QIB: Connecting Feynman Path Integrals to Biophysical Problems

Alexander Pease
Advisor: Robert Austin, PhD.
apease@princeton.edu

January 6, 2011

Abstract

Examines Quantum-Inspired Biology by connecting Feynman's path-integral formulation of quantum mechanics to the use of Feynman-Kac and Fisher-Kolomogorov equations in biophysics. Attempts to show how fundamental concepts in quantum mechanics are carried over via mathematics and formalism to be applied nonliterally in biological contexts.

This paper represents my own work in accordance with University regulations.

1 Introduction

The intersection of quantum mechanics and biophysics has been a field of research for decades, yet results have been mixed. The historical focus of effort has been on the potential for quantum mechanics to give rise to quantum weirdness at the level of macroscopic biology. This has been in part motivated by the marked discrepancy between physicists deep understanding of classical and thermal physics and our limited understanding of complex biological systems. While theoretical breakthroughs have been presented, experimental verification has been lacking. This has led to attempts with the converse approach: the non-literal application of theoretical, mathematical, and formal constructs in quantum mechanics to studies in biology. This we shall call QIB: Quantum-Inspired Biology. Just as certain phenomena were never (classically) understood until the rise of quantum mechanics in the early 20th century, the use of quantum mechanics in biology may prove to be the missing link in our attempts to get at the heart of biological systems.

This paper examines the ramifications of one such application in detail: Feynman path integrals. Starting from Feynman's formulation of quantum mechanics, we shall show how underlying physical concepts have motivated further work, work that connects directly to biophysics. The purpose is two-fold: to show the natural connection between quantum path integrals and other mathematical tools used in biophysics, namely the Feynman-Kac and Fisher-Kolomogorov equations; and to argue that the concepts underpinning Feynman's formulation of quantum mechanics give insight into the underlying forces at work in biology. This nonliteral extension of concepts is QIB.

This paper is both a response to the recent influx of biophysics papers on the topic and a wish to clearly introduce and motivate further QIB to the scientific community at large. Our hope is that our paper presents concepts and mathematics at a level to motivate any budding young biophysicist (pun intended), as well as reveal the untapped potential of this interesting synthesis of fields to mature theoreticians.

2 Path Integral Formulation of Quantum Mechanics

At the heart of quantum mechanics is probability. In all formulations, it is rightfully recognized that quantum systems exhibit a wave-particle duality; that is, the system is indeterministic (to a certain degree between conjugate coordinates). The complex wave function $\Psi(\vec{x}, t)$ is the square-integrable probability amplitude of finding the particle at position $\vec{x}$ and time t. Traditionally, $\Psi(\vec{x}, t)$ is solved via the Schrödinger equation

$$i\hbar \frac{\partial \Psi(\vec{x}, t)}{\partial t} = H\Psi(\vec{x}, t) \tag{1}$$

and observable quantities are mathematically derived via operators. These operators also inform upon the formulation of the differential Hamiltonian operator H; this creates a direct link between the quantum system and "the classical Hamiltonian $H(\vec{p}, \vec{x})$ of the associated classical system" [1]. So, while exact quantum behavior is inherently indeterministic, the Schrödinger equation gives us a solvable, deterministic equation for the probable evolution of the quantum system, ie. for any knowledge of $\Psi(\vec{x}, t = t_0)$ we know $\Psi(\vec{x}, t)$ for $t > t_0$.

Published in 1948, Feynman's interpretation of the quantum world was divergent from the historically older formulations of Schrödinger and Heisenberg. Feynman showed that the wave function $\Psi(\vec{x}, t)$ is truly "a probability amplitude associated with an entire motion of a particle as a function of time, rather than simply with a position of the particle at a particular time" [1].

Suppose we know a particle travels from point a to point c with probability P_{ac}. As the wave function is a probability amplitude, there exists a wave function ϕ_{ac} such that

$$P_{ac} = |\phi_{ac}|^2 \tag{2}$$

As this particle must have passed through an intermediate point b in going from a to c, it must be that

$$\phi_{ac} = \sum_b \phi_{ab}\phi_{bc} \tag{3}$$

The potential values of b represent mutually exclusive possible intermediate points B, ie. the different paths the particle may travel. Therefore, the summation over b represents *all possible paths* the particle may take from a to c. However, when dealing with quantum systems, it is important to understand that (3) does not correspond to the intuitive sum of independent probabilities

$$P_{ac} = \sum_b P_{ab}P_{bc} \tag{4}$$

The above equation is only true for quantum systems that are actually measured at the point b; such a measurement collapses the wave function, and we are certain for which b the particle passed through. If no measurement is made, then we cannot assume "that to get from a to c the system had to go through a condition such that [the intermediate step] B had to have some definite value, b" [1]. Thus, equation (3) represents the fundamental wave nature of the particle in question: the particle may travel from a to c through *several different routes at once*. These distinct routes are represented by distinct b, and the total probability of getting from a to c via any b is a sum of their complex quantities.

Generalizing to a path from point a to z for which the particle must travel through infinitely many exact points of the sequence $a, b, c, d...z$

$$P_{abcd...z} = |\phi_{abcd...z}|^2 \tag{5}$$

This represents an exact path for the particle to travel through from a to z. As stated before, it is theoretically possible to determine the actual path that a particle takes from a to z by measuring the particle at every point along the way. This collapses the wave function, and we know the exact probability for a to b, b to c, etc. For a general path from a to z, in which we are unconcerned with the intermediate steps, we must sum over all possible points $b, c, d..z-1$ that the particle may pass through in between a and z, given as

$$P_{az} = |\sum_{b}\sum_{c}\cdots\sum_{z-1}\phi_{abcd...z}|^2 \tag{6}$$

Given this, we assume that the distance between points a and b, and b and c, etc. is of a fixed distance for which it takes the particle a fixed time ϵ to travel between. Taking the limit $\epsilon \to 0$ for the path a to z, the probability P^R_{az} becomes an integral over possible positions x_i defining R

$$\int_{a-\delta a}^{a+\delta a}\int_{b-\delta b}^{b+\delta b}\cdots\int_{z-\delta z}^{z+\delta z} P(x_a, x_b, \cdots x_z)\cdots \mathrm{d}x_a \mathrm{d}x_b \cdots \mathrm{d}x_z \tag{7}$$

$$= \int_R P(\cdots x_i, x_{i+1}, \cdots)\cdots \mathrm{d}x_i \mathrm{d}x_{i+1}\cdots \tag{8}$$

in which R is the region of space through which the particle may travel. More specifically, region R is such that the particle passes through it's first point x_a between $a-\delta a$ and $a+\delta a$, it's second point x_b between $b-\delta b$ and $b+\delta b$, etc. for which δi may vary. This is a more advanced notion of position; instead of saying that the particle is or is not exactly at point b an intermediate step, we say that it is within a range of space.

As in quantum mechanics the probability that a particle is found in region R is given by a complex number, we let $P^R_{az} = |\phi(R)|^2$ for

$$\phi(R) = \lim_{\epsilon\to 0}\int_R \Phi(\cdots x_i, x_{i+1}, \cdots)\cdots \mathrm{d}x_i \mathrm{d}x_{i+1}\cdots \tag{9}$$

The complex number $\Phi(\cdots x_i, x_{i+1}, \cdots)$ is "the probability amplitude functional of paths x(t)"[1], ie. simply the square-integrable complex value of the probability the particle travels through R. Thus, we may summarize that the probability that "a particle has a path lying in a region [R] of space-time...is the absolute square of a sum of complex contributions, one from each [possible] path in the region." [1]

For such a formulation to be useful requires a great conceptual leap, one that allows us to calculate the relative contributions of each possible path to the transition probability between two points. Feynman gives us the key postulate:

"Paths contribute equally in magnitude, but the phase of their contribution is the classical action...i.e., the time integral of the Lagrangian taken along the path." [1]

So, the contribution $\Phi[x(t)]$ to the transition amplitude is proportional to $\exp(i/\hbar)S[x(t)]$, for the action S over a possible path $x(t)$. While in classical mechanics it is the minimization of the action that completely determines the path taken, Feynman recognized that in quantum mechanics it is the minimization of the classical action that determines the probability that a path is taken. The genius behind this is the recognition that the minimization of the action gives "how much" a particular quantum path varies from the true, classical path. The action may then be substituted for each path $\Phi[x(t)]$ as the probability of a quantum path taken is proportional to how deviant it is from the classical limit. So, Feynman gives us the mechanics to relate the energies of a path to the probability that this path occurs. To minimize the action between two successive points in space-time

$$S(x_{i+1}, t_{i+1}; x_i, t_i) = \text{Min.} \int_{t_i}^{t_{i+1}} L(\dot{x}(t), x(t))\mathrm{d}t \tag{10}$$

for an entire path

$$S = \sum_i S(x_{i+1}, t_{i+1}; x_i, t_i) \tag{11}$$

in which $x(t_k) = x_k$.

The phase and amplitude of the action are substituted in for the contribution of path $\Phi[x(t)]$. This gives us the final equation for a Feynman path-integral:

$$\phi(R) = \lim_{\epsilon \to 0} \int_R \exp[\frac{i}{\hbar} \sum_i S(x_{i+1}, x_i)] \cdots \frac{\mathrm{d}x_{i+1}}{A_{i+1}} \frac{\mathrm{d}x_i}{A_i} \cdots \tag{12}$$

$$= \int_R D_A[x(t)] \cdot e^{\frac{i}{\hbar} \sum_R [\text{Min.} \int_t L(\dot{x}(t), x(t))\mathrm{d}t]} \cdot \mathrm{d}x \tag{13}$$

in which we have calculated the probability that a quantum particle travels in a path through a region R in space-time.[1] The encompassing integral is taken over all possible paths in R. The sum of actions is for each one of these paths. The integral of the Lagrangian is for the entire time for the particle path to be completed.

For a path through region R, our mathematics utilizes all possible paths through the region. If we remove the restriction of the path being in R, we are asking for the probability that the quantum particle travels between two points. Such a question requires, in this formulation, a knowledge of *all possible paths through all of space-time*, and not just all paths through R. Inherent in such a conception is the quantum phenomenon that there is a probability, however

[1] $D_A[x(t)]$ is the normalization of each path. I shall not discuss how it is obtained, as it does not lend any conceptual relevance to my discussion of the above equation to biophysics.

small, that a particle may take any path through space-time. The main integral in Eq. (12) is then for all space. Because of Feynman's connection between the energies of the system and it's path, such a calculation is not as formidable as it sounds. Use of the Lagrangian makes the calculation of the weight of any path possible. It is also easy to make approximations by only taking the integral over the "most-classical" paths, ie. those paths that will contribute greatest to the transition amplitude.

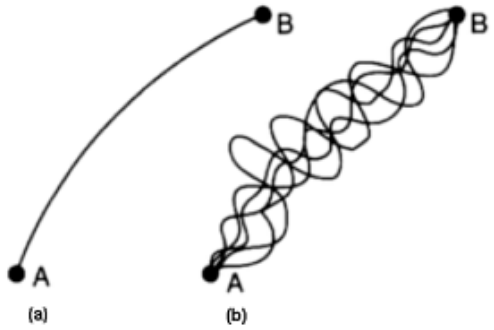

Figure 1: Two ways of physical transition from A to B. (a) gives the classical trajectory, which minimizes the action. (b) shows the quantum case, in which all possible routes from A to B are viable, and for which each route contributes to the total transition probability. The routes with the most weighted contributions are those which most minimize the action, ie. are closest to (a). Figure from [2], p.74.

2.1 Collapsing the Wave Function

To generalize such a path-integral formation, it is worth noting that there are inherently quantum concepts at work. The wave-particle duality of systems at the quantum level means that multiple paths may be taken simultaneously. To generalize to larger scale systems, we have been hard pressed to find any that exhibit this kind of "quantum weirdness."

However, the key concept behind path integrals, that each possible path contributes a probability, is still valid. Quantum systems are not the only inherently stochastic systems. Instead, if we apply Feynman's path integral conception directly to non-quantum systems, we need simply treat it as a quantum system that has been measured. This ensures that we may use the mathematical apparatus of assigning probabilistic values to all possible paths, but that the system is always "forced" to choose one particular path. Which path it takes is a function of the inherent probabilities.

3 Action

3.1 Examining Quantum Action

To generalize from quantum systems, we must take a closer look at the concept of action. In classical mechanics, action is an absolute concept; it always determines the exact path that a system takes. In quantum mechanics, action is still exact. There is a minimum to the action, a stationary point, and this is determined in the classical limit. The quantum system *will*, in the classical limit, take the path of least action.

But Feynman begins to hint at the possibility of stochasticity in action. For a wave function that is left undisturbed, the quantum system will travel via multiple pathways simultaneously. The standard interpretation must then be that the action for a given path gives "how much" the quantum system traveled down that particle path.

However, if we do collapse the wave function, as discussed above, the particle takes a definite path. And this path *is not always the classical path.* In this scenario, then, the action functional gives a probabilistic value of a disturbed quantum system taking the given path.

This can be intuitively applied nonliterally to other systems. If we are to examine the dynamical behavior of a stochastic system, we cannot definitely say how it will evolve. If, however, we have action values for its possible pathways, we can calculate the probabilities of the system taking certain paths. Taken nonliterally, we may use path integrals as a mathematical tool if we consider action as a stochastic concept for certain systems. The question then becomes, what can be thought of as an action?

3.2 What is an Action...

Looking back in history, before the formulation of Hamiltonian mechanics physicists did not yet understand how the dynamical behavior of a classical system is completely determined by the energies of that system. Systems have always held this information, but it took the work of a few brilliant 19th century physicists to realize the connection. Feynman made the next leap 150 years later, utilizing the energies of a quantum system to show its probabilistic behavior in a similarly path-based approach. This field has made extensive use of path integrals because, for the first time, the energies of a quantum system were useful in determining the evolution of the system's path through both time and space.

Action is an attribute of the dynamics of a physical system, a mathematical functional that examines trajectories in terms of energy and time. It is a fundamental view of a path in terms of the energies of the system. Taken another way, it is a tool that allows us to determine the evolution of a system as a function of the systems energies. It is this second idea that will be most useful to us, and

that we will use in nonliteral contexts. First, we review action in the original, mechanical context.

To fully understand the power of action, let's start from first principles. If we imagine a classical system, such as a ball being thrown through the air, it is intuitive that this ball must follow a certain trajectory. It will arc, and eventually hit the ground; there is no chance that it will start spinning figure-eights around your head. How the ball arcs depends on how you throw it. If you throw it harder, you give it more energy, and the ball will fly farther.

For a classical system, the energies of the system completely determine the systems dynamical behavior. Put another way, a knowledge of the relation of a systems energies implies complete information of the evolution of the system. This mathematical relation of energies is the Lagrangian. For a simple system, this is the kinetic and potential energies,

$$L = T - V. \tag{14}$$

To understand the dynamical behavior, we need to examine this relation of energies over the course of an entire trajectory. The action is the required functional that integrates the Lagrangian over the course of the trajectory,

$$S = \int_{t_{initial}}^{t_{final}} L(q, \dot{q}, t)\mathrm{d}t \tag{15}$$

where we have assumed that the energies of most systems only rely on their spatial coordinate (potential term) and its first derivative (kinetic term). [2]

Thinking back to trajectories, it is intuitive that the "correct" trajectory will have a different value of action from the hypothetical, figure-eights-around-your-head trajectory. The principle of stationary action gives this idea mathematical utility: the exact dynamics of a physical system correspond to a stationary point of the system's action. More simply, a classical system takes the path of "least resistance." [3] Using variational calculus, we can satisfy the principle of stationary action such that $\delta S = 0$ for deviations from the system's true path. The perturbation of a system from this true path can be written as $\delta q(t) = q_{perturbation}(t) - q_{true}(t)$, in which the endpoints are always the same, ie. $\delta q(t_{final}) = \delta q(t_{initial}) = 0$. We use (15) to examine these perturbations:

$$\delta S = \int_{t_{initial}}^{t_{final}} [L(q_{true} + \delta q, \dot{q}_{true} + \delta\dot{q}_{true}) - L(q_{true}, \dot{q}_{true}, t)]\mathrm{d}t \tag{16}$$

$$= \int_{t_{initial}}^{t_{final}} \left(\delta q \frac{\partial L}{\partial q} + \delta\dot{q}\frac{\partial L}{\partial \dot{q}}\right)\mathrm{d}t = \int_{t_{initial}}^{t_{final}} \left(\delta q \frac{\partial L}{\partial q} - \delta q \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\right)\mathrm{d}t \tag{17}$$

[2] This is potentially troubling, but it has been found that a vast majority of systems satisfy such a requirement

[3] This has nothing to do with real resistance. It is merely a phrase to try and coax an intuitive understanding of some paths being "easier" than others for a system to follow.

in which

$$\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = 0 \tag{18}$$

must be true to minimize the action $\delta S = 0$. These are the Euler-Lagrange equations (18) which give the exact equations of motion for the system. This is the Hamiltonian formulation of classical mechanics.

3.3 ...and What Can Be an Action?

This is action in the original sense, but we shall argue that it can be expanded for non-mechanical applications. Again, the fundamental question is how can action be generalized? This will allow a link between Feynman and biophysical problems.

For there to be an action, a system must be evolving along a dimension X. The action is then an attribute of the dynamics in X. These dynamics must be the direct result of some driving force f, and so the action is a functional of f over a path in X. The action must also be descriptive over the entire trajectory. For our mechanical system, we examined trajectories through space-time X as being determined by the energies f of the system.

A nonliteral example can be found in the gene toggle switch network presented by Wang, et. al. [3]. The state of the lambda phage CI/Cro switch is a function of the chemical kinetics of the production of these two proteins, as well as inherent noise from cellular fluctuations. Wang argues this cellular network has an action in terms of these protein concentrations and the diffusion of cellular fluctuations, and continues to analyze cellular dynamics utilizing path integrals.

As shall be discussed in more detail in section 5, Maeshiro and Kimura argue that in the evolutionary process, the action is the fitness between genetic sequences in adjacent generations. The Lagrangian is the product of the fitness of a genetic sequence and the probability of attaining this sequence. This makes up the driving force f over genetic evolution X. Taken over the trajectory of multiple generations, the action is in terms of $time \cdot fitness$. These authors are both using action in a stochastic sense. Their dynamical systems are probabilistic, their actions give the probabilities of evolution through a particular dimension.

The weighting of paths is also a function of the action. In Feynman, it is the phase of the action that gives the probability of a quantum trajectory. For certain non-mechanical systems (in which conservation of energy is not applicable), it may be that the calculation of the action itself is the relevant weighting, perhaps with a normalization over all paths in question. Or it may be that the action gives a value that is to be further attenuated to give the probability of following the trajectory in question.

This is also a function of whether the most weighted paths have maximal or minimal action. It shall be clarified that the principle of stationary action is often referred to as the principle of least action, yet this is a misnomer. The path

of least resistance has an action with a first order variation of zero; it can be a maximum or minimum. Depending on how we define the system's action, the system may either tend to follow paths of maximum action or to follow paths of minimal action. In section 4, the action gives a paths fitness, and it is intuitive that those paths with greater fitness are more likely to succeed.

3.4 Action vs. Lagrangian

Why have we focused on a systems action, when the more fundamental quantity is the Lagrangian? First, our argument for non-literal path integrals is (we hope) focused on path integration. Action, as a concept, is more directly linked to paths. Each path has a value of action; it makes no sense to speak of the action of a system without relation to a possible trajectory of that system. Second, as we make the argument to generalize path integration, it may be necessary to approximate the action of certain possible paths. While the Lagrangian connotes a complete knowledge of the energies of the system, the action only refers to a particular path. We feel that it's more kosher to say that we have approximated the action rather than the more fundamental Lagrangian. That is not to say a knowledge of the Lagrangian is unwanted; a knowledge of the Lagrangian is invaluable.

4 Feynman-Kac

4.1 Potential Landscapes

The beauty of path integrals is that it allows us to conceive of a system literally evolving through, say, space-time. The system takes a path through a landscape.

This landscape is the potential function, and it can be the driving force of dynamical behavior. If we consider simple quantum systems, the action is a function of both kinetic and potential energies; but the kinetic energy of a particle is generally simple, and similar across many systems. What differentiates problems in quantum mechanics is when systems must navigate different forms of potential functions. This is literally the landscape that paths are taken through.

The concept that potential functions make up the landscape of path-space is central to Mark Kac's generalization of Feynman's work to diffusion problems. In this case, the potential is easily broadened in conception. As with the previous examples, the potential can be the selective pressure of evolution, cellular fluctuations, or any other landscape through which a biophysical problem must evolve. Essentially all biological systems must navigate a complex environment; by acknowledging that this environment is the driving potential of the systems dynamical behavior, we may use path-integral based tools.

4.2 The Feynman-Kac Equation

As with all great ideas, Feynman's revolutionary paper found a life outside of its original realm. A colleague at Cornell University, Mark Kac, was one of the first to build off of Feynman's work. Even before it was published in Reviews of Modern Physics, Kac was able to see the genesis of a path integral formulation of quantum mechanics in Feynman's then-unpublished Princeton doctoral thesis. His main interest being probability theory, Kac quickly grasped the innovative approach to solving the Schrodinger equation and realized the same concept could be applied to other stochastic processes. It is historically noted that Kac's inspiration came from the heat equation, with an external cooling term V:

$$\frac{\partial u}{\partial t} = \frac{1}{2}\nabla^2 u - V(x)u, \qquad u(x,t) = u_0(x) \tag{19}$$

It is easy to see the similarities to the Schrödinger equation (1), for which both equations contain a time derivative and a second-order derivative in space.

Just a year later, Kac published his proof of what would become known as the Feynman-Kac formula. The partial differential equation (19) admits a solution given by:

$$u(x,t) = Expect.\left\{u_0 e^{-\int_0^t V(x)\mathrm{d}s}\right\} \tag{20}$$

This paper, "On the Distributions of Certain Wiener Functionals" [4], ties path integral formulations to the solution of partial differential equations involving diffusion processes. As Feynman's proof went from the assumed discrete probability amplitudes (15) to an integral over continuous paths, Kac made the mathematical leap from Wiener processes to PDEs.

As the rigorous mathematics in [4] are beyond the scope of this paper, and of less interest to a physicist, I shall attempt to elucidate the key concepts and power behind it's formulation.

4.3 Primacy of the Potential

A standard diffusion equation,

$$\frac{\partial u}{\partial t} = D\nabla^2 u \tag{21}$$

can be solved quite simply by separation of variables. But most systems have a driving force, the potential V,

$$\frac{\partial u}{\partial t} = D\nabla^2 u + V \tag{22}$$

It is this term that differentiates a standard diffusion process from anything "more interesting."

The formulation of Feynman-Kac is based on an understanding of this potential. The landscape of the potential directly influences what paths the system may take; intuitively, the system is less likely to take a path through a part of the potential surface in which the potential is high. Feynman-Kac measures take this potential landscape as the driving force of how the system will evolve. By exponentially weighting the paths based on the potentials they travel through, we come to an expectation of the systems behavior. In this sense, "the potential is rather regarded as a likelihood function of the states with respect" [5] to the path space.

The potential function is then an energy function, or the action, for the paths. This is a similar setup to Feynman's path integrals (12). The Feynman-Kac formula thus ties two worlds together through an understanding of potential landscapes. Regardless of the system in question, it is the potential function that acts as the action, the driving force for dynamical behavior.

Feynman-Kac is therefore widely applicable. It may be used as a path-based approach to solve the Schrödinger equation (1), or any other diffusion equation of a similar nature. These types of equations are widely found in biology, such as for allelic diffusion, chemical interactions, or directed polymers. It is amazing that such varied phenomena can be approached with the same tool. While it is stretch to say that, for instance, chemical diffusion across a cell membrane exhibits any quantum-like behavior, it is true that dynamically their behavior exhibits remarkable similarities.

4.4 Fisher-Kolomogorov

The Fisher-Kolomogorov equation was developed "as a deterministic version of a stochastic model for the spatial spread of a favored gene in a population" [6]. It is perhaps the best example of a PDE used in biology that admits traveling wave solutions. For a habitat spatially distributed in dimension x, if a mutant arises with a selective advantage, it is expected that this mutation will diffuse through the population over time, at the expense of other alleles at the same locus. Fisher found that the frequency of the advantageous gene satisfies the following equation of allelic diffusion

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + u(1-u) \tag{23}$$

While this equation predates quantum mechanics, and therefore the Feynman-Kac equation as well, it is remarkably similar. We can see that the above equation (23) satisfies the form of Feynman-Kac PDEs (20). Despite the fact that allelic diffusion is a stochastic process, the Fisher-Kolomogorov equation is decidedly deterministic. The selective pressure $u(1-u)$ is not a potential landscape, and therefore does not necessitate a path integral approach. It is also a not a linear problem, complicating the standard requirement for the Lagrangian, or the

potential, or a path based approach to be first-order.

Still, it is worth bringing to our attention. Kac's goal was to connect Feynman's multi-path approach to diffusion equations in a general form. Diffusion is a multi-path problem itself; it asks, given a complex environment, how is the system expected to evolve? Kac's solution is to treat the multi-path approach similarly to the principle of stationary action used by Feynman. The Fisher-Kolomogorov equation represents the limit in which the potential function for allelic diffusion is not a function of outside pressure, but rather is fully tied to the spread of the allele itself. Even in this case, we see the Fisher-Kolomogorov PDE is remarkably close to those diffusion problems that Feynman-Kac is outfitted for. In a more complex case, a potential landscape could be added to Fisher-Kolomogorov, and the expectation of paths taken over the landscape could be solved via Feynman-Kac. This would give Fisher-Kolomogorov gene dynamics a more advanced treatment.

5 Mathematical Analysis of Evolutionary Process

The above discussion has been quite limited to mathematical tools, but it may help to present some examples of biological problems for which this paper has been relevant. We will examine a paper on population genetics that attempts to use path integrals directly, without the medium of Feynman-Kac. While perhaps not the most advanced approach, this makes clear exactly how the underlying concepts in path integrals are nonliterally applied to a biological problem.

Population genetics is the study of how a population's phenotypic proportions change over time. It has been a powerful theory, giving important insights into the evolution of life. Yet as genetics-based biology has exploded, there has been increased interest in dealing with the evolution of genotypes in population dynamics. Such an approach treats the actual genetic information, the DNA, as the building block of evolutionary progress. This could allow geneticists to better understand *how*, at the molecular level, the replication and propagation of DNA leads to the evolving fitness of organisms.

By looking at the string of nucleic acids directly, we can overcome the complication of mutation "distance" between amino acids. The 64 combinations of three nucleotides make up any of 20 amino acids, yet there is not a direct correlation in between mutating individual nucleotides, and getting different amino acids. For instance, one, two, or all three of arginine's nucleotides may mutate, and the result be histidine in each case.

Maeshiro & Kimura [7] make an interesting assumption in their study of this field: that there is a "principle of least action in evolutionary mechanism at genetic sequence level"[sic] [7]. By this assumption, the evolutionary process follows the *path of greatest fitness*. This is not purely genetic fitness, but rather the concept of genetic fitness within a potential landscape of selective pressures. The survival of a species is determined by both mutation and selection, which is a reality of

the species' natural habitat. If we ignore selection, then calculation is simply finding the optimal path of genetic mutation. With selection, it may be that the path of evolution is suboptimal; it may be that the realities of the fitness landscape, coupled with the possibility that the most evolutionary advantageous paths are reached with a low genetic probability, gives higher probability to a more moderate evolution.

An evolutionary action is then a function of the fitness of the species' evolution over successive generations. Maeshiro & Kimura thus define the evolutionary Lagrangian as the fitness in evolving to the next generation, a function of mutation rate and fitness landscape. A path integral formulation gives the probability of reaching "a genetic sequence after T replications in a given environment" [7]. This approach is well-founded because of their recognition that species evolve in a selective landscape. This landscape is the path-space for all genetic mutations.

Take a haploid, asexual population in which we only consider point mutations. As individual cells replicate independently, we assume that the population starts either with a single cell, or with a culture of identical cells. Then we need only consider the possible paths of a single organism. The evolution of the population is then easily generalizable.

Consider that our organism is DNA-based, with a genetic sequence m base pairs long. This sequence can be represented as a one-dimensional vector $\vec{x} = \{x_1, x_2, \cdots, x_m\}$, in which the pairing between DNA strands is assumed to be exact. Each point in the sequence can be one of four possible nucleic acids, denoted as $n_e = 4$, thus the number of possible genetic sequences is $n_e^m = 4^m$. (insert possible m? look up) For the replication of each base there is a constant probability μ of mutation.[4] The number of mutations per replication is n_μ, with an expected value of $\mu \cdot m$. In determining how our population evolves, the idea is to determine and assign probabilities to all paths between a starting genetic sequence $\vec{x}_a$ and a hypothetical future sequence $\vec{x}_b$ over a fixed number of generations T.[5] This is a function of the number of point mutations ν that must occur between adjacent generations $\vec{x}_a$ and $\vec{x}_b$,

$$\nu_{ij} = m - \sum_{k=1}^{m} \delta(x_i^k - x_j^k). \tag{24}$$

The probability that mutations p^{mut} cause an initial sequence $\vec{x}_j$ to generate $\vec{x}_i$ in the next generation is the product of independent probabilities that ν_{ij} mutations do occur and that $m - \nu_{ij}$ mutations do not occur, divided by the probability that each mutation is to the correct nucleic acid,

[4] We therefore for assume that a mutation changes both bases, such that they remain perfectly matched

[5] The question is rarely "given infinite time, will the organism evolve to $\vec{x}_b$. Rather, "will the organism evolve fast enough to be fit enough to survive?"

$$p_{ij}^{mut} = \mu^{\nu_{ij}}(1-\mu)^{m-\nu_{ij}}\frac{1}{(n_e-1)^{\nu_{ij}}}. \tag{25}$$

By the reasonable assumption that the sequence is long, $m >> 1$, and the mutation rate small, $\mu << 1$, which can be approximated to

$$p_{ij}^{mut} \to \frac{e^{-\mu m}(\mu m)^{\nu_{ij}}}{\nu_{ij}! C_m^{\nu_{ij}}(n_e-1)^{\nu_{ij}}}, \qquad C_m^{\nu_{ij}} = \frac{m!}{\nu_{ij}!(m-\nu_{ij})!} \tag{26}$$

Taking this as our probability that sequence $\vec{x}_j$ evolves to $\vec{x}_i$ in the next generation,

The other side of our calculation is natural selection, ie. the probability p^{sel} that our haploid organism will survive due to the natural environment and the relative fitness of competing species. The genetic fitness $W(\vec{x}_i)$ of a given organism is purely a function of it's DNA $\vec{x}_i$. It's relative fitness is influenced by the mean fitness of the entire population N of it's habitat $\bar{w}$,

$$\bar{w} = \sum_k^N W(\vec{x}_k)q_k \tag{27}$$

The probability that q_i organisms of sequence $\vec{x}_i$ is given by,

$$p^{sel} = \frac{W(\vec{x}_i)q_i}{\bar{w}} \tag{28}$$

p^{sel} requires a knowledge, or at least an approximation, of the fitness landscape over multiple generations. For an isolated petri dish, this is of little concern. For a tree of life, this becomes a formidable task, perhaps currently prohibitive. As our collective database of genetic information continues to grow, this will hopefully become more realistic.

The action is becomes the sum [6] of the Lagrangian over the intervening generations. If we are interested in the probability of our asexual haploid evolving from $\vec{x}_a$ to $\vec{x}_b$,

$$S = \sum_{t_a}^{t_b} \alpha \cdot p_{i,j}W \tag{29}$$

This is the action for a single path; the "total probability of the change from $\vec{x}_a$ to $\vec{x}_b$ is the sum of contributions over all possible paths" [7], ie. the discrete path integral,

$$P(b,a) = \sum_{allpaths} \phi[\vec{x}(t)] \tag{30}$$

for each possible path normalized by A.

[6] Not an integral, because generations are discrete.

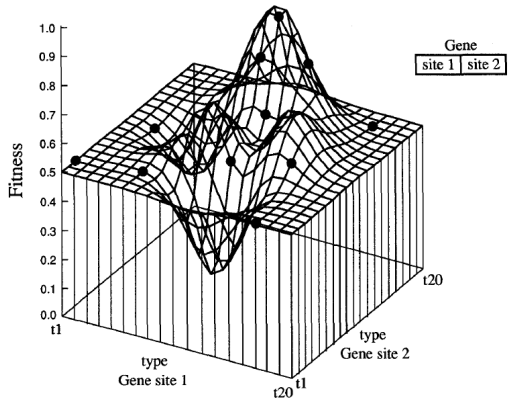

Figure 2: A fitness potential landscape for a genetic sequence of m=2. Assume that each of the two genes has 20 possible alleles. The genetic code "walks" along the surface, finding the optimal genetic code for the given fitness landscape. Figure from [7].

5.1 Critique

Maeshiro & Kimura continue to work out the "isolated limit," as we shall call it, for which

$$p_{ij} = p_{ij}^{mut} \cdot p_{ij}^{sel} \rightarrow p_{ij} = p_{ij}^{mut} \tag{31}$$

This is simple to solve, but not particularly useful for our investigation.

Our interest in the problem stems from how selective forces influence evolution. In this case, the key idea for Maeshiro & Kimura is that genetic paths may "walk" along a potential landscape of selective pressures, as in Fig. 2. It is this potential landscape that will determine the strongest genetic paths as a combination of selection and mutation. This concept is similar in spirit to the potential landscapes of Feynman and Feynman-Kac.

Maeshiro & Kimura also exhibit some of the ideas of Section 3. They have given a biophysical problem a non-mechanical, probabilistic action. This action is a function they have determined from first principles, from answering the question of what factors influence the selection of a genetic code. It is this kind of leap that must necessarily be taken for path integrals to be used in a non-literal context. We are accustomed to the idea of mechanical action, a Lagrangian relating energies. But for an abstract system, as this genetic system over fitness space is, it is up to the investigating authors to determine what determines the viability of paths. In this sense, the mathematics of this paper have presented some strong tools for diffusion problems in biology. The difficulty for some lines of inquiry may then be more conceptual, in seeing how certain systems abstractly fit into a path-integral formulation.

6 Conclusion

Part of the beauty of our mathematical understanding of nature are the relationships between vastly different phenomena. This paper started with Feynman's path integral formulation, a mathematical framework for understanding systems at the subatomic level. Yet we have shown that much of this framework is still relevant to physical systems on a vastly different scale, such as in biophysics. There are conceptual similarities as well as related mathematical tools. All physicists are given an education that runs the gamut of physical phenomena, from classical and statistical physics, to electromagnetism, and all the way down, in this current age, to string theory. For some, it may be easy to lose sight of the astounding range of Nature that physics encompasses; it may be easy to overlook the astonishing feat of tying together seemingly distant pieces, all through mathematics. Our particular fascination of this sort is QIB, and in closing, we would like to remind the reader that Feynman's path integral formulation is but a part of quantum mechanics. Quantum physics has been a focus of the physical community for decades; the field is rife with ideas and tools that could be the genesis for further QIB.

References

[1] R. Feynman, "Space-time approach to non-relativistic quantum mechanics," *Reviews of Modern Physics*, vol. 20, pp. 367–387, April 1948.

[2] V. G. Ivancevic and T. T. Ivancevic, *Quantum Leap: From Dirac and Feynman, Across the Universe, to Human Body and Mind*. World Scientific Publishing Co. Pte. Ltd., 2008.

[3] J. Wang, K. Zhang, and E. Wang, "Kinetic paths, time scale, and underlying landscapes: A path integral framework to study global natures of nonequilibrium systems and networks," *The Journal of Chemical Physics*, vol. 133, 2010.

[4] M. Kac, "On distributions of certain weiner functionals," *Transactions of the American Mathematical Society*, vol. 65, pp. 1–13, Jan 1949.

[5] P. del Moral, *Feynman-Kac Formulae: Genealogical and Interacting Particle Systems*. Springer, 2004.

[6] J. D. Murray, *Mathematical Biology: An introduction*. Springer-Verlag, 2002.

[7] T. Maeshiro and M. Kimura, "Mathematical analysis of evolutionary process," in *Proceedings of the Third European Conference on Advances in Artificial Life*, (London, UK), pp. 290–301, Springer-Verlag, 1995.

[8] R. A. Fisher, "The wave of advance of advantageous genes," *Annals of Eugenics*, vol. 7, pp. 355–369, 1937.

[9] L. S. Schulman, *Techniques and Applications of Path Integration*. John Wiley and Sons, 1981.