Journal of
Computational and Theoretical Nanoscience
Vol. 5, 1–35, 2008

REVIEW

Several Issues in Modeling Molecular Motors

Hongyun Wang

Department of Applied Mathematics and Statistics, University of California, Santa Cruz, CA 95064, USA

Protein motors play a central role in many cell functions. Understanding the operating principles of molecular motors is crucial to comprehending intracellular protein transport and cell motility. Due to the small size and the negligible inertia of molecular motors, the motor motion is damped by high viscous friction and is subject to large thermal excitations form the surrounding fluid environment. These properties determine that molecular motors behave drastically different from macroscopic motors. There are many levels of modeling for molecular motors: from simple chemical kinetic models with a few discrete states to all atom molecular dynamics simulations. Models at each level represent a compromise between being able to accommodate all aspects of motor operation and being easy to analyze/compute. All these models of different levels are valuable and are contributing to the understanding of the operating mechanism of molecular motors. In most of our studies of molecular motors, we adopt a mathematical framework of an intermediate level. In this framework, the major conformational changes of the motor are treated as continuous motions and changes in the occupancy state at the catalytic sites are modeled as discrete Markov transitions. In this framework, a motor system is described by a system of coupled Fokker-Planck equations, in which each equation corresponds to a chemical occupancy state. In this review, we discuss several issues in modeling molecular motors under this mathematical framework: (1) detailed balance and breaking of detailed balance; (2) motor potential profile and reconstruction of motor potential profile; (3) thermodynamic efficiency, Stokes efficiency and other efficiencies; and (4) numerical solutions of motor systems. In this review, we also examine the derivation of randomness parameter. We like to make it clear that this review is far from being comprehensive. The issues discussed are simply the ones we encounter in our modeling of molecular motors.

Keywords: Markov Fokker-Planck Formulation, Detailed Balance, Motor Potential Profile, Stokes Efficiency, Numerical Solutions of Fokker-Planck Equations, Randomness Parameter.

CONTENTS

1. INTRODUCTION

Protein motors play a central role in many cellular functions. For example, an F_0 motor utilizes the trans-membrane proton gradient to power the ATP synthesis in the F_1 portion of the ATP synthase; a kinesin dimer hydrolyzes ATP to drive intracellular vesicle transportation; and a

V-ATPase hydrolyzes ATP to pump protons against a large proton gradient to regulate intracellular acidity. Due to the small size of protein motors, the inertia of the motor is negligible. As a result, the motor operation is dominated by high viscous friction and large thermal fluctuations from the fluid environment {Berg, 1993 #1}. Because of these properties of molecular motors, the operating mechanisms of macroscopic motors, in general, do not necessarily extend to molecular motors. In both macroscopic motors and molecular motors, a chemical reaction can be used to generate a conformational change and an active force at the reaction site. The conformational change along with the active force is then mechanically delivered to drive the motor motion. This mechanism of generating directed motion is called a power stroke motor.[1, 2] However, for molecular motors, a unidirectional motion can be generated by a completely different mechanism. For simplicity, we consider the case where a motor is restricted to moving in a single spatial dimension, such as moving along a polymer track. Suppose we use the free energy released from the chemical reaction to block thermal fluctuations toward one direction. Then the motor will move toward the other direction. Here the unidirectional motion is not directly driven by an active force produced at the reaction site. Rather it is directly driven by the thermal forces caused by the bombardments of surrounding fluid molecules. Of course, the free energy for blocking the backward fluctuations comes from the chemical reaction. This mechanism of generating a unidirectional motion is called a Brownian ratchet.[3–9] The power stroke motor and Brownian ratchet are not exactly mutually exclusive, and it is quite possible that biological molecular motors use a combination of these two mechanisms.

There are many levels at which the operations of molecular motors are modeled. These models of different levels range from simple chemical kinetic models with a few discrete states to all atom molecular dynamics simulations. Basically, models at each level represent a compromise between the two requirements we put on models: (1) the model should be able to accommodate all aspects of the motor operation; and (2) the model should be simple enough so that analytic solutions can be derived and/or numerical solutions can be easily computed. Our opinion is that all models of different levels are valuable. Each model is best suitable for studying a particular problem or in a particular situation. All models are contributing, in some aspects, to the understanding of the operating mechanism of molecular motors. Currently there is no single model satisfying the two requirements. It is unlikely that such a perfect model will arise in the near future. Simple kinetic models are simple to solve numerically, and in certain situations can be solved analytically. Kinetic models are also easy to fit to biochemical data. But kinetic models are not very good at capturing the mechanical aspects of motors, such as, the viscous drag on the motor and the active force driving the motor. On the other hand, an all atom molecular dynamic simulation, in principle, can capture all details of the motor molecule in very fine time resolution.[10] But even for molecules of moderate size, with today's computing facilities, it is still very far from practical to run molecular dynamic simulations over a physical time period of milliseconds, which is the time scale of chemical reaction in molecular motors.[11, 12] Simple kinetic models and all atom molecular dynamic simulations are the two extreme ends on the spectrum of models of various levels. Models between these two extreme ends can be constructed. As we go from simple kinetic models all the way toward all atom molecular dynamic simulations, more features of the motor operation can be accommodated/captured. At the same time the model gets mathematically and computationally

Hongyun Wang (born in 1964 in Beijing, China) is Associate Professor in the department of Applied Mathematics and Statistics at University of California, Santa Cruz (UC Santa Cruz). He obtained his B.S. and M.S. degrees in Mathematics at Beijing University, China, and obtained his Ph.D. in Mathematics at University of California, Berkeley (UC Berkeley), in 1996, under the supervision of Professor Alexandre J. Chorin and with a dissertation in Fluid Dynamics. While finishing his Ph.D., he got interested in Biophysics, and in particular, the force generation mechanism of molecular motors. After finishing his Ph.D., he changed his research into the area of modeling molecular motors. In 1996–1999, he worked as a postdoctoral fellow in the group of Professor George, F. Oster, in the department of Molecular and Cellular Biology at UC Berkeley. While working with Oster's group, he proposed the binding zipper model, which explains how ATP binding free energy is utilized gradually and efficiently to generate a nearly constant force at the catalytic site. In 1999, he joined the School of Engineering at UC Santa Cruz, as Assistant Professor. At UC Santa Cruz, he continued his research on molecular modeling. He proposed the Stokes efficiency, which measures how efficiently a motor is using the chemical free energy to drive through the viscous media. Recently, he proposed using motor potential profile to represent the overall motor operation and designed a robust formulation for extracting motor potential profile from experimental date. At UC Santa Cruz, he was a founding member of, and was actively involved in establishing the department of Applied Mathematics and Statistics. In 2005, he was promoted to Associate Professor with tenure. His research interests also include material sciences and complex fluids.

more complicated. In most of our studies, we adopt a modeling approach of intermediate level. We model the major mechanical motion of the molecular motor as a continuous movement and model changes in the chemical occupancy state at the reaction sites as discrete Markov transitions. For molecular motors, the length scales over which inertial effects are important are much shorter than the characteristic length scales of the motor motion. So we can safely ignore inertial effects. The effect of thermal fluctuations, however, must be included in the model. The resulting mathematical framework is systems of coupled Fokker-Planck equations, each equation of which governs the time evolution of the spatial probability density of the motor in one particular chemical occupancy state. We like to make it clear that we do not claim this modeling approach of intermediate level is the only viable approach for modeling molecular motors. We do believe that this approach is one of good compromises that are both practical to solve and able to accommodate most current experimental setups and experimental results. In many single molecule experiments, mechanical aspects of motor operation, such as, the stochastic motor position and the loading force on the motor, can be directly observed/recorded. To better utilize these new experimental measurements, we like to incorporate the mechanical aspects of motor operation directly into the model. That is why we adopted this intermediate modeling approach.

Because of their small size, molecular motors have some peculiar characters that are fundamentally different from those of macroscopic motors:

(1) The time scale of inertia of molecular motors is much smaller than the time scale of reaction cycle. In contrast, the time scale of inertia of macroscopic motors is much larger than the time scale of reaction cycle.

(2) For a molecular motor, the instantaneous velocity is mainly caused by thermal excitations from the fluid environment, and thus, is highly stochastic. In contrast, the instantaneous velocity of a macroscopic motor is barely affected by thermal excitations.

(3) The instantaneous velocity of a molecular motor is several orders of magnitude larger than the average velocity over one reaction cycle. In contrast, the instantaneous velocity of a macroscopic motor is nearly a constant over many reaction cycles.

(4) The kinetic energy of a molecular motor calculated using the average velocity is much smaller than the thermal energy $k_{\mathrm{B}}T$ (the kinetic energy calculated using the instantaneous velocity is comparable to the thermal energy). In contrast, the kinetic energy of a macroscopic motor is much larger than the thermal energy.

(5) Both molecular motors and macroscopic motors can overcome energy barriers. But they overcome energy barriers in very different ways. Molecular motors wait for thermal fluctuations to carry them over energy bumps. In contrast, macroscopic motors use the stored kinetic energy to get over energy bumps.

These peculiar characters of molecular motors demonstrate that we should not simply extend results for macroscopic motors directly to molecular motors without careful examination. The difference between molecular motors and macroscopic motors is manifested in the issue of efficiency. For a macroscopic motor, the efficiency is well defined no matter it is working again a conservative force or a friction force. For a molecular motor, we do need to distinguish these two cases. When a molecular motor is working against a conservative force, the thermodynamic efficiency is well defined and is the ratio of energy output to the external agent exerting the conservative force and the chemical free energy consumption by the motor. When a molecular motor is working against a viscous drag, the situation is completely different. It has no real energy output at all. One way to define efficiency for this case is to proceed with the apparent energy output based on the average velocity. The efficiency defined this way, called Stokes efficiency, is different from the thermodynamic efficiency. First of all, these two efficiencies are for two different cases: thermodynamic efficiency is for a motor working against a conservative force; and Stokes efficiency is for a motor working against a viscous drag. But this is not the only difference. In single molecule experiments, a motor can be put to work again a conservative force. In a different experimental setup, the same motor can be put to work again a viscous drag. For a molecular motor, both the thermodynamic efficiency and the Stokes efficiency can be measured but in two different experimental setups. There is no guarantee, however, that for a given motor, these two measured efficiencies will match each other. As we will see, each of these two efficiencies tells us a different aspect about the motor. When these two measured efficiencies are different from each other, it does not necessarily imply that one of the two experiments is wrong. To the contrary, it is valuable to carry out both experiments and measure both efficiencies.

The rest of the paper is organized as follows. In Section 2, we review the Markov Fokker-Planck formulation of molecular motors. In Section 3, we clarify the issue of detailed balance. In particular, we describe the kinetic formulation in which detailed balance is still obeyed even when the chemical reaction in the motor is away from equilibrium. The key is that the corresponding occupancy states in different reaction cycles have different free energy and should be distinguished. A chemical reaction in equilibrium is described by a closed loop of states. This is justified because in equilibrium each occupancy state has a well-defined free energy. Away from equilibrium, the free energy of an occupancy state depends on which cycle it is in. Thus, A chemical reaction away from equilibrium should be described by a helical sequence of cycles instead of by a closed loop of states. The transitions among states in the helical sequence still obey detailed balance. In Section 4, we discuss the motor potential profile.

The motor potential profile is the integral of the motor force profile, which is the motor force at each spatial position, averaged over all chemical occupancy states. The introduction of motor force profile is another example of adapting modeling framework to accommodate experiments. In the Markov Fokker-Planck framework, the motor is driven by switching among a set of potentials, each corresponding to a particular chemical occupancy state. It is desirable to determine the full set of potentials. However, in most of the present single molecule experiments, only motor position is measured. The chemical occupancy state of the motor is still not directly observable. Instead of waiting for the advance of experimental technologies and determining the full set of potentials when necessarily data is available, we like to use a single potential to represent the motor operation and use the currently available experimental data to determine that potential. This motivation leads to the introduction of the motor potential profile. The biggest advantage of considering the motor potential profile is that it can be reconstructed from measured time series of motor positions. In Section 5, we examine various efficiencies for molecular motors, including the Stokes efficiency for motors working against viscous drag instead of working against an external conservative loading force. As we will see, the concept of motor potential profile allows us to decompose the Stokes efficiency into the chemical efficiency and the mechanical efficiency. We will discuss these efficiencies and their implications for motor mechanism in Section 5. In Section 6, we study numerical methods for solving Fokker-Planck systems. The key quantity we like to preserve in numerical solutions is detailed balance. Detailed balance ensures that when the motor system is at equilibrium a numerical method does not artificially produce a net drift. Detailed balance is important for numerical solutions of molecular motors since in molecular motors the physical drift is small in comparison with thermal fluctuations and any artificial drift may easily ruin the true solution. For a molecular motor moving in steps, a quantity called randomness parameter can be defined. Actually, the randomness parameter can be defined in two ways: it measures the randomness in the number of steps for long time; or it measures the randomness in the cycle time of one step. To have a unified randomness parameter, we need to address the relation between these two randomnesses. For Markov processes, these two randomnesses are equivalent. For semi-Markov processes, the equivalence of these two randomnesses has been argued using the approach of Laplace transform. However, the approach of Laplace transform breaks down when the cycle time distribution has a discrete component. In Section 7, we review and discuss the randomness and the approach of Laplace transform for deriving the equivalence between the two randomnesses. We present a new approach based on semi-Markov processes with age-structure to derive the equivalence of the two randomnesses. The new approach works for all semi-Markov processes with arbitrary cycle time distribution.

2. MARKOV FOKKER-PLANCK FRAMEWORK FOR MODELING MOLECULAR MOTORS

In this section, we review a modeling framework of intermediate level in which the continuous motor motion is described explicitly by a continuous variable in the model while the change of chemical occupancy is described by a discrete Markov processe.

We start by reasoning why most of the time we follow the motor motion only along one spatial dimension. A molecular motor, in general, has many degrees of freedom. Of these degrees of freedom, there is one that is associated with the unidirectional motion of the motor. For example, a kinesin dimer walks along a microtubule,[13–16] the γ shaft of the F_1 ATPase rotates with respect to the $\alpha_3\beta_3$ hexamer,[1, 17–20] and a flagellar motor rotates the flagellar filament with respect to the cell body.[21–23] In many studies of molecular motors, the mechanical motion is followed only along the dimension of the unidirectional motion.[4, 6, 24, 25] Usually the time scales associated with the other degrees of freedom are much shorter than the time scale set by the average motor velocity. As a result, the effects of the fast motions along these other degrees of freedom on the motor's behavior along the dimension of the unidirectional motion can be approximated by a mean field potential. Another important reason of adopting models of one spatial dimensional is that we like to make it possible to connect models to experiments. In most of the current single molecule experiments, the motor motion is only observed/recorded along one spatial dimension. Adopting models of high spatial dimensions not only increases the complexity of analysis and computation but also makes it difficult to connect to the current experimental results. Of course, the situation will change when the single molecule experimental technologies allow us to track the motor motion accurately in more than one spatial dimensions. Below, we introduce the Langevin equation with inertia, which is based on the Newton's second law. We will discuss the time scale of inertia, over-damped Langevin equation, the Fokker-Planck equation, and Fokker-Planck equation with discontinuous potentials.

2.1. Langevin Equation with Inertia

We consider the simple case of a small particle in a fluid environment, restricted to moving in one spatial dimension, and subject to a static potential, $\phi(x)$, where x is the coordinate along the spatial dimension. This situation is actually very close to that of a molecular motor. The main difference is that a molecular is driven by switching among a set of static potentials, each corresponding to one chemical occupancy state. The particle experiences the force derived from the potential, a viscous drag force from the fluid environment and a Brownian force also form the fluid environment. Both the drag force and the Brownian

force arise from collisions of the particle with the surrounding fluid molecules. The drag force is the mean of this stochastic bombarding force and the Brownian force is the fluctuating part of this bombarding force. The drag force on the particle is always opposing the motion and is proportional to the velocity

$$\text{Drag} = -\zeta u$$

where u is the velocity of the particle and ζ is the drag coefficient. The drag coefficient of the particle depends on the shape and size of the particle, and depends on the properties of the surrounding fluid. In general, the drag coefficient is proportional to the size of the particle.[26] The Brownian force on the particle has zero mean and is modeled as a Gaussian white noise, which is the formal derivative of the Weiner process (the Brownian motion). The magnitude of the Brownian force is related to the drag coefficient as

$$\text{Brownian force} = \sqrt{2k_B T\zeta}\frac{dW(t)}{dt}$$

where k_B is the Boltzmann constant, and T is the absolute temperature. This is a result of the fluctuation-dissipation theorem.[27, 28] The stochastic motion of the particle is governed by the Langevin equation (Newton 's second law):

$$m\frac{du}{dt} = -\zeta u - \phi'(x) + \sqrt{2k_B T\zeta}\frac{dW(t)}{dt} \qquad (1)$$

where m is the mass of the particle, x is the position and u is the velocity. In the above equation, $W(t)$ is the standard Weiner process, which satisfies

- $W(0) = 0$
- $W(t+s) - W(t) \sim N(0,\ s)$ for $s > 0$
- $W(t+s) - W(t)$ and $W(t) - W(t-\tau)$ are independent for $s > 0$ and $\tau > 0$.

Here $N(a,\ \text{s})$ denotes a Gaussian random variable with mean $= a$ and variance $= s$.

2.2. Time Scale of Inertia and Reduction to the Overdamped Langevin Equation

Equation (1) has three different time scales: the time scale for the motor to forget about its initial velocity (time scale of inertia), the time scale associated with motion derived from the potential $\phi(x)$, and the time scale associate with thermal diffusion. Because of the small size of molecular motors, the time scale of inertia is the shortest time scale in the system. To illustrate this, we consider the special case where the static potential is zero: $\phi(x) \equiv 0$. In this case, Eq. (1) simplifies to

$$m\frac{du}{dt} = -\zeta u + \sqrt{2k_B T\zeta}\ \frac{dW(t)}{dt}$$

Dividing both sides by m, and noticing that $t_0 = m/\zeta$ has the dimension of time, we get

$$\frac{du}{dt} = -\frac{1}{t_0}u + \frac{1}{t_0}\sqrt{2D}\frac{dW(t)}{dt} \qquad (2)$$

where $D = k_B T/\zeta$ is the diffusion coefficient of the particle.[26] Equation (2) can be solved analytically. Multiplying both sides by the integration factor $\exp(t/t_0)$ yields

$$\frac{d}{dt}\left[\exp\left(\frac{t}{t_0}\right)u(t)\right] = \frac{1}{t_0}\exp\left(\frac{t}{t_0}\right)\sqrt{2D}\frac{dW(t)}{dt}$$

Integrating with respect to time, we have

$$u(t) = \exp\left(\frac{-t}{t_0}\right)u(0) + W_1(t) \qquad (3)$$

where

$$W_1(t) = \exp\left(\frac{-t}{t_0}\right)\int_0^t \frac{1}{t_0}\exp\left(\frac{s}{t_0}\right)\sqrt{2D}\ \frac{dW(s)}{ds}\,ds$$

Since $W(s)$ is the Weiner process, $dW(s)$ is a Gaussian random variable with mean $= 0$ and variance $= ds$. It follows that

$$\exp\left(\frac{-t}{t_0}\right)\frac{1}{t_0}\exp\left(\frac{s}{t_0}\right)\sqrt{2D}dW(s) \sim N\left(0,\ \exp\left(\frac{-2t}{t_0}\right)\frac{1}{t_0}\exp\left(\frac{2s}{t_0}\right)2Dds\right)$$

Thus, $W_1(t)$ is a Gaussian random variable, independent of the initial velocity, with mean zero and variance given by

$$\text{var}\,[W_1(t)] = \int_0^t \exp\left(\frac{-2t}{t_0}\right)\frac{1}{t_0}\exp\left(\frac{2s}{t_0}\right)2D\,ds = \frac{D}{t_0}\left[1 - \exp\left(\frac{-2t}{t_0}\right)\right]$$

In Eq. (3), as t increases, the effect of the initial velocity $u(0)$ decays exponentially with time scale t_0. Furthermore, as t increases, the second term on the right hand side of (3) converges exponentially with time scale t_0 to a Gaussian random variable with mean zero and variance

$$\lim_{t\to\infty}\text{var}[W_1(t)] = \lim_{t\to\infty}\frac{D}{t_0}\left[1 - \exp\left(\frac{-2t}{t_0}\right)\right] = \frac{D}{t_0} = \frac{k_B T}{m}$$

That is, as t increases, the velocity distribution converges exponentially with time scale t_0 to the Maxwell-Boltzmann distribution. For these reasons, we shall call t_0 the time scale of inertia. We now look at how the time scale t_0 changes with the particle size. For simplicity, we consider a spherical particle with radius r. The mass and the drag coefficient of a spherical particle are, respectively, given by,[26]

$$m = \frac{4}{3}\pi\rho r^3, \quad \zeta = 6\pi\eta r$$

REVIEW

where ρ is the density of the particle, and η is the viscosity of the surrounding fluid. Substituting these expressions into the time scale of inertia, we have

$$t_0 = \frac{m}{\zeta} = \frac{2\rho}{9\eta} r^2$$

Although the above expression is derived based on spherical particles, it is true that the time scale of inertia t_0 is, in general, proportional to the square of the particle size whereas the coefficient varies with the shape of particle. Thus, for small particles, the time scale of inertia t_0 is very small. For example, for a latex bead of diameter 1 μm in water, we have $\rho = 1$ g/cm^3 $= 10^{-21}$ g/nm^3, $\eta = 0.01$ poise $= 10^{-9}$ pNnm^{-2} s, $r = 500$ nm, and the time scale of inertia is $t_0 = 56 \times 10^{-9}$ s $= 56$ ps. A time period of 56 ps is much smaller than the typical time scales of chemical reaction cycles in molecular motors. Notice that in the worlds of molecular motors, a bead of 1 μm is a big object. For smaller objects, the time scale of inertia is even smaller.

The short time dynamics of the particle is full of fast decaying of the instantaneous velocity in time scale of t_0 and fast increasing of the instantaneous velocity excited by the stochastic Brownian force. These short time dynamical behaviors present significant difficulties for analysis and computations. To capture relatively long time behavior without resolving these short time details, we consider the mathematical limit of Eq. (1) as time scale t_0 goes to zero. We first rewrite Eq. (1) as

$$\frac{du}{dt} = -\frac{1}{t_0}\left[u + \frac{D}{k_B T}\phi'(x) - \sqrt{2D}\,\frac{dW(t)}{dt}\right]$$

The mathematical limit of this equation as t_0 going to zero can be seen by considering a simple model equation

$$\frac{du}{dt} = -\frac{1}{t_0}[u - g(t)]$$

Given function $g(t)$, the exact solution of the model equation has the expression

$$u(t) = g(t) + \exp\left(\frac{-t}{t_0}\right)(u(0) - g(0)) + \frac{1}{t_0}\int_0^0 \exp\left(\frac{-(t-s)}{t_0}\right)(g(s) - g(t))\,ds$$

Taking the limit as t_0 goes to zero yields

$$\lim_{t_0 \to 0} u(t) = g(t)$$

Formally identifying

$$g(t) = -\frac{D}{k_B T}\phi'(x) + \sqrt{2D}\frac{dW(t)}{dt}$$

we see that as t_0 goes to zero, the solution of Eq. (1) satisfies

$$u(t) = -\frac{D}{k_B T}\phi'(x) + \sqrt{2D}\frac{dW(t)}{dt}$$

Writing it as a differential equation for $x(t)$, we obtain

$$\frac{dx}{dt} = -\frac{D}{k_B T}\phi'(x) + \sqrt{2D}\frac{dW(t)}{dt} \tag{4}$$

which is the over-damped Langevin equation. The reduction from Eqs. (1) to (4) in the limit of small t_0 is called the Einstein-Smoluchowski limit.[29] The Einstein-Smoluchowski limit can also be illustrated intuitively by considering a special case where

$$\frac{-\phi'(x)}{\zeta} = u_0 = \text{const}$$

that is, the velocity induced by the static potential is a constant. In this special case, Eq. (1) becomes

$$\frac{du}{dt} = -\frac{1}{t_0}u + \frac{1}{t_0}u_0 + \frac{1}{t_0}\sqrt{2D}\,\frac{dW(t)}{dt} \tag{5}$$

This equation can be solved analytically. We first solve for the particle velocity $u(t)$. Multiplying by the integration factor and integrating from 0 to τ, we get

$$u(\tau) = u_0 + \exp\left(\frac{-\tau}{t_0}\right)[u(0) - u_0] + \sqrt{2D}\exp\left(\frac{-\tau}{t_0}\right)\int_0^\tau \frac{1}{t_0}\exp\left(\frac{s}{t_0}\right)\frac{dW(s)}{ds}\,ds$$

Integrating $u(\tau)$ from 0 to t, we have

$$x_{\text{Equation (5)}}(t) = x(0) + u_0 t + [u(0) - u_0]\,t_0 \times \left[1 - \exp\left(\frac{-t}{t_0}\right)\right] + W_2(t) \tag{6}$$

where random variable $W_2(t)$ is given by

$$\begin{aligned} W_2(t) &= \sqrt{2D}\int_0^t \int_0^\tau \frac{1}{t_0}\exp\left(\frac{s-\tau}{t_0}\right)\frac{dW(s)}{ds}\,ds\,d\tau \\ &= \sqrt{2D}\int_0^t \frac{dW(s)}{ds}\int_s^t \frac{1}{t_0}\exp\left(\frac{s-\tau}{t_0}\right)d\tau\,ds \\ &= \sqrt{2D}\int_0^t \frac{dW(s)}{ds}\left[1 - \exp\left(\frac{s-t}{t_0}\right)\right]ds \end{aligned}$$

Again, since $W(s)$ is the Weiner process, $dW(s)$ is a Gaussian random variable with mean $= 0$ and variance $= ds$. It follows that $W_2(t)$ is a Gaussian random variable with mean $= 0$ and variance given by

$$\begin{aligned} \text{var}[W_2(t)] &= 2D\int_0^t \left[1 - \exp\left(\frac{s-t}{t_0}\right)\right]^2 ds \\ &= 2D\int_0^t \left[1 - \exp\left(\frac{-s}{t_0}\right)\right]^2 ds \\ &= 2D\left\{t - 2t_0\left[1 - \exp\left(\frac{-t}{t_0}\right)\right] + \frac{t_0}{2}\left[1 - \exp\left(\frac{-2t}{t_0}\right)\right]\right\} \end{aligned} \tag{7}$$

In the solution of Eq. (5) given in (6), taking the limit as t_0 going to zero, we have

$$\lim_{t_0\to 0} \text{var}[W_2(t)] = 2Dt$$

$$\lim_{t_0\to 0} x_{\text{Equation (5)}}(t) = x(0) + u_0\, t + W_3(t)$$

where $W_3(t)$ is a Gaussian random variable with mean $= 0$ and variance $= 2Dt$.

On the other hand, for the case of constant driving force, Eq. (4) becomes

$$\frac{dx}{dt} = u_0 + \sqrt{2D}\,\frac{dW(t)}{dt} \tag{8}$$

The exact solution of Eq. (8) is

$$x_{\text{Equation (8)}}(t) = x(0) + u_0 t + W_4(t)$$

where $W_4(t)$ is a Gaussian random variable with mean $= 0$ and variance $= 2Dt$. Thus, we arrive at

$$\lim_{t_0\to 0} x_{\text{Equation (5)}}(t) = x_{\text{Equation (8)}}(t)$$

which justifies the reduction from Eqs. (1) to (4) in the limit of small t_0 for the special case of constant driving force. For the general case, a more sophisticated analysis of the reduction from (1) to (4) can be found in Ref. [29].

The convergence of the solution of Eq. (5) given in (6) for small but finite t_0 depends also on the initial velocity $u(0)$. In (6), for $t \gg t_0$, the contribution of the instantaneous initial velocity to the displacement is approximately $u(0)t_0$ (the term $u_0 t_0$ is small and is ignored). If the initial velocity comes from thermal excitations, then we have $u(0) \sim \sqrt{k_B T/m}$, and the contribution of the initial velocity to the displacement is $u(0)t_0 \sim \sqrt{Dt_0}$, which is much smaller than the diffusion in a reaction cycle: $\sqrt{Dt_0} \ll \sqrt{Dt_{\text{Cycle}}}$. However, when the instantaneous velocity of a particle is much larger than $\sqrt{k_B T/m}$, the contribution of the initial velocity may not be ignored. In the process of ballistospore discharge in mushrooms and basidiomycete yeasts,[30, 31] the ballistospore is approximately a spherical particle of radius 5 μm; the typical velocity caused by thermal excitations is $\sqrt{k_B T/m} \approx 0.88 \times 10^{-4}$ m s^{-1}; and the initial launch velocity is about 1.3 m s^{-1}, which is at least 4 orders of magnitude larger than the thermal excitations. In a time interval of 2 millisecond after the initial launch, the displacement due to the initial launch velocity is about 0.4 mm $= 400$ μm while the displacement caused by Brownian diffusion is less than 0.1 μm. In this case, it is clear that the effect of the initial launch velocity dominates the motion of particle for a time period of a few t_0. During this time period, the effect of the initial velocity does decay exponentially but it is still larger than the effect of thermal excitations.

2.3. Modeling Changes of Occupancy

Now we modify Langevin Eq. (4) to accommodate molecular motors driven by chemical reactions. The key difference is that for a molecular motor, it is no longer driven by one static potential. Instead, it is driven by switching among a set of static periodic potentials, $\phi_S(x)$, each corresponding to a chemical occupancy state. In addition to making the internal motor force dependent on the chemical state, we add an external loading force, F, to better accommodate the single molecule experiments. The stochastic motion of a molecular motor is described by the set of Langevin equations below.

$$\frac{dx}{dt} = D\frac{[F - \phi'_S(x)]}{k_B T} + \sqrt{2D}\frac{dW(t)}{dt}, \quad S(t) = 1, 2, \ldots, N \tag{9}$$

(9) is a set of over-damped Langevin equations. At any given time instance, the motor system is in a particular chemical occupancy state, say $S(t) = k$. While the motor system is in state k, the stochastic motor motion is only governed by the Langevin equation with $S(t) = k$ in the set. The other equations in the set do not apply as long as the motor system is still in state k. When the chemical reaction carries the motor system to a different occupancy state, say $S(t) = k_2$, the stochastic motor motion will be governed by the Langevin equation with k_2 in the set. Basically, at any given time, only one Langevin equation in the set is governing the stochastic motor motion, and the chemical reaction switches the governing equation among the set. If we view (9) as one equation, then the periodic potential $\phi_S(x)$ changes with the chemical occupancy state $S(t)$ of the motor system.[4, 24, 32]

Equation (9) describes only the spatial motion of the motor. To model the stochastic evolution of the whole motor system, we also need to follow the change of chemical occupancy state. In a transition from occupancy state A to occupancy state B, the commitment time is generally much smaller than the residence time. The commitment time is the time of the actual transition process. More precisely, the commitment time is the time period during which the system has left state A and but has not yet arrived at state B. If we define a transition state between state A and state B, the commitment time can also be viewed as the residence time of the transition state. Because the commitment time of transition is small in comparison with the residence time, we can approximate the transition as instantaneous. Thus, we model the stochastic evolution of the motor's chemical state as a continuous time discrete Markov process, also called a jump process. Mathematically, a continuous time discrete Markov process is described as

$$\Pr[S(t+\Delta t) = j \mid S(t) = i] = \Delta t\, k_{i\to j} + o(\Delta t)$$

where $k_{i\to j}$ is the transition rate from occupancy state i to occupancy state j. In molecular motors, the chemical

REVIEW

reaction is coupled with mechanical motion. In Eq. (9), the spatial motion is affect by the change of $S(t)$, which switches the motor from one potential to another. Conversely, the evolution of the occupancy state should also be affected by the motor position x. Thus, a more precise description for the chemical reaction in a motor should be

$$\Pr[S(t+\Delta t)=j \mid S(t)=i, X(t)=x]=\Delta t k_{i\to j}(x)+o(\Delta t)$$

Here we have adopted the standard notation, i.e., using $X(t)$ to denote the stochastic motor position and using x to denote a particular value of $X(t)$. Alternatively, the change of occupancy state can be described in terms of the time evolution of probabilities of occupancy states. Let

$$p_i(t) \equiv \text{Prob}[S(t)=i]$$

$$\mathbf{P}(t) \equiv [p_1(t), p_2(t), \ldots, p_N(t)]$$

The vector of probabilities $\mathbf{P}(t)$ satisfies the differential equation

$$\frac{d\mathbf{P}(t)}{dt} = \mathbf{K}(x)\cdot\mathbf{P}(t) \tag{10}$$

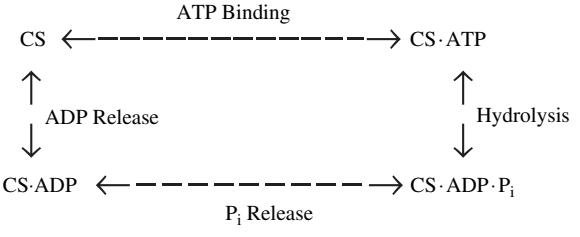
The matrix $\mathbf{K}(x)=\{k_{ij}(x)\}$ is called the transition matrix. For $j\neq i$, $k_{ij}(x)=k_{j\to i}(x)$. The diagonal elements of $\boldsymbol{K}(x)$ are defined as:

$$k_{jj}(x) = -\sum_{i\neq j} k_{j\to i}(x)$$

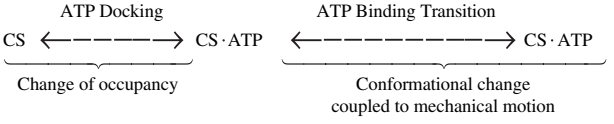
A key property of the transition matrix is

$$\sum_{i=1}^{N} k_{ij}(x) = 0 \tag{11}$$

It is important to point out that in this mathematical framework, the motor operation is divided into two kinds of processes: those that involve changes in the occupancy state of the catalytic sites (chemical transitions) and those that involve physical movement of the motor system. Changes in occupancy states are governed by the Markov process given in Eq. (10). Changes in positions of the motor and/or its subunits are modeled as continuous motion along the potential $\psi_S(x)$ of the corresponding occupancy state S and are governed by Eq. (9). For an ATPase motor, the motor operation is powered by the ATP hydrolysis at one or more catalytic sites. For each catalytic site, there are four occupancy states in the ATP hydrolysis cycle:[17, 33–35]



In a molecular motor, the chemical reaction needs to provide a direct or indirect driving force for the motor motion. As a result, the chemical reaction usually involves more than just changes in occupancy states. In particular, the ATP hydrolysis cycle involves more than just changes in the occupancy of the catalytic site. The ATP binding step in the above diagram involve both a change of occupancy and a continuous conformational change. The sub-step of ATP diffusing into a catalytic site and becoming weakly bound involves a change of occupancy of the catalytic site and is modeled as a chemical transition (a jump) in the discrete Markov process (10). On the other hand, the sub-step that the ATP goes from being weakly bound to being tightly bound does not involve an occupancy change in the catalytic site, but rather it involves a conformational change of the complex containing both the catalytic site and the ATP molecule inside it. This conformational change is modeled as a continuous motion along the potential curve corresponding to the ATP state in Eq. (9). So the ATP binding step consists of the initial ATP docking (ATP diffusing into a catalytic site and becoming weakly bound) and the subsequent ATP binding transition (from being weakly bound to being tightly bound). A more precise reaction diagram for the ATP binding step should be



A general feature of molecular motors is that a change of occupancy induces a conformational change and the conformational change makes the motor system ready for further change of occupancy. In the ATP hydrolysis cycle, the ATP docking induces the subsequent ATP binding transition, which delivers a power stroke. At the end of ATP binding transition and only at the end of ATP binding transition, the ATP is tightly bound and is well coordinated with the attacking water molecule inside the catalytic site so it can be hydrolyzed at a fast rate. Without the coordination brought on by the tight binding, the hydrolysis can only occur at a very slow rate (in water and without a catalytic site ATP can still hydrolyze at a very slow rate). The completion of the ATP binding power stroke is a necessary condition for the ATP hydrolysis reaction to proceed at a fast rate. Thus, in molecular motors, the change of occupancy state and the mechanical motion are interlocked. One cannot study the chemical reaction cycle at a catalytic site without considering the conformational change.

The stochastic evolution of a motor system (mechanical motion and change of occupancy) is governed by Langevin Eq. (9) coupled with the discrete Markov process (10). This type of coupling between mechanical motion and chemical reactions is often referred to as mechanochemistry. Sample time series of motor positions and motor occupancy states can be generated by running Monte Carlo simulations based on Eqs. (9) and (10). From sample time

series, average quantities, such as the average velocity and the effective diffusion coefficient can be calculated. From computational point of view, Monte Carlo simulations based on Eqs. (9) and (10) are straightforward to program. But averaging over sample time series suffer from statistical error. In general, the statistical error decreases as the square root of the number of samples. It takes a large number of samples to reduce the statistical error. Fortunately, all statistical properties of molecular motors can be calculated efficiently and reliably from the probability density of the motor system.

2.4. Fokker-Planck Equations

To introduce the probability density of a motor system, we consider an ensemble of identical motors, each evolving independently in time according to Eqs. (9) and (10). Let $\rho_j(x, t)$ be the probability density that the motor system is at position x and in occupancy state j at time t. Mathematically $\rho_j(x, t)$ is defined as

$$\rho_j(x,t) \equiv \lim_{\Delta x \to 0} \frac{\text{Prob}[S(t) = j, x \le X(t) < x + \Delta x]}{\Delta x}$$

The time evolution of $\rho_j(x, t)$ is governed by the Fokker-Planck equation:[29, 36]

$$\frac{\partial \rho_i}{\partial t} = D \frac{\partial}{\partial x}\left(\frac{-F + \phi_i'(x)}{k_B T}\rho_i + \frac{\partial \rho_i}{\partial x}\right) + \sum_{j=1}^{N} k_{ij}(x)\rho_j \qquad i = 1, 2, \ldots, N \quad (12)$$

which is the Fokker-Planck equation corresponding to Langevin Eq. (9) and discrete Markov process (10). Equation (12) is the mathematical framework often used to study the mechanochemistry of molecular motors.[7, 24, 32, 37–43] It is important to point out that although (12) is a linear differential equation in terms of the probability density $\rho_j(x, t)$, it is nonlinear in terms of the vector $(\rho_j, \phi_j, F, k_{i\to j})$. In general, the motor system does not respond linearly to changes in $(\rho_j, \phi_j, F, k_{i\to j})$. For example, the diagram of the average velocity versus the external loading force is usually non-linear. So deciphering motor mechanism is definitely not a linear problem.

Before we end the discussion of Markov Fokker-Planck framework, we examine the mathematical issue of Fokker-Planck equations with discontinuous potentials. While real physical potentials do not have discontinuities, discontinuous stylized potentials often serve as informative and useful approximations. If the potential $\phi_i(x)$ in Eq. (12) contains discontinuities, then at a discontinuity $\phi_i'(x)$ does not make sense. Consequently, the solution of (12) cannot be defined as a smooth function that satisfies the partial differential equation everywhere. Away from the discontinuities, $\phi_i'(x)$ is well defined and the solution of Eq. (12) should satisfy the differential equation in the classical sense. However, satisfying the differential equation in regions away from discontinuities is not enough to make the solution unique. To specify the solution of Eq. (12), we need to impose certain conditions at the discontinuities of potentials.

For the simplicity of discussion below, we consider the Fokker-Planck equation corresponding to Langevin Eq. (4) for the case of a single static potential:

$$\frac{\partial \rho}{\partial t} = D \frac{\partial}{\partial x}\left(\frac{\phi'(x)}{k_B T}\rho + \frac{\partial \rho}{\partial x}\right) \quad (13)$$

Our goal is to find out the conditions we need to impose at the discontinuities of the potential. We study the case where $\phi(x)$ is a tilted periodic potential: $\phi(x + l) = \phi(x) + \Delta\phi$ where l is the period and $\Delta\phi$ is a constant. Without loss of generality, we assume that potential $\phi(x)$ has only one discontinuity in $[0, l)$. Let d denote the position of discontinuity. Mathematically, $\phi(x)$ is smooth, respectively, in $[0, d)$ and in $(d, l]$. At the discontinuity d, $\phi(x)$ has limits:

$$\phi(d^-) = \lim_{x \to d^-} \phi(x) \quad \text{and} \quad \phi(d^+) = \lim_{x \to d^+} \phi(x)$$

Basically, potential $\phi(x)$ is two smooth functions connected at the discontinuity d. At the discontinuity d, $\phi'(x)$ is not a regular function. If the system is brought to equilibrium by placing reflecting boundaries at $x = 0$ and $x = l$, then the solution is given by the Boltzmann distribution:

$$\rho(x) = \frac{1}{\int_0^l \exp(-\phi(x)/k_B T)\, dx} \exp\left(\frac{-\phi(x)}{k_B T}\right)$$

which is discontinuous at d. Thus, we expect the time-dependent solution $\rho(x, t)$ also to be discontinuous at d. Away from the discontinuity, the equilibrium solution is smooth and satisfies the differential equation. So it is reasonable to expect $\rho(x, t)$ to be smooth and satisfy the Fokker-Planck equation in the classical sense away from the discontinuity. To derive the conditions for the solution $\rho(x, t)$ at the discontinuity d, we use conservation of probability to rewrite Eq. (13) as

$$\frac{\partial \rho}{\partial t} = -\frac{\partial J(x)}{\partial x}, \quad J(x) = -D\left(\frac{\phi'(x)}{k_B T}\rho + \frac{\partial \rho}{\partial x}\right)$$

where $J(x)$ is the probability flux at x and is defined to be positive for flow moving toward the positive end of the x-axis. The first term in $J(x)$ is due to transport driven by the potential $\phi(x)$ and the second term is caused by Brownian diffusion. The first condition on $\rho(x, t)$ is that the probability be conserved at the discontinuity d. That is, the probability flux into the discontinuity must be the same as the probability flux out of the discontinuity:

$$\left(\frac{\phi'(x)}{k_B T}\rho + \frac{\partial \rho}{\partial x}\right)\Big|_{d^-} = \left(\frac{\phi'(x)}{k_B T}\rho + \frac{\partial \rho}{\partial x}\right)\Big|_{d^+} \quad (14)$$

REVIEW

Unfortunately, condition (14) alone is still not enough to uniquely determining the solution. For example, the function

$$\rho(x)=\begin{cases}\dfrac{1}{\int_0^d \exp\left(-\phi(x)/k_{\mathrm{B}}T\right)dx}\exp\left(\dfrac{-\phi(x)}{k_{\mathrm{B}}T}\right), & x<d\\ 0, & x>d\end{cases} \tag{15}$$

satisfies the differential equation away from the discontinuity and satisfies condition (14) at the discontinuity. But it is clear that $\rho(x)$ given in (15) is not the correct equilibrium solution. To uniquely determine the exact solution, we need to impose another condition on $\rho(x,t)$.

In modeling molecular motors, a discontinuous potential is simply a mathematical abstraction. In reality, the discontinuity represents a very narrow transition region in which the potential is smooth but changes rapidly from the value on one side of the discontinuity to that on the other side. Thus, the solution for the discontinuous potential is interpreted as the limit of solutions corresponding to a sequence of smooth potentials converging to the discontinuous potential. In the sequence of smooth potentials, the discontinuity at d is approximated by smooth transitions over smaller and smaller transition regions. Thus, we replace the discontinuous transition of $\phi(x)$ at d by a smooth transition $\phi_\varepsilon(x)$ over $[d-\varepsilon,\ d+\varepsilon]$. The smooth potential $\phi_\varepsilon(x)$ satisfies

$$\phi_\varepsilon(d-\varepsilon)=\phi(d^-)\,,\quad \phi_\varepsilon(d+\varepsilon)=\phi(d^+)$$

The smooth solution corresponding to potential $\phi_\varepsilon(x)$ satisfies

$$\lim_{\varepsilon\to 0}\rho_\varepsilon(d-\varepsilon,t)=\rho(d^-,t)\,,\quad \lim_{\varepsilon\to 0}\rho_\varepsilon(d+\varepsilon,t)=\rho(d^+,t)$$

In the transition region $[d-\varepsilon,\ d+\varepsilon]$, the probability flux satisfies

$$-D\left(\frac{\phi'_\varepsilon(x)}{k_{\mathrm{B}}T}\rho_\varepsilon+\frac{\partial\rho_\varepsilon}{\partial x}\right)=J_d$$

Multiplying both sides by the integration factor and integrating over $[d-\varepsilon,\ d+\varepsilon]$, we get

$$\exp\left(\frac{\phi_\varepsilon(x)}{k_{\mathrm{B}}T}\right)\rho_\varepsilon\Big|_{d-\varepsilon}^{d+\varepsilon}=-\frac{1}{D}\int_{d-\varepsilon}^{d+\varepsilon}\exp\left(\frac{\phi_\varepsilon(s)}{k_{\mathrm{B}}T}\right)J_d\,ds$$

Taking the limit as ε goes to zero, we arrive at

$$\exp\left(\frac{\phi(d^+)}{k_{\mathrm{B}}T}\right)\rho(d^+,t)=\exp\left(\frac{\phi(d^-)}{k_{\mathrm{B}}T}\right)\rho(d^-,t) \tag{16}$$

Conditions (14) and (16) are the two conditions that the solution $\rho(x,t)$ must satisfy at the discontinuity. As we will see later in the discussion of numerical methods for solving the Fokker-Planck equation, condition (14) is satisfied automatically if the numerical method is based on conservation of probability. Condition (16) is related to detailed balance. Therefore, to design a good numerical method capable of solving Fokker-Planck equations with discontinuous potentials, the method must preserve detailed balance.

3. DETAILED BALANCE

In this section, we discuss the issue of detailed balance. In the modeling framework described in the previous section for molecular motors, the transition rates among different occupancy states cannot be specified arbitrarily. Even though the motor operates away from equilibrium, these transition rates are still constrained by detailed balance. Strictly speaking, detailed balance is a condition on transition rates at equilibrium, which ensures that the probability density is given by the Boltzmann distribution and there are no net physical or chemical fluxes. While detailed balance is an equilibrium property, molecular motors operate in non-equilibrium mode. If one insists on avoid using detailed balance for non-equilibrium situations, one can say the transition rates in molecular motors are constrained by conditions similar to detailed balance. In the discussion below, we shall refer to these conditions simply as detailed balance. Indeed, at equilibrium, these conditions become detailed balance. From the point of view of connecting models to experiments, it is desirable to introduce the condition of detailed balance because more constraints lead to fewer undetermined parameters. From the point view of thermodynamics, introducing the condition of detailed balance will ensure that at equilibrium no artificial drift is produced by the model.

We impose detailed balance on the transition rates because in the mathematical framework described above, the transition rates do not know whether or not the system is in equilibrium. Indeed, the transition rates between states A and B at position x in Eq. (12) depend only on position x and states A and B. The transition rates between states A and B at position x are not affected if transition rates at other locations or transition rates among other states are changed. Thus, we may block some transitions at other locations and/or among other states to bring the system to equilibrium without changing the transition rates between states A and B at position x. Notice that by placing reflecting boundaries at a position other than x and/or placing blocking barrier among states other than states A and B, the free energies of states A and B at position x are not changed. Let $G_A(x)$ and $G_B(x)$ denote the free energies of, respectively, states A and B at position x. The transition rates between states A and B at position x satisfy detailed balance

$$\frac{k_{A\to B}(x)}{k_{B\to A}(x)}=\exp\left(\frac{G_A(x)-G_B(x)}{k_{\mathrm{B}}T}\right) \tag{17}$$

Assigning free energy to a chemical state is a subtle issue. Actually the subtlety results from the ambiguity in defining what is a chemical state. More precisely the free energy of an occupancy state depends on the number of cycles completed. To illustrate the problem, we consider a hypothetical example in which there are $N>2$ chemical states in one reaction cycle and the motor goes through the N

chemical states sequentially. The case of $N = 2$ is even more tricky and will be discussed later.

$$\begin{array}{l} \vdots \\ \updownarrow \\ \underbrace{S_1 \leftrightarrow S_2 \leftrightarrow \cdots \leftrightarrow S_N}_{\text{Cycle } k-1} \\ \qquad\qquad \updownarrow \\ \qquad\qquad \underbrace{S_1 \leftrightarrow S_2 \leftrightarrow \cdots \leftrightarrow S_N}_{\text{Cycle } k} \\ \qquad\qquad\qquad\qquad \updownarrow \\ \qquad\qquad\qquad\qquad \underbrace{S_1 \leftrightarrow S_2 \leftrightarrow \cdots \leftrightarrow S_N}_{\text{Cycle } k+1} \\ \qquad\qquad\qquad\qquad\qquad\qquad \updownarrow \\ \qquad\qquad\qquad\qquad\qquad\qquad \vdots \end{array} \tag{18}$$

In the diagram above, three cycles are shown. Suppose we denote the free energy of state j in cycle k by G_j. Let ΔG be the free energy change of the system in one reaction cycle. When $\Delta G < 0$, the chemical reaction proceeds forward spontaneous. For a reaction not in equilibrium, we can always make $\Delta G < 0$ by selecting an appropriate positive direction for the chemical reaction. In a chemical reaction, $(-\Delta G) > 0$ is the free energy released in one reaction cycle, and is called the chemical affinity.[44,45] In reaction diagram (18), state j in cycle $k-1$ has free energy $G_j - \Delta G$; state j in cycle $k+1$ has free energy $G_j + \Delta G$. The transition rates are constrained by the condition of detailed balance:

$$\begin{aligned} \frac{k_{j\to j+1}(x)}{k_{j+1\to j}} &= \exp\left(\frac{G_j(x) - G_{j+1}(x)}{k_B T}\right), \quad 1 \le j \le N-1 \\ \frac{k_{N\to 1}(x)}{k_{1\to N}} &= \exp\left(\frac{G_N(x) - (G_1(x) + \Delta G)}{k_B T}\right), \quad j = N \end{aligned} \tag{19}$$

For $1 \le j \le N-1$, $k_{j\to j+1}(x)$ is the rate of transition from state j to state $j+1$ within the same cycle (say, in cycle k). $k_{N\to 1}(x)$, however, is the rate of transition from state N of one cycle to state 1 of the next cycle (say, from state N in cycle k to state 1 in cycle $k+1$). In mathematical formulation (12), only N occupancy states are followed and occupancy states in different cycles are not distinguished explicitly. This is done by projecting over all cycles. State j in formulation (12) is a composite state containing states j of all cycles in diagram (18). In terms of composite states, reaction diagram (18) becomes

$$\begin{array}{l} S_1 \leftrightarrow S_2 \leftrightarrow \cdots \leftrightarrow S_N \\ \uparrow\!\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\uparrow \end{array} \tag{20}$$

The biggest disadvantage of using composite states in reaction diagram is that free energy of a composite state is not well defined since it contains an infinite sequence of states and the probability of each state is not specified. If we artificially assign G_j as the free energy of composite state j (which is thermodynamically incorrect), then we can say that when the system is away from the equilibrium (when $\Delta G \neq 0$) detailed balance is broken:

$$\frac{k_{N\to 1}(x)}{k_{1\to N}} \neq \exp\left(\frac{G_N(x) - G_1(x)}{k_B T}\right)$$

For a molecular motor operating in a non-equilibrium mode, we can say detailed balance is obeyed or we can say detailed balance is broken. Both statements may be correct and each statement simply implies different definitions of (composite) states and different reaction diagrams for the same chemical reaction. We prefer diagrams in which detailed balance is obeyed. In such a diagram, the corresponding occupancy states of different cycles are distinguished; the free energies are defined correctly; and more important, all transition rates are still constrained.

When a reaction cycle has only $N = 2$ chemical states, the situation is most confusing. When $N = 2$, reaction diagram (18) becomes

$$\cdots \underbrace{S_1 \underset{k^-_{2\to 1}}{\overset{k^+_{1\to 2}}{\rightleftarrows}} S_2}_{\text{Cycle } k-1} \underset{k^-_{1\to 2}}{\overset{k^+_{2\to 1}}{\rightleftarrows}} \underbrace{S_1 \underset{k^-_{2\to 1}}{\overset{k^+_{1\to 2}}{\rightleftarrows}} S_2}_{\text{Cycle } k} \underset{k^-_{1\to 2}}{\overset{k^+_{2\to 1}}{\rightleftarrows}} \underbrace{S_1 \underset{k^-_{2\to 1}}{\overset{k^+_{1\to 2}}{\rightleftarrows}} S_2}_{\text{Cycle } k+1} \cdots \tag{21}$$

Notice, in particular, that transition $1 \to 2$ and transition $2 \to 1$ are not well defined. For example, transition $1 \to 2$ may mean the transition from state 1 forward to state 2 within the same cycle, or it may mean the transition from state 1 of one cycle backward to state 2 of the previous cycle (say, from state 1 in cycle k back to state 2 in cycle $k-1$). This ambiguity is caused by that there are only 2 states in one reaction cycle. In reaction diagram (21), to get rid of this ambiguity, we use $k^+_{1\to 2}(x)$ and $k^-_{1\to 2}(x)$ to denote, respectively, the transition rates of these two $1 \to 2$ transitions. Once the states and transition rates are precisely specified, the condition of detailed balance can be enforced:

$$\begin{aligned} \frac{k^+_{1\to 2}(x)}{k^-_{2\to 1}(x)} &= \exp\left(\frac{G_1(x) - G_2(x)}{k_B T}\right) \\ \frac{k^-_{1\to 2}(x)}{k^+_{2\to 1}(x)} &= \exp\left(\frac{(G_1(x) + \Delta G) - G_2(x)}{k_B T}\right) \end{aligned} \tag{22}$$

As before, we define composite state j as the collection of state j of all cycles. For a two-state reaction cycle, the reaction diagram in terms of composite states is deceptively simple:

$$S_1 \underset{k_{2\to 1}}{\overset{k_{1\to 2}}{\rightleftarrows}} S_2 \tag{23}$$

REVIEW

where the transition rates $k_{1\to2}(x)$ and $k_{2\to1}(x)$ for composite states are given by

$$k_{1\to2}(x) = k^{+}_{1\to2}(x) + k^{-}_{1\to2}(x)$$
$$k_{2\to1}(x) = k^{+}_{2\to1}(x) + k^{-}_{2\to1}(x)$$

It follows from (22) that detailed balance is not satisfied for transition rates $k_{1\to2}(x)$ and $k_{2\to1}(x)$

$$\frac{k_{1\to2}(x)}{k_{2\to1}(x)} \neq \exp\left(\frac{G_1(x) - G_2(x)}{k_{\rm B}T}\right)$$

It is clear that the breaking of detailed balance is a direct consequence of viewing the chemical reaction as a system jumping between N fixed composite states, and inappropriately assigning each composite state a free energy. In the more precise description (18) and (21), the number of reaction cycles completed is tracked explicitly and the free energy of a state depend on the number of reaction cycles completed. Therefore, for the discussion of free energy, it is more appropriate to view the chemical reaction as a system going along an infinite sequence of cycles as shown in diagrams (18) and (21). This view is especially important for avoiding confusion in the case where the reaction cycle has only two states.

Now we derive the detailed-balance-like-condition for a non-equilibrium reaction in a more concrete setting relevant for molecular motors. We consider a simple example: an ATPase motor with only one catalytic site. An ATP catalytic site has four occupancy state: **E** (empty), **T** (ATP), **DP** (ADP·P_i), and **D** (ADP). The free energy change in one ATP hydrolysis cycle is

$$\Delta G = \Delta G^0 - k_{\rm B}T \ln\left(\frac{[\rm ATP]}{[\rm ADP]\cdot[\rm P_{\it i}]}\right)$$

where $\Delta G^0 = -12.3 k_{\rm B}T$ is the standard free energy change of ATP hydrolysis (when all reactant and product concentrations are one molar).[46–48] At physiological conditions, $\Delta G \approx -20 k_{\rm B}T$.[46] Let $[\rm ATP]^{(EQ)}$, $[\rm ADP]^{(EQ)}$, and $[\rm P_{\it i}]^{(EQ)}$ be a set of equilibrium concentrations. At equilibrium $\Delta G = 0$. So $[\rm ATP]^{(EQ)}$, $[\rm ADP]^{(EQ)}$, and $[\rm P_{\it i}]^{(EQ)}$ satisfy

$$k_{\rm B}T \ln\left(\frac{[\rm ATP]^{(EQ)}}{[\rm ADP]^{(EQ)}\cdot[\rm P_{\it i}]^{(EQ)}}\right) = \Delta G^0$$

Let $\psi_E^{(EQ)}(x)$, $\psi_T^{(EQ)}(x)$, $\psi_{DP}^{(EQ)}(x)$ and $\psi_D^{(EQ)}(x)$ be the free energies of, respectively, the four occupancy states **E**, **T**, **DP** and **D**, at equilibrium. At equilibrium, detailed balance is satisfied so we have

$$\frac{k^0_{E\to T}(x)[\rm ATP]^{(EQ)}}{k_{T\to E}(x)} = \exp\left(\frac{\psi_E^{(EQ)}(x) - \psi_T^{(EQ)}(x)}{k_{\rm B}T}\right)$$
$$\frac{k_{T\to DP}(x)}{k_{DP\to T}(x)} = \exp\left(\frac{\psi_T^{(EQ)}(x) - \psi_{DP}^{(EQ)}(x)}{k_{\rm B}T}\right)$$
$$\frac{k_{DP\to D}(x)}{k^0_{D\to DP}(x)[\rm P_{\it i}]^{(EQ)}} = \exp\left(\frac{\psi_{DP}^{(EQ)}(x) - \psi_D^{(EQ)}(x)}{k_{\rm B}T}\right)$$
$$\frac{k_{D\to E}(x)}{k^0_{E\to D}(x)[\rm ADP]^{(EQ)}} = \exp\left(\frac{\psi_D^{(EQ)}(x) - \psi_E^{(EQ)}(x)}{k_{\rm B}T}\right)$$

where $k^0_{E\to T}(x)$, $k^0_{D\to DP}(x)$ and $k^0_{E\to D}(x)$ are first order rate constants, and others are zeroth order rate constants. In a first order reaction (such as ATP binding), the reaction rate is proportional to the concentration of reactant and the coefficient is called the first rate constant. In a zeroth order reaction, the reaction rate is independent of any concentration and is directly given by the zeroth order rate constant. In mathematical framework (12), none of the first order rate constants or the zeroth order rate constants is affected by changes in concentrations. In particular, for a set of non-equilibrium concentrations [ATP], [ADP], and $[\rm P_{\it i}]$, these rate constants stay the same. We define the free energies for the four occupancy states for non-equilibrium concentrations [ATP], [ADP], and $[\rm P_{\it i}]$

$$\psi_E(x) = \psi_E^{(EQ)}(x) + k_{\rm B}T \ln\left(\frac{[\rm ATP]}{[\rm ATP]^{(EQ)}}\right)$$
$$\psi_T(x) = \psi_T^{(EQ)}(x)$$
$$\psi_{\rm DP}(x) = \psi_{\rm DP}^{(EQ)}(x)$$
$$\psi_D(x) = \psi_D^{(EQ)}(x) + k_{\rm B}T \ln\left(\frac{[\rm P_{\it i}]}{[\rm P_{\it i}]^{(EQ)}}\right)$$

The transition rates at non-equilibrium concentrations [ATP], [ADP], and $[\rm P_{\it i}]$, satisfy detailed balance

$$\frac{k^0_{E\to T}(x)[\rm ATP]}{k_{T\to E}(x)} = \exp\left(\frac{\psi_E(x) - \psi_T(x)}{k_{\rm B}T}\right)$$
$$\frac{k_{T\to DP}(x)}{k_{DP\to T}(x)} = \exp\left(\frac{\psi_T(x) - \psi_{DP}(x)}{k_{\rm B}T}\right)$$
$$\frac{k_{\rm DP\to D}(x)}{k^0_{D\to \rm DP}(x)[\rm P_{\it i}]} = \exp\left(\frac{\psi_{\rm DP}(x) - \psi_D(x)}{k_{\rm B}T}\right)$$
$$\frac{k_{D\to E}(x)}{k^0_{E\to D}(x)[\rm ADP]} = \exp\left(\frac{\psi_D(x) - \psi_E(x) - \Delta G}{k_{\rm B}T}\right)$$

where we have used the result:

$$k_{\rm B}T\left\{\ln\left(\frac{[\rm ATP]^{(EQ)}}{[\rm ATP]}\right) - \ln\left(\frac{[\rm P_{\it i}]^{(EQ)}}{[\rm P_{\it i}]}\right) - \ln\left(\frac{[\rm ADP]^{(EQ)}}{[\rm ADP]}\right)\right\}$$
$$= \Delta G^0 - k_{\rm B}T \ln\left(\frac{[\rm ATP]}{[\rm ADP]\cdot[\rm P_{\it i}]}\right) = \Delta G$$

In summary, when we are *given* the transition rates and we are solving for the time evolution of the motor system, we only need to follow the occupancy state of the motor system; we do not need to track the number of cycles completed. This is reflected in the Markov Fokker-Planck formulation (12) in that the corresponding occupancy states

in different cycles are not distinguished explicitly in (12) and the reaction diagram used in (12) is based on composite states. An ATP hydrolysis reaction diagram based on composite states is shown in Figure 1(A). When we follow only the occupancy state, a chemical reaction goes around a closed loop it does not matter the reaction is at equilibrium or not. However, when we are studying constraints on transition rates and studying the free energies of different states, we do need to distinguish occupancy states in different cycles. As a result, in studies of energetics, a chemical reaction away from equilibrium does not go around a closed loop. Instead, it goes along an infinite helical sequence of cycles. Figure 1(B) shows an ATP hydrolysis reaction diagram which distinguishes occupancy states of different cycles and thus is suitable for studying transition rates and free energies.

It is clear that the breaking of detailed balance is caused by using the simple diagram shown in Figure 1(A) instead of using the full diagram in Figure 1(B). The diagram in Figure 1(A) is the project of the diagram in Figure 1(B) over all cycles. This projection is mathematically convenient when all transition rates are given. When we study the free energies and study the constraints on transition rates, we have to resort to the full reaction diagram in Figure 1(B).

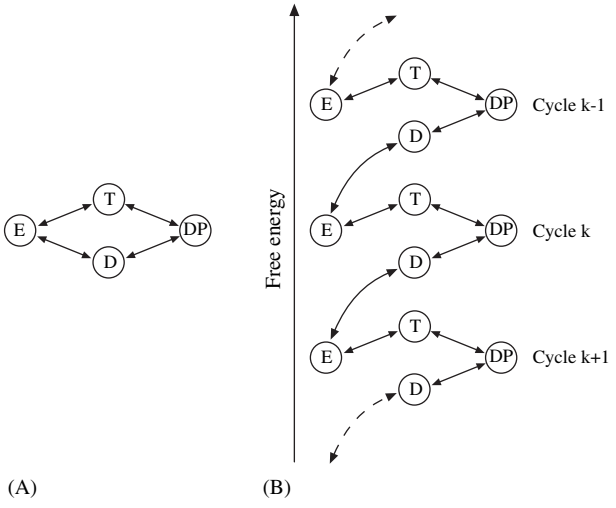


Fig. 1. Two kinds of reaction diagrams for the ATP hydrolysis cycle. (A) A diagram based on the occupancy state, showing the ATP hydrolysis as a closed loop. Indeed, if we only focus on the occupancy state of the catalytic site, at the end of each ATP hydrolysis cycle, it goes back to the **E** (empty) state. In this diagram, the number of cycles completed is not followed and the occupancy states in different cycles are not distinguished. For example, the composite state **E** consists of all **E** states in all cycles. When the system is not at equilibrium, it is inappropriate to assign a free energy to a composite state, and along such a closed loop of composite states, detailed balance is not obeyed. (B) A diagram based on the occupancy state and the number of cycles completed, showing the ATP hydrolysis as an infinite helical sequence of cycles. In this diagram, the occupancy states in different cycles are distinguished and are shown as different states. As a result, the free energy of each state can be properly defined and detailed balance is obeyed. The reaction diagram in (B) is suitable for studying the free energies and for studying the constraints on transition rates. The reaction diagram in (A) is simply a project of the diagram in (B). When the transition rates are given, the diagram in (A) is mathematically more convenient since it has finite number of (composite) states. But the diagram in (A) is not suitable for discussing energetics.

4. POTENTIAL PROFILE OF A MOLECULAR MOTOR

The small size of molecular motors has hindered, in many ways, the studies of motor mechanism. For a macroscopic motor, it is possible to observe/record directly the internal operation details of the motor. It is not yet possible to observe/record the internal motion of a single molecule motor. Nor is it possible to observe/record the chemical occupancy state of a single molecule motor. Because of the small size, molecular motors have several features that distinguish themselves from macroscopic motors. The most prominent feature of molecular motors is that the time scale of inertia is much smaller than the time scale of chemical reaction cycle in the motor. On the time scale of chemical reaction cycles, the effect of inertia is negligible. It turns out that this feature of molecular motors can be exploited to our advantage in studying motor mechanism. As we will see, negligible inertia implies that the local dynamics at a position is completely determined by the local driving force at the position since the driving force at other positions will not be able to affect the local dynamics at the current position via inertia. This give us the possibility of recovering the local driving force at any given position from the time series of motor positions measured in single molecule experiments. The local driving force at a fixed position is still stochastic for two reasons. The chemical reaction in a molecular motor has many occupancy states, each occupancy state exerting a different motor force. So as long as the chemical occupancy state is stochastic, the motor force is stochastic. Besides the motor force, the local driving force contains the Brownian force, which is also stochastic. Our long-term goal is to recover the motor force at any given position for any given occupancy state from experimental data. In most of current single molecule experiments, the chemical occupancy state cannot be observed/recorded. This severely limits our ability of recovering the motor force for each individual occupancy state. To accommodate the current experimental data, we introduce a single potential to represent the overall effect of the chemical reaction on the motor motion. We call it the motor potential profile. The significance of recovering the motor potential profile is that the potential profile reveals how the motor force changes with motor position in a motor step, which may lead to insights into how the chemical reaction is coupled to force generation. Interestingly, it is the small size (negligible inertia) of molecular motors that makes it possible to recover the potential profile from time series of motor positions. For macroscopic motors, the variation of driving force within a cycle is smoothed out by the large inertia.

REVIEW

The current experimental technologies allow us to measure forces and motions of a single motor to the precision of piconewtons and nanometers.[49–53] Time series of motor positions have been measured for various molecular motors at various mechanical loads and chemical concentrations.[15, 54–57] In the past, these measured time series of motor positions have not been fully utilized to yield all possible information about motor mechanism. Usually, only the average velocity and sometimes the randomness parameter of the motor were extracted from the measured time series of motor positions.[15, 58] In Ref. [59], we proposed the concept of motor potential profile. The chemical reaction in a molecular motor has many occupancy states, each having a different effect on the motor motion. The overall effect of the chemical reaction on the motor motion can be characterized by the motor potential profile. The extraction of the randomness parameter from time series of motor positions is a significant addition to the average velocity in deciphering motor mechanism. In the framework of kinetic models and under certain assumptions, the reciprocal of the randomness parameter tells us the number of rate-limiting chemical transitions per motor step.[15, 58] There is much more information about motor mechanism still buried in these time series of motor positions. The first step in digging out more information is to extract the motor potential profile from measured time series of motor positions. Once the reconstruction method is designed, the motor potential profile becomes a measurable entity (in addition to average velocity and randomness parameter). We hope the motor potential profile will yield more information about motor operation. In particular, the potential profile reveals how the motor force changes with position in a motor step, which provides insights into how the chemical reaction is coupled to force generation. In a power stroke motor, the chemical reaction generates an active driving force, which corresponds to a gradually decreasing potential. In a Brownian ratchet, the chemical reaction establishes a free energy barrier, which corresponds to a vertical drop followed by a flat step in free energy. Thus, the potential profile of a Brownian ratchet is a sequence of vertical free energy drops rectifying forward fluctuations. On the other hand, the potential profile of a power stroke motor is a gradually decreasing function of the motor position, generating an active force to drive the motor. Brownian ratchet and power stroke motor are just two extreme situations. The potential profile of a motor may have both vertical free energy drops and gradual down hill slopes. The motor potential profile is the link between the chemical reaction and the motor motion. We can conceptually divide the motor into two parts: first the chemical reaction generates the potential profile; then the potential profile produces the unidirectional motion. Below we will first introduce the concept of motor potential profile. Then we will discuss a robust mathematical/numerical formulation for reconstructing the motor potential profile from time series of motor positions. Basically, we use the framework of Markov Fokker-Planck equations to write the motor potential profile in terms of the steady state probability flux and the steady state probability density. The steady state probability flux is related to the average velocity, which can be calculated reliably from data. As a result, we only need to focus on how to recover the steady state probability density and how to estimate the statistical errors associated with the estimated probability density. To achieve a good balance between spatial resolutions and statistical errors, we use an adaptive spatial grid based on the data. The statistical errors are estimated, and based on the estimated errors the spatial resolution can be tuned to achieve an optimal result. The adaptive spatial grid is especially important when we have only a limited amount of experimental data.

4.1. Motor Potential Profile and a Robust Formulation for Reconstructing Motor Potential Profile

At any given position, the motor is driven by (1) the external loading force, (2) the motor force derived from potential $\phi_S(x)$ corresponding to the current chemical occupancy state $S(t)$, and (3) the Brownian force from the surrounding fluid. The motor is also subject to the passive viscous drag, which in the limit of high viscous friction (negligible inertia) balances the sum of all other forces. The over-damped Langevin Eq. (9) is rewritten below in the form of force balancing:

$$\underbrace{\zeta\frac{dx}{dt}}_{\substack{\text{Viscous}\\ \text{drag}}} = \underbrace{F}_{\substack{\text{Load}\\ \text{force}}} - \underbrace{\phi'_S(x)}_{\substack{\text{Motor}\\ \text{force}}} + \underbrace{\sqrt{2k_BT\zeta}\frac{dW(t)}{dt}}_{\substack{\text{Brownian}\\ \text{force}}} \tag{24}$$

The external loading force is usually a known constant independent of time, position and chemical state of the motor. The Brownian force is stochastic. The motor force, $\phi'_S(x)$, is also stochastic. Let us consider the sum of these three active driving forces on the right side of (24). In principle, samples of the stochastic total active force on the motor at any given position can be measured/calculated by differentiating the time path of motor position. The average of the total active force at a given position is calculated by averaging over all samples regardless of chemical states.

$$\text{Average total active force} \equiv \left\langle \zeta\frac{dx}{dt}\right\rangle_{\text{Over all states}}$$

The calculation of average total active force does not require distinguishing different chemical occupancy states in data. If individual occupancy states can be resolved completely or partially in the data, then the average total active force for each individual occupancy state may be extracted. In averaging the total active force, the stochastic Brownian force disappears. The average motor force

$\langle -\phi_S'(x)\rangle$ over all chemical state can be calculated as the average total active force over all chemical states minus the external loading force F. The average motor force is a periodic function of motor position, and is called the motor force profile.

$$\text{Motor force profile} \equiv \langle -\phi_S'(x)\rangle_{\text{Over all states}} = \left\langle \zeta \frac{dx}{dt} \right\rangle_{\text{Over all states}} - F$$

The integral of the motor force profile is a tilted periodic function of motor position, and is called the motor potential profile.

For recovering the potential profile, there are practical difficulties with the approach of differentiating and averaging the time paths. For molecular motors, the stochastic instantaneous velocity caused by the Brownian force is usually much larger than the average velocity. Consider the situation where a spherical bead with radius r, drag coefficient ζ and mass m, is driven by a constant force F. The deterministic velocity caused by the constant force F is $\langle u\rangle = F/\zeta$. In addition to the constant driving force F, the bead is constantly bombarded by the surrounding water molecules. At equilibrium, the one-dimensional root-mean-square of the instantaneous bead velocity is given by the equi-partition of energy:

$$\sqrt{\text{var}(u)} = \sqrt{\frac{k_B T}{m}}$$

Using the fact that the drag coefficient ζ is proportional to r and the mass m is proportional to r^3, we obtain an estimate on the magnitude of fluctuations relative to the mean:

$$\frac{\sqrt{\text{var}(u)}}{\langle u\rangle} = \frac{\zeta}{F}\sqrt{\frac{k_B T}{m}} \sim \frac{1}{\sqrt{r}}$$

Thus, for small beads, the fluctuations in the instantaneous velocity are larger than the average velocity. For a latex bead of 1 μm in diameter and for $F = 1$ pN, we have $\sqrt{\text{var}(u)}/\langle u\rangle \approx 26.7$. In other words, the fluctuations are 27 times as large as the mean. In such a situation, it takes roughly 700∼800 samples per motor position just to reduce the uncertainty to the magnitude of the mean value. It requires more samples to reduce the uncertainty to a small fraction of the mean value. This difficulty is compounded by other difficulties listed below:

- Differentiating the time path of motor position is sensitive to small errors in position measurements. Mathematically, since the time path of motor position is almost nowhere differentiable, differentiating it is not a well-posed mathematical problem.
- In the above, only the fluctuations caused by the Brownian force are included. For molecular motors, there are additional fluctuations caused by the stochastic chemical reaction.

To overcome the difficulties listed above, we define the motor potential profile in the mathematical framework of Fokker-Planck equation. Then in the framework of Fokker-Planck equation, we express the motor potential profile in terms of the steady state probability flux and the steady state probability density.

Consider the steady state of Fokker-Planck equation (12). Summing over all chemical states, and using the property $\sum_{i=1}^{N} k_{ij}(x) = 0$, we have

$$0 = D\frac{\partial}{\partial x}\left(\frac{-F+\psi'(x)}{k_B T}\rho + \frac{\partial \rho}{\partial x}\right) \tag{25}$$

where $\rho(x) = \sum_{i=1}^{N} \rho_i(x)$ is the steady state probability density regardless of the chemical occupancy state. The motor potential profile, $\psi(x)$, is defined as

$$\psi'(x) = \frac{1}{\rho(x)}\sum_{i=1}^{N} \rho_i(x)\phi_i'(x) \tag{26}$$

In Ref. [24] an effective potential for a two-state model was considered in a similar way. In Eq. (25), the steady state probability density $\rho(x)$ behaves as if the motor were driven by a single potential $\psi(x)$. In this sense, the motor potential profile represents the overall effect of the chemical reaction on the motor motion. Since $\psi'(x)$ defined in (26) is periodic with period l, $\psi(x)$ is a tilted periodic function and can be written as

$$\psi(x) = \phi(x) + \Delta\psi \cdot \frac{x}{l}$$

where $\phi(x)$ is a periodic function with period l and $\Delta\psi = \psi(l) - \psi(0)$. $(-\Delta\psi)$ can be viewed as the potential energy made available per displacement l from the chemical reaction to driving the motor motion. Equation (26) is a mathematical definition. (26) is not operational in constructing the potential profile $\psi(x)$ from the data since neither of $\phi_j(x)$ nor $\rho_j(x)$ can be recovered from the data. If we can recover $\phi_j(x)$ and $\rho_j(x)$ from the data, we will not bother to reconstruct the potential profile $\psi(x)$. To design a more robust formulation for reconstructing $\psi(x)$, we rewrite steady state Eq. (25) in the form of constant flux:

$$D\left(\frac{-F+\psi'(x)}{k_B T}\rho + \frac{\partial \rho}{\partial x}\right) = -J \tag{27}$$

where J is the steady state probability flux, a constant independent of motor position x. Dividing (27) by $\rho(x)$ and integrating with respect to x yields

$$\frac{\psi(x)}{k_B T} = \frac{F\cdot x}{k_B T} - \ln(\rho(x)) - \frac{J}{D}\int_0^x \frac{1}{\rho(s)}\,ds \tag{28}$$

Here, we have dropped the integration constant because the motor potential profile, like other potentials, can only be determined up to an additive constant. In expression (28), the external loading force F is known and the probability flux is related to the average velocity as $J = \langle u\rangle/l$. The average velocity $\langle u\rangle$ can be calculated reliably from the data. Therefore, to construct the motor potential profile

REVIEW

$\psi(x)$, we only need to recover the steady state probability density $\rho(x)$ from the time series of motor positions. Mathematically and numerically, recovering $\rho(x)$ from the time series is certainly more viable and reliable than differentiating the time path to calculate the stochastic force.

4.2. Reconstruction of Steady State Probability Density and Motor Potential Profile from Experimental Data

In the above, we just reduced the problem of constructing motor potential profile to that of recovering steady state probability density from the time series of motor positions. Here we review the numerical procedure proposed in Ref. [60] for recovering the steady state probability density. Consider a time series of M motor positions: $\{x_1, x_2, \ldots, x_M\}$. The average velocity and the probability flux are estimated as

$$\langle u\rangle \approx \frac{x_M - x_1}{t_M - t_1}, \quad J = \frac{\langle u\rangle}{l}$$

We shift each motor position x_j by an integer multiple of l to make it fall in the interval $[0,\ l)$:

$$y_j = x_j - k_j l \in [0,\ l), \quad k_j = \text{integer},\ j = 1, 2, \ldots, M$$

A straightforward way of recovering probability density is the histogram method. Let us examine the behavior of the histogram method in constructing $\psi(x)$. The findings will motivate us to adopt an adaptive grid instead of a uniform grid. Suppose $\rho(x)$ is the exact probability density, and $\rho(x)+\Delta\rho(x)$ is the probability density estimated using the histogram method form the data. Here $\Delta\rho(x)$ is the absolute error and $\Delta\rho(x)/\rho(x)$ is the relative error in the estimated probability density. In expression (28), $\psi(x)$ contains a term $\ln(\rho(x))$. When the exact probability density $\rho(x)$ is replaced by the estimated probability density $\rho(x)+\Delta\rho(x)$, the error caused by the term $\ln(\rho(x))$ is

$$\ln(\rho(x)+\Delta\rho(x)) - \ln(\rho(x)) = \ln\left[1 + \frac{\Delta\rho(x)}{\rho(x)}\right] \approx \frac{\Delta\rho(x)}{\rho(x)} \tag{29}$$

Thus, the relative error in the estimated probability density $\rho(x)$ contributes to the absolute error in the estimated potential profile $\psi(x)$. So we need to control the relative error in the estimated $\rho(x)$. The relative error in the estimated $\rho(x)$ comes from two sources: the spatial discretization error and the statistical error. The spatial discretization error is determined by the bin size we use in the histogram method, i.e., big bin size resulting in low spatial resolution and large spatial discretization error. The statistical error is affected by the number of samples in each bin. When the total number of samples is fixed, the statistical error is also determined by the bin size, i.e., big bin size resulting in large number of samples in each bin and small statistical error. So, when the total number of samples is fixed, the bin size should be selected to minimize both the spatial discretization error and the statistical error. In the histogram method, the relative statistical error in the estimated $\rho(x)$ is approximately given by

$$\frac{\Delta\rho(x)}{\rho(x)} \sim \frac{1}{\sqrt{N_x}} \tag{30}$$

where N_x is the number of samples (y_j's) in the bin surrounding x. If we use bins of equal size, then N_x is proportional to $\rho(x)$. As a result, the relative statistical error in the estimated $\rho(x)$ is large in regions where $\rho(x)$ is small. To keep the relative error uniformly small, we should use bins of variable sizes according to the data so that each bin contains about the same number of data points. This is how we are going to form our numerical grid.

A drawback of the histogram method is that the probability density estimated is a discontinuous step function. It is non-trivial to interpolate this step function to make it continuous while keeping it positive and integrating to one. One may set $\rho(x)$ to be continuous and piecewise linear, and use the maximum likelihood method to estimate $\rho(x)$. Given that the true value of $\rho(x)$ may differ by orders of magnitude at different locations, linear interpolation does not offer the best way for representing $\rho(x)$. A more robust and more efficient way of representing $\rho(x)$ is motivated by considering the Boltzmann distribution corresponding to potential $\psi(x)$: $\rho^{(\mathrm{EQ})}(x) \propto \exp(-\psi(x)/k_{\mathrm{B}}T)$. Although the steady state probability density $\rho(x)$ will not be an equilibrium, it is reasonable to set $\rho(x)$ to the form of Boltzmann distribution and set the exponent to be a continuous piecewise linear function. The corresponding $\rho(x)$ is a continuous piecewise exponential function, which guarantees automatically the positivity of $\rho(x)$. Another advantage of using piecewise exponential probability density is that the maximum likelihood formulation has a relatively simple form since multiplying probability densities corresponds to adding exponents. The step-by-step procedure for reconstructing steady state probability density $\rho(x)$ and motor potential profile $\psi(x)$ is described below:

Step 1. We divide the interval $[0,\ l)$ into cells according to the data $\{y_1, y_2, \ldots, y_M\}$:

$$0 = a_0 < a_1 < a_2 < \cdots < a_{m-1} < a_m = l$$

To keep the relative error uniformly small, we require that each cell $[a_{j-1},\ a_j)$ contains at least m_2 data points. For computational efficiency, we also require that the size of each cell $(a_j - a_{j-1})$ is not smaller than a prescribed size Δx. Specifically, we start with $a_0 = 0$. Once we have a_j, we select a_{j+1} as

$$a_{j+1} = \min\{a \mid a \geq a_j + \Delta x \text{ and } [a_j, a) \text{ contains at least } m_2 \text{ data points}\}$$

Note that the number of cells m is not prescribed. Rather, m is automatically determined in the process of sub-dividing $[0, l)$. The sub-division process is controlled by specifying parameters m_2 and Δx. We should choose m_2 and Δx to make the number of cells m much smaller than the number of data points M.

Step 2. For maintaining the positivity of $\rho(x)$ and for accurate spatial representation, we set $\rho(x)$ to be exponential in each cell and be continuous at cell boundaries. Mathematically, $\rho(x)$ is completely specified by $(b_0, b_1, \ldots, b_m)$, and has the form

$$\rho(x) = \exp\left[b_{j-1} + \frac{b_j - b_{j-1}}{a_j - a_{j-1}}(x - a_{j-1})\right] \qquad x \in [a_{j-1}, a_j],\ j = 1, 2, \ldots, m \quad (31)$$

where $b_j = \ln(\rho(a_j))$. Since $\rho(x)$ integrates to one, $(b_0, b_1, \ldots, b_m)$ must satisfy

$$\sum_{j=1}^{m} \int_{a_{j-1}}^{a_j} \exp\left[b_{j-1} + \frac{b_j - b_{j-1}}{a_j - a_{j-1}}(x - a_{j-1})\right] dx = 1$$

which leads to a constraint on $(b_0, b_1, \ldots, b_m)$:

$$S(b_1, b_2, \ldots, b_m) \equiv \sum_{j=1}^{m} \frac{a_j - a_{j-1}}{b_j - b_{j-1}} [\exp(b_j) - \exp(b_{j-1})] = 1 \quad (32)$$

Because $\rho(x)$ is periodic, we have $b_0 = b_m$. So we only need to determine $(b_1, b_2, \ldots, b_m)$. We use the maximum likelihood method to determine $(b_1, b_2, \ldots, b_m)$. To proceed with the maximum likelihood method, we separate data points $\{y_1, y_2, \ldots, y_M\}$ into m groups shown below, each group containing data points falling in each individual cell.

$$\begin{aligned} &\text{Group 1:} && y_{1,1}, y_{1,2}, \ldots, y_{1,M_1} \in [a_0, a_1) \\ &\text{Group 2:} && y_{2,1}, y_{2,2}, \ldots, y_{2,M_2} \in [a_1, a_2) \\ &&& \vdots \\ &\text{Group } m\text{:} && y_{m,1}, y_{m,2}, \ldots, y_{m,M_m} \in [a_{m-1}, a_m) \end{aligned}$$

The likelihood of observing the data $\{y_1, y_2, \ldots, y_M\}$ is given by

$$\begin{aligned} &L(b_1, b_2, \ldots, b_m) \\ &\quad \equiv \prod_{i=1}^{M} \rho(y_i \mid b_1, b_2, \ldots, b_m) \\ &\quad = \prod_{j=1}^{m} \prod_{i=1}^{M_j} \exp\left[b_{j-1} + \frac{b_j - b_{j-1}}{a_j - a_{j-1}}(y_{j,i} - a_{j-1})\right] \\ &\quad = \prod_{j=1}^{m} \exp[(M_j - q_j) b_{j-1} + q_j b_j] = \exp\left[M \sum_{j=1}^{m} c_j b_j\right] \end{aligned} \quad (33)$$

where quantity c_j is defined as

$$q_j = \frac{1}{a_j - a_{j-1}} \sum_{i=1}^{M_j} (y_{j,i} - a_{j-1}), \quad j = 1, 2, \ldots, m$$

$$c_j = \frac{1}{M}(M_{j+1} - q_{j+1} + q_j) > 0, \quad j = 1, 2, \ldots, m$$

$$c_m = \frac{1}{M}(M_1 - q_1 + q_m) > 0$$

In this approach, $\{c_1, c_2, \ldots, c_m\}$ along with the numerical grid $\{a_0, a_1, \ldots, a_m\}$ is a sufficient statistics of the data. $\{c_1, c_2, \ldots, c_m\}$ satisfies $\sum_j^m c_j = 1/M \sum_j^m M_j = 1$.

The logarithm of the likelihood function is

$$F(b_1, b_2, \ldots, b_m) \equiv \frac{1}{M} \ln[L(b_1, b_2, \ldots, b_m)] = \sum_{j=1}^{m} c_j b_j$$

Step 3. $(b_1, b_2, \ldots, b_m)$ is determined by maximizing the logarithm of the likelihood function.[61–63]

$$\arg\max_{S(b_1, b_2, \ldots, b_m) = 1} F(b_1, b_2, \ldots, b_m) \quad (34)$$

In the constrained maximization problem (34), the objective function $F(b_1, b_2, \ldots, b_m)$ is linear while the constraint $S(b_1, b_2, \ldots, b_m) = 1$ is nonlinear. We switch the roles of $S(b_1, b_2, \ldots, b_m)$ and $F(b_1, b_2, \ldots, b_m)$, and consider the minimization problem

$$\arg\min_{F(b_1, b_2, \ldots, b_m) = 0} S(b_1, b_2, \ldots, b_m) \quad (35)$$

It can be derived that if $(\beta_1, \beta_2, \ldots, \beta_m)$ is the solution of minimization problem (35), then

$$(\beta_1, \beta_2, \ldots, \beta_m) - \ln[S(\beta_1, \beta_2, \ldots, \beta_m)] \quad (36)$$

is the solution of maximization problem (34).[60] So we only need to solve problem (35). Since $F(b_1, b_2, \ldots, b_m)$ is linear, we rewrite problem (35) as an unconstrained minimization problem

$$\arg\min_{(b_1, b_2, \ldots, b_{m-1})} S\left(b_1, \ldots, b_{m-1}, \frac{-1}{c_m} \sum_{j=1}^{m-1} c_j b_j\right) \quad (37)$$

There are many good numerical methods for solving unconstrained non-quadratic minimization problems. For example, we can use non-linear conjugate gradient method[64, 65] to solve problem (37). Let $(\beta_1, \ldots, \beta_{m-1})$ be the solution of problem (37). Once $(\beta_1, \ldots, \beta_{m-1})$ is computed, we set

$$\beta_m = \frac{-1}{c_m} \sum_{j=1}^{m-1} c_j \beta_j$$

$$(b_1, b_2, \ldots, b_m) = (\beta_1, \beta_2, \ldots, \beta_m) - \ln[S(\beta_1, \beta_2, \ldots, \beta_m)]$$

Then $\mathbf{b} = (b_1, b_2, \ldots, b_m)$ is the solution of non-linearly constrained maximization problem (34).

The recovered probability density function is given by

$$\rho(x \mid \mathbf{b}) = \exp\left[b_{j-1} + \frac{b_j - b_{j-1}}{a_j - a_{j-1}}(x - a_{j-1})\right] \qquad x \in [a_{j-1}, a_j],\ j = 1, 2, \ldots, m$$

Once we know the probability density, it is not difficult to reconstruct the whole function of the motor potential profile using expression (28). But the statistical error of a function is a bit more difficult to deal with. To simplify the

REVIEW

estimation of statistical error, we focus on the values of the motor potential profile at numerical grid points. Mathematically, we use the vector

$$\psi(\mathbf{b}) \equiv \{\psi(a_j \mid \mathbf{b}), j = 0, 1, \ldots, m\}$$

to represent the motor potential profile. $\psi(a_j \mid \mathbf{b})$ has the expression:

$$\begin{aligned}\frac{\psi(a_j \mid \mathbf{b})}{k_B T} &= \frac{f \cdot a_j}{k_B T} - \log[\rho(a_j \mid \mathbf{b})] - \frac{J}{D}\int_0^{a_j} \frac{1}{\rho(s \mid \mathbf{b})}\, ds \\ &= \frac{f \cdot a_j}{k_B T} - b_j - \frac{J}{D}\sum_{l=1}^{j} \frac{a_l - a_{l-1}}{(-b_l) - (-b_{l-1})} \\ &\quad \times [\exp(-b_l) - \exp(-b_{l-1})] \end{aligned} \tag{38}$$

Step 4. In this step, we estimate the statistical error in vector $\mathbf{b}$ and the statistical error in vector $\psi(\mathbf{b})$. We first define several shorthand notations.

Let $\mathbf{y} = \{y_1, y_2, \ldots, y_M\}$ denote the given data, shifted into one period.

Let $\mathbf{b}(\mathbf{y}) = (b_1(\mathbf{y}), b_2(\mathbf{y}), \ldots, b_m(\mathbf{y}))$ denote the solution of (34) using data $\mathbf{y}$. $\mathbf{b}(\mathbf{y})$ is a vector.

Let $\rho(x \mid \mathbf{b}(\mathbf{y}))$ denote the probability density estimated using data $\mathbf{y}$. $\rho(x \mid \mathbf{b}(\mathbf{y}))$ is a function.

Let $\psi(\mathbf{b}) \equiv \{\psi(a_j \mid \mathbf{b}), j = 0, 1, \ldots, m\}$ denote the potential profile on the numerical grid, estimated using data $\mathbf{y}$. $\boldsymbol{\psi}\mathbf{b}(\mathbf{y}))$ is a vector.

Let $\mathbf{d}(\mathbf{y}) = \{d_1, d_2, \ldots, d_M\}$ be a set of M independently identically distributed samples drawn according to the estimated probability density $\rho(x \mid \mathbf{b}(\mathbf{y}))$. $\mathbf{d}(\mathbf{y})$ is a random vector.

Let $\mathbf{b}(\mathbf{d})$ denote the solution of (34) usings data $\mathbf{d}(\mathbf{y})$. $\mathbf{b}(\mathbf{d})$ is a vector.

Note that if we draw an infinite set of independently identically distributed samples according to $\rho(x \mid \mathbf{b}(\mathbf{y}))$ and use this infinite data set in solving (34), then $\mathbf{b}(\mathbf{y})$ will be recovered exactly. The statistical error in $\mathbf{b}(\mathbf{y})$, caused by the finite sample size, is estimated as the standard deviation of random vector $\mathbf{b}(\mathbf{d})$:

$$\text{Error}[b_j(\mathbf{y})] \sim \text{std}[b_j(\mathbf{d})]_{\mathbf{d}}, \quad j = 1, 2, \ldots, m$$

Here $\mathbf{d}$ is a random vector and the average is taken with respect to $\mathbf{d}$. Similarly, the statistical error in $\boldsymbol{\psi}(\mathbf{b}(\mathbf{y}))$ is estimated as the standard deviation of random vector $\boldsymbol{\psi}(\mathbf{b}(\mathbf{d}))$

$$\text{Error}[\psi(a_j \mid \mathbf{b}(\mathbf{y}))] \sim \text{std}[\psi(a_j \mid \mathbf{b}(\mathbf{d}))]_{\mathbf{d}}, \quad j = 1, 2, \ldots, m$$

In the numerical implementation, we draw n independent sets of random numbers, each set consisting of M independently identically distributed samples drawn according to $\rho(x \mid \mathbf{b}(\mathbf{y}))$:

$$\begin{aligned} &\text{Set 1:} \quad \mathbf{d}^{(1)} = \{d_1^{(1)}, d_2^{(1)}, \ldots, d_M^{(1)}\} \\ &\text{Set 2:} \quad \mathbf{d}^{(2)} = \{d_1^{(2)}, d_2^{(2)}, \ldots, d_M^{(2)}\} \\ &\qquad\qquad \vdots \\ &\text{Set } n\text{:} \quad \mathbf{d}^{(n)} = \{d_1^{(n)}, d_2^{(n)}, \ldots, d_M^{(n)}\} \end{aligned}$$

With each data set $\mathbf{d}^{(i)}$, $i = 1, 2, \ldots, n$, we solve (34) for $\mathbf{b}(\mathbf{d}^{(i)})$. In solving (34), we keep the numerical grid $\{a_0, a_1, \ldots, a_m\}$ unchanged. That is, we use the numerical grid calculated from the original data $\mathbf{y}$; we do not recalculate the numerical grid for each data set $\mathbf{d}^{(i)}$. The statistical error in $b_j(\mathbf{y})$ is estimated as

$$\begin{aligned} \langle b_j(\mathbf{d})\rangle_{\mathbf{d}} &\approx \frac{1}{n}\sum_{i=1}^{n} b_j(\mathbf{d}^{(i)}), \quad j = 1, 2, \ldots, m \\ \text{std}[b_j(\mathbf{d})]_{\mathbf{d}} &\approx \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(b_j(\mathbf{d}^{(i)}) - \langle b_j(\mathbf{d})\rangle_{\mathbf{d}})^2} \\ &\qquad j = 1, 2, \ldots, m \end{aligned} \tag{39}$$

The statistical error in $\psi(a_j \mid \mathbf{b}(\mathbf{y}))$ is estimated as

$$\begin{aligned} &\langle \psi(a_j \mid \mathbf{b}(\mathbf{d}))\rangle_{\mathbf{d}} \approx \frac{1}{n}\sum_{i=1}^{n}\psi(a_j \mid \mathbf{b}(\mathbf{d}^{(i)})), \quad j = 1, 2, \ldots, m \\ &\text{std}[\psi(a_j \mid \mathbf{b}(\mathbf{d}))]_{\mathbf{d}} \\ &\quad \approx \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\psi(a_j \mid \mathbf{b}(\mathbf{d}^{(i)})) - \langle \psi(a_j \mid \mathbf{b}(\mathbf{d}))\rangle_{\mathbf{d}})^2} \\ &\qquad\qquad j = 1, 2, \ldots, m \end{aligned} \tag{40}$$

Recently Walton proposed to use the hidden Markov model to analyze single molecule experimental data on kinesin.[66] Both Walton's study and our study share the same general approach of combining mathematical models for the motor with statistical methods to process the data. These two studies also share the same goal of extracting more information about the motor mechanism from experimental data. The specific goals and the specific mathematical models and statistical methods used are different. In Walton's study, the kinesin is described by a kinetic model with a few states; the statistical tool is the hidden Markov model; and the goal is to extract the time series of kinesin states from the measured time series of bead positions. Since the goal is to extract the stochastic time series of kinesin states instead of extracting an average quantity, it requires high time resolution (many data points near each time instance). In Walton's study, it was found that the time resolution of the current experimental data is not high enough for extracting the intermediate states.[66] This, however, should not be viewed as a defect of Walton's study. Rather, it should be viewed as a design guide for future experiments and for future experimental technologies. In our study, the motor is described by a Fokker-Planck equation; the statistical tool is the maximum likelihood method; and the goal is to extract the motor potential profile, which is an average entity. Both Walton's study and our study demonstrate that when the general statistical methods are combined with the specific mathematical models describing the underlying physical process, more information may be uncovered.

5. EFFICIENCIES OF MOLECULAR MOTORS

In this section, we discuss various efficiency measures for molecular motors. We start with the thermodynamic efficiency, which measures the efficiency of energy conversion.

5.1. Thermodynamic Efficiency

In single molecule experiments, if a motor works against a conservative force exerted by an external agent (such as the force from a laser trap or force clamp), then the energy input is the chemical free energy consumed by the motor and the energy output is the potential energy increase in the external agent that exerts the conservative force. In this case, the thermodynamic efficiency is defined as the energy conversion efficiency:

$$\eta_{\text{Thermodynamics}} = \frac{(-F)\langle u\rangle}{(-\Delta G)\langle r\rangle} \tag{41}$$

where $\Delta G < 0$ is the free energy change in one reaction cycle, F the external force, $\langle u\rangle$ the average velocity and $\langle r\rangle$ the average reaction rate (average number of reaction cycles per time). Here the external force is defined as positive if it is in the same direction as the motion of the unloaded motor. So a loading force that opposes the motor motion is negative. In (41), the numerator is the energy output per unit time to the external agent and the denominator is the energy consumed per unit time in the chemical reaction.

5.2. Stokes Efficiency

In many single molecule experiments it is only possible to load a molecular motor by manipulating the viscous drag from the fluid medium. For example, Yasuda et al.[54] measured the rotational velocities of the F_1 ATPase motor driving actin filaments of various sizes. In this situation, there is no energy output to increase the potential energy of an external agent, and consequently, the thermodynamic efficiency does not apply. It is tempting to define the heat generated via the motor motion divided by the chemical energy consumption as an efficiency measure. However, an efficiency measure must be less than 100%. As we will see, *the heat generated via the motor motion may be negative or may exceed the free energy consumption*. To illustrate this subtle issue, we show two examples. In the first example, the work done to the surrounding fluid via motor motion (the heat generated via the motor motion) is negative. That is, the kinetic energy flows from the motor motion to the fluid environment is negative. The kinetic energy actually flows from the fluid environment to the motor motion and then to the catalytic site and finally to the outside. The kinetic energy flowing out of the catalytic site may go back to the fluid environment via the interaction between the catalytic site and the fluid environment and/or it may go to reaction products in the form of increased enthalpy. The increased enthalpy happens when ions are driven by a large concentration gradient to diffuse against a small voltage difference. After the reaction, ions end up on the high voltage side with increased enthalpy. This is in sharp contrast with the situation of macroscopic motors. In a macroscopic motor, the kinetic energy flows from the catalytic site (combustion cylinder) to the motor and then to the surrounding environment in the form of heat. In the second example, the work done to the surrounding fluid via motor motion exceeds the free energy consumption of the motor. In this case, the kinetic energy originated at the catalytic site comes from two sources: enthalpy in the reactant and the kinetic energy captured (borrowed) by the catalytic site from fluid environment. So the total kinetic energy flowing from the catalytic site to the motor motion may be larger then the free energy change of the reaction cycle. This happens when ions are driven by a large voltage difference to go against a small concentration gradient. The free energy of an ion going from the low concentration side to the high concentration side is smaller than the voltage difference because part of the voltage difference is off-set by the concentration gradient. However, the heat generated via the motion of ion going from the low concentration side to the high concentration side is given by the voltage difference. Thus, the heat generated via the ion motion exceeds the free energy consumption.

In Figure 2, ① → ② → ③ represents one half of one reaction cycle. The second half, ③ → ④ → ⑤, is identical to the first half except that it is shifted in the spatial direction by half of the motor step and shifted in the free energy direction by half of ΔG. The motor system starts at ①. Potential $\phi_1(x)$ does not directly drive the motor forward. Rather, for the motor to move forward, it has to diffuse uphill along potential $\phi_1(x)$ from ① to ②. The kinetic energy for the motor to move from ① to ② comes from thermal bombardments by fluid molecules (the heat in the surrounding fluid). In the process of ① to ②, the surrounding fluid does work on the motor motion. In other words, the work done to the surrounding fluid via motor motion is negative. Of course, diffusing against an uphill potential is not sustainable. It is impossible to extract heat repeatedly from environment to do work for free. The work done by the surrounding fluid on the motor motion and the uphill fluctuation from ① to ② is rectified by the chemical transition from ② to ③. During the chemical transition from ② to ③, the catalytic site of the motor may exchange heat with the surrounding fluid. But no work is done via the motor motion during the chemical transition. In ① → ② → ③, the heat generated via the motor motion is negative.

In Figure 3, ① → ② → ③ represents one half of one reaction cycle. The motor system starts at ①. For the motor to move forward, it has to first make an uphill chemical

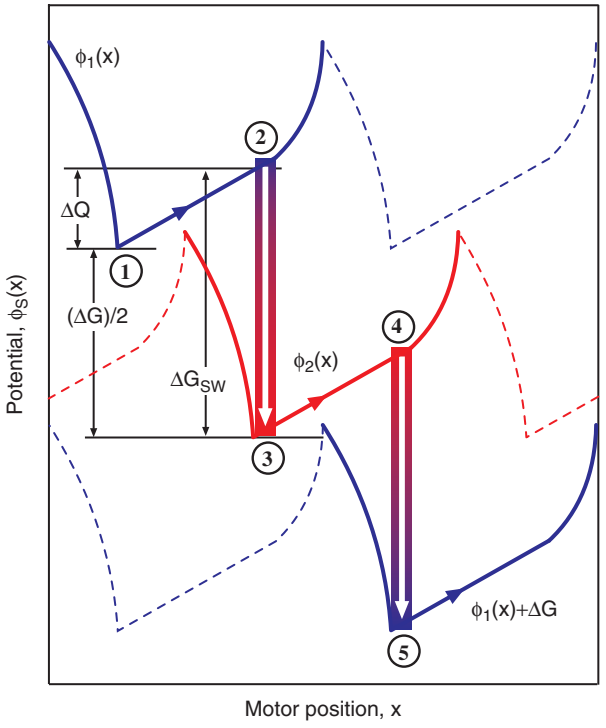


Fig. 2. A hypothetical motor system in which the work done to the surrounding fluid via motor motion is negative.

transition from ① to ②, a change of occupancy that will switch the motor system from potential $\phi_1(x)$ to potential $\phi_2(x)$. No work is done via the motor motion during the chemical transition. Once on potential $\phi_2(x)$, the motor is driven by $\phi_2(x)$ from ② to ③. In the process of downhill sliding from ② to ③, kinetic energy flows from potential $\phi_2(x)$ to the motor motion then as heat to the surrounding fluid. In this process, the amount of heat generated via the motor motion is determined by the positions of ② and ③ on potential $\phi_2(x)$, and is independent of the free energy difference between ① and ②. If we shift $\phi_1(x)$ vertically relative to $\phi_2(x)$, it will change, ΔG_{SW}, the free energy barrier for the transition from ① to ②. But the vertical shifting will not change, ΔQ, the amount of heat generated via the motor motion in the sliding from ② to ③. When ΔG_{SW} is negative, as shown in Figure 3, ΔQ is larger than $(-\Delta G)/2$, the free energy consumed in the half cycle ① → ② → ③. A possible realization of the motor system shown in Figure 3 is the F_1 ATPase motor operating in low ATP concentration. The power stroke driven by the ATP binding transition from weak binding to strong binding does a fixed amount of work to the surrounding fluid via motor motion.[2,54] This fixed amount of work is not affected by the ATP concentration in solution. The ATP concentration in solution affects the free energy change of the hydrolysis cycle and affects the frequency of ATP binding. As a result, at low ATP concentration, the ATP hydrolysis free energy is lower and the frequency of power stroke is also lower but the amount of work done via motor motion stays the same. When the ATP concentration is low, the amount of work done via motor motion exceeds the free energy consumption.

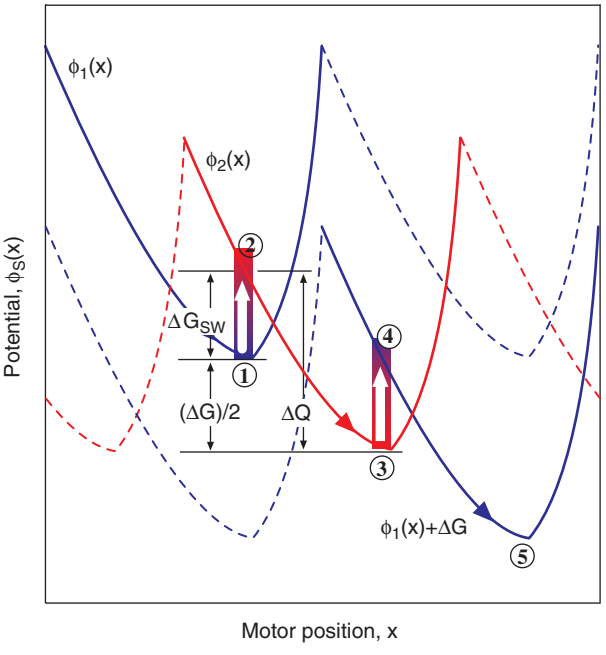


Fig. 3. A hypothetical motor system in which the work done to the surrounding fluid via motor motion exceeds the free energy consumption of the motor.

The above examples demonstrate that for molecular motors, the amount of work done to the surrounding fluid via motor motion is not a suitable candidate for replacing the numerator of the thermodynamic efficiency. Again, this is in sharp contrast with the situation of macroscopic motors. For macroscopic motors, the amount of work done to the surrounding fluid via motor motion is frequently used to replace the energy output to the external agent exerting the conservative loading force. This is justified for macroscopic motors. For macroscopic motors, the effects of thermal fluctuation are negligible and the large inertia makes the instantaneous velocity nearly constant over many reaction cycles. As a result, a macroscopic motor does not see any difference whether it is loaded by a conservative force or it is loaded by a friction force. For molecular motors, however, thermal excitations play an important role both in the mechanical motion and in the chemical reaction. Furthermore, the negligible inertia makes molecular motors very sensitive to changes in driving force and/or drag force. From the point of kinetic energy flow, a macroscopic motor has a large amount of stored kinetic energy. Consequently, the motion of a macroscopic motor is only affected by the total amount of kinetic energy sent from the reaction site (combustion cylinder) to the motor motion; the specific spatial/temporal pattern of the kinetic energy flow is not important. In contrast, a molecular motor has almost no stored kinetic

energy, and the motion of a molecular motor is dominated by thermal excitations and is highly affected by the specific spatial/temporal pattern of the kinetic energy flow.

In Refs. [39, 67], the Stokes efficiency was defined to measure how efficiently the motor is utilizing the chemical energy to generate a unidirectional motion in a viscous fluid medium:

$$\eta_{\text{Stokes}} = \frac{\zeta \langle u \rangle^2}{(-\Delta G)\langle r \rangle + F \langle u \rangle} \tag{42}$$

In Ref. [39], we proved rigorously that for a motor system governed by Fokker-Planck equation (12), the Stokes efficiency is between zero and one, and thus, it is a properly defined efficiency measure. Note that although the numerator of the Stokes efficiency, $\zeta \langle u \rangle^2$, has the dimension of energy per unit time, but $\zeta \langle u \rangle^2$ does not have a clear thermodynamic meaning. Therefore, it is not possible to argue directly from thermodynamic that the Stokes efficiency is bounded by 100%. The denominator of the Stokes efficiency, $(-\Delta G)\langle r \rangle + F \langle u \rangle$, is the free energy consumption per unit time by the motor system.

5.3. Relation Between the Thermodynamic Efficiency and the Stokes Efficiency

As specified in the definition, the thermodynamic efficiency measures the efficiency of energy conversion and is the ratio of the energy output to the external agent to the free energy consumption in chemical reaction. The energy output to the external agent is highly affected by the conservative loading force exerted on the motor by the external agent. Most macroscopic motors have gears so a different gear may be used to work efficiently in response to a different situation. Most molecular motors, however, do not have the ability of changing the coupling ratio. Here the coupling ratio is defined as the average displacement per reaction cycle. For example, a kinesin dimer moves forward along a microtubule filament by a step of about 8 nm in each ATP hydrolysis cycle regardless of the loading force. For a motor with a fixed coupling ratio, the free energy consumption per unit displacement in the chemical reaction is fixed, not affected by the loading force from the external agent. The energy output per unit displacement to the external agent, however, is proportional to the conservative loading force. As a result, for a motor with a fixed coupling ratio, the thermodynamic efficiency is proportional to the conservative loading force as long as the loading force is not too large so the motor can still move forward against the loading force. For a motor with a fixed coupling ratio in which a chemical reaction cycle is coupled to a displacement of l, we have $\langle u \rangle / \langle r \rangle = l$ and the thermodynamic efficiency has the simple form

$$\eta_{\text{Thermodynamics}} = \frac{(-F) \cdot l}{(-\Delta G)}$$

When the loading force is large enough to drive the motor backward or detach the motor from the polymer track, the thermodynamic efficiency either does not apply or has to be modified to measure the efficiency of converting the potential energy from the external agent to the chemical free energy. When the loading force is just large enough to stall the motor without driving it backward or detaching it from the polymer track, the thermodynamic efficiency attains its maximum and the maximum is given by

$$\eta_{\text{Thermodynamics}}(\text{at stall}) = \frac{(\text{conservative stall force}) \cdot l}{(-\Delta G)} \tag{43}$$

Since the thermodynamic efficiency is maximized by slowing down the motor to near stall, we like to examine what happens to the Stokes efficiency when the motor is slowed down to near stall by increasing the drag coefficient. First, in the absence of a conservative loading force, for a motor with a fixed coupling ratio in which a chemical reaction cycle is coupled to a displacement of l, we have $\langle u \rangle / \langle r \rangle = l$ and the Stokes efficiency has the form

$$\eta_{\text{Stokes}} = \frac{\zeta \langle u \rangle l}{(-\Delta G)}$$

As we increase the drag coefficient to stall the motor, the viscous stall drag is measured as

$$\text{viscous stall drag} = \lim_{\zeta \to \infty} \zeta \langle u(\zeta) \rangle$$

At stall, the Stokes efficiency has the expression:

$$\eta_{\text{Stokes}}(\text{at stall}) = \frac{(\text{viscous stall drag}) \cdot l}{(-\Delta G)} \tag{44}$$

Comparing (3) and (4), we see that, at stall, both the thermodynamic efficiency and the Stokes efficiency have the form

$$\eta(\text{at stall}) = (\text{conservative stall force OR viscous stall drag}) \cdot l/(-\Delta G)$$

Thus, for a motor with a fixed coupling ratio, to see how the thermodynamic efficiency is different from the Stokes efficiency, we only need to focus on how the conservative stall force is different from the viscous stall drag. For that purpose, we consider a simple situation where a motor is driven by a tilted periodic potential and by a possible external conservative loading force. A tilted periodic potential can be decomposed into a bumpy potential and a tilting force:

$$\psi(x) = \phi(x) + \alpha \cdot l$$

where $\phi(x)$ is a periodic bumpy potential and α is the tilting force. In this case, the potential change is tightly coupled with motor motion and the conservative stall force is given by

$$\text{conservative stall force} = \alpha$$

REVIEW

In the absence of a conservative loading force, the motor is only loaded by the viscous drag. The steady state probability density and the steady state probability flux of the motor satisfy the differential equation

$$D\left(\frac{\alpha+\phi'(x)}{k_{\mathrm{B}}T}\rho+\frac{\partial\rho}{\partial x}\right)=-J$$

which is derived from the steady state version of the Fokker-Planck equation. In addition, the probability density is constrained by $\int_0^l \rho(x)\,dx = 1$. The steady state probability flux can be solved from the differential equation coupled with the constraint. The average velocity has the expression:

$$\zeta\langle u(\zeta)\rangle = \alpha\cdot\underbrace{B(x;\{\phi(x)\})}_{\text{Effect of potential }\phi(x)}$$

where $B(x;\{\phi(x)\})$ contains the effect of potential $\phi(x)$ and is given by

$$B(x;\{\phi(x)\}) = \frac{\int_0^l \exp\left(\frac{-\alpha x}{k_{\mathrm{B}}T}\right)dx}{\int_0^l \exp\left(\frac{-\alpha x}{k_{\mathrm{B}}T}\right)\int_0^1 \exp\left(\frac{\phi(x+ql)-\phi(ql)}{k_{\mathrm{B}}T}\right)dq\,dx}$$

Notice that in this simple case, the quantity $\zeta\langle u(\zeta)\rangle$ does not depend on drag coefficient ζ any more. Mathematically, it can be shown that $B(x;\{\phi(x)\}) \le 1$ for any $\phi(x)$ and $B(x;\{\phi(x)\}) = 1$ only when $\phi(x) \equiv$ const. Thus, we obtain

$$(\text{viscous stall drag}) \le (\text{conservative stall force})$$

Indeed, the viscous stall drag measured in Ref. [50] for kinesin dimers is smaller than the conservative stall force measured in Refs. [15, 51]. Here we do not claim that the simple model we considered above fully explains the discrepancy in the stall force measured in these two labs with different experimental setups. The simple model above, however, does point out that in general we should not expect the viscous stall drag to be the same as the conservative stall force. When these two stall forces are observed to be different we should not argue which measurement is correct and which measurement is wrong. Instead, it is likely that both measurements are correct and the smaller value for the measured viscous stall drag may be caused by that the kinesin motor potential is not a constant slope or in other words the kinesin motor force is non-uniform. In general, the viscous stall drag of a motor is mainly affected by the motor force profile. The further a motor force profile is away from being uniform, the smaller is the viscous stall drag. When the viscous stall drag is close to the thermodynamic stall, the motor force profile must be uniform. On the other hand, the conservative stall force of a motor is mainly affected by how tightly the chemical reaction and the mechanical motion are coupled all the way to the stall point. When the conservative stall force is close to the thermodynamic stall, the chemical reaction and mechanical motion must be tightly coupled. The viscous stall drag (Stokes efficiency) and the conservative stall force (thermodynamic efficiency) inform us different aspects about the motor operation. Thus, it is valuable to measure both these two stall forces (both efficiencies), and we should not expect these two stall forces to be equal.

5.4. Motor Potential Profile and Decomposition of the Stokes Efficiency into Chemical Efficiency and Mechanical Efficiency

The motor potential profile $\psi(x)$ serves as a link between the chemical reaction and the motor motion. $(-\Delta\psi) = \psi(0)-\psi(l)$ can be viewed as the potential energy made available in the chemical reaction per motor step to driving the motor motion. $(-\Delta\psi)$ is generated by the chemical reaction and is used to drive the motor motion. We like to investigate how efficiently $(-\Delta\psi)$ is generated in the chemical reaction and how efficiently $(-\Delta\psi)$ is utilized in driving the motor motion. For that purpose, we decompose the Stokes efficiency into the product of two terms. Here we consider the case where the conservative loading force is zero: $F = 0$.

$$\begin{aligned}\eta_{\text{Stokes}} &= \frac{\zeta\langle u\rangle^2}{(-\Delta G)\langle r\rangle} = \left\{\frac{-\Delta\psi}{(-\Delta G)\langle r\rangle(l/\langle u\rangle)}\right\}\cdot\left\{\frac{\zeta\langle u\rangle l}{-\Delta\psi}\right\}\\ &\equiv \eta_{\text{Chemical}}\cdot\eta_{\text{Mechanical}}\end{aligned} \tag{45}$$

where η_{Chemical} and $\eta_{\text{Mechanical}}$ are called, respectively, the chemical efficiency and the chemical efficiency, and are defined as

$$\begin{aligned}\eta_{\text{Chemical}} &= \frac{(-\Delta\psi)}{(-\Delta G)\,\langle r\rangle(l/\langle u\rangle)}\\ eta_{\text{Mechanical}} &= \frac{\zeta\langle u\rangle l}{(-\Delta\psi)}\end{aligned} \tag{46}$$

As shown in Figure 4, conceptually, the reaction cycle first produces the motor potential profile. Then the motor potential profile drives the mechanical motion. The motor potential profile is something that links the chemical reaction and the mechanical motion. The mechanical motion is completely determined by the motor potential profile. The mechanical motion feels the effects of chemical reaction only via the motor potential profile. In the definition of η_{Chemical}, the numerator is $(-\Delta\psi)$ and the denominator is $-\Delta G\langle r\rangle(l/\langle u\rangle)$ which is the free energy consumed per motor step. The chemical efficiency η_{Chemical} measures how efficiently $(-\Delta\psi)$ is generated in chemical reaction per motor step.

In η_{Chemical}, the denominator is $(-\Delta\psi$ and the numerator is $\zeta\langle u\rangle l$, which is proportional to the average velocity

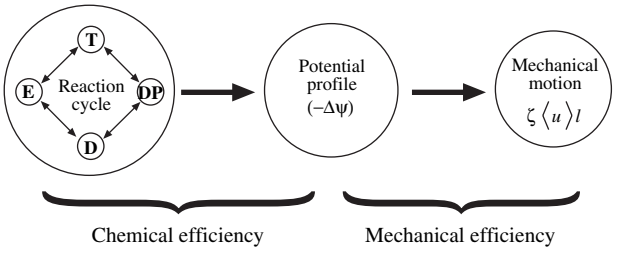


Fig. 4. A schematic diagram showing a conceptual view of motor operation. The chemical reaction produces the motor potential profile. The motor potential profile then drives the mechanical motion. The motor potential profile divides the overall process of force generation into two sub-processes. The chemical efficiency measures how efficiently the motor potential profile is produced in chemical reaction. The mechanical efficiency measures how efficiently the motor potential profile is used to driving the motion.

and measures the mechanical performance of the motor. In this sense, $\eta_{\text{Mechanical}}$ measures how efficiently $(-\Delta\psi)$ is utilized to drive the motor through the viscous media. However, for them to be properly defined efficiencies, they must be bounded by 100%. In Ref. [68], we derived that both the chemical efficiency and the mechanical efficiency defined above are indeed bounded by 100%. Thus, both the chemical efficiency and the mechanical efficiency defined above are legitimate efficiency measures. In Ref. [68], we studied the behavior of these two efficiencies in a model system where a motor is driven by switching between two potentials. We found that the mechanical efficiency is sole determined by the shape of the motor potential profile. If the motor potential profile is close to a constant slope the mechanical efficiency is close to 100%. When the motor potential profile deviates from a constant slope, the mechanical efficiency decreases. If we are given a fixed $(-\Delta\psi)$ but are allowed to change the shape of potential profile while fixing $(-\Delta\psi)$, then the most efficient way of utilizing $(-\Delta\psi)$ is to have a constant slope potential. This is consistent with our motivation in defining the mechanical efficiency. The chemical efficiency is a bit more complicated. First, the chemical efficiency is affected by the time scales of transitions relative to the time scales of mechanical motion. If chemical transitions are the rate-limiting steps in the overall motor operation, then the long waiting for chemical transitions will create bumps on the potential profile, lower $(-\Delta\psi)$ and thus reduce the chemical efficiency. So the first requirement for high chemical efficiency is that all chemical transitions are much faster than the mechanical motion. Since the chemical efficiency measures how efficiently the motor is utilizing the free energy $(-\Delta G)$ from a reaction cycle to generate $(-\Delta\psi)$ it is affected by of how much free energy of $(-\Delta G)$ is allocated to chemical transitions and how much is allocated to driving the motion. It is reasonable to expect that if almost all of (ΔG) is allocated to chemical transitions and very little is allocated to driving the motion, then $(-\Delta\psi)$ will be small. On the other hand, if a free energy of more than $(-\Delta G)$ is allocated to driving the motion and a negative amount of free energy is allocated to chemical transitions, then the long waiting and the frequent reversals of individual transitions caused by energetically unfavorable transitions will significantly reduce $(-\Delta\psi)$. Thus, for a given free energy $(-\Delta G)$, to maximize $(-\Delta\psi)$, the motor needs to allocated all of $(-\Delta G)$, no more, no less, to driving the motion and allocate zero to each individual chemical transition. In Ref. [68], it was found that the second condition for high chemical efficiency is that the equilibrium constants of individual chemical transitions are close to one. In terms of free energy, that means the free energy associated with each individual chemical transition is close to zero.

5.5. Stokes Efficiency for Motor-Cargo Systems

In many single molecules experiments, the motor is not observed/recorded directly. Instead a large latex bead is attached to the motor, the position of the bead is observed/recorded, and the loading force is controlled by pulling the latex bead with a laser trap or a force clamp.[15, 19, 50, 54, 56] This situation also occurs in biological settings. For example, a kinesin dimer walks on a microtubule filament, towing a vesicle toward the membrane for exporting. To accommodate these experimental and biological settings, we need to study the behaviors of motor-cargo systems.[69, 70] The derivation in Ref. [39] was for motor systems without cargos, described by Eq. (12). Below, we first describe the Markov Fokker-Planck framework for motor-cargo systems. A motor-cargo system has two spatial degrees of freedom: one for the motor and the other for the cargo. The motor and the cargo are linked by an elastic element.[71] Over long time, the cargo must follow the motor and they must have the same average velocity and the same effective diffusion coefficient. Under the framework of two-dimensional Fokker-Planck equations, we derive that the Stokes efficiency is bounded by 100% for motor-cargo systems. Once this conclusion is established, the Stokes efficiency is a justified efficiency measure for motor-cargo systems.

A motor-cargo system is shown in Figure 4, corresponding to the experimental setup in Refs. [50 and 15] where a kinesin dimer walks on a microtubule toward the plus end and a latex bead is linked to the kinesin dimer to visualize the motion of the kinesin dimier and to exert a force on the kinesin dimer. The force exerting on the kinesin dimer is controlled either by changing the size of the latex bead and changing the viscosity of the surrounding fluid[50] or by applying a force on the bead using a laser trap.[15] In the recently developed 2-D laser trap,[56] a force perpendicular to the direction of motor motion can also be exerted on the kinesin dimer.

Let x be the coordinate of the motor and y the coordinate of the cargo along the direction of motion as

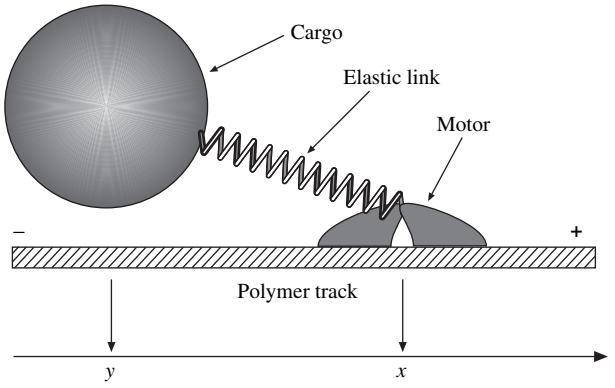


Fig. 5. A motor-cargo system in single molecule experiments, corresponding to the experimentals setup in Refs. [50 and 15]. A kinesin dimer walks on a microtubule filament toward the plus end and a latex bead is linked to the kinesin dimer to visualize the motion of the kinesin dimier and to exert a force on the kinesin dimer. The latex bead is either loaded by the force from a laser trap or loaded by the viscous drag from the fluid media.

illustrated in Figure 5. As we discussed in the previous sections, the motor is driven by switching among a set of N potentials, each corresponding to an individual occupancy state. The stochastic motion of the motor is governed by an over-damped Langevin equation

$$\frac{dx}{dt} = D_M \frac{[-E'(x-y-R) - \phi'_S(x)]}{k_B T} + \sqrt{2D_M}\frac{dW(t)}{dt} \qquad S(t) = 1, 2, \ldots, N \quad (47)$$

where D_M is the diffusion coefficient of the motor, and $-E'(x-y-R)$ is the elastic force on the motor exerted by the elastic link between the motor and the cargo. Here $E(x-y-R)$ is the elastic potential and R the rest length of the elastic link. $\phi_S(x)$ is the motor potential corresponding to the current occupancy state $S(t)$. As before, the stochastic evolution of the occupancy state is governed by a discrete Markov process. In the motor-cargo system, the external loading force no longer acts directly on the motor. The external loading force acts directly on the cargo. The stochastic motion of the cargo is governed by a Langevin equation in a similar form

$$\frac{dy}{dt} = D_C \frac{[F + E'(x-y-R)]}{k_B T} + \sqrt{2D_C}\frac{dW(t)}{dt} \quad (48)$$

where D_C is the diffusion coefficient of the cargo, and F is the external loading force acting on the cargo. Note that in the motor-cargo system, the motor does not directly feel the external loading force on the cargo; the motor sees the elastic force from the cargo. On the other hand, the cargo does not directly feel the internal motor force; the cargo sees the elastic force from the motor via the link. All communications between the motor and the cargo are done via the elastic link.

Let $\rho_j(x, y, t)$ be the probability density that the motor is at position x, the cargo is at position y, and the chemical reaction is in occupancy state j at time t. The probability density of the motor-cargo system is governed by the Fokker-Planck equation corresponding to Langevin equations (47) and (48). Here we focus on a simple case where the chemical reaction goes through N occupancy state sequentially, which corresponds to motors with one catalytic site. The analysis can be extended to motors with more than one catalytic site. For the convenience of derivation below, we write the Fokker-Planck equation in the form of conservation of probability. In a sequential reaction, each occupancy state has only one in-flux and one out-flux in the reaction direction. The conservation of probability takes the form:

$$\frac{\partial \rho_i}{\partial t} = \underbrace{-\frac{\partial}{\partial x} J_i^{(x)}(x, y)}_{\text{Probability flow along the } x\text{-direction}} - \underbrace{\frac{\partial}{\partial y} J_i^{(y)}(x, y)}_{\text{Probability flow along the } y\text{-direction}} - \underbrace{(I_{i+1/2}(x, y) - I_{i-1/2}(x, y))}_{\text{Probability flow along the reaction direction}}, \quad i = 1, 2, \ldots, N \quad (49)$$

where $J_i^{(x)}(x, y)$ is the probability flux along the x-direction in occupancy state i, and is given by

$$J_i^{(x)}(x, y) = -D_M \left(\frac{1}{k_B T} \frac{\partial \Phi_i(x, y)}{\partial x} \rho_i + \frac{\partial \rho_i}{\partial x} \right) \qquad i = 1, 2, \ldots, N$$

$J_i^{(y)}(x, y)$ is the probability flux along the y-direction in occupancy state i, and is given by

$$J_i^{(y)}(x, y) = -D_C \left(\frac{1}{k_B T} \left[-F + \frac{\partial \Phi_i(x, y)}{\partial y} \right] \rho_i + \frac{\partial \rho_i}{\partial y} \right) \qquad i = 1, 2, \ldots, N$$

$I_{i+1/2}(x, y)$ is the probability flux density (flux per unit area in the $x-y$ plane) along the reaction direction from state i to state $i+1$, and is given by

$$I_{i+1/2}(x, y) = k_{i \to i+1}(x)\rho_i - k_{i+1 \to i}(x)\rho_{i+1} \quad (50)$$

The transition rates $k_{i \to i+1}(x)$ and $k_{i+1 \to i}(x)$ satisfy detailed balance

$$\frac{k_{i \to i+1}(x)}{k_{i+1 \to i}(x)} = \exp\left(\frac{\phi_i(x) - \phi_{i+1}(x)}{k_B T} \right), \quad 1 \le i \le N$$
$$\phi_{N+i}(x) = \phi_i(x) + \Delta G \quad (51)$$

Note that the chemical transition rates are functions of motor position x only. The cargo does not directly affect the chemical reaction. Instead, the cargo affects the chemical reaction indirectly by changing the motor position.

$\Phi_i(x, y)$ is the total potential of the motor-cargo system in occupancy state i. $\Phi_i(x, y)$ includes both the motor internal potential energy caused by the chemical reaction and the elastic energy in the link connecting the motor and the cargo. $\Phi_i(x, y)$ does not contain the external loading force F, which is included separately in the model. $\Phi_i(x, y)$ is given by

$$\Phi_i(x, y) = \phi_i(x) + E(x - y - R), \quad i = 1, 2, \ldots, N$$

The rest length R can be eliminated using a change of variable $\tilde{y} = y + R$. So without loss of generality, we assume $R = 0$. Here we do not assume the elastic link between the motor and the cargo is a linear spring. We do assume that the elastic energy satisfy

$$E'(0) = 0 \quad \text{and} \quad E''(x) > 0 \quad \text{for all } x$$

The Stokes efficiency for a motor-cargo system is defined as

$$\eta_{\text{Stokes}} = \frac{(\zeta_M + \zeta_C)\langle u\rangle^2}{(-\Delta G)\langle r\rangle + F\langle u\rangle}$$

To show that the Stokes efficiency is bounded by 100%, we need to derive

$$(\zeta_M + \zeta_C)\langle u\rangle^2 \le (-\Delta G)\langle r\rangle + F\langle u\rangle \tag{52}$$

where ζ_M and ζ_C are, respectively, the drag coefficients of the motor and the cargo. In most motor-cargo systems, the cargo is much larger than the motor, and we have $\zeta_C \gg \zeta_M$. Drag coefficients ζ_M and ζ_C are related to diffusion coefficients D_M and D_C by the Einstein relation:

$$\zeta_M = \frac{k_B T}{D_M}, \quad \zeta_C = \frac{k_B T}{D_C}$$

Our goal is to derive (52). We start by introducing function $F_i(x, y)$ defined as

$$F_i(x, y) = \frac{\Phi_i(x, y)}{k_B T} + \ln(\rho_i(x, y)) \tag{53}$$

The two probability fluxes in the two spatial dimensions can be written as

$$\begin{aligned} J_i^{(x)}(x, y) &= -D_M \left(\frac{1}{k_B T} \frac{\partial \Phi_i(x, y)}{\partial x} \rho_i + \frac{\partial \rho_i}{\partial x} \right) \\ &= -D_M \rho_i \left(\frac{1}{k_B T} \frac{\partial \Phi_i(x, y)}{\partial x} + \frac{1}{\rho_i} \frac{\partial \rho_i}{\partial x} \right) \\ &= -D_M \rho_i \frac{\partial F_i(x, y)}{\partial x} \end{aligned} \tag{54}$$

$$\begin{aligned} J_i^{(y)}(x, y) &= -D_C \left(\frac{1}{k_B T} \left[-F + \frac{\partial \Phi_i(x, y)}{\partial y} \right] \rho_i + \frac{\partial \rho_i}{\partial y} \right) \\ &= -D_C \rho_i \left(\frac{\partial F_i(x, y)}{\partial y} - \frac{F}{k_B T} \right) \end{aligned} \tag{55}$$

The boundary conditions for Eq. (49) in the x-direction are periodic:

$$\begin{aligned} \rho_i(x + \ell, y, t) &= \rho_i(x, y - \ell, t) \\ \phi_i(x + \ell) &= \phi_i(x) \\ k_{i \to i+1}(x + \ell) &= k_{i \to i+1}(x) \\ k_{i+1 \to i}(x + \ell) &= k_{i+1 \to i}(x) \end{aligned}$$

In the y-direction, Eq. (49) extends from negative infinity to positive infinity. So no boundary conditions are required in the y-direction. The boundary conditions along the reaction direction are pseudo periodic:

$$\begin{aligned} \rho_{N+i}(x, y, t) &= \rho_i(x, y, t) \\ \phi_{N+i}(x) &= \phi_i(x) + \Delta G \\ k_{N+i \to N+i+1}(x) &= k_{i \to i+1}(x) \\ k_{N+i+1 \to N+i}(x) &= k_{i+1 \to i}(x) \end{aligned}$$

The average velocity and the average reaction rate are determined from the steady state, which satisfies Eq. (49) without the time derivative on the left side:

$$\begin{aligned} 0 = &-\frac{\partial}{\partial x} J_i^{(x)}(x, y) - \frac{\partial}{\partial y} J_i^{(y)}(x, y) - (I_{i+1/2}(x, y) \\ &- I_{i-1/2}(x, y)), \quad i = 1, 2, \ldots, N \end{aligned} \tag{56}$$

In terms of steady state fluxes, the average velocity and the average reaction rate have the form

$$\begin{aligned} \langle u\rangle &= \int_0^\ell dx \int_{-\infty}^{+\infty} dy \sum_{i=1}^N J_i^{(x)}(x, y) \\ &= \int_0^\ell dx \int_{-\infty}^{+\infty} dy \sum_{i=1}^N J_i^{(y)}(x, y) \\ \langle r\rangle &= \int_0^\ell dx \int_{-\infty}^{+\infty} dy\, I_{i+1/2}(x, y) \quad \text{(for any } i\text{)} \end{aligned}$$

In the above we have two expressions for $\langle u\rangle$, the first expression is the average velocity of the motor and the second expression is the average velocity of the cargo. Since the motor and the cargo are elastically linked, they must have the same average velocity, which is the average velocity of the motor-cargo system. Using the expression for the average velocity of the motor, using the expression for the probability flux in the x-direction given in (54), and then applying the Cauchy-Schwarz inequality, we get

$$\begin{aligned} \langle u\rangle^2 &= \left\{ \int_0^\ell dx \int_{-\infty}^{+\infty} dy \sum_{i=1}^N J_i^{(x)}(x, y) \right\}^2 \\ &= \left\{ \int_0^\ell dx \int_{-\infty}^{+\infty} dy \sum_{i=1}^N \rho_i^{1/2} \left[-\rho_i^{1/2} D_M \frac{\partial F_i(x, y)}{\partial x} \right] \right\}^2 \\ &\le \left\{ \int_0^\ell dx \int_{-\infty}^{+\infty} dy \sum_{i=1}^N \rho_i \right\} \end{aligned}$$

REVIEW

$$\cdot\left\{\int_0^\ell dx\int_{-\infty}^{+\infty}dy\sum_{i=1}^N\rho_i D_M^2\left(\frac{\partial F_i(x,y)}{\partial x}\right)^2\right\}$$

Using the fact that the total probability is one and converting back to the probability flux in the x-direction yields

$$\langle u\rangle^2\le\int_0^\ell dx\int_{-\infty}^{+\infty}dy\sum_{i=1}^N -J_i^{(x)}(x,y)D_M\frac{\partial F_i(x,y)}{\partial x}$$

Integrating by parts with respect to independent variable x and using the periodic boundary condition in the x-direction leads to

$$\langle u\rangle^2\le D_M\int_0^\ell dx\int_{-\infty}^{+\infty}dy\sum_{i=1}^N F_i(x,y)\frac{\partial J_i^{(x)}(x,y)}{\partial x}\qquad(57)$$

Now using the expression for the average velocity of the cargo, using the expression for the probability flux in the y-direction given in (55), and then applying the Cauchy-Schwarz inequality, we have

$$\begin{aligned}\langle u\rangle^2&=\left\{\int_0^\ell dx\int_{-\infty}^{+\infty}dy\sum_{i=1}^N J_i^{(y)}(x,y)\right\}^2\\&=\left\{\int_0^\ell dx\int_{-\infty}^{+\infty}dy\sum_{i=1}^N\rho_i^{1/2}\left[-\rho_i^{1/2}D_C\left(\frac{\partial F_i(x,y)}{\partial y}-\frac{F}{k_BT}\right)\right]\right\}^2\\&\le\left\{\int_0^\ell dx\int_{-\infty}^{+\infty}dy\sum_{i=1}^N\rho_i\right\}\\&\quad\cdot\left\{\int_0^\ell dx\int_{-\infty}^{+\infty}dy\sum_{i=1}^N\rho_i D_C^2\left(\frac{\partial F_i(x,y)}{\partial y}-\frac{F}{k_BT}\right)^2\right\}\end{aligned}$$

Using the fact that the total probability is one and converting back to the probability flux in the y-direction yields

$$\begin{aligned}\langle u\rangle^2&\le\int_0^\ell dx\int_{-\infty}^{+\infty}dy\sum_{i=1}^N -J_i^{(y)}(x,y)D_C\left(\frac{\partial F_i(x,y)}{\partial y}-\frac{F}{k_BT}\right)\\&=D_C\int_0^\ell dx\int_{-\infty}^{+\infty}dy\sum_{i=1}^N -J_i^{(y)}(x,y)\frac{\partial F_i(x,y)}{\partial y}\\&\quad+D_C\frac{F}{k_BT}\int_0^\ell dx\int_{-\infty}^{+\infty}dy\sum_{i=1}^N J_i^{(y)}(x,y)\end{aligned}$$

Converting the second term back to the average velocity, integrating the first term by parts with respect to independent variable y, and using the fact that for large y the elastic energy is large and the probability density is exponentially small, we obtain

$$\langle u\rangle^2\le D_C\int_0^\ell dx\int_{-\infty}^{+\infty}dy\sum_{i=1}^N F_i(x,y)\frac{\partial J_i^{(y)}(x,y)}{\partial x}+D_C\frac{F}{k_BT}\langle u\rangle\qquad(58)$$

Combining results (57) and (58) gives us

$$\begin{aligned}\langle u\rangle^2&=\frac{D_C}{D_M+D_C}\langle u\rangle^2+\frac{D_M}{D_M+D_C}\langle u\rangle^2\\&\le\frac{D_CD_M}{D_M+D_C}\int_0^\ell dx\int_{-\infty}^{+\infty}dy\sum_{i=1}^N F_i(x,y)\\&\quad\times\left[\frac{\partial J_i^{(x)}(x,y)}{\partial x}+\frac{\partial J_i^{(y)}(x,y)}{\partial y}\right]+\frac{D_CD_M}{D_M+D_C}\cdot\frac{1}{k_BT}F\langle u\rangle\end{aligned}$$

Using the steady state differential Eq. (56) and using

$$\frac{D_CD_M}{D_M+D_C}\cdot\frac{1}{k_BT}=\frac{1}{\zeta_M+\zeta_C}$$

we have

$$\begin{aligned}\langle u\rangle^2&\le\frac{-D_CD_M}{D_M+D_C}\int_0^\ell dx\int_{-\infty}^{+\infty}dy\sum_{i=1}^N F_i(x,y)\\&\quad\times(I_{i+1/2}(x,y)-I_{i-1/2}(x,y))+\frac{1}{\zeta_M+\zeta_C}F\langle u\rangle\end{aligned}$$

Using summation by parts with respect to index i and using the boundary conditions in the reaction direction, in particular, $F_{N+i}(x,y)=F_{N+i}(x,y)+\Delta G/k_BT$, we arrive at

$$\begin{aligned}\langle u\rangle^2&\le\frac{D_CD_M}{D_M+D_C}\int_0^\ell dx\int_{-\infty}^{+\infty}dy\\&\quad\times\sum_{i=1}^N I_{i-1/2}(x,y)(F_i(x,y)-F_{i-1}(x,y))\\&\quad+\frac{D_CD_M}{D_M+D_C}\frac{(-\Delta G)}{k_BT}\int_0^\ell dx\int_{-\infty}^{+\infty}dyI_{1/2}(x,y)\\&\quad+\frac{1}{\zeta_M+\zeta_C}F\langle u\rangle\equiv T_1+T_2+T_3\end{aligned}$$

We examine the three terms in the above inequality one by one. The third term is the simplest and it needs no further reduction. For the second term, we use the expression for the average reaction rate $\langle r\rangle=\int_0^\ell dx\int_{-\infty}^{+\infty}dy\,I_{i+1/2}(x,y)$. We write T_2 as

$$T_2=\frac{1}{\zeta_M+\zeta_C}(-\Delta G)\langle r\rangle$$

To simplify the first term, we use detailed balance to write

$$\begin{aligned}I_{i-1/2}(x,y)&=k_{i-1\to i}\rho_{i-1}\left(1-\frac{k_{i\to i-1}}{k_{i-1\to i}}\cdot\frac{\rho_i}{\rho_{i-1}}\right)\\&=k_{i-1\to i}\rho_{i-1}\left(1-\exp\left(\frac{\phi_i(x)-\phi_{i-1}(x)}{k_BT}\right)\cdot\frac{\rho_i}{\rho_{i-1}}\right)\\F_i(x,y)-F_{i-1}(x,y)&=\frac{\phi_i(x)-\phi_{i-1}(x)}{k_BT}+\ln\frac{\rho_i}{\rho_{i-1}}\end{aligned}$$

Combining these two results yields

$$T_1=\frac{D_CD_M}{D_M+D_C}\int_0^\ell dx\int_{-\infty}^{+\infty}dy\sum_{i=1}^N k_{i-1\to i}\rho_{i-1}(1-G_{i-1/2})\ln(G_{i-1/2})$$

where function $G_{i-1/2}$ is positive and is defined as

$$G_{i-1/2}(x,y) \equiv \exp\left(\frac{\phi_i(x)-\phi_{i-1}(x)}{k_B T}\right)\cdot\frac{\rho_i}{\rho_{i-1}}$$

The expression $(1-\alpha)\ln(\alpha)$ is always non-positive for any positive value of α, which implies

$$T_1 \le 0$$

Combining the results for the three terms, we arrive at

$$\langle u\rangle^2 \le \frac{1}{\zeta_M+\zeta_C}(-\Delta G)\langle r\rangle + \frac{1}{\zeta_M+\zeta_C}F\langle u\rangle$$

which leads immediately to that the Stokes efficiency is bounded by 100% for a motor-cargo system. Thus, the Stokes efficiency is a justified efficiency measure for motor-cargo systems.

6. NUMERICAL METHODS FOR SOLVING FOKKER-PLANCK EQUATIONS

In Ref. [72], we proposed a robust numerical method for solving Fokker-Planck equations. We shall refer to this numerical method as the WPE method. When the potential is smooth, the convergence of the WPE method was mathematically established in Ref. [72], but that analysis is not sufficient to completely explain the robust performance of the WPE method. One of the advantages of the WPE method is that it works well even if the potential is discontinuous. Recently, the convergence of the WPE method for discontinuous potentials was established mathematically.[73] Furthermore, in the case of discontinuous potentials, an improved version of the WPE method was proposed, which maintains the second order accuracy even when the potential is discontinuous.[73] Below, we first review the WPE method. We study the condition of detailed balance for numerical methods. When the potential is smooth, the exact solution satisfies the differential equation in classical sense and the numerical solution converges to the exact solution as we refine the numerical grid. When the potential is discontinuous, the exact solution satisfies the differential equation in classical sense away from discontinuities. At discontinuities, the exact solution is constrained by two conditions (14) and (16). Condition (14) follows automatically from the conservation of probability. Condition (16) is equivalent to detailed balance for numerical methods. Thus, detailed balance is a necessary condition for convergence when the potential is discontinuous.[73] The WPE method is based on approximating a spatially continuous Markov process (Langevin equation) with a spatially discrete Markov process (a jump process). In the jump process, the transition rates of the WPE method are determined by considering local exact solutions. Alternatively, the transition rates can also be determined by trying to preserve detailed balance and preserve the first two moments of the exact solution in a model example. Preserving detailed balance is important in simulating motor systems. It turns out that only one of the first two moments can be preserved. The WPE method can be interpreted as preserving the first moment and detailed balance. The improved WPE method can be interpreted as preserving the second moment and detailed balance. The improved WPE has second order accuracy even when the potential is discontinuous.[73] Although the second order accuracy of the improved WPE method has been derived mathematically, the intrinsic connection (if any) between preserving the second moment (but giving up the first moment) and achieving second order accuracy for discontinuous potentials is still unknown. Both the WPE method and the improved WPE method have been extended to solving two-dimensional Fokker-Plack equations, which govern the evolution of motor-cargo systems.[74, 75] The main difficulty of simulating motor-cargo systems is not the extension of numerical methods to the two-dimensional case. The main difficulty is how to accommodate the full range of elasticity for the link between the motor and the cargo. In particular, when the link is stiff, the motor and the cargo move together, and tracking the motor and cargo by their individual coordinates presents a big numerical difficulty. This numerical difficulty is solved by reformulating the system using two new coordinates: the coordinate measuring the motion of the motor-cargo complex and the coordinate measuring the relative motion between the motor and the cargo.[75]

6.1. The WPE Method

We review the WPE method developed in Ref. [72]. To illustrate the spatial discretization of the WPE method, we consider the Fokker-Planck equation below, which does not include chemical reactions. The extension of the spatial discretization to the case of Fokker-Planck equation with chemical reactions is straightforward.

$$\frac{\partial\rho}{\partial t} = D\frac{\partial}{\partial x}\left(\frac{\phi'(x)}{k_B T}\rho + \frac{\partial\rho}{\partial x}\right) \tag{59}$$

We divide the period $[0,\ l]$ into M cells of equal size $h = l/M$. The j-th cell is

$$[x_{j-1/2}, x_{j+1/2}],\quad x_j = \frac{-h}{2}+jh,\quad x_{j+1/2} = \frac{-h}{2}+\left(j+\frac{1}{2}\right)h$$

Each cell is represented by its center $x_j = (j-0.5)h$. The Fokker-Planck equation (59) governs the evolution of probability density for a spatially continuous Markov process (Langevin equation). The basic idea of the WPE method is that we approximate the continuous Markov process with a jump process (discrete Markov process). This idea was originated in Ref. [76] and in an unpublished result by Peskin. The sites in the jump process are the centers of cells: $\{x_j\}$. As shown in Figure 6, the system is restricted to the discrete sites $\{x_j\}$. In a single jump, it

REVIEW

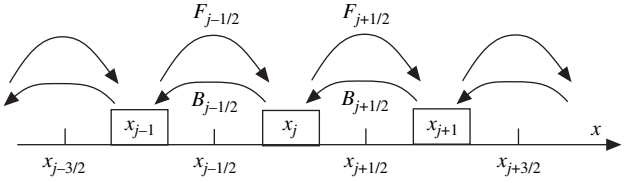


Fig. 6. Spatial discretization of a Fokker-Planck equation without reaction. The spatially continuous Markov process is approximated using a jump process. In the jump process, the system is restricted to the set of discrete sites $\{x_j\}$ and is allowed to jump only to an adjacent site in a single jump.

can only jump to an adjacent site. For a one-dimensional Fokker-Planck equation without chemical reactions, each site x_j has two adjacent sites: x_{j-1} and x_{j+1}.

Let $p_j(t)$ be the probability that the system is at site x_j at time t in the jump process. Since site x_j represents cell $[x_{j-1/2}, x_{j+1/2}]$, probability $p_j(t)$ and probability density $\rho(x_j, t)$ are related by

$$p_j(t) \approx \int_{x_{j-1/2}}^{x_{j+1/2}} \rho(x,t)\,dx \approx h \cdot \rho(x_j, t)$$

Let $F_{j+1/2}$ be the transition rate from x_j to x_{j+1}, and $B_{j+1/2}$ be the transition rate from x_{j+1} to x_j. In the jump process, the numerical probability flux through $x_{j+1/2}$ toward the right side is given by

$$J_{j+1/2}(t) = F_{j+1/2}p_j(t) - B_{j+1/2}p_{j+1}(t) \tag{60}$$

The time evolution of $p_j(t)$ is governed by the master equation, which is based on the conservation of probability:

$$\begin{aligned}\frac{dp_j}{dt} &= J_{j-1/2} - J_{j+1/2} \\ &= (F_{j-1/2}p_{j-1} - B_{j-1/2}p_j) - (F_{j+1/2}p_j - B_{j+1/2}p_{j+1})\end{aligned} \tag{61}$$

In Ref. [72], the jump rates were determined by considering local steady state solutions of the Fokker-Planck equation. In calculating $F_{j+1/2}$ and $B_{j+1/2}$, we make two assumptions:

(1) In the interval $[x_{j-1/2}, x_{j+3/2}]$, which consists of cell j and cell $j+1$, potential $\phi(x)$ is approximated by a linear function going through $(x_j, \phi(x_j))$ and $(x_{j+1}, \phi(x_{j+1}))$:

$$\tilde{\phi}(x) = C + \frac{\Delta\phi_{j+1/2}}{h} \cdot x, \quad \Delta\phi_{j+1/2} = \phi(x_{j+1}) - \phi(x_j)$$

(2) Let $\tilde{\rho}(x)$ be the steady state solution of Eq. (59) with linear potential $\tilde{\phi}(x)$ given above and satisfying the conditions

$$\int_{x_{j-1/2}}^{x_{j+1/2}} \tilde{\rho}(x)\,ds = p_j, \quad \int_{x_{j+1/2}}^{x_{j+3/2}} \tilde{\rho}(x)\,ds = p_{j+1}$$

We assume that the probability flux through $x_{j+1/2}$ in Fokker-Planck equation (59) is well approximated by the probability flux in the jump process. This is a key step in constructing the WPE method. Instead of using the standard numerical procedure of Taylor expansion, we use local steady state solutions. The direct consequence of this approach is that detailed balance is preserved exactly.

The probability flux associated with $\tilde{\rho}(x)$ is derived in Ref. [72] and is given by

$$\begin{aligned}\tilde{J} = \frac{D}{h^2} \cdot \frac{\Delta\phi_{j+1/2}/k_{\mathrm{B}}T}{\exp(\Delta\phi_{j+1/2}/k_{\mathrm{B}}T) - 1} \\ \times \left(p_j - \exp\left(\frac{\Delta\phi_{j+1/2}}{k_{\mathrm{B}}T}\right) p_{j+1}\right)\end{aligned}$$

The jump rates are determined by comparing the numerical flux given in (60) with the exact probability flux associated with $\tilde{\rho}(x)$

$$\begin{aligned}F_{j+1/2} &= \frac{D}{h^2} \cdot \frac{\Delta\phi_{j+1/2}/k_{\mathrm{B}}T}{\exp(\Delta\phi_{j+1/2}/k_{\mathrm{B}}T) - 1} \\ B_{j+1/2} &= \frac{D}{h^2} \cdot \frac{-\Delta\phi_{j+1/2}/k_{\mathrm{B}}T}{\exp(-\Delta\phi_{j+1/2}/k_{\mathrm{B}}T) - 1} \\ \Delta\phi_{j+1/2} &= \phi(x_{j+1}) - \phi(x_j)\end{aligned} \tag{62}$$

A more accurate formula for calculating the transition rates without assuming the linear potential can be found in Ref. [77]. However, this more accurate formula requires the numerical evaluation of integrals. It is important to notice that the jump rates given in (62) are always positive. Thus, the numerical approximation is a proper jump process.

6.2. Detailed Balance for Numerical Method

The numerical approximation to the Fokker-Planck equation is a jump process in which each site x_j has a well-defined free energy $\phi(x_j)$. At equilibrium, the probability density of Fokker-Planck equation (59) has the Boltzmann form and the probability flux of Fokker-Planck equation (59) vanishes everywhere. To extend this thermodynamic property of Fokker-Planck equation (59) to the numerical method, we require that the jump rates in the numerical method satisfy the condition of detailed balance:

$$\frac{F_{j+1/2}}{B_{j+1/2}} = \exp\left(\frac{\phi(x_j) - \phi(x_{j+1})}{k_{\mathrm{B}}T}\right) \tag{63}$$

It is straightforward to verify that the jump rates given in (62) satisfy detailed balance.

To demonstrate the importance of preserving detailed balance in numerical methods, we show that when the potential is discontinuous, detailed balance is a necessary condition for the convergence of any numerical method. Suppose the discontinuity is at $d = x_{j+1/2}$ (of course, the index j will change as the numerical grid is refined and

the numerical grid size h decreases). The numerical probability flux through $x_{j+1/2}$ is

$$\begin{aligned} J_{j+1/2} &= F_{j+1/2}p_j - B_{j+1/2}p_{j+1} \\ &= \frac{1}{h}(h^2 B_{j+1/2})\frac{p_j}{h}\left(\frac{F_{j+1/2}}{B_{j+1/2}} - \frac{p_{j+1}}{j}\right) \end{aligned} \tag{64}$$

Suppose the numerical solution converges to the exact solution that satisfies conditions (14) and (16) at the discontinuity. We have

$$\begin{gathered} \lim_{h\to 0}\frac{p_j}{h} = \rho\left(d^-, t\right), \quad \lim_{h\to 0}\frac{p_{j+1}}{h} = \rho\left(d^+, t\right) \\ \lim_{h\to 0} J_{j+1/2} = \text{exact flux} = \text{finite} \end{gathered} \tag{65}$$

Fokker-Planck equation contains both the transportation driven by potential $\phi(x)$ and the Brownian diffusion. The time scale of diffusing by a distance of h is $t_{\text{Diffusion}} \sim h^2/D = (h^2)$. When using a force of magnitude f and in the absence of diffusion, the time scale of transporting a particle by a distance of h is $t_{\text{Transport}} \sim h/D \cdot k_{\text{B}}T/f = O(h)$. It is clear that for small h we have $t_{\text{Diffusion}} \ll t_{\text{Transport}}$. Thus, when the numerical grid size h is small, the jump process is dominated by diffusion jumps. Since the jump rates are inversely proportional to the time scales, we have

$$\begin{gathered} F_{j+1/2} = O\left(\frac{1}{t_{\text{Diffusion}}}\right) = O\left(\frac{1}{h^2}\right) \\ B_{j+1/2} = O\left(\frac{1}{h^2}\right) \end{gathered} \tag{66}$$

Dividing both sides of (64) by $(1/h)(h^2 B_{j+1/2})(p_j/h)$, taking the limit as $h \to 0$, and using results (22) and (66), we obtain

$$\lim_{h\to 0}\left(\frac{F_{j+1/2}}{B_{j+1/2}} - \frac{p_{j+1}}{p_j}\right) = \lim_{h\to 0} h\frac{J_{j+1/2}}{(h^2 B_{j+1/2})(p_j/h)} = 0$$

which, when combined with result (65) and condition (16) on the exact solution, leads to

$$\lim_{h\to 0}\frac{F_{j+1/2}}{B_{j+1/2}} = \lim_{h\to 0}\frac{p_{j+1}}{p_j} = \frac{\rho(d^+,t)}{\rho(d^-,t)} = \exp\left(\frac{\phi(d^-)-\phi(d^+)}{k_{\text{B}}T}\right)$$

This corresponds to the condition of detailed balance on spatial jump rates (63). Therefore, in the presence of discontinuities, detailed balance is a necessary condition for the convergence of the numerical method.

The standard central difference method can also be formally cast into the framework of jump process (61) with jump rates

$$\begin{aligned} F^{(\text{CD})}_{j+1/2} &= \frac{D}{h^2}\cdot\left[1 - \frac{\Delta\phi_{j+1/2}}{2k_{\text{B}}T}\right] \\ B^{(\text{CD})}_{j+1/2} &= \frac{D}{h^2}\cdot\left[1 + \frac{\Delta\phi_{j+1/2}}{2k_{\text{B}}T}\right] \end{aligned} \tag{67}$$

It is straightforward to show that the central difference method does not preserve detailed balance. As a matter of fact, at discontinuities of magnitude larger than $2k_{\text{B}}T$, one of the jump rates is negative. For example, when $\Delta\phi_{j+1/2} = 3k_{\text{B}}T$, we have $F^{(\text{CD})}_{j+1/2} = -D/2h^2 < 0$. That explains why the numerical solution of the central difference method does not converge to the correct exact solution as observed in Ref. [73].

6.3. A Different Approach for Calculating the Jump Rates and the Improved WPE Method

The framework of jump process (61) is very broad. Even the standard central difference method, which is constructed based on Taylor expansion, can be cast into the framework of jump process (61). We discuss numerical methods for solving Fokker-Planck equations in this framework. Within framework (61), different numerical methods have different jump rates. A new numerical method is constructed by finding a new set of jump rates $(F_{j+1/2}, B_{j+1/2})$.

As we discussed previously, to ensure that the numerical solution converge to the exact solution even when the potential is discontinuous, we should enforce the condition of detailed balance on numerical methods. To determine the jump rates $(F_{j+1/2}, B_{j+1/2})$, we need two equations. The condition of detailed balance gives us one equation. To find additional equations on the jump rate, let us consider the simple case where the jump process is spatially homogeneous (i.e., translationally invariant):

$$F_{j+1/2} = r_F, \quad B_{j+1/2} = r_B \quad \text{for all } j$$

which corresponds to $\phi'(x) = \Delta\phi/h = \text{const}$ in Fokker-Planck equation (59). For a spatially homogeneous jump process with sites located at $x_j = x_0 + jh$, the average velocity and the effective diffusion are given by

$$\begin{aligned} h\frac{\langle N(t) - N(0)\rangle}{t} &= (r_F - r_B)h \\ h^2\frac{\text{var}(N(t) - N(0))}{2t} &= \frac{(r_F + r_B)}{2}h^2 \end{aligned} \tag{68}$$

On the other, for Fokker-Planck equation (59) with $\phi'(x) = \Delta\phi/h = \text{const}$, the average velocity and the effective diffusion are given by

$$\begin{aligned} \frac{\langle X(t) - X(0)\rangle}{t} &= \frac{D}{k_{\text{B}}T}\cdot\frac{\Delta\phi}{h} \\ \frac{\text{var}(X(t) - X(0))}{2t} &= D \end{aligned} \tag{69}$$

Comparing (68) with (69), and identify $hN(t)$ in the jump process with $X(t)$ in the continuous process, we obtain two

REVIEW

more conditions on jump rates. Thus, including detailed balance, we now have three conditions on the jump rates

$$\text{(Detailed balance):} \quad \frac{r_F}{r_B} = \exp\left(\frac{\alpha h}{k_B T}\right)$$

$$\text{(First moment):} \quad r_F - r_B = \frac{D}{h^2} \cdot \frac{\Delta\phi}{k_B T}$$

$$\text{(Second moment):} \quad r_F + r_B = \frac{2D}{h^2}$$

In general, the two jump rates (r_F, r_B) cannot satisfy all three conditions. We have to select two conditions out of three. If we select the condition of detailed balance and the condition of first moment, we obtain a set of jump rates:

$$r_F = \frac{D}{h^2} \cdot \frac{\Delta\phi/k_B T}{\exp(\Delta\phi/k_B T) - 1}$$

$$r_B = \frac{D}{h^2} \cdot \frac{-\Delta\phi/k_B T}{\exp(-\Delta\phi/k_B T) - 1}$$

which are the jump rates of the WPE method given in (21). If we abandon the condition of detailed balance and select the conditions of the first two moments, we have

$$r_F = \frac{D}{h^2} \cdot \left[1 - \frac{\Delta\phi}{2k_B T}\right], \quad r_B = \frac{D}{h^2} \cdot \left[1 + \frac{\Delta\phi}{2k_B T}\right]$$

which are the jump rates of the standard central difference method given in (24). It is not surprising that the central difference method does not obey detailed balance. If we select the condition of detailed balance and the condition of second moment (abandoning the condition of first moment), we arrive at a new set of jump rates:

$$r_F = \frac{D}{h^2} \cdot \frac{2}{\exp(\Delta\phi/k_B T) + 1}, \quad r_B = \frac{D}{h^2} \cdot \frac{2}{\exp(-\Delta\phi/k_B T) + 1}$$

which are the jump rates of the improved WPE method proposed in Ref. [73]. The improved WPE method has second order accuracy even when the potential is discontinuous. In comparison, when the potential is discontinuous, the standard central difference method has zeroth order accuracy (it converges to a wrong solution) while the WPE method has first order accuracy. The second order accuracy of the improved WPE method was derived mathematically in Ref. [73].

7. RANDOMNESS PARAMETER

Many molecular motors move in steps. For example, a kinesin dimer walks on a microtubule filament in 8 nm steps.[14, 50, 51] For a kinesin dimer, each motor step is coupled to one ATP hydrolysis cycle.[51] Assuming at the end of each reaction cycle, the kinesin dimer returns to the same configuration as the one at the beginning of the cycle but with the roles of the two heads switched, then statistically each reaction cycle is independent of other reaction cycles. Thus, we may model the kinesin dimer as a stochastic stepper. A stochastic stepper is a special case of semi-Markov processes in which the cycle time of each motor step is independent and has the same distribution. The distribution of the cycle time can be arbitrary and is not necessarily exponential. The framework of stochastic stepper is a very general one. It can accommodate backward motor steps and variable step sizes.[78] Here we study the simple case where the stochastic stepper goes only forward and the step size is uniform. To get information about the chemical reaction inside the motor, we would like to have the statistics of the cycle time. A key statistical quantity is the randomness of the cycle time, defined as the variance of the cycle time divided by the square of the average cycle time.[79, 80] However, in the current single molecule experiments, it is very difficult (if not impossible) to observe/record samples of the cycle time. Fortunately, the statistical quantities of the motor position can be measured reliably. This allows us to consider another randomness: the randomness in the number of motor steps over long time, which is defined as the variance of number of motor steps divided by the average number of steps. In Ref. [79] it was established that the randomness in the number of motor steps is the same as the randomness in the cycle time. This equivalence result is significant because it relates something that we can measure to something that we want to know. The equivalence between the two randomnesses was derived using the approach of Laplace transform.[79] As we will show below, when the cycle time distribution has a discrete component the approach of Laplace transform breaks down but amazingly the equivalence between the two randomnesses still holds. Below, we will first review the equivalence proof using the approach of Laplace transform. Then we present a counter example to show how the approach of Laplace transform breaks down. After that we present a new derivation for the equivalence between the two randomnesses. The new derivation is based on the framework of semi-Markov processes with age structure.

Let l be the motor step size and T be the random cycle time of a motor step. Let $p(t)$ denote the probability density of the random cycle time T. The randomness in the cycle time T is denoted by R_T and is mathematically defined as

$$R_T \equiv \frac{\text{var}(T)}{\langle T \rangle^2}$$

where $\langle T \rangle$ denotes the mean and $\text{var}(T) \equiv \langle T^2 \rangle - \langle T \rangle^2$ denotes the variance of the random cycle time T. R_T informs us about the internal motor operation but R_T is not directly measurable in experiments. A quantity that can be measured in experiments is the randomness in the number of steps over long time. Let $N(t)$ be the stochastic number of steps that the motor has completed by time t. The motor position at time t is $X(t) = N(t)l$. The randomness in the

number of steps is defined below and can be calculated from $X(t)$

$$R_N \equiv \lim_{t\to\infty} \frac{\text{var}(N(t))}{\langle N(t)\rangle} = \lim_{t\to\infty} \frac{\text{var}(X(t))}{\langle X(t)\rangle \cdot l}$$

If each motor step is simply a Markov step with transition rate r, then the cycle time has the exponential distribution, $p(t) = r\cdot\exp(-rt)$, and the number of steps completed by time t has the Poisson distribution, $\Pr[N(t)=n] = \lambda^n \exp(-\lambda)/n!$ where $\lambda = rt$. It follows that

$$\langle T^2\rangle = \text{var}(T) = \frac{1}{r^2}, \quad R_T = 1$$
$$\langle N(t)\rangle = \text{var}(N(t)) = \lambda, \quad R_N = 1$$

Thus, we have $R_T = R_N$ for this special case. In Ref. [79], $R_T = R_N$ was established for the general case using the approach of Laplace transform. Let us first review the approach of Laplace transform used in Ref. [79].

Suppose the motor has just arrived at step 0 at time 0. Let $P(n,t) \equiv \Pr[N(t)=n]$. For $n \geq 0$, $P(n,t)$ satisfies

$$P(0,t) = \int_t^\infty p(\tau)\, d\tau$$
$$P(n+1,t) = \int_0^t P(n, t-\tau)p(\tau)\, d\tau$$

Let $\tilde{f}(s)$ denote the Laplace transform of $f(t)$. Taking the Laplace transform of the above equations yields

$$\tilde{P}(0,s) = \frac{1-\tilde{p}(s)}{s}$$
$$\tilde{P}(n+1,s) = \tilde{P}(n,s)\tilde{p}(s)$$

Solving this recursive equation, we get

$$\tilde{P}(n,s) = \frac{1-\tilde{p}(s)}{s}\tilde{p}^n(s)$$

The first two moments of $N(t)$ can be expressed in terms of $P(n,t)$:

$$\langle N(t)\rangle = \sum_{n=0}^\infty nP(n,t)$$
$$\langle N^2(t)\rangle = \sum_{n=0}^\infty n^2P(n,t)$$

Taking Laplace transform of the first two moments gives us

$$L[\langle N(t)\rangle] = \sum_{n=0}^\infty n\tilde{P}(n,s)$$
$$= \frac{1-\tilde{p}(s)}{s}\sum_{n=0}^\infty n\tilde{p}^n(s) = \frac{\tilde{p}(s)}{s(1-\tilde{p}(s))}$$
$$L[\langle N^2(t)\rangle] = \sum_{n=0}^\infty n^2\tilde{P}(n,s)$$
$$= \frac{1-\tilde{p}(s)}{s}\sum_{n=0}^\infty n^2\tilde{p}^n(s) = \frac{\tilde{p}(s)+\tilde{p}^2(s)}{s(1-\tilde{p}(s))^2}$$

We expand $\tilde{p}(s) = \int_0^\infty \exp(-st)p(t)\, dt$ for small s,

$$\left.\frac{d^k\tilde{p}(s)}{ds^k}\right|_{s=0} = \int_0^\infty (-t)^k\exp(-st)p(t)\, dt = (-1)^k\langle T^k\rangle$$
$$\tilde{p}(s) = 1 - \langle T\rangle s + \frac{\langle T\rangle^2}{2}s^2 + \cdots$$

The Laplace transforms of the first two moments have the expansions:

$$L[\langle N(t)\rangle] = \frac{1}{\langle T\rangle}\frac{1}{s^2} - \left(\frac{2\langle T\rangle^2 - \langle T^2\rangle}{2\langle T\rangle^2}\right)\frac{1}{s} + \cdots$$
$$L[\langle N^2(t)\rangle] = \frac{2}{\langle T^2\rangle}\frac{1}{s^3} - \left(\frac{3\langle T\rangle^2 - 2\langle T^2\rangle}{\langle T\rangle^3}\right)\frac{1}{s^2} + \cdots$$

The key assumption in the approach of Laplace transform is that the first two moments as function of time have Laurent expansions as time goes to infinity:

$$\text{Assumption:} \quad \begin{aligned}\langle N(t)\rangle &= \sum_{k=-\infty}^\infty c_k^{(1)}t^k \\ \langle N^2(t)\rangle &= \sum_{k=-\infty}^\infty c_k^{(2)}t^k\end{aligned} \tag{70}$$

Under this key assumption, the coefficients in the Laurent expansions are calculated from their Laplace transforms.

$$\begin{aligned}\langle N(t)\rangle &= \frac{1}{\langle T\rangle}t - \left(\frac{2\langle T\rangle^2 - \langle T^2\rangle}{2\langle T\rangle^2}\right) + \cdots \\ \langle N^2(t)\rangle &= \frac{2}{\langle T^2\rangle}t^2 - \left(\frac{3\langle T\rangle^2 - 2\langle T^2\rangle}{\langle T\rangle^3}\right)t + \cdots\end{aligned} \tag{71}$$

which leads to

$$R_N = \lim_{t\to\infty}\frac{\langle N^2(t)\rangle - \langle N(t)\rangle^2}{\langle N(t)\rangle} = \frac{\langle T^2\rangle - \langle T\rangle^2}{\langle T\rangle^2} = R_T$$

Assumption (70) is the key in the approach of Laplace transform. Note that the existence of Laurent expansion as $s\to 0$ for the Laplace transform does not imply the existence of Laurent expansion as $t\to\infty$ for the original function. For example, function $f(t) = t - \sin(t)$ does not have a Laurent expansion as $t\to\infty$ while its Laplace transform $\tilde{f}(s) = 1/s^2 - (1/s^2+1)$ has a Laurent expansion as $s\to 0$. When the cycle time distribution has a discrete component, assumption (27) breaks down but the conclusion $R_N = R_T$ is still true. To illustrate this, we consider the discrete cycle time distribution given below:

$$\Pr[T=t] = \begin{cases} 0.5, & t=1 \\ 0.5, & t=2 \\ 0, & \text{otherwise}\end{cases}$$

The first two moments of T and the randomness in T are

$$\langle T\rangle = \frac{3}{2}, \quad \langle T^2\rangle = \frac{5}{2}, \quad \text{var}(T) = \frac{1}{4}, \quad R_T = \frac{1}{9}$$

REVIEW

To derive the first two moments of $N(t)$, we notice that a step can only be completed at an integer value of time. So we only need to calculate the first moments of $N(k)$ where k is an integer. The specific form of the cycle time distribution tells us that the first step is either completed at time 1 or at time 2. From this fact, we obtain the recursive relation

$$\langle N(k)\rangle = \frac{1}{2}[1+\langle N(k-1)\rangle]+\frac{1}{2}[1+\langle N(k-2)\rangle]$$

Using the initial condition $\langle N(0)\rangle = 0$ and $\langle N(1)\rangle = 0.5$, we solve the recursive relation to obtain an expression for the first moment. The second moment is calculated in a similar way. Thus, we have

$$\langle N(t)\rangle = \frac{2}{3}k - \frac{1}{9}\left[1-\left(\frac{-1}{2}\right)^k\right], \quad k \le t < k+1$$

$$\langle N^2(t)\rangle = \frac{4}{9}k^2 - \frac{2}{27}k\left[1-\left(\frac{-1}{2}\right)^k\right]+\frac{1}{2}\left[1-\left(\frac{-1}{2}\right)^k\right] \quad k \le t < k+1$$

It is clear that Laurent expansions (71), predicted from the Laplace transform, are invalid for this discrete cycle time distribution. The equivalence between R_N and R_T, however, still holds.

Now we present a new derivation that is valid for arbitrary cycle time distribution as long as the first moments of the cycle time exist. The new derivation is based on semi-Markov process with age structure. In a semi-Markov process, the system does not remember how it got to the current state but it does remember its age. The age of the system is defined as the time elapsed since the latest arrival at the current state. At the moment when the system moves to a new state, the age is reset to zero. After that the age increases with the time until the next jump. Let $A(t)$ be the age of the system at time t. Let $\rho(n,\tau,t)$ be the probability density that the system has age τ in state n at time t. Here n refers to the number of motor steps completed. For a stochastic stepper, $\rho(n,\tau,t)$ is governed by the conservation of probability:

$$\frac{\partial\rho(n,\tau,t)}{\partial t} = -\frac{\partial\rho(n,\tau,t)}{\partial\tau} - \rho(n,\tau,t)\beta(\tau)$$
$$\beta(\tau) \equiv \frac{p(\tau)}{\int_\tau^\infty p(s)\,ds} \tag{72}$$

where $\beta(\tau)$ is called the hazard function and is the conditional rate of completing the current cycle at age τ given that the cycle is not completed in age $[0,\tau)$. In the above equation, the first term on the right hand side corresponds to the fact that if no transition occurs then the age of the system increases with the time. The second term reflects that a transition out of a state reduces the probability of that state. The hazard function $\beta(\tau)$ satisfies

$$-\int_0^\tau \beta(s)\,ds = \ln\left(\int_\tau^\infty p(s)\,ds\right)$$

This relation between $\beta(s)$ and $p(s)$ will be very useful in our new derivation. Since all new arrivals have age 0, the boundary condition for $\rho(n,\tau,t)$ is

$$\rho(0,0,t) = 0, \quad t > 0$$
$$\rho(n,0,t) = \int_0^\infty \rho(n-1,\tau,t)\beta(\tau)\,d\tau, \quad t > 0,\ n > 0$$

To determine the evolution of $\rho(n,\tau,t)$, an initial state-age distribution needs to be specified. In experiments, a more realistic situation is that the system has a random age at time 0. This is especially true if in the statistical calculation the ensemble is created by cutting one or a few long time series into many time series. Mathematically, the age distribution at time t, regardless of the state, is

$$\rho(\tau,t) = \sum_{n=0}^\infty \rho(n,\tau,t)$$

The governing equation and the boundary condition for the age distribution are obtained by summing the evolution Eq. (72) for $\rho(n,\tau,t)$ over all states

$$\frac{\partial\rho(\tau,t)}{\partial t} = -\frac{\partial\rho(\tau,t)}{\partial\tau} - \rho(\tau,t)\beta(\tau)$$
$$\rho(0,t) = \int_0^\infty \rho(\tau,t)\beta(\tau)\,d\tau$$

Let $\rho^{(S)}(\tau)$ be the stationary age distribution. Solving for $\rho^{(S)}(\tau)$ from the above equation, we get

$$\rho^{(S)}(\tau) = \frac{\int_\tau^\infty p(s)\,ds}{\int_0^\infty \tau p(\tau)\,d\tau} = \frac{\int_\tau^\infty p(s)\,ds}{\langle T\rangle}$$

Thus, to be consistent with experiments, we use the stationary age distribution as the initial condition.

Initial condition:
$$\rho(0,\tau,0) = \rho^{(S)}(\tau)$$
$$\rho(n,\tau,0) = 0, \quad n > 0$$

With the stationary age distribution as the initial condition, the age distribution will remain stationary. Expressing the first moment of $N(t)$ in terms of $\rho(n,\tau,t)$, differentiating with respect to t, using the governing Eq. (72) for $\rho(n,\tau,t)$, and using the boundary condition, we obtain

$$\frac{d\langle N(t)\rangle}{dt} = \sum_{n=0}^\infty n\int_0^\infty \frac{\partial\rho(n,\tau,t)}{\partial t}\,d\tau$$
$$= -\sum_{n=0}^\infty n\int_0^\infty \left(\frac{\partial\rho(n,\tau,t)}{\partial\tau} + \rho(n,\tau,t)\beta(\tau)\right)d\tau$$
$$= \sum_{n=0}^\infty n(\rho(n,0,t)-\rho(n+1,0,t))$$
$$= \sum_{n=0}^\infty \rho(n,0,t) = \rho^{(S)}(0) = \frac{1}{\langle T\rangle}$$

Similarly for the second moment of $N(t)$, we have

$$\begin{aligned}\frac{d\langle N^2(t)\rangle}{dt} &= \sum_{n=0}^{\infty} n^2 \int_0^{\infty} \frac{\partial \rho(n,\tau,t)}{\partial t}\, d\tau \\ &= -\sum_{n=0}^{\infty} n^2 \int_0^{\infty} \left(\frac{\partial \rho(n,\tau,t)}{\partial \tau} + \rho(n,\tau,t)\beta(\tau) \right) d\tau \\ &= \sum_{n=0}^{\infty} n^2 (\rho(n,0,t) - \rho(n+1,0,t)) \\ &= \sum_{n=0}^{\infty} (2n-1)\rho(n,0,t) \\ &= \frac{2}{\langle T\rangle} q(0,t) + \frac{2t}{\langle T\rangle^2} - \frac{1}{\langle T\rangle}\end{aligned}$$

where the function $q(\tau, t)$ is defined as

$$q(\tau,t) \equiv \langle T\rangle \sum_{n=0}^{\infty} n\rho(n,\tau,t) + (\tau - t)\rho^{(S)}(\tau)$$

Integrating the derivatives of the first two moments obtained above gives us

$$\langle N(t)\rangle = \frac{t}{\langle T\rangle}$$

$$\langle N^2(t)\rangle = \frac{2}{\langle T\rangle}\int_0^t q(0,s)\, ds + \frac{t^2}{\langle T\rangle^2} - \frac{t}{\langle T\rangle}$$

Using these results to calculate R_N, we arrive at

$$\begin{aligned}R_N &= \lim_{t\to\infty} \frac{\langle N^2(t)\rangle - \langle N(t)\rangle^2}{\langle N(t)\rangle} = \lim_{t\to\infty} \frac{2}{t}\int_0^t q(0,s)\, ds - 1 \\ &= 2\left(\lim_{t\to\infty} \frac{1}{t}\int_0^t q(0,s)\, ds - \frac{\langle T^2\rangle}{2\langle T\rangle^2} \right) + R_T\end{aligned}$$

It can be shown that the function $q(\tau, t)$ defined above satisfies

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t q(0,s)\, ds = \frac{\langle T^2\rangle}{2\langle T\rangle^2}$$

Combining these results leads immediately to the desired conclusion: $R_N = R_T$. Therefore, the two randomnesses: the randomness in the cycle time and the randomness in the number of motor steps, are equivalent to each other for arbitrary cycle time distribution provided that the first two moments of the cycle time exist. In particular, the equivalence holds even when the cycle time distribution has a discrete component, in which case the approach of Laplace transform breaks down. It is unclear whether and how the approach of Laplace transform can be rescued. The derivation above is based on the framework of semi-Markov processes with age structure.

8. CONCLUDING REMARKS

In this review, we have discussed and clarified a few issues in modeling molecular motors. We started with a modeling framework of intermediate level: Markov Fokker-Planck equations. The modeling framework is one of good compromises that are sophisticated enough to accommodate the current experimental settings and at the same time are simple enough for theoretical analysis and computer simulations. This framework is not the only good compromise. Depending on the particular experimental setup and specific goal of modeling, we need either to modify this framework or to adopt a different framework. Modeling frameworks at all levels are valuable. Each is useful in a different situation. In the framework we adopt, the unidirectional motion of the motor is described explicitly by a one-dimensional continuous variable while chemical reaction is modeled as changes of occupancy state. The continuous description of mechanical motion allows us to make connections to measurements of mechanical properties in single molecule experiments and allows us to study mechanical behaviors theoretically.

A subtle issue closely associated with the chemical reaction is detailed balance. Traditionally, detailed balance is only applied to systems in equilibrium. Also traditionally, a chemical reaction cycle is schematically represented as a closed loop. When the system is away from equilibrium we say detailed balance is broken. Indeed, for a system described by a close loop and satisfying detailed balance, it must be at equilibrium. So a chemical reaction away from equilibrium either breaks detailed balance or should not be described by a closed loop. We prefer the later. When we consider the free energy, we do have to distinguish the corresponding occupancy states in different cycles. Thus, from energetic point of view, a chemical reaction away from equilibrium is not a closed loop. Instead it is a helical sequence of cycles. When the system completes a cycle, it does not go back to the state it started with. It goes to the next cycle, which, in the coordinate of free energy, is below the current cycle. More precisely, while the occupancy state is repeated going from one cycle to the next, the free energy value is nor repeated. In the diagram of an infinite sequence of cycles, a detailed balance like condition is still satisfied. Enforcing this detailed balance like condition away from equilibrium has two benefits: it ensures that detailed balance is automatically obeyed at equilibrium; and it provides constraints on chemical transition rates. In contrast, when we say a system breaks detailed balance. It contains very little information on the transition rates.

The motor potential profile is a single potential that characterizes the overall effect of the chemical reaction on motor motion when the motor is at the steady state. The introduction of motor potential profile is another example of adapting modeling framework to accommodate experimental setup and measurements. In the Markov Fokker-Planck framework, the motor is driven by switching among a set of N potentials, each corresponding to an individual chemical occupancy state. While it is desirable to construct all N potentials, the measurements in the current

REVIEW

single molecule experiments do not contain enough information on occupancy state. As a result, it is still not realistic to extract potentials for individual states. The motor force profile is the average of motor force over all occupancy states weighted by steady state probability densities of these states. The motor potential profile is the integral of the motor force profile. In averaging over all states, the motor operation is now represented by a single potential. Although the motor potential profile does not contain all information about the motor operation, it tells us how the motor force varies with the motor position, which may reveal the coupling between the chemical reaction and motor motion. The biggest motivation for the motor potential profile is that, in principle, it can be recovered from measured time series of motor positions. We proposed a robust mathematical formulation for constructing the motor potential profile and for estimating the associated statistical error.

Efficiency of molecular motors is another subtle issue. When a motor is working against a conservative loading force, the thermodynamic efficiency is well defined and is the energy conversion efficiency. The energy input is the chemical free energy consumed in reaction and the energy output is the increase of potential energy in the external agent exerting the conservative loading force. When a motor is working against a viscous drag, there is no energy output and the thermodynamic efficiency does not apply. In this situation, the Stokes efficiency was introduced to measure how efficiently the motor is utilizing the chemical free energy to drive the motor in the viscous media. Because of the small size, for a molecular motor, the motion is dominated by large thermal fluctuations and the effect of inertia is negligible. The motor motion is highly stochastic, full of fast decay and fast excitation of the instantaneous velocity. This is in sharp contrast with the case of macroscopic motors where the instantaneous velocity is nearly constant over many reaction cycles. In a molecular motor, kinetic energy can flow into the motor motion from the fluid environment by thermal excitations. It is also possible that the amount of kinetic energy flowing from the motor motion to the fluid environment is more than the free energy consumption. These behaviors of molecular motors tell us that the heat generated via the motor motion is not a good candidate for the numerator of an efficiency definition. In the Stokes efficiency, the numerator measures the mechanical performance. But it is not the heat generated via the motor motion, and it does not have a clear thermodynamic meaning. Therefore, the Stokes efficiency is different from the thermodynamic efficiency and the Stokes efficiency being bounded by 100% cannot be argued simply from thermodynamics. The Stokes efficiency being bounded by 100% has to be derived mathematically in a modeling framework. The concept of motor potential profile allows us to divide the process from chemical reaction to mechanical motion into two sub-processes. First, the motor potential profile is produced in the chemical reaction. This sub-process is measured by the chemical efficiency. Second, the motor potential profile is utilized to driving the motor motion. This sub-process is measured by the mechanical efficiency. The overall Stokes efficiency is the product of the chemical efficiency and the mechanical efficiency. The mechanical efficiency is completely determined by the shape of the motor potential profile. High mechanical efficiency means the motor potential profile is close to being a constant slope. The chemical efficiency may be reduced by slow chemical transitions and/or large free energy changes associated with transitions. High chemical efficiency means none of the transitions is rate limiting and equilibrium constants of the transitions are all near 1. The difference between the thermodynamic efficiency and the Stokes efficiency is also manifested in stall force. The conservative stall force and the viscous stall drag are measured in different experimental setups and they tell us different aspects about the motor. It is valuable to measure both of these two stall forces and we should not expect them to be the same.

Solving Fokker-Planck equations efficiently and accurately is an important part of modeling molecular motors. Since the motion of a molecular motor is dominated by large thermal fluctuations, it is important to ensure that in the numerical process, these large fluctuations do not create any artificial drift. The best way to achieve this goal is to preserve detailed balance. We have demonstrated that when the potential is discontinuous, detailed balance is a necessary condition for converging to the correct solution. We also proposed a new method that preserves the second order accuracy even when the potential is discontinuous.

Acknowledgments: The author was partially supported by the US National Science Foundation.

References

1. H. Wang and G. Oster, *Nature* 396, 279 (**1998**).
2. G. Oster and H. Wang, *Biochimica et Biophysica Acta (Bioenergetics)* 1458, 482 (**2000**).
3. C. S. Peskin, G. M. Odell, and G. Oster, *Biophys. J.* 65, 316 (**1993**).
4. T. Elston, H. Wang, and G. Oster, *Nature* 391, 510 (**1998**).
5. A. Mogilner and G. Oster, *Eur. J. Biophys.* 28, 235 (**1999**).
6. R. Astumian, *Science* 276, 917 (**1997**).
7. P. Reimann, *Physics Reports* 361, 57 (**2002**).
8. R. D. Astumian and I. Derenyi, *Eur. Biophys. J.* 27, 474 (**1998**).
9. A. Parmeggiani, F. Julicher, A. Ajdari et al., *Phys. Rev. E* 60, 2127 (**1999**).
10. T. Schlick, Molecular Modeling and Simulation, Springer (**2002**).
11. Y. Q. Gao, W. Yang, and M. Karplus, *Cell* 123, 195 (**2005**).
12. I. Antes, D. Chandler, H. Wang et al., *Biophys. J.* 85, 695 (**2003**).
13. R. Vale, *TIBS* 11, 464 (**1986**).
14. J. Howard, A. J. Hudspeth, and R. D. Vale, *Nature* 342, 154 (**1989**).
15. K. Visscher, M. Schnitzer, and S. Block, *Nature* 400, 184 (**1999**).
16. C. Coppin, D. Pierce, L. Hsu et al., *Proc. Natl. Acad. Sci. USA* 94, 8539 (**1997**).
17. J. Abrahams, A. Leslie, R. Lutter et al., *Nature* 370, 621 (**1994**).
18. D. Sabbert, S. Engelbrecht, and W. Junge, *Nature* 381, 623 (**1996**).
19. H. Noji, R. Yasuda, M. Yoshida et al., *Nature* 386, 299 (**1997**).
20. R. I. Menz, J. E. Walker, and A. G. Leslie, *Cell* 106, 331 (**2001**).

21. S. M. Block and H. C. Berg, *Nature* 309, 470 **(1984)**.
22. H. C. Berg, *Annu. Rev. Biochem.* 72, 19 **(2003)**.
23. J. Xing, F. Bai, R. Berry et al., *Proc. Natl. Acad. Sci.* 103, 1260 **(2006)**.
24. J. Prost, J. Chauwin, L. Peliti et al., *Phys. Rev. Lett.* 72, 2652 **(1994)**.
25. F. Julicher, A. Ajdari, and J. Prost, *Reviews of Modern Physics* 69, 1269 **(1997)**.
26. H. C. Berg, Random Walks in Biology, Princeton University Press, Princeton, NJ **(1993)**.
27. F. Reif, Fundamentals of Statistical and Thermal Physics, McGraw Hill, New York **(1965)**.
28. L. D. Landau, E. M. Lifshitz, and L. P. Pitaevskii, Statistical Physics, Pergamon Press, Oxford, New York **(1980)**.
29. H. Risken, The Fokker-Planck Equation, Springer-Verlag, New York **(1989)**.
30. M. Fischer, J. Cox, D. J. Davis et al., *Fungal Genet. Biol.* 41, 698 **(2004)**.
31. N. P. Money, *Mycologia* 90, 547 **(1998)**.
32. P. Dimroth, H. Wang, M. Grabe et al., *Proc. Natl. Acad. Sci. USA* 96, 4924 **(1999)**.
33. P. Boyer, *Biochim. Biophys. Acta* 1140, 215 **(1993)**.
34. J. Weber and A. E. Senior, *Biochim. Biophys. Acta* 1319, 19 **(1997)**.
35. P. Boyer, *Annu. Rev. Biochem.* 66, 717 **(1997)**.
36. C. Gardiner, Handbook of Stochastic Methods, Springer-Verlag, New York **(1985)**.
37. C. S. Peskin, G. B. Ermentrout, and G. F. Oster, Interplay of Genetic and Physical Processes in the Development of Biological Form, edited by F. Dykstra, D. Beysens, M.-A. Felix, G. Gorgacs, and F. Gail, Springer-Verlag, Les Houches **(1994)**.
38. C. Doering, W. Horsthemke, and J. Riordan, *Phys. Rev. Lett.* 72, 2984 **(1994)**.
39. H. Wang and G. Oster, *Europhys. Lett.* 57, 134 **(2002)**.
40. G. Oster and H. Wang, *Trends in Cell Biology* 13, 114 **(2003)**.
41. M. O. Magnasco, *Phys. Rev. Lett.* 71, 1477 **(1993)**.
42. M. O. Magnasco, *Philosophical Magazine B* 69, 39 **(1994)**.
43. D. Keller and C. Bustamante, *Biophys. J.* 78, 541 **(2000)**.
44. A. Katchalsky and P. Curran, Nonequilibrium Thermodynamics in Biophysics, Harvard University Press, Cambridge, MA **(1965)**.
45. D. K. Kondepudi and I. Prigogine, Modern Thermodynamics, John Wiley, New York **(1998)**.
46. L. Stryer, Biochemistry, W. H. Freeman, New York **(1995)**.
47. A. Lehninger, D. Nelson, and M. Cox, Biochemistry, Worth, New York **(1993)**.
48. C. Matthews, K. V. Holde, and K. Ahern, Biochemistry, Benjamin Cummings, San Francisco, California **(2000)**.
49. E. A. Abbondanzieri, W. J. Greenleaf, J. W. Shaevitz et al., *Nature* 438, 460 **(2005)**.
50. A. J. Hunt, F. Gittes, and J. Howard, *Biophys. J.* 67, 766 **(1994)**.
51. M. J. Schnitzer and S. M. Block, *Nature* 388, 386 **(1997)**.
52. H. Itoh, A. Takahashi, K. Adachi et al., *Nature* 427, 465 **(2004)**.
53. Y. Hirono-Hara, K. Ishizuka, K. Kinosita et al., *Proc. Natl. Acad. Sci. USA* 120, 4288 **(2005)**.
54. R. Yasuda, H. Noji, K. Kinosita et al., *Cell* 93, 1117 **(1998)**.
55. Y. Sambongi, Y. Iko, M. Tanabe et al., *Science* 286, 1722 **(1999)**.
56. S. M. Block, C. L. Asbury, J. W. Shaevitz et al., *Proc. Natl. Acad. Sci. USA* 100, 2351 **(2003)**.
57. R. Yasuda, H. Noji, M. Yoshida et al., *Nature* 410, 898 **(2001)**.
58. A. Samuel and H. Berg, *Proc. Natl. Acad. Sci.* 92, 3502 **(1995)**.
59. H. Wang, *IEE Proceedings Nanobiotechnology* 150, 127 **(2003)**.
60. H. Wang, *J. Theo. Biol.* 242, 908 **(2006)**.
61. J. W. Harris and H. Stocker, Handbook of Mathematics and Computational Science, Springer-Verlag, New York **(1998)**.
62. P. G. Hoel, Introduction to Mathematical Statistics, Wiley, New York **(1962)**.
63. R. A. Fisher, *Philos. Trans. Roy. Soc. London Ser. A* 222, 309 **(1922)**.
64. R. Fletcher and C. M. Reeves, *Computer Journal* 7, 148 **(1964)**.
65. E. Polak and G. Ribiere, *Revue. Francaise d'Informatique et de Recherche Operationnelle* 16, 35 **(1969)**.
66. D. B. Walton, Ph.D. Dissertation, University of Arizona **(2002)**.
67. I. Derenyi, M. Bier, and D. Astumian, *Phys. Rev. Lett.* 83, 903 **(1999)**.
68. H. Wang, *J. Phys.: Condens. Matter* 17, S3997 **(2005)**.
69. T. Elston and C. Peskin, *SIAM Journal of Applied Math.* 60, 842 **(2000)**.
70. T. Elston, D. You, and C. Peskin, *SIAM Journal of Applied Math.* 61, 776 **(2000)**.
71. A. Raj and C. S. Peskin, *Proc. Natl. Acad. Sci.* 103, 5349 **(2006)**.
72. H. Wang, C. Peskin, and T. Elston, *J. Theo. Biol.* 221, 491 **(2003)**.
73. H. Wang, C. Peskin, and T. Elston, *SIAM Journal on Numerical Analysis* 45, 1425 **(2007)**.
74. J. Fricks, H. Wang, and T. Elston, *J. Theo. Biol.* 239, 33 **(2006)**.
75. H. Wang, *J. Comput. Theor. Nanosci.* 3, 922 **(2006)**.
76. T. Elston and C. Doering, *J. Stat. Phys.* 83, 359 **(1996)**.
77. J. Xing, H. Wang, and G. Oster, *Biophys. J.* 89, 1551 **(2005)**.
78. J. W. Shaevitz, S. M. Block, and M. J. Schnitzer, *Biophys. J.* 89, 2277 **(2005)**.
79. M. J. Schnitzer and S. M. Block, *Cold Spring Harb. Symp. Quant. Biol.* 60, 793 **(1995)**.
80. K. Svoboda, P. P. Mitra, and S. M. Block, *Proc. of the National Academy of Sciences* 91, 11782 **(1994)**.

Received: 14 July 2007. Accepted: 18 August 2007.

REVIEW