mathbf{T} , is not a stochastic matrix. To preserve probability, the sum of each columns of \mathbf{T} is zero, and the diagonal elements are ≤ 0 .⁴

⁴This may be obtained by letting $p_i = (1, 0, 0, \dots)$, substituting p_i into eq. (2.1), and solving for the condition $\sum_i p_i = 0$.

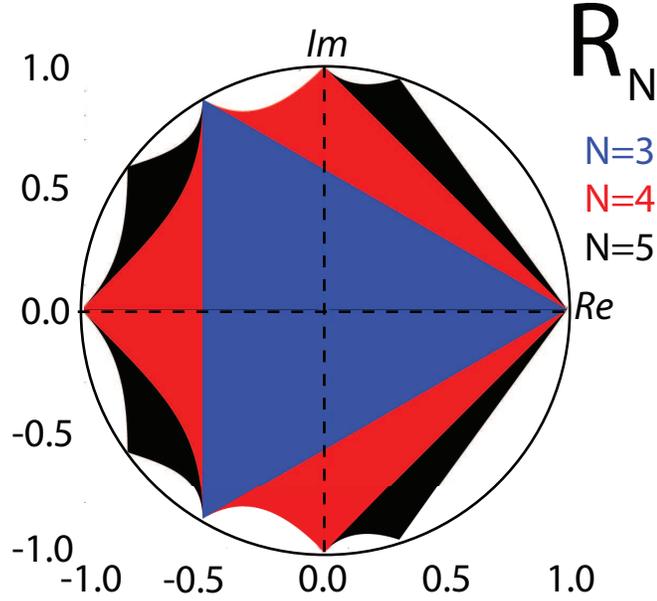


Figure 2.1: The region R_N contains all possible eigenvalues of N -dimensional stochastic matrices with unit spectral radius. Region R_{N+1} contains R_N . This region is symmetric to the real axis and circumscribed by the unit circle. The curves defining each region are given by eq. (2.4), due to Karpelevich's Theorem.

To transform \mathbf{T} into a stochastic matrix, denoted \mathbf{T}' , divide \mathbf{T} by the sum of its largest diagonal element and largest eigenvalue, and add the identity matrix. This transformation allows us to write the eigenvalues of \mathbf{T}' in terms of the eigenvalues of \mathbf{T} :

$$\lambda'_i = \frac{\lambda_i}{\max_j |\mathbf{T}_{jj}| + \max_j |\lambda_j|} + 1. \quad (2.5)$$

Because the most positive eigenvalues of the original \mathbf{T} are 0 and all others have negative real parts, the most positive eigenvalue of \mathbf{T}' is 1. This unique normalization technique ensures all other eigenvalues are less than 1 and fit within the region R_N on the complex plane. Therefore, all of the eigenvalues of \mathbf{T} , will fit within the region $(\max_i |\mathbf{T}_{ii}| + \max_i |\lambda_i|) \times (R_N - 1)$, where these operations on R_N denote scaling and translation, respectively. Within this transformed region, the maximum number of

oscillations will be produced by eigenvalues on the line $\lambda \propto (e^{\pm 2\pi i/N} - 1)$, giving

$$\mathcal{Q}_{\max} = \frac{1}{2} \left| \frac{\sin(2\pi/N)}{\cos(2\pi/N) - 1} \right| = \frac{1}{2} \cot(\pi/N) < \frac{N}{2\pi}. \quad (2.6)$$

We can further refine the limit in eq. (2.6) using a result from Kellogg and Stephens [20], giving

$$\mathcal{Q}_{\max} = \frac{1}{2} \cot \frac{\pi}{\ell_{cyc}} < \frac{\ell_{cyc}}{2\pi}, \quad (2.7)$$

where ℓ_{cyc} is the longest cycle in the system.

Up to this point in our proof, we have restricted ourselves to systems without any degeneracy in the eigenvalues of \mathbf{T} . With degeneracy, as shown in eq. (2.2), the time dependence of eigenvector j may pick up an extra polynomial factor, $a_j(t)$, with degree less than the degeneracy of λ_j , which is always less than $N - 1$. Fortuitous balancing of coefficients could allow a p^{th} -order polynomial to add an additional $p/2$ oscillations. Examining eq. (2.2), we see that the total maximum oscillation quality can be

$$\mathcal{Q}_{\max} < \frac{\ell_{cyc}}{\pi} + \frac{N - 1}{2} < N, \quad (2.8)$$

where the second term is strictly due to degeneracy⁵

2.3 Oscillations in Macrostates: Chemical Clocks

Oscillations which consist of cycles on the discrete state-space are only possible when the system in question violates detailed balance [17], which would be the case in a

⁵Such degeneracy will usually emerge only in hypothetical systems where rates balance perfectly.

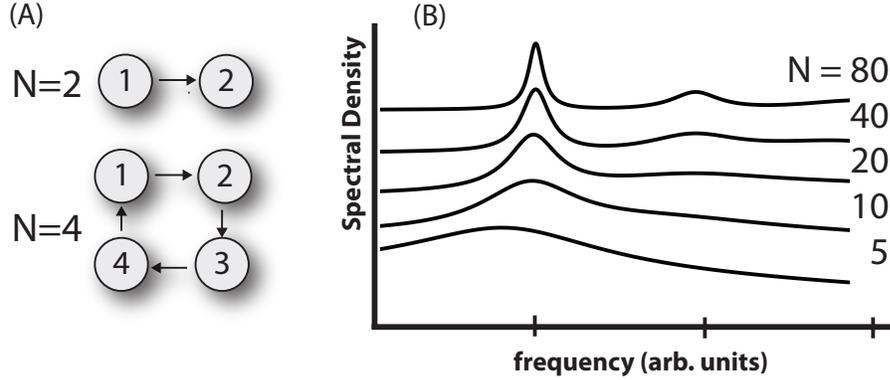


Figure 2.2: (a) When the energy landscape has barriers much larger than $k_B T$, the system will spend most of its time in the minima of the environment. Approximating the continuous landscape by discrete states gives the familiar master equation kinetics. Here we document two examples of systems with unidirectional transition rates. This cyclical system produces the maximum \mathcal{Q} for any given N . (b) As shown in eq. (2.6), a system with only two states cannot coherently oscillate. It produces only random jumps. As the number of states in the unidirectional cycle increases (in the same family as shown in (a)), oscillations become more coherent and more persistent. The spectral density of the unidirectional cycle shows a distinct peak which becomes sharper as N increases. The transition rates have been normalized by the number of states. The \mathcal{Q} of the systems are, from bottom to top: 0.75, 1.46, 3.09, 6.40, 12.7, obtained by fitting Lorentzian functions to the peaks.

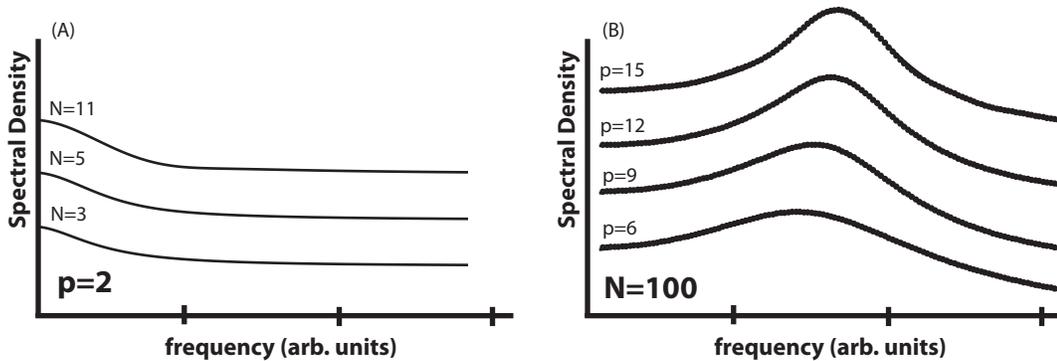


Figure 2.3: Although the linear set of states given by eq. (2.9) does not have any imaginary eigenvalues, macrostates can oscillate. Macrostates are defined as $\langle A \rangle = \sum_i^N A_i^p p_i(t)$. In this case, $A_i^p = \{1 \text{ if } \text{mod}_p i = 0; 0 \text{ otherwise}\}$. (a) Dynamics for different values of N for fixed $p = 2$ (inset) Identical dynamics, but with scaled rates so expected traversal times are the same. (b) Increasing p increases the number and coherence of oscillations given fixed $N = 61$, demonstrating the limit in eq. (2.7).

system which is driven. Otherwise, conservation of energy prevents the system from completing a cycle without encountering significant energy barriers. The microstates of the system, represented by the instantaneous values of $\mathbf{p}(t)$, would not show any oscillatory dynamics or peaks in the spectral density, shown in figure 2.2(a). On the other hand, oscillations in macrostates, which are the linear superpositions $\langle A(t) \rangle = \sum_i A_i p_i(t)$, do not require the underlying microstates to oscillate. The microstate probabilities only need to evolve such that $\langle A(t) \rangle$ oscillates. For example, consider a hypothetical chemical clock⁶ described by a cycle of states [22, 10]

$$s_1 \rightarrow \cdots \rightarrow s_N \rightarrow s_1. \quad (2.9)$$

If the clock advances each time the cycle is traversed, adding states to the cycle improves the quality of the clock, as shown in Fig 2.2(b). A more abstract implementation could be to consider a clock cycle and a fuel reservoir. We define that the clock consumes 1 unit of fuel during the transition from s_N back to s_1 , so the cycle is an open system. This accounting method, in effect, unrolls the cycle into a linear chain of microstates enumerated by the dyad $\{f, s_i\}$, where f is the amount of fuel remaining. When f is effectively infinite, the dynamics of the open cycle and the linear chain are equivalent. As the system moves from one state to another, we count time by keeping track of the evolution of the macrostate denoted $A(t)$. The macrostate of the clock is $\langle A(t) \rangle = \sum_{i,f} A_i p(f, s_i, t) = \sum_i A_i \langle p(s_i, t) \rangle$. Therefore, the quality of oscillations in

⁶The oscillating chemical clock is distinct from the traditional “clock reaction,” where an autocatalytic reaction causes a sudden one-time change in state on a distinct timescale, such as the iodate-bisulphite system [21].

the infinite linear chain is bounded by \mathcal{Q}_s , regardless of the precise amount of fuel. When $A_i = \{1 \text{ if } \text{mod}_2 i = 0; 0 \text{ otherwise}\}$, as shown in figure 2.3(a), the total number of states, N does not effect the quality of the clock. However, if we change A_i to

$$A_i^p = \{1 \text{ if } \text{mod}_p i = 0; 0 \text{ otherwise}\},$$

figure 2.3(b) shows increasing oscillation in $\langle A(t) \rangle$ with increasing p . That is, adding more states to the clock directly increases its accuracy.

Texts exploring chemical oscillations state that nonlinearity is a requirement for oscillations. In fact, nonlinearity is a shorthand for describing extremely large systems [13]. Under conditions of detailed balance, systems must consume some sort of fuel to sustain oscillations. If we consider the fuel-free states as being an abstract engine with N states, the combined engine-fuel system will be in one of the N different states and have f units of fuel remaining. Therefore, a fuel reservoir can allow a total number of oscillations $\approx fN/\pi$. eq. (2.8) implies that the number of inherently unique states, absent fuel consumption, will constrain the possible regularity of reciprocal motion [10]. Thus, the quality of oscillations appears to be bounded by the smallest irreducible cycle in the system, although this is not proven. That is, the topology of the network is inherently related to the ability of the system to sustain oscillations, and this will be explored in future work.

Similarly, if the system is driven by oscillations in multiple parameters or species, we can again parameterize the state of the system based on the population of each component. However, most chemical dynamics are modeled using continuous vari-

ables, not discrete numbers of states. The microscopic description of the system, comprised of a discrete number of states, is connected to the continuous mass-action approximation of chemical dynamics by a system size expansion described by Van Kampen [13]. Take, as an example, the multidimensional oscillating chemical reaction called the Brusselator. By expanding the mass-action Brusselator into a discrete state-space, the system size-expansion parameter determines the length of the largest cycle [23, 24]. Indeed, multiple authors have observed that the quality of the Brusselator limit cycle scales with system size, consistent with this work [23, 22, 25, 26, 27, 24]. This observation is not merely coincidence, but a fundamental efficiency limit of the master equation.

This efficiency limit of oscillations has obvious implications on how well a high-dimensional system can be numerically approximated by a smaller system. The approximation will only be successful if the relevant eigenvalues of the larger system lie within the allowed region of the smaller system. However, the inverse stochastic eigenvalue problem has not yet been solved, so we cannot know *a priori* if a stochastic matrix exists for a given set of eigenvalues, even if they all reside within the allowed region [28]. This fact prevents us from constructing the opposite bounds, the conditions for a *minimum* number of oscillations. Hopefully, future results will further constrain the present bounds, and we may gain deeper insight into the necessary conditions for creating oscillations.

The bounds on oscillations can play a key role in interpreting experimental observations by determining a minimum number of underlying states. For example,

the oscillation of fluorescence wavelength in fluorescent protein GFPmut2 remains unexplained [29, 30, 31]. After application of a denaturant, the ionic state of the fluorophore can switch up to $\mathcal{Q} \sim 50$ times with high regularity, observed as oscillations in the emission wavelength [29]. Because eq. (2.7) bounds the number of states involved in the oscillation to be at least 3 times larger than \mathcal{Q} , this predicts that the oscillations are driven by large-scale rearrangement of the numerous hydrogen bonds in the β -barrel, not merely exchange between the few amino acids directly connected to the fluorophore. If the protein were to be mutated to alter the number of bonds in the β -barrel, we predict that we should see a corresponding alteration in the number and quality of observed oscillations.

2.4 Oscillations in Green Fluorescent Protein GFPmut2

In a series of recent experiments, a mutant of Green Fluorescent Protein, GFPmut2, was encapsulated in silica gel and observed under denaturing conditions [32, 33, 31]. Ordinarily, when folded or even during unfolding, GFPmut2 is stable in the anionic green state, with stochastic transitions to the neutral blue state. At the very end of the denaturing process, just prior to complete fluorescence quenching, the fluorescence oscillates between green and blue [33]. This resonant oscillation is unique in fluorescence behavior and unobserved in single-protein dynamics except for the slow oscillations in activity of the ECTO-NOX protein [34, 35]. In addition to being

a fascinating window into denaturing dynamics, the observed GFPmut2 oscillations prompt the question of how a single molecule can be driven to autonomously oscillate.

Although the GFPmut2 oscillations are fascinating, they have not been fully explored experimentally. Although further experiments could cast new light onto this unique dynamics, the current body of work suggests these oscillations are autonomous, meaning that no laser or mechanical driving occurs. Somehow oscillations spontaneously emerge late in the denaturing process, and they persist far longer than the picosecond timescale of natural underdamped motion in protein bonds [36]. No other groups have reported independent observation of fluorescence observations from GFPmut2 as of early 2011, although we have tried, both at Caltech, and with the help of Jau Tang at Academia Sinica. The largest hurdle has been avoiding photobleaching. Additional experimental evidence will of course further inform the accuracy of the results below.

Because the oscillations take place on the millisecond timescale, and they do not begin for up to an hour after denaturing starts, MD simulation is impossible. Thus, there is no hope of brute force replication of the experiment *in silico*. Furthermore, the oscillations are only apparent on the single molecule level, so NMR cannot directly access the chain of events. We are left with an approach where modeling can suggest new experimental variables and observables to probe.

The timescale of oscillation in GFPmut2 is too long to be attributed to most normal processes associated with protein dynamics, including bond vibrations, torsional modes, and isolated residue rotation [37]. For example, in wildtype GFP, Agmon

has observed that the stochastic blinking in wildtype GFP is due to the rotation of Thr204, but it is characterized by a switching time of tens of nanoseconds [38]. Most protein dynamics on the millisecond timescale are characterized as two-step processes, indicating only a single degree of freedom dominates the folding process. However, Langevin dynamics indicate that self-sustained oscillations in a single degree of freedom in a protein would be impossible. In fact, a single degree of freedom cannot produce any oscillations at all without some external driving, as demonstrated in the section 2.2.

Given that the oscillations in GFPmut2 are not associated with any known periodic driving, there must be some coordinated interplay between ordinarily unobservable degrees of freedom. Even the nature of the experiment suggests this, because the fluctuation of any single hydrogen bond normally cannot be observed via the fluorescence of the molecule. The denaturants used, urea and guanidinium HCl, attack the barrel in slightly different ways [39, 40, 41], but the resulting oscillations are identical [42]. The fluorescence photophysics does not deviate from normal throughout the vast majority of the denaturing process, except for the moments before quenching. Because the fluorescence oscillates between anionic and neutral up to 20 times [33], there must be at least 120 separate internal states coordinating the oscillations. The only source of this many states within a single protein would be the hydrogen bonds of the β -barrel. Somehow, the denaturant sets off a cascade of hydrogen bond breaking and reforming that is observed in the experiment as the ionization state of the fluorophore. There is evidence that the oscillations are not due to rearrangements of

precise bond networks, because these networks are sensitive to salt concentrations. Protonation rates vary continuously with GdnHCl, but we do not see oscillation periods vary [43].

Here, we explore our hypothesis of hydrogen bond fluctuations by suggesting a two variable system where the ionic state of the chromophore alters the stability of the β -barrel. Without a direct crystal structure of GFPmut2, we cannot know if there are unique structural features in GFPmut2 to focus on as a starting point. However, we can draw analogies from other GFP mutants, such as S65T [44]. All mutants in the green fluorescent protein family fold into the distinctive β -barrel conformation, shown in figure 2.4. The β -barrel is held tightly closed by dense array of hydrogen bonds running up and down the sides of the barrel. This protects the chromophore, shown in figure 2.5, which is quenched by water.

By considering these facts, we attempt to synthesize a model to increase our understanding of this system.

2.4.1 Limit Cycles

Long term dynamics in closed systems are always driven to equilibrium. This stable state quenches any oscillations, leading to what Lord Kelvin profoundly called “Heat Death,” where no more free energy is available to sustain nonstochastic motion [50]. In the case of GFPmut2, we observe this as the fully denatured state where all fluorescence is quenched. Prior to oscillations, and toward the end of the series of oscillations, the ionization state of the fluorophore is stochastic. Only during a brief

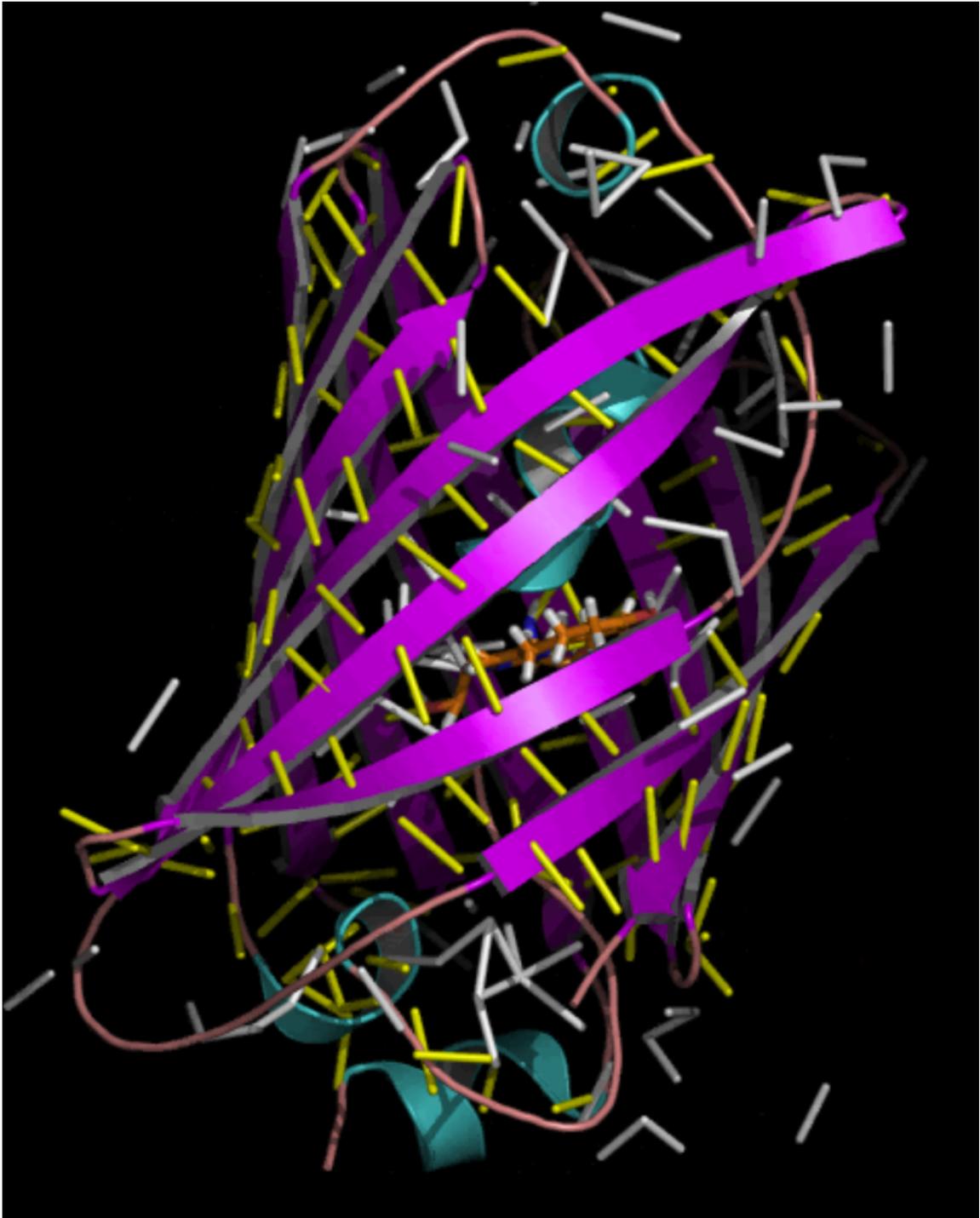


Figure 2.4: Ribbon representation of GFP, made using PyMol [45]. The β -barrel is colored purple. Unstructured regions are pink. α -helix loops are colored blue. White and yellow rods are hydrogen bonds between residues, which are not shown. The chromophore is orange, and can be seen edge-on. Original model was PDB entry 2HPW [46, 47, 48].

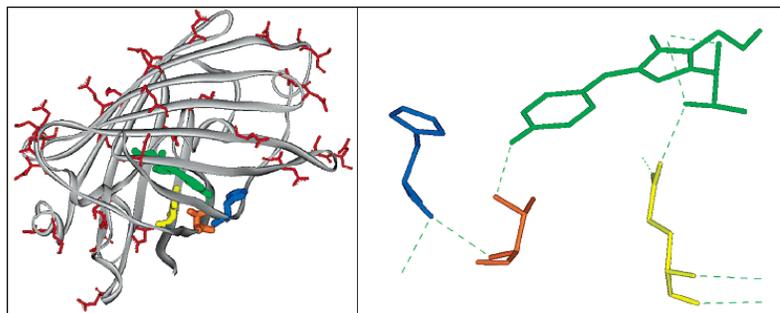


Figure 1. (Left) Three-dimensional structure of S65T GFP at pH 8 (PDB entry 1EMG). The chromophore is shown in green, His148, in blue, Thr203, in orange, and Glu222, in yellow. The carboxylic acids present on the protein are shown in red. (Right.) Closeup of the chromophore with the mutated residues.

Figure 2.5: The chromophore is in green, shown in the center of the β -barrel in left, and in its native H-bond arrangement in the right. Taken from [49].

sub-100 ms window does the ionization state oscillate somewhat deterministically.

A proton determining the ionic state of the chromophore is transported to and from the fluorescent ligand via proton channels determined by the orientation of amino acids in the binding pocket and the barrel. Proton channels transport protons on a picosecond timescale, similar to other breathing modes of amino acids. Therefore, these local vibrations and rotations average out over the course of an oscillation. Large scale rearrangements of the hydrogen bond network are much slower. These “proteinquakes” transpire over milliseconds, the same timescale as the oscillation dynamics [51, 52].

Matching timescales indicates that the oscillations could be dominated by two parameters, the ionization of the fluorophore and a proxy for the number of hydrogen bonds holding the protein together. Because the proton traverses proton-wires that open and close via rotational modes of amino acids along the wire [53], the proton’s position will be stochastic over the experimental timescale. We therefore wish to calculate the time-dependent probability of the proton being on the fluorophore (an-

ionic), $p_A(t)$, as a function of a chemical coordinate of hydrogen bonds D , which shall act as our proxy measure of the hydrogen bond network.

The internal charge distribution can determine the stable configuration of the hydrogen bond network, as was also seen in [53]. After exposure to denaturants for some time, a number of stabilizing hydrogen bonds break down, and denaturant begins to penetrate into the internal binding cavity [54]. After sufficient time, the entire β -barrel is broken from the inside-out, leading to a general two-state picture for denaturing. The precise chain of events during the unfolding process may depend on the employed denaturant [39], but the resulting oscillation dynamics do not appear to have any such dependency [31].

Because the observed persistent oscillations are so unusual, we first attempt to find conditions for totally self-sustained oscillations, i.e., a limit cycle. If one assumes that the internal charge distribution alters the local equilibrium of the hydrogen bond network, and that the internal charge distribution is reported through the ionic state of the fluorophore, the dynamical equation describing their interdependence must take the general form:

$$\dot{p}_A(t) = -f_1(D)p_A(t) + f_2(D)(1 - p_A(t)) \quad (2.10a)$$

$$\dot{D}(t) = g_1(D)p_A(t) + g_2(D)(1 - p_A(t)). \quad (2.10b)$$

Oscillations would be indicated by $p_A(t)$ swinging between 0 to 1 and back on a fixed timescale, representing near certainty of the fluorophore being neutral and anionic, respectively, and then returning. We see that Eq (2.10a) has the form of a

Fokker-Planck equation, while eq. (2.10b) describes the ensemble average dynamics of D . Because eq. (2.10a) is a master equation, detailed balance would suggest f_1 and f_2 to be

$$f_1(D) = k_1 e^{-D/d_0}$$

$$f_2(D) = k_2 e^{D/d_0}.$$

This functional form indicates that the parameter D is affine in the free energy of the protein. It is also consistent with experimentally observed reprotonation rates in GFP as a function of denaturant concentration [43]. When in the anionic state, the charged state favors D growing, consistent with denaturing occurring spontaneously. In the neutral state, small D becomes unfavorable. Hence, the free energy difference between the anionic and neutral states would be $\Delta G = \ln(k_1/k_2) - 2D/d_0$. Research by Saxena et al. has shown that the deprotonation process in native EGFP, a mutant similar to GFPmut2, is nearly barrierless (0.3 kcal/mol), while the reprotonation has an activation energy of 14.8 kcal/mol [43]. Thus, we posit that denaturation alters the free energy balance between the two states, ultimately determining the charge state of the chromophore.

We use eq. (2.10b) to describe the ensemble average of $D(t)$, making g_1 and g_2 the recruitment rate of newly denatured bonds in the anionic and neutral state, respectively. When the protein is well folded, denaturant attacks bonds and opens up the barrel, disrupting cooperative folding. If the state of the chromophore directly or indirectly determines the local equilibrium fold, denaturant may be squeezed out of

the β -barrel under certain conditions. We take

$$g_1(D) = k_3 D$$

$$g_2(D) = -k_4 \frac{D}{k_5 + D}.$$

That is, under the anionic state, hydrogen bonds decay as a first order process. Upon the transition from neutral to the anionic charge state, the fold is now no longer in equilibrium, so the protein undergoes a proteinquake [51]. In the neutral state, the equilibrium structure is one that catalyzes reforming of hydrogen bonds and restabilizing the barrel, leading to a Michaelis-Menten expression due to neighboring amino acids effectively catalyzing bonds with their neighbors. In the neutral state, amino acids no longer have the central anion competing with the other amino acids for preferential alignment due to the anion-induced dipole coupling, diagrammed in figure 2.6. In total, Eqs. (2.10a) and (2.10b) become

$$\dot{p}_A(t) = -k_1 e^{\frac{D}{d_0}} p_A(t) + k_2 e^{-\frac{D}{d_0}} (1 - p_A(t)) \quad (2.11a)$$

$$\dot{D}(t) = k_3 D p_A(t) - k_4 \frac{D}{k_5 + D} (1 - p_A(t)). \quad (2.11b)$$

With a model of this form, we successfully produce oscillations in both charge state. p_A . and structural stability, D . Tools such as Mathematica's EquationTrekker allow a rapid visual search of parameter space. figure 2.7 shows a typical set of phase-space trajectories produced from one combination of parameters, showing that sustained oscillations can be readily produced when the two charge states favor two

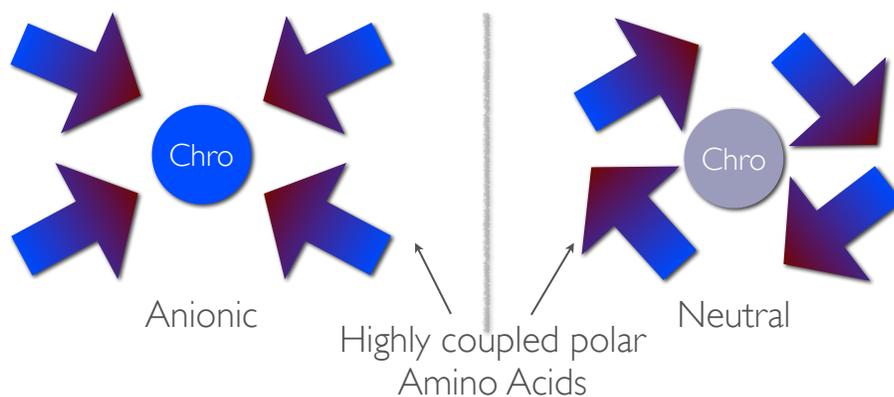


Figure 2.6: When the chromophore is anionic, left, its excess charge draws its tightly coupled neighbors into alignment with the chromophore. When neutral, the amino acids are now free to preferentially align with their neighbors, increasing the structural rigidity of the β -barrel.

different folds, as hypothesized. figure 2.7 shows the GFP oscillating between p_A between approximately 0.9 and 0.1, corresponding to a 90% probability of being in the anionic state when highly structured (D small). The anionic state allows denaturing, and p_A swings to 0.1 as D increases, ultimately reaching a 90% chance of being in the neutral state.

Because eq. (2.11) produce limit cycles, these equations suggest a mechanism for the dynamics. However, limit cycles that produces the desired oscillations cannot explain the complete dynamics of GFPmut2, because no system can oscillate indefinitely. The key assumption is that the energy of the protein spontaneously increases when the charge state flips, favoring a more robust (or delicate) fold. Although it would not be unusual for an internal degree of freedom to shift the energy landscape [52], a physically motivated description of the protein dynamics could not produce indefinite oscillations. Furthermore, GFPmut2 oscillations do not appear

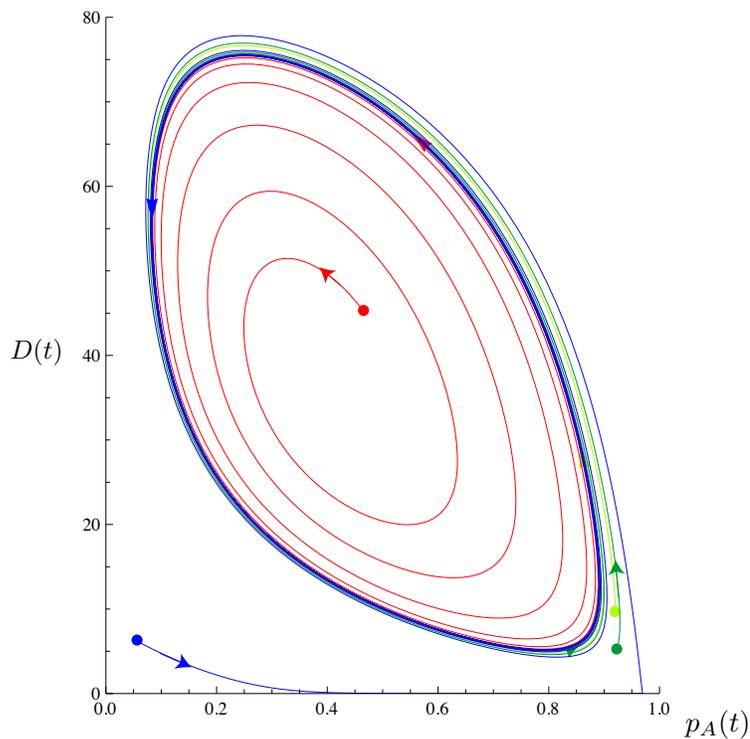


Figure 2.7: Given the parameters $d_0 \rightarrow 19.0, k_1 \rightarrow 0.12, k_2 \rightarrow 3.8, k_3 \rightarrow 4.8, k_4 \rightarrow 124.0, k_5 \rightarrow 1.6$, eq. (2.11) shows that the expected charge state of the chromophore can develop strong oscillations. These parameters are not physically determined, but indicate that the functional form of eq. (2.10) likely corresponds to some underlying dynamics.

the instant denaturant is applied. It can take an hour of denaturing before any oscillations are observed [30, 31]. Our model should be able to suggest an explanation for this as well as predict a finite resolution to the oscillations. To address this, we consider more closely the dynamics of our model system and how the limit cycle forms from eq. (2.11).

Consider the nullclines of the dynamics described by eq. (2.11), which are determined by fixing $\dot{p}_A = 0$ or $\dot{D} = 0$. Each condition will give a different curve, corresponding to the path in phase-space which satisfies each condition. Both curves

can be written as a function of $D(p_A)$:

$$D(p_A) = \frac{d_0}{2} \log \left(\frac{k_2}{k_1} \frac{1 - p_A}{p_A} \right) \quad \dot{p}_A = 0 \quad (2.12a)$$

$$D(p_A) = \frac{k_4}{k_3} \frac{1}{p_A} - \frac{k_4}{k_3} - k_5 \quad \dot{D} = 0. \quad (2.12b)$$

The mutual intersections of the nullclines determines the character of the phase-space dynamics, as shown in Fig 2.4.1. Altering any of the rates adjusts the position of the nullclines, such that they do not intersect (figure 2.8(a)), intersect once (figure 2.8(b)), or intersect twice (figure 2.8(c)). A limit cycle only emerges when these nullclines intersect at one point. In fact, altering only a single rate, k_3 , is sufficient to drive the system from a stable fixed point into a limit cycle by altering the position of the $\dot{D} = 0$ nullcline.

If the rate k_3 were time dependent such that k_3 started small at $t = 0$ and grew monotonically larger over time, the system would initially exhibit the experimentally observed nonoscillatory behavior until k_3 reached a critical value (which is dependent on all other parameters in a nontrivial way). The curves would first intersect twice, as shown in figure 2.8(c). The phase-space dynamics still display a single stable node, but the system entering the onset of instability. This is consistent with the experimentally observed steady increase in the blinking rate in the minutes leading up to the onset of GFPmut2 oscillations [30]. Once k_3 grows a bit large, the nullclines then intersect only once, causing the system to bifurcate into a limit-cycle, corresponding to the experimentally observed oscillations. As k_3 continues to increase, eq. (2.12b) shows the D nullcline monotonically shifts towards the $p_A = 0$ axis. The center of the limit

cycle shifts toward this axis, as well. As a result, the dynamics are shifted to favor the $p_A = 0$ state. As the center of the limit cycle drifts to $p_A = 0$, the probability of returning to the anionic state decreases, and the oscillations begin to look like chaotic blinking, as has been observed [33]. This entire series of events are played out in figure 2.9 by assuming an exponential time dependence for k_3 .

What could produce such a time dependence? The GFPmut2 chromophore is tightly coupled to a handful of neighboring amino acids, as shown in figure 2.5. These amino acids coordinate the local field and help determine the equilibrium ionization state of the chromophore, which has been found to have a pKa of 6.2 [55]. However, the amino acids most tightly coupled with the chromophore are also held in place by hydrogen bonds from their neighbors, and so on. As the bonds start to break up, fluctuations in the local pKa grow in proportion to the degree of stabilization. The loss of mutual stabilization increases local fluctuations in the β -barrel, and the amino acids take larger excursions from their local equilibria. This in turn allows the denaturant better access to further destabilize bonds, resulting in a denaturing rate that increases with time. Therefore, k_3 may have some dependence on additional degrees of freedom, approximated to first order by an exponential growing function of time.

2.4.2 Loop Dynamics

Although the limit-cycle model in the previous section established that the denaturing increases as the oscillations evolve, it used a continuous variable, $D(t)$, not the number

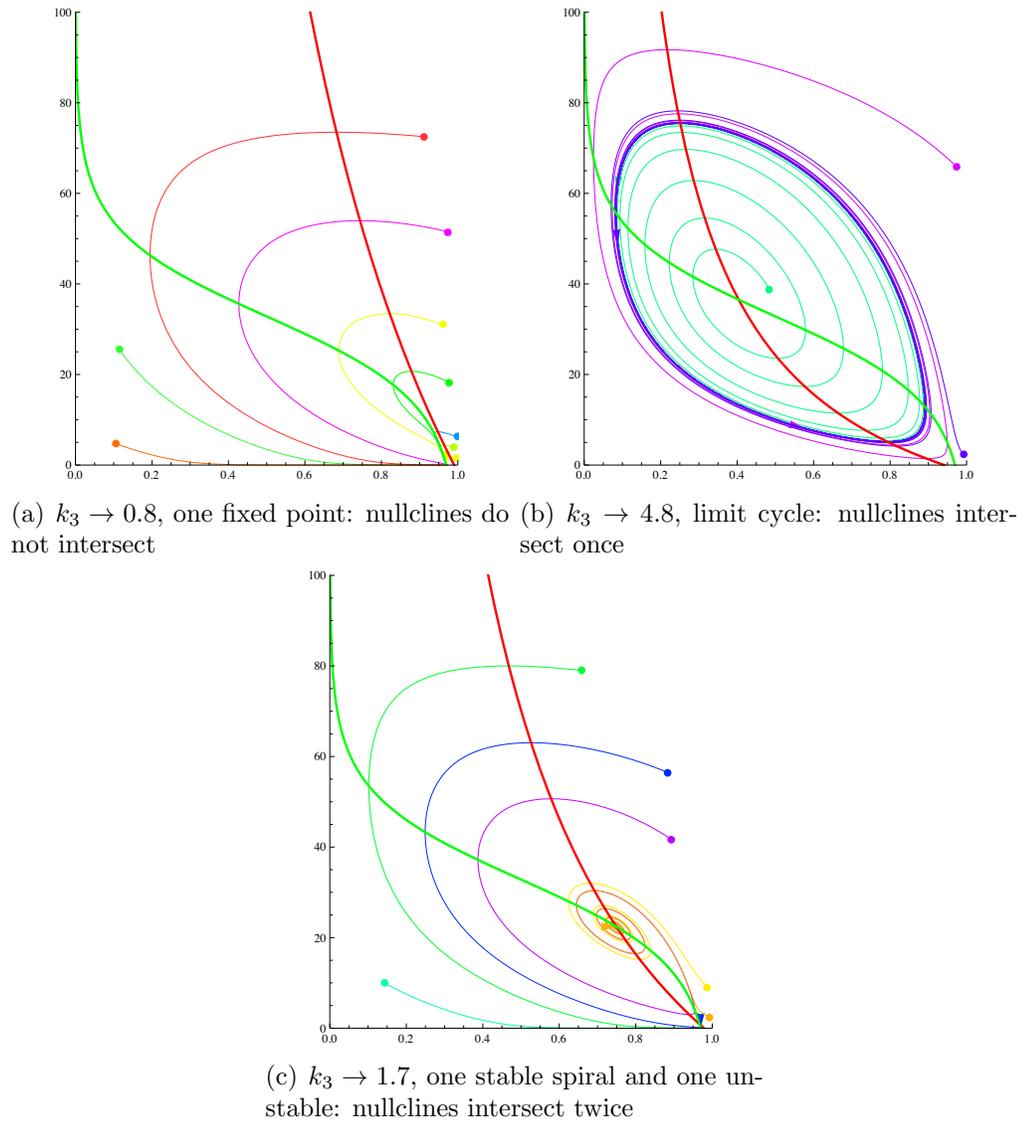


Figure 2.8: Green curve: $p_A(t) = 0$ nullcline; Red curve: $D(t) = 0$ nullcline. Above the green curve, $\dot{p}_A < 0$ and below $\dot{p}_A > 0$. Above the red curve, $\dot{D} > 0$, and below $\dot{D} < 0$. $d_0 \rightarrow 19.0, k_1 \rightarrow 0.12, k_2 \rightarrow 3.8, k_4 \rightarrow 124, k_5 \rightarrow 1.6$

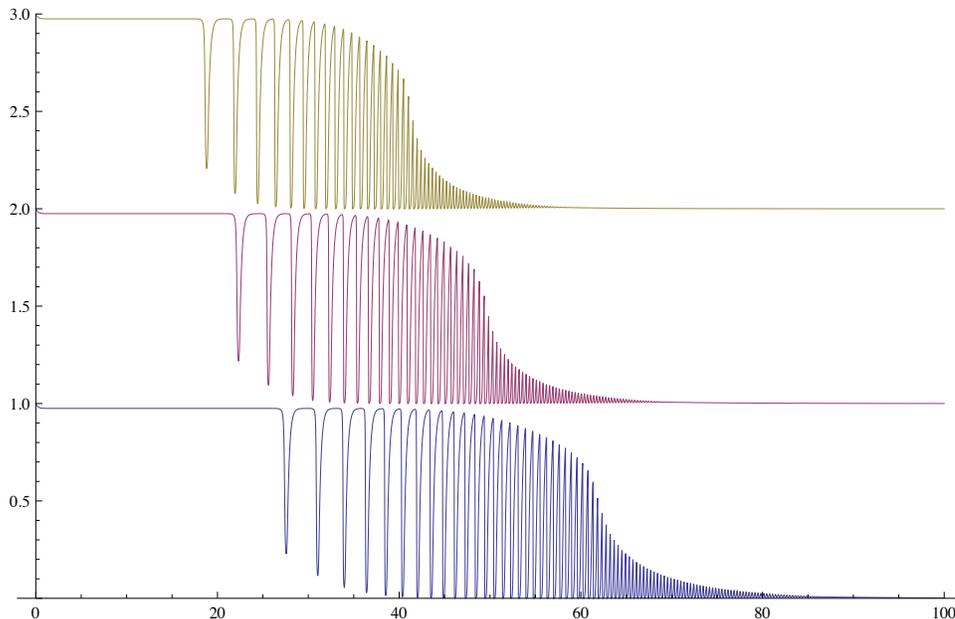


Figure 2.9: Time-dependent rates create a transition from a static system to a limit cycle. In this case, we transform $k_3 \rightarrow k_3 e^{k_6 t}$. k_6 is different between the three plots. top: $k_6 \rightarrow 0.12$, middle: $k_6 \rightarrow 0.10$, and bottom: $k_6 \rightarrow 0.08$. Each range runs from $p_A = [0, 1]$ versus time, and the curves are offset for clarity. $d_0 \rightarrow 5.29, k_1 \rightarrow 0.11, k_2 \rightarrow 4.29, k_3 \rightarrow 1.0, k_4 \rightarrow 177.18, k_5 \rightarrow 1.34$.

of hydrogen bonds. This allowed for better examination of the dynamics, but does not explain the how the bonds break. The speed of bond breakage determines rate of increase in $D(t)$, and thus, also determines the oscillation rate, because a faster change in D directly translates into a faster change in p_A , the observed fluorescent state. We also postulated that the central destabilizing force was a collective breakdown in the mutually stabilizing hydrogen bond network, leading to a time dependent k_3 . What determines the precise value of k_3 ? This may only be hypothesized, but a consideration of the GFP crystal structure shows the β -barrel to be remarkably uniform, with no obvious lines of weakness in the arrangement of hydrogen bonds. However, the β -barrel is constructed by looping the backbone to reach the final fold. Three loops are needed to properly fold GFP, as shown in figure 2.10. Loops are

entropically unfavorable [56], acting like springs pulling on the β -barrel [57, 58].

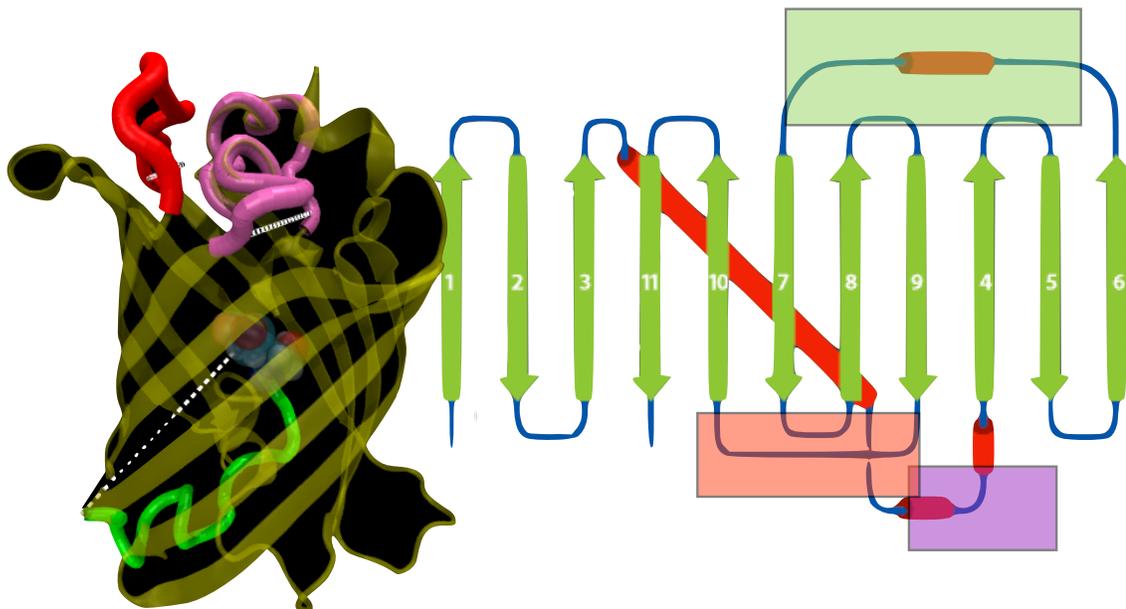


Figure 2.10: left: 3D representation of GFP, with backbone colored yellow. The three loops are colored red, purple, and green. Dotted white lines indicate the distance from the start to end of the loop. right: depiction of the secondary structure of GFP, with strands of the β -barrel numbered 1 – 11. The loops are the red cylinders, and colored red, purple, and green to correspond to the 3D image on the left.

Because of the large damping, the collisions between the loops and the bath molecules cause the loop to rapidly explore its configuration space, subject to the constraints of its end-to-end distance. This results in an entropic force being exerted on the amino acids at the ends of the loop. Loops are analogous to a spring under tension [59]. Of course, under native conditions the loops have little or no effect on the extremely stable β -barrel. Over the course of the denaturing process, the loops will eventually dominate the weakened hydrogen bond network and begin to pull apart the β -barrel.

Because of the denaturant, the α -helices in the loops will break down, and there will be very little structure in the loops. This allows the loop to be modeled as

a wormlike chain (WLC)[60, 61, 57], a model defined by the correlation relation $\langle \hat{r}(s) \cdot \hat{r}(s + \Delta s) \rangle = \exp(-\Delta s/p)$, where $\hat{r}(s)$ is the orientation vector of polymer a distance s from the end, and p is called the persistence length. This model works well for short loops of amino acids [62], and under the WLC model, the entropic force is a function of end-to-end distance is given by

$$F(R) = \frac{k_b T}{p} \left(\frac{1}{4} \left(1 - \frac{R}{N\ell} \right)^{-2} - \frac{1}{4} + \frac{R}{N\ell} \right), \quad (2.13)$$

where R is the end-to-end distance, p is the persistence length, N is the number of amino acids, and ℓ is the effective loop length contribution of each amino acid. Although the loops pulls continuously, the bonds break discretely and regularly. How does one map a continuous linear transduction into a periodic signal?

Consider a common baby’s toy, the bead maze, shown in Fig 2.11. Even if a train of beads are moved over the apex of the curve at a continuous rate, only one falls over at a time.⁷ If the bead train is moved at constant velocity, v , the time between beads moving over the apex is ℓ/v , where ℓ is the length of the bead. Hence, the audible “click” produced when the beads collide at the bottom has a period of ℓ/v seconds. Linear motion has been transformed into a periodic signal.

Analogously, the drag and damping by the water cause the loop to move without any inertia, leaving a Stokes drag condition, $F(R) = \eta v$, where η is the drag

⁷Because there is no coupling between beads, the bead that is balanced on the apex has an adjacent bead one bead-width, ℓ , lower on the wire track. The bead at the apex becomes unstable and slides over the top, but the adjacent bead must move another ℓ before it reaches the apex and slides over. This takes a time ℓ/v .

coefficient.⁸ Under denaturing conditions, the weakened β -barrel is able to be deformed by the loop, pulling the amino acid adjacent to the loop at a constant velocity $v = F(R)/\eta$, and the expected time for the amino acids to be pulled out of their hydrogen bonds with their neighbors is $\tau = F(R)/\eta\ell$. We observe three different loops in figure 2.10, with the following pairs of parameters, $(N, R) = (17, 2.1 \text{ nm}), (11, 1.7 \text{ nm}), \text{ and } (8, 1.7 \text{ nm})$. Based on force-induced denaturing experiments, we have $p = \ell = 0.39 \text{ nm}$ [62]. It is reasonable to assume that any one of the three loops may dominate the final unfolding during a single unfolding cycle, and the resulting oscillation dynamics will depend on that particular driving force, just as a single fracture dominates the initial failure of a brittle object. Experiments observe that GFPmut2 oscillates with one of three frequencies: 930, 720, or 440 s^{-1} . The same protein will oscillate at only one frequency during one unfolding cycle, but it may oscillate at a different frequency during subsequent cycles. Even though we do not know η , the resulting calculated ratios of oscillation rates, 1:0.72:0.41, are in good agreement compared to the experimentally observed ratios, 1:0.72:0.47 [30].

This agreement provides a key insight into the dynamics of GFPmut2 unfolding. If the preceding analysis holds, the unfolding dynamics appear to be primed by the destabilization of the loops in a systematic way. The loops deform the β -barrel, popping out hydrogen bonds one-by-one. As discussed in the previous section, the charge state of the chromophore helps to coordinate the surrounding amino acids in such a way as to loosen the β -barrel sufficiently to allow the loop dynamics to

⁸The drag coefficient η will be a function both of the drag due to moving through the water and the energy absorbed by distortions of the β -barrel, making a calculation of η beyond the scope of the present work.



Figure 2.11: The bead maze, a common children’s toy, is an example of linear motion producing a periodic signal. As baby Leo pushes the train of beads over the top of a loop, one bead at a time slides down the other side as it passes the apex. Although Leo moves the beads at a constant velocity, v , he hears a periodic click of frequency v/ℓ as each bead slides to the bottom, where ℓ is the length of the bead.

dominate, but the neutral state strengthens the barrel, neutralizing the effect of the loops. This analysis predicts the oscillation rates are determined by the loop lengths. If experiments are able to observe the oscillations in fluorescence, altering the loop lengths should change the ratio of oscillation rates in systematic way, confirming the present predictions.

2.5 Conclusion

We have examined how system size relates to the dynamic behavior of overdamped systems. Because most systems on the nanoscale operate at very low Reynolds number, they satisfy this condition. When those states are enumerable, systems with more states have the ability to oscillate longer and more coherently. This carries

implications not only for designing nanoscale systems, but also the results apply to any system that may be described as dynamics on a network, such as social networks. The quality bounds proved here are universal. Because the master equation is used in nearly every branch of science, the dynamics being modeled need not be physical. For example, it could be money held by a bank [8], packets of data on the internet [63], agents traversing a network [64], or the populations in an ecosystem [65]. The oscillation limit could also be probed experimentally with sculpted landscapes using optical tweezers [66]. As a probe bead jumps from trap to trap, the energy landscape in unoccupied traps is sculpted to simulate an arbitrarily large designer network of discrete states. The current results are a fresh approach to analyzing the dynamics of discrete systems, and it serves as a new design principle for those seeking to engineer oscillations.

This theorem also has direct application to explaining the curious dynamic of GFP-Pmut2 oscillations. Using the theorem, we were able to construct a model based on the rearrangement of hydrogen bonds in the β -barrel, as opposed to oscillations in the dynamic conditions around the protein or in the underdamped motion of protons. Using a two component model to track the probability of ionization of the chromophore and degree of denaturing of the β -barrel, we are able to analyze how the oscillations emerge and why they dissipate. Further phase-space analysis predicts a cooperative effect maintains the β -barrel. Once the barrel begins to fully destabilize, the denaturing occurs rapidly, and the barrel's ability to repair itself degrades until finally the protein is fully denatured. We hypothesize that the loops of the GFP have a central

role in determining the rate of unfolding due to their destabilizing nature, and an experiment which alters the length of these loops should see a corresponding change in the oscillation time.

This chapter has shown how linear state spaces may be mapped into nonlinear systems, allowing us to better understand oscillations on the nanoscale. Further work will be able to refine both parts of the work in this chapter. First, identifying motifs that give rise to oscillations or other chaotic behavior in reaction networks will uncover more applications for the theorem. Second, connecting the predicted consequences of GFP oscillations with further experimental observations may provide more evidence to refine our understanding of the oscillations. The next chapter will explore how optical nonlinearity, instead of nonlinearity in state-space, may be used to probe the geometry of proteins inside muscle using nonlinear optics, extracting more information than possible using traditional fluorescence techniques.