MOLECULAR MOTORS AND STOCHASTIC NETWORKS

REINHARD LIPOWSKY and STEFFEN LIEPELT

Max-Planck-Institute of Colloids and Interfaces
Science Park Golm, 14424 Potsdam, Germany
E-mail: reinhard.lipowsky@mpikg-golm.mpg.de, liepelt@mpikg.mpg.de
http://www.mpikg.mpg.de/theory/

Abstract. Molecular motors are nano- or colloidal machines that keep the living cell in a highly ordered, stationary state far from equilibrium. This self-organized order is sustained by the energy transduction of the motors, which couple exergonic or 'downhill' processes to endergonic or 'uphill' processes. A particularly interesting case is provided by the chemomechanical coupling of cytoskeletal motors which use the chemical energy released during ATP hydrolysis in order to generate mechanical forces and to perform mechanical work. We describe a general network theory for molecular motors which leads to local and nonlocal balance conditions that are valid far from equilibrium and generalize the well-known detailed balance conditions for equilibrium states. The nonlocal balance conditions may also be viewed as generalizations of the classical law of mass action to small systems that transduce chemical energy into mechanical work. As a pedagogical example, we discuss the simple case of a single motor head. We also review the application to two-headed motors such as kinesin, which serves as a paradigmatic example for energy transduction via chemomechanical coupling.

1. Introduction and outline. All biological systems are built up from the same kind of building blocks. In fact, these building blocks or 'modules' form a whole hierarchy of structures that cover a wide range of length and time scales. Bottom–up, i.e., as one moves from small scales to larger ones, these structural levels are: aqueous solutions, i.e., water and ions, small molecules or monomers, polymers, supramolecular assemblies, organelles, bacterial or procaryotic cells, eucaryotic cells of animals and plants, populations of cells, tissues etc.

The most basic levels of this hierarchy correspond to nanostructures and colloids that extend from molecules to cells as shown in Fig. 1. The first level contains the small

molecules or monomers that provide the precursors for most macromolecules in the cell. These monomers can be divided into several subsets or 'alphabets'. Different elements of one alphabet can be connected by covalent bonds which leads to the formation of four classes of polymers: nucleic acids (RNA and DNA), proteins, polysaccharides, and lipids, see second level (or column) of Fig. 1.
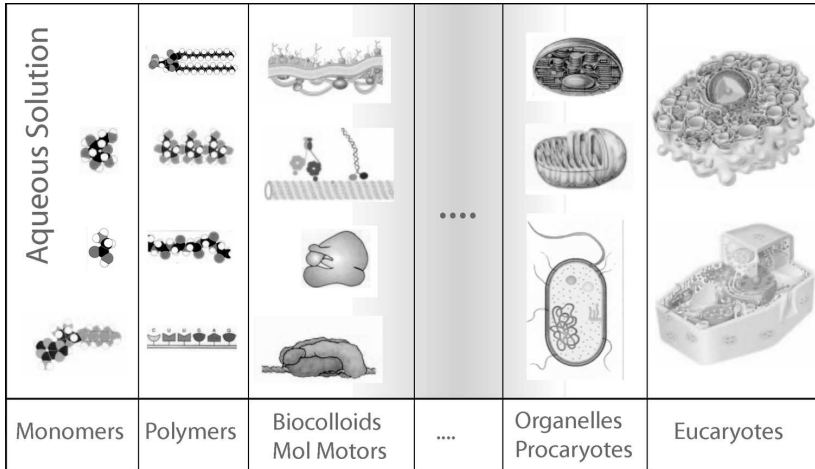

Fig. 1. Hierarchy of biological nanostructures and colloids: The first three levels on the left correspond to aqueous solutions, polymers (nucleic acids, proteins, polysaccharides, and lipids) as well as biocolloids and molecular motors (polymerases, ribosomes, cytoskeletal filaments and motors, membranes), respectively. The last two levels on the right correspond to bacterial or procaryotic cells and to eucaryotic cells. The purple regime between biocolloids and procaryotes represents the twilight zone between matter and life.

These polymers assemble into biocolloids and molecular machines, see third level of Fig. 1. The four examples shown for this latter level are in ascending order (i) a RNA polymerase that moves along a DNA strand and transcribes it into RNA, (ii) a ribosome, (iii) a cytoskeletal filament with two motor proteins, and (iv) a biomembrane. The integration of different types of biocolloids and molecular machines leads to the simplest form of life, i.e., to bacterial or procaryotic cells that have no interior subcompartments. The corresponding level in Fig. 1 also displays certain organelles such as mitochondria and chloroplasts, which are believed to have evolved from procaryotic precursor cells. Finally, the combination and integration of different types of organelles leads to eucaryotic cells as shown in the last level (or column) of Fig. 1.

Biopolymers are dead objects whereas procaryotic cells represent the simplest form of life. Thus, the regime between polymers and bacteria corresponds to the twilight zone between matter and life as indicated by the purple regime in Fig. 1. At the lower boundary of this zone, we find many different molecular motors and colloidal machines that perform various tasks such as assembly of polymers, creation of concentration gradients across membranes, transport of vesicles, cell division, and cell locomotion. These motors and machines act like little demons or nanorobots that keep the living cell in a highly ordered,

stationary state far from equilibrium, see Fig. 2(b). This self-organization is based on the energy transduction of the motors, which couple exergonic or 'downhill' processes to endergonic or 'uphill' processes. A particularly interesting case is provided by energy transduction based on chemomechanical coupling. In this latter case, the motors use the chemical energy released during ATP hydrolysis in order to generate mechanical forces and to perform mechanical work.
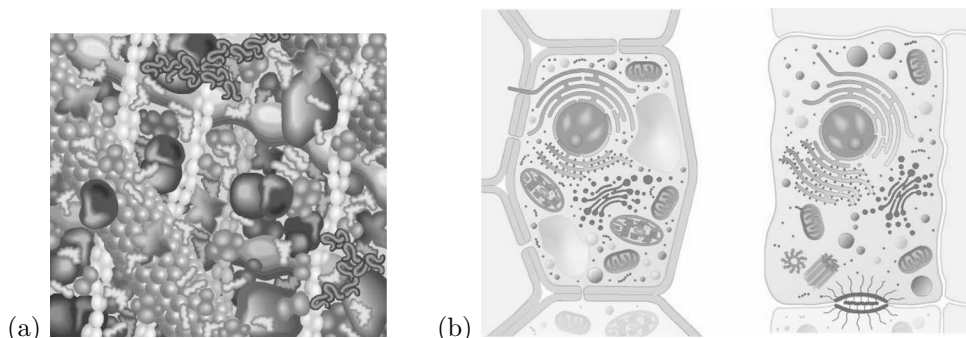

Fig. 2. Spatial disorder and order at the molecular and colloidal scale, respectively: (a) Nanoscopic view into a living cell which reveals many different molecules and molecular assemblies that appear to be spatially disordered. The width of this image corresponds to 125 nm; and (b) Microscopic view of eucaryotic cells which exhibit soft colloidal order on the scale of hundreds of nanometers. The linear dimension of these cells is tens of micrometers.

This review is organized as follows. In Section 2, the main features of molecular motors are discussed in a qualitative manner. It is emphasized that molecular motors are colloidal machines that (i) either involve a molecular complex that interacts with a much larger supramolecular assembly or (ii) are based on the supramolecular assembly process itself. One class of motors that has been studied in much detail during the last decade consists of cytoskeletal motors such as kinesin that walk along cytoskeletal filaments by transforming the chemical energy of ATP hydrolysis into mechanical work. The thermodynamics of the motor system and its ATP hydrolysis is discussed in Section 3. As explained in this latter section, the motor is viewed as a small system that is always taken to be in thermal equilibrium but may be far from chemical equilibrium with its enviroment.

The understanding of molecular motors provides a challenge since these motors are capable of chemomechanical coupling and energy transduction even though they exhibit several aspects of randomness [1]. This behavior can be addressed within a general network theory [2, 3, 4] as explained in Sections 4, 5, and 6. Section 4 introduces network models for a single motor head or catalytic domain and for two-headed motors with two catalytic domains. The steady state fluxes of these models are discussed in Section 5, the steady state entropy production and energy balance in Section 6, in which we review our recent results as obtained in [2]. Our theory represents an extension of the theoretical framework developed by T. L. Hill and coworkers for biochemical networks [5], which do not involve mechanical forces, to the case of chemomechanical coupling as present in molecular motors. In particular, the thermodynamic forces as considered by Hill and

coworkers are shown to be equal to the entropy produced during the energy transduction. The latter relation is useful since the produced entropy can be identified with the heat released by the motor, and the latter quantity is accessible to experiment, at least in principle. A detailed comparison of our network theory with other theoretical studies is provided in [4].

As described in Section 6, our theory leads to energy balance conditions that provide generalizations of the classical law of mass action to small systems that can perform mechanical work. We also discuss a local form of the balance conditions that involves the internal energies of the motor states and the entropy produced during a single transition. The latter entropy follows directly from the decomposition of the entropy production rate in the steady state and agrees with the more general proposal of U. Seifert [6] when applied to steady states. The identification of the entropy produced by the network model with the heat released by the system relies on an appropriate level of description which includes the relevant states and transitions between these states. In particular, one should include those transitions, that describe the coupling of the motor to the different reservoirs, which are essential for energy transduction. This latter aspect is illustrated in Section 7 where we discuss energy transduction of a single motor head as described (i) by a 4-state network as introduced by T. L. Hill [5] and (ii) by a reduced 3-state network. The final Section 8 gives a brief summary of the application of our theoretical framework to the cytoskeletal motor kinesin [3], which provides a case study for chemomechanical coupling.

2. Multiscale motility of molecular motors. In each living cell, there are several types of molecular motors and machines which fulfill different functions [7]. Prominent examples are: (i) DNA and RNA polymerases which move along the strands of DNA in order to replicate it and to transcribe it into RNA; (ii) ribosomes that attach to mRNA and translate the nucleotide sequence into proteins; (iii) membrane pumps which transport ions and small molecules across membranes; the resulting concentration gradients may be used in order to drive (iv) rotary motors such as the bacterial flagellar motor and the $F_0 F_1$–ATPase which are used for cell locomotion and ATP synthesis, respectively; (v) myosins in muscles which work in ensembles and collectively displace actin filaments; (vi) cytoskeletal motors which bind to the filaments of the cytoskeleton and then walk along these filaments in a directed fashion; and (vii) Cytoskeletal filaments that generate pushing forces during their assembly and growth.

In most cases, the molecular motor consists of a protein complex that interacts with another colloidal structure such as a macromolecule, membrane, or filament. In case (vii), the molecular machine is provided by the assembly of many proteins into a filament. In all cases, the molecular motor involves the interaction of at least two different species of molecules which bind together to form a larger colloidal complex that acts as a nanomachine.

Apart from case (iv), the molecular motors just mentioned are powered by the hydrolysis of adenosine triphosphate (ATP). Thus, all of these motors represent ATPases, i.e., catalysts or enzymes for the hydrolysis of ATP. For the concentrations which prevail in

living cells, the ATP hydrolysis is strongly exergonic or 'downhill' but it is also quite slow in the absence of any enzymatic activity. The motors act as enzymes for this chemical reaction which leads to much faster hydrolysis rates. In addition, these motors are also able to transform the energy released from the ATP hydrolysis into useful work. In fact, this energy transduction occurs even at the level of *single* hydrolysis events. In this way, they consume quantized amounts of fuel.

In this article, we will focus on cytoskeletal motors such as kinesin, see Fig. 3(a), that are essential for intracellular transport, cell division, and cell locomotion [8, 9]. Three superfamilies of cytoskeletal motors have been identified: kinesins, dyneins and myosins [9, 10]. Kinesins and dyneins bind to microtubuli as shown in Fig. 3(a) whereas myosins bind to actin filaments.

The motor properties of kinesin have been determined in a long series of sophisticated experiments [11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22]. Its processive walks consist of about 100 steps and take a few seconds [11]. In its bound state, kinesin walks in a 'hand-over-hand' fashion, i.e., by alternating steps in which one head moves forward while the other one remains bound to the filament [12, 19]. Each step corresponds to a motor displacement of 8 nm corresponding to the lattice constant of the microtubule. If there is no shortage of ATP, the motor makes about 100 steps in one second which leads to a velocity of about one micrometer per second. The absolute value of this velocity is not very impressive, but relative to its size, the motor molecule moves very fast: On the macroscopic scale, its movement would correspond to an athlete who runs 200 meters in one second! This is even more surprising if one realizes that the motor moves in a very viscous and noisy environment since it steadily undergoes thermally excited collisions with a large number of water molecules.
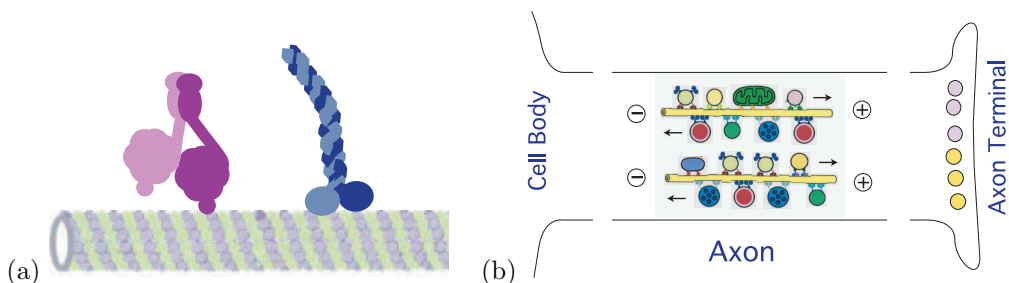

Fig. 3. Multiscale motility of molecular motors: (a) Microtubule with one dynein (violett) and one kinesin (blue) motor. The tubulin dimers form a lattice of binding sites with a lattice parameter of 8 nm; and (b) Bidirectional transport of vesicles and other types of cargo within an axon. The longest axons in our body have a length of about half a meter.

Indeed, because of this thermal noise, a single motor molecule falls off from the filament after it has made about a hundred steps. It has then covered a walking distance or run length of about one micrometer [11] which is rather small compared to the long distances – centimeters or even meters – over which cargo is transported in cells and axons, see Fig. 3(b). One rather effective way to increase the run length is via coopera-

tive transport of cargo particles by several motor molecules [23]. The corresponding run length distribution has been recently measured experimentally for an in vitro assay [24]. On length scales that are large compared to a single step along the filament, cytoskeletal motors undergo peculiar random walks that consist of both directional movement along the filament and diffusion in the surrounding solution [25, 26].

In biological cells, vesicles and other cargo particles, which are transported along microtubules, often carry both kinesins and dyneins which leads to bi-directional transport along the filaments [27, 28]. The presumably simplest mechanism for this kind of movement is provided by the competition or 'tug-of-war' between the two motor species. A recent theory for this tug-of-war leads to several coexisting motility states and is found to be consistent with the experimental data currently available [29]. In addition, many interesting transport phenomena occur if one includes the mutual exclusion between different motors or between different filaments. Indeed, both the traffic of motors on immobilized filaments [25, 26] and the traffic of filaments on immobilized motors [30, 31] leads to crowding, pattern formation, and phase transitions.

3. Thermodynamics of ATP hydrolysis. In the following, we will focus on the most basic level of motor motility as provided by its energy transduction via chemomechanical coupling. We will use the framework of statistical mechanics and view the motor as a small system that is coupled to a heat reservoir at temperature T and to particle reservoirs for the chemical species $X = $ ATP, ADP, and inorganic phosphate P. We will always assume thermal equilibrium, i.e., the motor is always characterized by the same temperature as the surrounding aqueous solution. However, the concentrations or activities of the chemical species involved in ATP hydrolysis and synthesis may or may not correspond to chemical equilibrium, which is defined as follows.

The activity of the chemical species X will be denoted by $[X]$ and has the dimension of a molar concentration. For dilute solutions, the activity is equal to the molar concentration. In general, the activity $[X]$ is defined in such a way that the chemical potential μ_X for the chemical species X has the simple form

$$\mu_X \equiv \mu_X^o + k_B T \, \ln([X]/[X]^o) \tag{3.1}$$

with Boltzmann constant k_B where the superscript o refers to some standard or reference activity $[X]^o$. The relation (3.1) implies that

$$[X]^o e^{-\mu_X^o/k_B T} = [X] e^{-\mu_X/k_B T} \equiv [X]^* \tag{3.2}$$

where the activity scale $[X]^*$ has a unique value that does not depend on the choice of the standard or reference activity $[X]^o$. Therefore, the chemical potential may also be written as

$$\mu_X = k_B T \, \ln([X]/[X]^*) \tag{3.3}$$

The activity scale $[X]^*$ can be determined using the grand-canonical ensemble. One then has to consider a large volume V of water and calculate the partition functions Z_0 and Z_1 corresponding to the situations in which this volume contains either no or a single molecule of species X, respectively. The activity scale $[X]^*$ is then given by $[X]^* = Z_1/Z_0 V N_{\mathrm{Av}}$ with the Avogadro number N_{Av} [32].

When the motor hydrolyzes a single ATP molecule, it binds one such molecule and releases one inorganic phosphate P and one ADP molecule. According to the Gibbs fundamental form of thermodynamics, the corresponding change in internal energy of the motor is given by

$$\Delta\mu = \mu_{\text{ATP}} - \mu_{\text{P}} - \mu_{\text{ADP}} \tag{3.4}$$

which also represents the chemical energy input from the aqueous solution to the motor molecule. Using the expression (3.1) for the three chemical potentials, we then obtain

$$
\begin{aligned}
\Delta\mu &= \Delta\mu^o + k_B T \ln\left(\frac{[\text{ATP}]}{[\text{ADP}]\,[\text{P}]}\frac{[\text{ADP}]^o\,[\text{P}]^o}{[\text{ATP}]^o}\right) \\
&= k_B T \ln\left(\frac{[\text{ATP}]}{[\text{ADP}]\,[\text{P}]}\frac{[\text{ADP}]^*\,[\text{P}]^*}{[\text{ATP}]^*}\right),
\end{aligned}
\tag{3.5}
$$

with $\Delta\mu^o \equiv \mu^o_{\text{ATP}} - \mu^o_{\text{P}} - \mu^o_{\text{ADP}}$.

Chemical equilibrium between ATP hydrolysis and ATP synthesis corresponds to $\Delta\mu = 0$ which implies

$$\left.\frac{[\text{ADP}]\,[\text{P}]}{[\text{ATP}]}\right|^{\text{eq}} = \frac{[\text{ADP}]^*\,[\text{P}]^*}{[\text{ATP}]^*}. \tag{3.6}$$

The right hand side of this equation defines the equilibrium (dissociation) constant K^{eq}. Therefore, the relation (3.6) is equivalent to

$$K^{\text{eq}} \equiv \left.\frac{[\text{ADP}]\,[\text{P}]}{[\text{ATP}]}\right|^{\text{eq}}. \tag{3.7}$$

For dilute solutions, the activities of the three chemical species are equal to their molar concentrations and can be directly measured (after the system has relaxed into equilibrium for a sufficiently long time). For ATP hydrolysis, the precise value of the equilibrium constant depends on the ionic conditions but a typical value is given by $K^{\text{eq}} = 4.9 \times 10^{11}\ \mu\text{M}$ [33, 18].

The three activities or concentrations [ATP], [ADP], and [P] represent three control parameters that can be varied independently. Since all concentrations have to be zero or positive, they span the first octant of the three dimensional parameter space with $x = [\text{ADP}]$, $y = [\text{P}]$, and $z = [\text{ATP}]$. The equilibrium states, which satisfy the relation (3.7), form the 2-dimensional surface as given by $z = (1/K^{\text{eq}})xy$. On the other hand, all parameter values that do not lie on this surface and do not satisfy (3.7) represent states of chemical nonequilibrium.

For a molecular such as kinesin that performs mechanical work against an external load force, the complete thermodynamic description of the motor depends on the three activities or concentrations [ATP], [ADP], and [P] as well as on the force component F parallel to the filament [2, 3]. An extended discussion of this thermodynamic description is given in [4].

4. Network theory for molecular motor cycles. In the main body of this article, we will address the energy transduction of molecular motors which underlies their ability to perform mechanical work. This coupling is based on their chemomechanical motor cycles.

Such a cycle consists of a cyclic sequence of conformational transitions that enables the motor to hydrolyze ATP and to perform useful work.

Molecular motors perform their cycles in a rather noisy environment. Even in the absence of ATP, they can attain many conformational states because of thermal collisions with adjacent water molecules. Kinesin, for example, is a rather large molecule, compare Fig. 3(a), and it will bend and rotate in various ways as a result of such collisions. In addition, when bound to a filament, a motor can undergo 1-dimensional diffusion or 'passive sliding' as has been observed both for dynein at microtubules [34] and for RNA polymerase at DNA [35]. The delivery of the ATP molecules represents another stochastic process since these molecules diffuse through the surrounding solution and will have many collisions with the motor molecule before they stick to the catalytic motor domains. For kinesin, this implies that the adsorption of ATP is a random process, and that the time for this adsorption is governed by some probability distribution which depends on the ATP concentration.

In order to describe these different stochastic processes, it is necessary to construct motor models in which the motor can undergo transitions between different internal states and to incorporate the basic property that some of these transitions are coupled to ATP hydrolysis. Such a description is provided by stochastic networks as reviewed in the following. In order to be specific, we will first introduce network representations for two specific cases, namely for single motor heads and for two-headed motors. The subsequent theory is, however, rather general and can be applied to any other type of molecular motor or colloidal machine.

4.1. *Network description of single motor head.* Let us first consider a single motor head of kinesin or any other motor component with a single catalytic domain for ATP hydrolysis. In the latter process, a single molecule of ATP is broken up into a single ADP molecule and inorganic phosphate P. This process involves several substeps: adsorption of ATP, hydrolysis proper, i.e., the transformation of bound ATP into bound ADP/P, the release of P, and the release of ADP. As far as the chemical composition of the catalytic domain is concerned, one may then distinguish four different states of the motor head denoted by E, T, Θ, and D as shown in Fig. 4(a): [36] State E corresponds to the empty head, state T to the state with bound ATP, state Θ to bound ADP/P, and state D to bound ADP. Such a representation was previously used by T. L. Hill for a generic ATPase [5].

The edges between the different chemical states in Fig. 4(a) represent forward and backward transitions. The edge between state i and state j will be denoted by $\langle ij \rangle$. It consists of two directed edges or di-edges, $|ij\rangle$ and $|ji\rangle$, corresponding to the forward transition from i to j and to the backward transition from j to i, respectively. Thus, the di-edge or transition $|ET\rangle$ corresponds to ATP binding to the motor head, whereas the transition $|TE\rangle$ represents ATP release from this head.

Likewise, the transitions $|\Theta D\rangle$, $|D\Theta\rangle$, $|DE\rangle$, and $|ED\rangle$ describe P release, P binding, ADP release, and ADP binding, respectively. Finally, the transition $|T\Theta\rangle$, corresponds to ATP hydrolysis proper and the transition $|\Theta T\rangle$ to ATP synthesis from ADP and P.

The three edges $\langle ET \rangle$, $\langle \Theta D \rangle$, and $\langle DE \rangle$ involve the binding and release of a certain molecular species from the aqueous solution. In contrast, the edge $\langle T\Theta \rangle$ in the 4-state
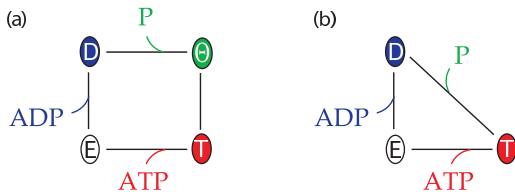
Fig. 4. Single motor head as an ATPase: (a) Network with 4 states corresponding to the motor head being empty (E), occupied by ATP (T), by ADP/P (Θ), and by ADP (D); and (b) Reduced network with 3 states in which the hydrolysis transition and the P release transition have been combined into the single transition $|31\rangle$.

network does not involve such an interaction of the motor head with the particle reservoir, see Fig. 4(a). Therefore, one may combine the two edges $\langle T\Theta \rangle$ and $\langle \Theta D \rangle$ of the 4-state network into the edge $\langle TD \rangle$ as shown in Fig. 4(b). The latter representation involves only 3 states: the motor head is occupied by ADP in state D, empty in state E, and occupied by ATP in state T. This reduced representation is useful since it eliminates some parameters but is still consistent with the fundamental constraints arising from the energetics of ATP hydrolysis, see Section 7.2 below.

4.2. *Two-headed motors with two catalytic domains.* Next, we consider a two-headed motor such as kinesin or myosin V. If each motor head can attain three different chemical states as in Fig. 4(b), the two-headed motor can attain $3^2 = 9$ different states as in Fig. 5(a). All of these states represent two-headed motors that are bound to a polar filament. We use the convention that the right head is the leading head whereas the left head is the trailing head with respect to the prefered direction of the motor movement.

If the motor walks via the 'hand-over-hand' mechanism, the leading and the trailing head interchange their positions during each *mechanical* step. If one assumes that this step is fast on the timescale of the chemical transitions, one has, in general, three possible mechanical steps: from state (E,D) to state (D,E), from state (E,T) to state (T,E), and from state (D,T) to state (T,D). These possible mechanical transitions are shown in Fig. 5(a) as broken edges. Recent experiments by Carter and Cross [21] provide strong evidence that such a separation of time scales does indeed apply to the cytoskeletal motor kinesin, see Section 8 below.

Inspection of Fig. 5(a) shows that each state of the two-headed motor is connected to four other states via four solid edges. As in Fig. 4, each solid edge between two states i and j represents both the forward chemical transition $|ij\rangle$ and the backward transition $|ji\rangle$. Thus, the 9-state network in Fig. 5(a) contains 18 solid edges corresponding to 36 *chemical* transitions.

Let us now assume that the processive motion of the two-headed motor is 'out-of-phase' and involves primarily those states for which the two heads have different chemical composition. This assumption applies to kinesin [2, 3], see section 8 below. In such a situation, one can ignore the three states (E,E), (D,D), and (T,T) in Fig. 5(a). In addition, a detailed comparison with experimental data for kinesin reveals that the only mechanical transition that is compatible with these data is provided by the transition from (D,T)
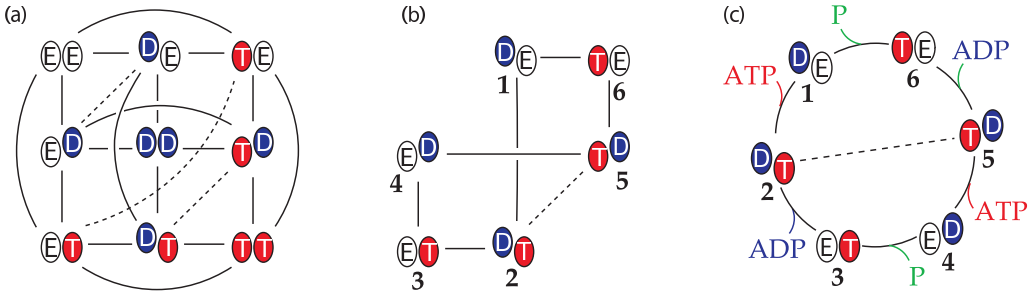
Fig. 5. Stochastic networks for a molecular motor with two catalytic domains or motor heads. Each head can be empty (E), occupied by ATP (T), or occupied by ADP (D): (a) State space with $3^2 = 9$ chemical states, each of which is connected to four neighboring states via solid edges (or lines). Each such edge represents both the forward and backward chemical transition as in Fig. 4; (b,c) Reduced state space with 6 chemical states, each of which is connected to two neighboring states via solid edges. We use the convention that the left and right head represent the trailing and leading head, respectively. Broken edges (or lines) correspond to mechanical step transitions in which the two heads interchange their relative position. In (a), three possible step transitions have been included but only one of those transitions is kept in (b) and (c). The latter transition applies to kinesin [3], see section 8 below.

to (T,D). If we delete the three states (E,E), (D,D), and (T,T) and the two mechanical transitions from (D,E) to (E,D) and from (E,T) to (T,E) from the 9-state network, we obtain the reduced 6-state network as shown in Fig. 5(b). As shown in this figure, we have labeled the six remaining states by $i = 1, 2, \ldots, 6$. A more transparent representation of the 6-state network is shown in Fig. 5(c): the six states form a cycle consisting of six (forward and backward) chemical transitions; the broken edge corresponds to the mechanical step transition. According to our convention, the mechanical forward step is given by $|25\rangle$, the mechanical backward step by $|52\rangle$.

As shown in Fig. 5(c), the cycle of chemical transitions consists (i) of the two ATP binding transitions $|12\rangle$ and $|45\rangle$, (ii) of the two transitions $|61\rangle$ and $|34\rangle$, which both represent ATP hydrolysis and P release, as well as (iii) of the two ADP release transitions $|23\rangle$ and $|56\rangle$.

4.3. *Cycles and dicycles.* The two previous subsections provided two specific examples for the description of molecular motors in terms of a discrete state space. These states are represented as the vertices of a network graph, $\mathbb{G}$, and are labeled by $i = 1, 2, \ldots, |\mathbb{G}|$. If two states i and j are connected by an edge $\langle ij \rangle$, the motor can undergo the forward transition $|ij\rangle$ from i to j and the backward transition $|ji\rangle$ from j to i. Inspection of Fig. 4 and Fig. 5 shows that these edges may form cycles. These cycles are particularly important in the present context since they are intimately related to fluxes and nonequilibrium states [5].

In order to be precise, we will distinguish (undirected) cycles from directed cycles or *dicycles*. The smallest dicycle consists of three states and three di-edges. In the mathematical literature, di-edges are often called arcs and dicycles are sometimes called circuits. An (undirected) cycle $\mathcal{C}_\nu$ is given by a closed sequence of vertices and edges in which

each vertex and each edge occurs only once. The notation $\mathcal{C}_\nu = \langle i_1 i_2 \ldots i_n i_1 \rangle$ implies that the cycle $\mathcal{C}_\nu$ contains the edges $\langle i_1 i_2 \rangle$, $\langle i_2 i_3 \rangle$, ..., and $\langle i_n i_1 \rangle$. Each cycle $\mathcal{C}_\nu$ leads to two dicycles $\mathcal{C}_\nu^d$ with $d = \pm$ that differ in their orientation.

First, we choose a certain arbitrary but fixed orientation of cycle $\mathcal{C}_\nu$ to correspond to the positive direction $d = +$, which defines the dicycle $\mathcal{C}_\nu^+$. When we pass through $\mathcal{C}_\nu$ in the opposite direction $d = -$, we obtain the dicycle $\mathcal{C}_\nu^-$. These two dicycles will be denoted by $\mathcal{C}_\nu^+ = |i_1 i_2 \ldots i_n i_1 \rangle$ and $\mathcal{C}_\nu^- = |i_1 i_n \ldots, i_2 i_1 \rangle$. The dicycle $\mathcal{C}_\nu^+$ contains the di-edges $|i_1 i_2 \rangle$, $|i_2 i_3 \rangle$, ..., and $|i_n i_1 \rangle$; the dicycle $\mathcal{C}_\nu^-$ consists of the di-edges $|i_1 i_n \rangle$, $|i_n i_{n-1} \rangle$, ..., and $|i_2 i_1 \rangle$.

The network description of a single motor head, see Fig. 4, involves only a single cycle and, thus, two dicycles. Analogous unicycle models have also been frequently used for two-headed motors, see, e.g., [37, 38]. Inspection of Fig. 5 shows, however, that these motors will, in general, exhibit several motor cycles. The 9-state network in Fig. 5(a) involves a rather large number of cycles (more than 200). In contrast, the 6-state network in Fig. 5(c) contains only three cycles: the forward cycle $\mathcal{F} = \langle 25612 \rangle$, the backward cycle $\mathcal{B} = \langle 52345 \rangle$, and the chemical cycle $\mathcal{D} = \langle 1234561 \rangle$.

We will now distinguish different types of cycles by the following terminology:

(i) *Chemomechanical cycles* that contain both an ATP hydrolysis transition and a mechanical step. Examples are given by the forward cycle $\mathcal{F}$ and the backward cycle $\mathcal{B}$ in the 6-state model, see Fig. 5(c). In general, a chemomechanical cycle couples the hydrolysis of at least one ATP molecule to a net mechanical displacement;

(ii) *Dissipative slip cycles* that contain hydrolysis transitions but no mechanical step transition. One example is provided by the chemical cycle $\mathcal{D} = \langle 1234561 \rangle$ of the 6-state model;

(iii) *Mechanical slip cycles* that contain mechanical step transitions but no net hydrolysis of ATP. Examples without any hydrolysis transitions are the cycles $\langle (E, E)(E, D)(D, E) (E, E) \rangle$ and $\langle (E, T)(D, T)(T, D)(T, E)(E, T) \rangle$ of the 9-state model, see Fig. 5(a). An example for a mechanical slip cycle with one ATP hydrolysis and one ATP synthesis transition is provided by the cycle $\langle (T, D)(D, D)(D, T)(T, D) \rangle$; and

(iv) *Thermal slip cycles* that involve neither a mechanical step nor a hydrolysis transition. Examples are provided by $\langle (E, E)(E, D)(D, D)(D, E)(E, E) \rangle$ and $\langle (E, E)(E, T)(D, T) (D, E)(E, E) \rangle$.

It is interesting to note that the reduction of the 9-state model to the 6-state model eliminates all mechanical and thermal slip cycles from the network.

5. Motor dynamics as a Markov process. The dynamics of the motor is now described by a continuous-time Markov process [39, 40] on the network graph $\mathbb{G}$. Such a process involves two stochastic ingredients, the sojourn (or dwell) times and the transition probabilities. When the system arrives in the state i, it occupies this state for a certain sojourn (or dwell) time τ_i. This time is a random variable governed by the exponential probability distribution

$$P(\tau_i) = \frac{1}{\langle \tau_i \rangle} \ e^{-\tau_i / \langle \tau_i \rangle} \tag{5.1}$$

where $\langle \tau_i \rangle$ denotes the average sojourn time [40]. When the motor leaves the state i, it jumps to state j with transition probability π_{ij}. By definition, one has $\pi_{ii} \equiv 0$ for all i and $\sum_j \pi_{ij} = 1$. From the mathematical point of view, the latter transitions occur instantaneously; from the physical point of view, they occur on a time scale that is small compared to all average sojourn times $\langle \tau_i \rangle$.

The probability $P_i(t)$ to find the motor in state i at time t is then governed by the loss-and-gain equation

$$\frac{d}{dt} P_i = -\sum_j \left(P_i \omega_{ij} - P_j \omega_{ji} \right) \tag{5.2}$$

with the transition rates

$$\omega_{ij} = \pi_{ij}/\langle \tau_i \rangle \quad \text{and} \quad \sum_j \omega_{ij} = 1/\langle \tau_i \rangle. \tag{5.3}$$

In the physical literature, this equation is known as the Master equation [39]; in the mathematical literature, it is called the forward equation of the continuous-time Markov process [40] .

This description of the motor dynamics is convenient since it allows explicit calculation of the steady state fluxes and the steady state entropy production of the motor. These quantities are related to various experimentally accessible quantities such as the motor velocity and the ratio of forward to backward mechanical steps. In this and the following section, we will discuss the steady state fluxes and steady state entropy production, respectively.

It is convenient to define the local fluxes

$$J_{ij} \equiv P_i^{\mathrm{st}} \omega_{ij} \tag{5.4}$$

and the local excess fluxes (or currents)

$$\Delta J_{ij} \equiv P_i \omega_{ij} - P_j \omega_{ji} = -\Delta J_{ji} \tag{5.5}$$

from state i to state j. The *steady state* is characterized by probabilities P_i^{st} with $dP_i^{\mathrm{st}}/dt = 0$ which implies

$$\sum_j \Delta J_{ij}^{\mathrm{st}} = \sum_j (J_{ij}^{\mathrm{eq}} - J_{ji}^{\mathrm{eq}}) = \sum_j \left(P_i^{\mathrm{st}} \omega_{ij} - P_j^{\mathrm{st}} \omega_{ji} \right) = 0. \tag{5.6}$$

This relation provides N_{s} linear equations for the probabilities P_i^{st} which can be solved by linear algebra, see, e.g., [41, 42] or, more conveniently, by a graph-theoretic method [43, 44, 45, 46, 1].

The steady state equation (5.6) has a simple intuitive interpretation. The term $\sum_j J_{ij}^{\mathrm{eq}} = \sum_j P_i^{\mathrm{st}} \omega_{ij}$ represents the sum of all fluxes out of the state i; the term $\sum_j J_{ji}^{\mathrm{eq}} = \sum_j P_i^{\mathrm{st}} \omega_{ji}$ represents the sum of all fluxes into the state i. Furthermore, the local flux $J_{ij}^{\mathrm{eq}} = P_i^{\mathrm{st}} \omega_{ij}$ is equal to the number of transitions that the system performs from state i to state j per unit time in the steady state. Thus, this steady state flux is also equal to the transition frequency

$$\Omega_{ij}^{\mathrm{st}} \equiv P_i^{\mathrm{st}} \omega_{ij}, \tag{5.7}$$

i.e., to the frequency with which the system undergoes the transition $|ij\rangle$ in the steady state.

5.1. *Equilibrium, detailed balance, and dicycle ratios.* In thermal and chemical equilibrium, all local excess fluxes vanish simultaneously , i.e. $\Delta J_{ij}^{\text{eq}} = 0$ for all di-edges $|ij\rangle$. In such a situation, the transition rates ω_{ij} must satisfy certain conditions, the so-called detailed balance conditions. Indeed, the relations $\Delta J_{ij}^{\text{eq}} = 0$ for all di-edges can be rewritten as

$$P_i^{\text{eq}} \omega_{ij} = P_j^{\text{eq}} \omega_{ji} \tag{5.8}$$

for all edges $\langle ij \rangle$. These conditions can be expressed in terms of the transition rates ω_{ij} alone. In order to do so, it is convenient to define, for each dicycle $\mathcal{C}_\nu^d$ and arbitrary transition rates ω_{ij}, the *dicycle ratio*

$$\Xi(\mathcal{C}_\nu^d) \equiv \prod_{|ij\rangle \in \mathcal{C}_\nu^d} \omega_{ij} / \prod_{|ij\rangle \in \mathcal{C}_\nu^{-d}} \omega_{ij} = \prod_{|ij\rangle}^{\nu,d} (\omega_{ij}/\omega_{ji}) \tag{5.9}$$

where the superscipt ν, d at the product sign indicates a product over all di-edges $|ij\rangle$ of the dicycle $\mathcal{C}_\nu^d$. This definition implies that

$$\Xi(\mathcal{C}_\nu^-) = 1/\Xi(\mathcal{C}_\nu^+). \tag{5.10}$$

It can be shown that the detailed balance conditions as given by (5.8) are equivalent to the relations [47]

$$\Xi^{\text{eq}}(\mathcal{C}_\nu^d) = \prod_{|ij\rangle}^{\nu,d} (\omega_{ij}/\omega_{ji}) = 1 \quad \text{for all dicycles } \mathcal{C}_\nu^d \tag{5.11}$$

for the dicycle ratios $\Xi^{\text{eq}}(\mathcal{C}_\nu^d)$. In this form, the detailed balance conditions are no longer local (since large cycles involve widely separated edges of the network) but they are expressed only in terms of the transition rates ω_{ij}, which are the basic parameters of the model. In the context of chemical kinetics, these conditions are also known as Wegscheider conditions [48].

Away from equilibrium, some of the dicycle ratios $\Xi(\mathcal{C}_\nu^d)$ are different from one. The corresponding numerical values can be interpreted in a rather intuitive way. Indeed, as discussed in the next subsection, $\Xi(\mathcal{C}_\nu^+)$ is equal (i) to the ratio of the steady state fluxes through the dicycles $\mathcal{C}_\nu^+$ and $\mathcal{C}_\nu^-$; and (ii) to the ratio of the (unnormalized) probabilities or frequencies to complete the dicycles $\mathcal{C}_\nu^+$ and $\mathcal{C}_\nu^-$ in the steady state. In addition, we have recently shown [2] that the dicycle ratio $\Xi(\mathcal{C}_\nu^d)$ is directly related to the entropy $\Delta S(\mathcal{C}_\nu^d)$ produced during the completion of the dicycle $\mathcal{C}_\nu^d$ via the equation $\Delta S(\mathcal{C}_\nu^d) = k_B \ln(\Xi(\mathcal{C}_\nu^d))$, see next section.

In a long series of studies on biochemical networks, Hill and coworkers have previously discussed the quantity $k_B T \ln(\Xi(\mathcal{C}_\nu^d))$, which they interpreted as the "thermodynamic force" that drives the system out of equilibrium. Our results imply that this thermodynamic force is, in fact, equal to temperature T times entropy $\Delta S(\mathcal{C}_\nu^d)$ produced per completed dicycle $\mathcal{C}_\nu^d$ [2]. This identification is useful since the produced entropy can be identified with the heat released by the motor, and the latter quantity is accessible to experiment, at least in principle.

5.2. *Dicycle fluxes and frequencies.* In the steady state, each dicycle $\mathcal{C}_\nu^d$ with $d = \pm$ can be characterized by a dicycle flux $J^{\text{st}}(\mathcal{C}_\nu^d)$ which represents a global property of the

stochastic network. These fluxes can be explicitly calculated (for a finite network) using a graph-theoretic or diagrammatic method. They are then found to have the general form [45, 49, 1]

$$J(\mathcal{C}_\nu^d) = \left(\prod_{|ij\rangle}^{\nu,d} \omega_{ij}\right) \Upsilon(\mathcal{C}_\nu)/\Omega \quad \text{with } d = \pm \tag{5.12}$$

where the superscript ν, d at the product sign indicates a product over all di-edges $|ij\rangle$ that are contained in the dicycle $\mathcal{C}_\nu^d$. The factor $\Upsilon(\mathcal{C}_\nu)$ is identical for both dicycles but depends, in general, on the cycle $\mathcal{C}_\nu$; this factor is given by a multilinear polynomial of transition rates which belong to di-edges *not* contained in the two dicycles $\mathcal{C}_\nu^+$ and $\mathcal{C}_\nu^-$. As explained in Ref. [1], these di-edges correspond to the edges of the spanning trees that are not contained in the cycle $\mathcal{C}_\nu$. The factor Ω is a normalization factor which is given by a multilinear polynomials that depends on all transition rates of the network.

Even though the dicycle fluxes $J^{\mathrm{st}}(\mathcal{C}_\nu^d)$ are global properties of the network process, they have a rather intuitive interpretation since they determine the frequencies of dicycle completion and the average completion times. Thus, let us follow the network process for a long time period Δt. The system will then pass successively through many dicycles and we can count the number of times, $N_{\Delta t}(\mathcal{C}_\nu^d)$, that it passed through the dicycle $\mathcal{C}_\nu^d$. Thus, $N_{\Delta t}(\mathcal{C}_\nu^d)$ represents the absolute frequency of completed dicycles $\mathcal{C}_\nu^d$ within the time interval Δt.

It is intuitively clear that $N_{\Delta t}(\mathcal{C}_\nu^d)$ behaves as

$$N_{\Delta t}(\mathcal{C}_\nu^d) \approx \Omega^{\mathrm{st}}(\mathcal{C}_\nu^d) \, \Delta t \quad \text{for large} \quad \Delta t \tag{5.13}$$

which defines the frequencies (or unnormalized probabilities) $\Omega^{\mathrm{st}}(\mathcal{C}_\nu^d)$. This frequency is equal to the average number of dicycles $\mathcal{C}_\nu^d$ that are completed in the steady state per unit time. It can be shown that [49, 50]

$$\Omega^{\mathrm{st}}(\mathcal{C}_\nu^d) = J^{\mathrm{st}}(\mathcal{C}_\nu^d), \tag{5.14}$$

i.e., that the dicycle completion frequencies are equal to the dicycle fluxes $J^{\mathrm{st}}(\mathcal{C}_\nu^d)$ which implies that the ratio of the dicycle completion frequencies is equal to the ratio of the dicycle fluxes. Furthermore, the explicit expression (5.12) for the dicycle fluxes implies that

$$\frac{\Omega^{\mathrm{st}}(\mathcal{C}_\nu^+)}{\Omega^{\mathrm{st}}(\mathcal{C}_\nu^-)} = \frac{J^{\mathrm{st}}(\mathcal{C}_\nu^+)}{J^{\mathrm{st}}(\mathcal{C}_\nu^-)} = \prod_{|ij\rangle}^{\nu,+} (\omega_{ij}/\omega_{ji}) = \Xi(\mathcal{C}_\nu^+) \tag{5.15}$$

i.e., both ratios are equal to the dicycle ratio $\Xi(\mathcal{C}_\nu^+)$.

The relation (5.15) for the dicycle frequencies can be generalized if one uses the action functional A of Lebowitz and Spohn [51]. Thus, consider a directed walk or trajectory $\mathcal{T}^+ \equiv |i_1 i_2 \ldots i_m\rangle$ on the network graph together with the time-reversed walk $\mathcal{T}^- \equiv |i_m i_{m-1} \ldots i_1\rangle$. Each walk consists of a connected sequence of di-edges $|ij\rangle$. The action functional A is then defined by [51]

$$\exp[A(\mathcal{T}^+)] \equiv \prod_{|ij\rangle \in \mathcal{T}^+} \omega_{ij} \Big/ \prod_{|ij\rangle \in \mathcal{T}^-} \omega_{ij} = \exp[-A(\mathcal{T}^-)] \tag{5.16}$$

where the product in the numerator and denominator runs over all di-edges of the walk $\mathcal{T}^+$ and $\mathcal{T}^-$, respectively.

Now, consider a closed trajectory $\mathcal{T}^+ = \mathcal{T}_\nu^+$ that goes through the dicycle $\mathcal{C}_\nu^+$ only once and does not go through any other dicycle. In general, such a trajectory consists of (i) segments of the dicycle $\mathcal{C}_\nu^+$, (ii) "back-and-forth" segments of the cycle $\mathcal{C}_\nu$ that contain segments of the dicycle $\mathcal{C}_\nu^-$ together with the corresponding reverse segments of the dicycle $\mathcal{C}_\nu^+$, and (iii) tree-like excursions away from the cycle $\mathcal{C}_\nu$. The corresponding action functional satisfies

$$\exp[A(\mathcal{T}_\nu^+)] \equiv \prod_{|ij\rangle \in \mathcal{T}_\nu^+} \omega_{ij} \Big/ \prod_{|ij\rangle \in \mathcal{T}_\nu^-} \omega_{ij} = \prod_{|ij\rangle \in \mathcal{C}_\nu^+} \omega_{ij} \Big/ \prod_{|ij\rangle \in \mathcal{C}_\nu^-} \omega_{ij} = \Xi(\mathcal{C}_\nu^+). \tag{5.17}$$

Thus, in this ratio, all transition rates from the segments of type (ii) and (iii) of the trajectories $\mathcal{T}_\nu^+$ and $\mathcal{T}_\nu^-$ cancel out, and one is simply left with the transition rates corresponding to one completion of $\mathcal{C}_\nu^+$ and $\mathcal{C}_\nu^-$. Therefore, the dicycle ratio $\Xi(\mathcal{C}_\nu^+)$ applies, in fact, to the transition rate ratios of all pairs of closed trajectories $\mathcal{T}_\nu^+$ and $\mathcal{T}_\nu^-$.

As mentioned in Ref. [51], the frequencies $\Omega^{st}(\mathcal{T}^+)$ and $\Omega^{st}(\mathcal{T}^-)$ for the two walks $\mathcal{T}^+$ and $\mathcal{T}^-$ are related to the action functional A via

$$\Omega^{st}(\mathcal{T}^+)/\Omega^{st}(\mathcal{T}^-) = (P_1^{st}/P_m^{st}) \exp[A(\mathcal{T}^+)]. \tag{5.18}$$

This relation simplifies for closed walks $\mathcal{T}_c^+$ which are characterized by $i_m \equiv i_1$ and, thus, by $P_1^{st} = P_m^{st}$. If we now consider again a closed trajectory $\mathcal{T}_c^+ = \mathcal{T}_\nu^+$ that goes through the dicycle $\mathcal{C}_\nu^+$ only once and does not go through any other dicycle, we obtain the frequency ratio

$$\frac{\Omega^{st}(\mathcal{T}_\nu^+)}{\Omega^{st}(\mathcal{T}_\nu^-)} = \exp[A(\mathcal{T}_\nu^+)] = \prod_{|ij\rangle}^{\nu,+} (\omega_{ij}/\omega_{ji}) = \Xi(\mathcal{C}_\nu^+). \tag{5.19}$$

As before, the superscript $\nu, +$ at the product sign indicates a product over all di-edges $|ij\rangle$ contained in the dicycle $\mathcal{C}_\nu^+$.

5.3. *Cycle decomposition of fluxes.* Using the dicyle fluxes $J(\mathcal{C}_\nu^d)$ with $d = \pm$ as in (5.12), one may define the cycle flux

$$J^{st}(\mathcal{C}_\nu) \equiv J^{st}(\mathcal{C}_\nu^+) - J^{st}(\mathcal{C}_\nu^-) \tag{5.20}$$

which characterizes the flux through the (undirected) cycle $\mathcal{C}_\nu$. Furthermore, each excess flux ΔJ_{ij} through the di-edge $|ij\rangle$ can be decomposed according to

$$\Delta J_{ij}^{st} = \sum_\nu \epsilon_{ij,\nu} J^{st}(\mathcal{C}_\nu) = \sum_\nu \epsilon_{ij,\nu} \left(J^{st}(\mathcal{C}_\nu^+) - J^{st}(\mathcal{C}_\nu^-) \right) \tag{5.21}$$

where the antisymmetric symbol $\epsilon_{ij,\nu} = -\epsilon_{ji,\nu}$ is given by

$$\epsilon_{ij,\nu} \equiv \begin{cases} +1 & \text{for } |ij\rangle \in \mathcal{C}_\nu^+ \\ -1 & \text{for } |ij\rangle \in \mathcal{C}_\nu^- \\ 0 & \text{otherwise.} \end{cases} \tag{5.22}$$

Thus, the sum in (5.21) contains contributions from all cycles that go through the edge $\langle ij \rangle$, and the sign of these contributions is positive, if the di-edge $|ij\rangle$ is contained in (or parallel to) the dicycle $\mathcal{C}_\nu^+$, and negative, if $|ij\rangle$ is contained in (or parallel to) the reverse dicycle $\mathcal{C}_\nu^-$.

6. Entropy production and balance conditions. In this section, we will study the entropy that is produced by the motor dynamics. The entropy production rate can be decomposed in terms of separate contributions arising (i) from the different di-edges and (ii) from the different dicycles of the network graph. The dicycle entropy $\Delta S(\mathcal{C}_\nu^d)$ produced by completion of dicycle $\mathcal{C}_\nu^d$ is found to be given by $\Delta S(\mathcal{C}_\nu^d) = k_B \ln(\Xi(\mathcal{C}_\nu^d))$, i.e., by the logarithm of the dicycle ratio $\Xi(\mathcal{C}_\nu^d)$.

If one identifies the entropy produced by the motor dynamics with the heat that the motor releases into its environment, one obtains energy balance conditions that replace the detailed balance conditions for equilibrium states and hold for any state of chemical and/or mechanical nonequilibrium.

6.1. *Production of statistical or Shannon entropy.* First, we will discuss the dicycle entropies, $\Delta S(\mathcal{C}_\nu^d)$, which are produced on average, during the completion of dicycle $\mathcal{C}_\nu^d$ in the steady state, and show that these entropies satisfy the simple relations [2]

$$\Delta S(\mathcal{C}_\nu^d) \equiv k_B \ln(\Xi(\mathcal{C}_\nu^d)) = k_B \ln\left(\prod_{|ij\rangle}^{\nu,d} (\omega_{ij}/\omega_{ji})\right), \tag{6.1}$$

with the dicycle ratios $\Xi(\mathcal{C}_\nu^d)$. Since these ratios satisfy $\Xi(\mathcal{C}_\nu^-) = 1/\Xi(\mathcal{C}_\nu^+)$, the entropies of the two dicycles $\mathcal{C}_\nu^+$ and $\mathcal{C}_\nu^-$ are related via

$$\Delta S(\mathcal{C}_\nu^-) = -\Delta S(\mathcal{C}_\nu^+) \tag{6.2}$$

which reflects the fact that the dicycle $\mathcal{C}_\nu^-$ corresponds to the time-reversed dicycle $\mathcal{C}_\nu^+$.

T. L. Hill and coworkers have previously discussed the quantity $k_B T \ln(\Xi(\mathcal{C}_\nu^d))$, which they viewed as the "thermodynamic force" that drives the system out of equilibrium. Our work shows that this thermodynamic force is, in fact, equal to temperature T times dicycle entropy $\Delta S(\mathcal{C}_\nu^d)$.

It is interesting to note that one may combine the relations (5.15) and (6.1) in order to express the dicycle frequency ratio $\Omega^{\mathrm{st}}(\mathcal{C}_\nu^+)/\Omega^{\mathrm{st}}(\mathcal{C}_\nu^-)$ in terms of the dicycle entropies via

$$\frac{\Omega^{\mathrm{st}}(\mathcal{C}_\nu^+)}{\Omega^{\mathrm{st}}(\mathcal{C}_\nu^-)} = e^{\Delta S(\mathcal{C}_\nu^+)/k_B} = e^{-\Delta S(\mathcal{C}_\nu^-)/k_B} \tag{6.3}$$

which is reminiscent of the various relations that have been derived in the context of entropy fluctuations [52, 53, 54, 6]. However, the entropies $\Delta S(\mathcal{C}_\nu^d)$ as considered here do not fluctuate but have a certain, fixed value for each dicycle of the network. [55]

In the following, we will give two derivations of the relations (6.1) which are both based on the decomposition of the entropy production rate in the steady state. We start from the statistical or Shannon entropy

$$S\{P_i\} \equiv -k_B \sum_i P_i \ln(P_i) \tag{6.4}$$

which one may define for any probability distribution $\{P_i\}$ on a discrete state space $\mathbb{G}$. This entropy provides a well-defined measure by which one can compare different probability distributions. If the probability distribution changes with time t, so does the

statistical entropy $S\{P_i\}$. Its time derivative can be written in the form [46, 56, 51, 57]

$$\frac{d}{dt}S\{P_i\} = \sigma_{\mathrm{pr}} + \sigma_{\mathrm{fl}} \tag{6.5}$$

with the entropy production rate

$$\sigma_{\mathrm{pr}} \equiv \tfrac{1}{2}\, k_B \sum_i \sum_j{}' \Delta J_{ij}\, \ln\left(\frac{P_i\omega_{ij}}{P_j\omega_{ji}}\right) \tag{6.6}$$

and the entropy flux term

$$\sigma_{\mathrm{fl}} \equiv -\tfrac{1}{2}\, k_B \sum_i \sum_j{}' \Delta J_{ij}\, \ln\left(\frac{\omega_{ij}}{\omega_{ji}}\right) \tag{6.7}$$

where the prime at the summation sign indicates that there are no terms with $j = i$. Thus, the double sum represents a summation over all di-edges or transitions $|ij\rangle$ of the network.

In the steady state, the system's entropy does not change and $\frac{d}{dt}S\{P_i\} = 0$, which leads to the entropy production rate

$$\sigma_{\mathrm{pr}}^{\mathrm{st}} = \tfrac{1}{2} \sum_i \sum_j{}' \Delta J_{ij}^{\mathrm{st}}\, k_B \ln(\omega_{ij}/\omega_{ji}). \tag{6.8}$$

We now rewrite this expression in two different ways which leads to two different derivations of the dicycle entropy $\Delta S(\mathcal{C}_\nu^d)$ as given in (6.1). The first derivation is based on the decomposition of the entropy production rate in terms of di-edge contributions. The second derivation corresponds to our original calculation in [2] which is based on the decomposition of the entropy production rate in terms of dicycle contributions.

6.1.1. *Di-edge decomposition of entropy production rate.* In the expression (6.8) for the entropy production rate $\sigma_{\mathrm{pr}}^{\mathrm{st}}$, the di-edge or transition $|ji\rangle$ gives the same contribution as the transition $|ij\rangle$ which implies

$$\sigma_{\mathrm{pr}}^{\mathrm{st}} = \sum_{|ij\rangle} P_i^{\mathrm{st}}\omega_{ij}\, k_B \ln(\omega_{ij}/\omega_{ji}) = \sum_{|ij\rangle} \Omega_{ij}^{\mathrm{st}}\, k_B \ln(\omega_{ij}/\omega_{ji}), \tag{6.9}$$

where the sum represents a summation over all di-egdes or transitions $|ij\rangle$ of the network graph and $\Omega_{ij}^{\mathrm{st}}$ is the transition frequency as defined in (5.7). Thus, the second equation in (6.9) has the form

$$\sigma_{\mathrm{pr}}^{\mathrm{st}} = \sum_{|ij\rangle} \Omega_{ij}^{\mathrm{st}}\, \Delta S_{ij} \tag{6.10}$$

with the steady state transition entropies ΔS_{ij} defined via

$$\Delta S_{ij} \equiv k_B \ln(\omega_{ij}/\omega_{ji}) = -\Delta S_{ji}. \tag{6.11}$$

This definition is rather natural if one assumes that the entropy production rate is an additive quantity just like the entropy. Indeed, the expression (6.10) gives the total entropy production rate as a sum over contributions from all di-edges or transitions, and each of these contributions is equal to the transition frequency $\Omega_{ij}^{\mathrm{st}}$ times the transition entropy ΔS_{ij}. This conclusion agrees with the more general proposal of Seifert [6], who considered arbitrary trajectories of Markov processes. The dicycle entropy $\Delta S(\mathcal{C}_\nu^d)$ as

given by (6.1) can now be obtained from

$$\Delta S(\mathcal{C}_\nu^d) = \sum_{|ij\rangle}^{\nu,d} \Delta S_{ij}, \qquad (6.12)$$

i.e., by a summation of the transition entropies ΔS_{ij} over all di-edges contained in the dicycle $\mathcal{C}_\nu^d$.

6.1.2. *Dicycle decomposition of entropy production rate.* Alternatively, let us insert the dicycle decomposition (5.21) of the local excess fluxes $\Delta J_{ij}^{\text{st}}$ into the relation (6.8) for the entropy production rate $\sigma_{\text{pr}}^{\text{st}}$. We then obtain the expression

$$\sigma_{\text{pr}}^{\text{st}} = \sum_\nu J^{\text{st}}(\mathcal{C}_\nu) \, \tfrac{1}{2} \sum_{i,j}' \epsilon_{ij,\nu} \, k_B \ln(\omega_{ij}/\omega_{ji}) \qquad (6.13)$$

The summation over the di-edges may be divided up into three contributions corresponding to (i) those di-edges which are contained in the dicycle $\mathcal{C}_\nu^+$, (ii) those in $\mathcal{C}_\nu^-$, and (iii) those that are neither contained in $\mathcal{C}_\nu^+$ nor $\mathcal{C}_\nu^-$. The latter di-edges do not contribute since the antisymmetric symbol $\epsilon_{ij,\nu}$ vanishes in this case. Furthermore, the contributions from (i) and (ii) are equal which implies

$$\sigma_{\text{pr}}^{\text{st}} = \sum_\nu J^{\text{st}}(\mathcal{C}_\nu) \sum_{|i,j\rangle}^{\nu,+} k_B \ln(\omega_{ij}/\omega_{ji}) = \sum_\nu J^{\text{st}}(\mathcal{C}_\nu) \, k_B \ln(\Xi(\mathcal{C}_\nu^+)) \qquad (6.14)$$

with the dicycle ratio $\Xi(\mathcal{C}_\nu^+)$ as given by (5.9). Finally, using the cycle flux decomposition $J^{\text{st}}(\mathcal{C}_\nu) \equiv J^{\text{st}}(\mathcal{C}_\nu^+) - J^{\text{st}}(\mathcal{C}_\nu^-)$ as in (5.20) and the relation $\Xi(\mathcal{C}_\nu^-) = 1/\Xi(\mathcal{C}_\nu^+)$, one obtains

$$\sigma_{\text{pr}}^{\text{st}} = \sum_\nu \sum_{d=\pm} J^{\text{st}}(\mathcal{C}_\nu^d) \, k_B \ln(\Xi(\mathcal{C}_\nu^+)) = \sum_\nu \sum_{d=\pm} \Omega^{\text{st}}(\mathcal{C}_\nu^d) \, k_B \ln(\Xi(\mathcal{C}_\nu^+)) \qquad (6.15)$$

where we have replaced the dicycle fluxes $J^{\text{st}}(\mathcal{C}_\nu^d)$ by the dicycle frequencies $\Omega^{\text{st}}(\mathcal{C}_\nu^d)$.

In this way, the entropy production rate $\sigma_{\text{pr}}^{\text{st}}$ is decomposed into separate contributions arising from all dicycles of the network. This decomposition may be rewritten in the form

$$\sigma_{\text{pr}}^{\text{st}} = \sum_\nu \sum_{d=\pm} \Omega^{\text{st}}(\mathcal{C}_\nu^d) \, \Delta S(\mathcal{C}_\nu^d) \qquad (6.16)$$

with the dicycle entropies $\Delta S(\mathcal{C}_\nu^d)$ as given by (6.1).

6.2. *Energy conservation for motor cycles.* In general, each state i of the motor represents an ensemble of molecular conformations that differ in thermally excited vibrational modes. The corresponding internal energy will be denoted by U_i. For equilibrium states, the internal energy is a state function (or state property) that does not depend on the system's history. We will now assume that we can associate such a history-independent quantity with the steady states of the motor system as well. This assumption is justified in the next section by an explicit calculation of the internal energies U_i for the network models of a single motor head as introduced in section 4.1.

In general, the internal energy U_i can change during the transition $|ij\rangle$ because of (i) energy input $E_{\text{in},ij}$ during the transition $|ij\rangle$, (ii) work W_{ij}, which the motor performs during this transition, and (iii) heat Q_{ij}, which the motor exchanges with the reservoir.

Conservation of energy now implies

$$\Delta U_{ij} \equiv U_j - U_i = E_{\text{in},ij} - W_{ij} - Q_{ij}. \tag{6.17}$$

The energy input $E_{\text{in},ij}$ may arise from the coupling to an exergonic chemical reaction and/or from photon absorption. The work W_{ij} may consist of mechanical work against an external force and/or against an electrochemical potential gradient.

The energy input $E_{\text{in},ij}$ may be partially *stored* in the motor molecule after the transition $|ij\rangle$, which implies $\Delta U_{ij} > 0$, and this stored energy may be transformed into work and/or heat during a later transition $|i'j'\rangle$, which implies $\Delta U_{i'j'} < 0$. However, the situation simplifies if we sum the local energy balance as given by (6.17) over any dicycle $\mathcal{C}_\nu^d$ which leads to

$$\sum_{|ij\rangle}^{\nu,d} \Delta U_{ij} = 0 = \sum_{|ij\rangle}^{\nu,d} [E_{\text{in},ij} - W_{ij} - Q_{ij}] \tag{6.18}$$

or the dicycle heat

$$Q(\mathcal{C}_\nu^d) \equiv \sum_{|ij\rangle}^{\nu,d} Q_{ij} = E_{\text{in}}(\mathcal{C}_\nu^d) - W(\mathcal{C}_\nu^d) \tag{6.19}$$

with the dicycle energy input $E_{\text{in}}(\mathcal{C}_\nu^d)$ and the dicycle work $W(\mathcal{C}_\nu^d)$ as given by

$$E_{\text{in}}(\mathcal{C}_\nu^d) \equiv \sum_{|ij\rangle}^{\nu,d} E_{\text{in},ij} \quad \text{and} \quad W(\mathcal{C}_\nu^d) \equiv \sum_{|ij\rangle}^{\nu,d} W_{ij}. \tag{6.20}$$

6.3. *Nonlocal and local balance conditions.* We now want to combine the entropy produced during the motor dynamics, see subsection 6.1, with the energy conservation law as given by (6.19). In order to do so, we identify the dicycle entropy $\Delta S(\mathcal{C}_\nu^d)$ as given by (6.1) with the reduced heat $Q(\mathcal{C}_\nu^d)/T$. A combination of (6.19) and (6.1) then leads to the nonlocal balance conditions [2]

$$k_B T \ln(\Xi(\mathcal{C}_\nu^d)) = Q(\mathcal{C}_\nu^d) = E_{\text{in}}(\mathcal{C}_\nu^d) - W(\mathcal{C}_\nu^d) \tag{6.21}$$

or

$$k_B T \sum_{|ij\rangle}^{\nu,d} \ln\left(\frac{\omega_{ij}}{\omega_{ji}}\right) = E_{\text{in}}(\mathcal{C}_\nu^d) - W(\mathcal{C}_\nu^d) \tag{6.22}$$

between the transition rates ω_{ij}, the energy input E_{in}, and the work W for any dicycle $\mathcal{C}_\nu^d$ of the network. In equilibrium, all terms on the right hand side of these equations vanish, and we recover the detailed balance conditions $\Xi(\mathcal{C}_\nu^d) = 1$ as in (5.11).

The nonlocal balance conditions as given by (6.22) are valid for any stochastic network with an arbitrary number of cycles and extend the theoretical framework for biochemical networks as developed by T. L. Hill and coworkers [5], for which $W(\mathcal{C}_\nu^d) \equiv 0$. Note that these balance conditions relate kinetic parameters as given by the transition rates to thermodynamic controll parameters. In this sense, they represent a generalization of the law of mass action to nonequilibrium situations which involve various processes of energy transduction. Since we derived these nonlocal balance conditions from the steady state properties of the system, we called them 'steady state balance conditions' in our previous work [2, 3]. It is important to note, however, that these equations do not contain any quantities that are restricted to the steady state.

We may also derive a local form of the balance conditions if we start from the local energy balance as given by (6.17) and identify the heat Q_{ij} with the transition entropy $\Delta S_{ij} = k_B \ln(\omega_{ij}/\omega_{ji})$ as given by (6.11). This leads to the local balance conditions

$$k_B T \ln(\omega_{ij}/\omega_{ji}) = U_i - U_j + E_{\text{in},ij} - W_{ij}. \tag{6.23}$$

which involve the internal energies U_i and U_j. The latter quantities should be considered as system parameters that cannot be directly controlled in an experiment. A special case of the local conditions (6.23) was obtained in Ref. [58] for simplified motor models with a single cycle and $U_i = U_j$.

The local balance conditions as given by (6.23) can be rewritten in the form

$$\frac{\omega_{ji}}{\omega_{ij}} = \frac{\exp[-U_i/k_B T]}{\exp[-U_j/k_B T]} \, e^{(-E_{\text{in},ij}+W_{ij})/k_B T}. \tag{6.24}$$

As explained in [4], this relation may be interpreted as a constrained equilibrium between the motor states i and j *provided* one properly distinguishes the terms U_i and U_j, which characterize the single motor states or vertices i and j, from the terms $E_{\text{in},ij}$ and W_{ij}, which are related to the transition or di-edge $|ij\rangle$.

6.4. *Practical value of balance conditions.* If one wants to apply our theory to specific motors, the nonlocal balance conditions have the advantage that they provide direct relations between two types of variables that can both be determined experimentally: the transition rates can be measured, at least in principle, via kinetic experiments whereas the energy input E_{in} and the work W represent thermodynamic control parameters. Indeed, for motors that are powered by ATP hydrolysis, the energy input can be expressed in terms of the chemical energy change $\Delta\mu$ per ATP hydrolysis as given by (3.4) and the performed work depends on the external load force F that acts parallel to the filament and on the size of the mechanical step [2, 3, 4].

Therefore, one may use the nonlocal balance conditions in two different ways. If, for a certain motor, all transition rates were known from experiments, one could use the balance conditions to estimate the experimental accuracy. In practise, some of the transition rates will be difficult to measure, and the balance conditions can then be used to estimate the values of the unknown rates. We have recently applied this latter strategy to the cytoskeletal motor kinesin [2] as summarized in section 8 below.

6.5. *Some qualifications.* The derivation of the nonlocal conditions (6.22) was based on two important physical assumptions: (i) the stochastic process that describes the motor dynamics can be represented as a continuous-time Markov process and (ii) the statistical (or Shannon) entropy produced by this process can be identified with the heat that the motor releases into its environment.

These two assumptions will only apply to a specific physical system if we chose an appropriate level of description which includes the (most) relevant states and transitions between these states. This limitation can be understood if we coarse-grain our network models. If we start from the Markov process on the 6-state model shown in Fig. 5, e.g., we can derive the effective stochastic process that governs the mechanical steps alone but this latter process is no longer Markovian [59].

Furthermore, if the statistical entropy produced in a certain network model can be identified with the heat released by the motor, this identification is lost in a coarse-grained model unless the transition rates of this coarse-grained model satisfy certain relations with the transition rates of the original model. This latter point will be illustrated in the next section for the case of a single motor head.

7. Balance conditions for single motor heads. In this section, we will consider the unicycle networks for single motor heads, which contain a single catalytic domain for ATP hydrolysis, see Fig. 4(a) and (b), and provide the presumbly simplest examples for our balance conditions. In fact, these conditions ensure that both the 4-state model and the reduced 3-state model describe the same energy transduction process. We will also show that the internal energies U_i can be expressed in terms of the transition rate constants κ_{ij} and the activity scales $[X]^*$. It turns out that analogous relations apply for network representations in general. Indeed, as shown in Ref. [4], the landscape of internal energies U_i can always be expressed in terms of the transition rate constants and the activity scales for arbitrarily complex networks.

7.1. *Balance conditions for the 4-state model.* First, let us consider again the 4-state network in Fig. 4(a). As previously explained, this network has four states $i =$ E, T, Θ, and D. The state E corresponds to an empty head, the state T to a head occupied by one ATP molecule, the state Θ to bound ADP/P, and the state D to bound ADP. These four states are connected by forward and backward chemical transitions with transition rates ω_{ij}. Thus, the motor head can undergo chemical transitions between the four states but does not perform any mechanical work.

Three of the chemical transitions involve the binding of a chemical species: The ATP binding transition $|ET\rangle$, the P binding transition $|D\Theta\rangle$, and the ADP binding transition $|ED\rangle$. Therefore, the corresponding transition rates ω_{ET}, $\omega_{D\Theta}$, and ω_{ED} must depend on the activities [ATP], [P], and [ADP], respectively.

The 4-state network consists of the single cycle $\mathcal{H}_{4s} \equiv \langle ET\Theta DE \rangle$. As the motor completes the hydrolysis dicycle $\mathcal{H}_{4s}^+ = |ET\Theta DE\rangle$, it binds one ATP molecule during the transition $|ET\rangle$, hydrolyzes this ATP into ADP/P during $|T\Theta\rangle$, releases P during $|\Theta D\rangle$, and releases ADP during $|DE\rangle$. The corresponding chemical energy inputs $E_{in,ij} \equiv E_{ch,ij}$ are given by $E_{ch,ET} = \mu_{ATP}$, $E_{ch,T\Theta} = 0$, $E_{ch,\Theta D} = -\mu_P$, and $E_{ch,DE} = -\mu_{ADP}$.

If we sum these energy inputs over the dicycle $\mathcal{H}_{4s}^+$ and use the definition (3.4) of the overall chemical energy $\Delta\mu$, we obtain the dicycle energy input

$$E_{ch}(\mathcal{H}_{4s}^+) = \sum_{|ij\rangle}^+ E_{ch,ij} = \Delta\mu = k_B T \ln\left(K^{eq} \frac{[ATP]}{[ADP]\,[P]} \right). \tag{7.1}$$

Since this dicycle does not perform work, the steady state balance condition for this dicycle is simply given by

$$\Xi(\mathcal{H}_{4s}^+) = k_B T \ln\left(\frac{\omega_{ET}\,\omega_{T\Theta}\,\omega_{\Theta D}\,\omega_{DE}}{\omega_{TE}\,\kappa_{\Theta T}\,\kappa_{D\Theta}\,\kappa_{ED}} \right) = \Delta\mu. \tag{7.2}$$

This relation must hold for any values of the three activities [ATP], [ADP], and [P]. The latter requirement implies that the transition rates ω_{ET} for ATP binding, ω_{ED} for ADP

binding, and $\omega_{\mathrm{D\Theta}}$ for P binding are given by

$$\omega_{\mathrm{ET}} = \kappa_{\mathrm{ET}}\,[\mathrm{ATP}], \quad \omega_{21} = \kappa_{21}\,[\mathrm{ADP}], \quad \text{and} \quad \omega_{\mathrm{D\Theta}} = \kappa_{\mathrm{D\Theta}}\,[\mathrm{P}], \tag{7.3}$$

whereas all the other rates ω_{ij} are independent of these activities and equal to the corresponding rate constants κ_{ij}. If these dependencies are inserted into (7.2), one obtains the relation

$$\frac{\kappa_{\mathrm{ET}}\,\kappa_{\mathrm{T\Theta}}\,\kappa_{\mathrm{\Theta D}}\,\kappa_{\mathrm{DE}}}{\kappa_{\mathrm{TE}}\,\kappa_{\mathrm{\Theta T}}\,\kappa_{\mathrm{D\Theta}}\,\kappa_{\mathrm{ED}}} = K^{\mathrm{eq}} \tag{7.4}$$

between the rate constants and the equilibrium constant K^{eq}.

The relations (7.4) and (7.2) were previously derived by T. L. Hill for a generic ATPase [5]. In fact, he first derived the equilibrium relation (7.4) using detailed balance and then observed that this latter relation must remain valid even if $\Delta\mu \neq 0$. In contrast, our derivation does not use detailed balance but provides a direct derivation of the general steady state balance condition (7.2) from the energy balance of the dicycle. Detailed balance is then recovered as a special case of this general condition.

Next, we will show that the local form (6.17) of the energy balance leads to explicit expressions (i) for the internal energies U_i and (ii) for the probability distribution P_i^{eq} that the system attains in chemical equilibrium. If we apply the local energy balance relation (6.17) to the four edges of the 4-state network and identify the transition heat Q_{ij} with temperature T times the transition entropy ΔS_{ij} as given by (6.11), we obtain the internal energy changes

$$\Delta U_{\mathrm{ET}} = U_{\mathrm{T}} - U_{\mathrm{E}} = \mu_{\mathrm{ATP}} - k_B T\,\ln(\omega_{\mathrm{ET}}/\omega_{\mathrm{TE}}), \tag{7.5}$$

$$\Delta U_{\mathrm{T\Theta}} = U_{\Theta} - U_{\mathrm{T}} = -k_B T\,\ln(\omega_{\mathrm{T\Theta}}/\omega_{\mathrm{\Theta T}}), \tag{7.6}$$

$$\Delta U_{\mathrm{\Theta D}} = U_{\mathrm{D}} - U_{\Theta} = -\mu_{\mathrm{P}} - k_B T\,\ln(\omega_{\mathrm{\Theta D}}/\omega_{\mathrm{D\Theta}}), \tag{7.7}$$

and

$$\Delta U_{\mathrm{DE}} = U_{\mathrm{E}} - U_{\mathrm{D}} = -\mu_{\mathrm{ADP}} - k_B T\,\ln(\omega_{\mathrm{DE}}/\omega_{\mathrm{ED}}) \tag{7.8}$$

for the 4-state network of a single motor head.

At first sight, these energy differences ΔU_{ij} seem to depend on the chemical potentials and on the activities [ATP], [ADP], and [P]. Such a dependence is, however, not present. In order to show that all internal energy differences are independent of the various activities, we now use the definition (3.2) for the activity scales $[X]^* = [X]\exp[-\mu_X/k_B T]$ that have a certain fixed value for each chemical species X. In this way, the four relations (7.5) – (7.8) are transformed into

$$\Delta U_{\mathrm{ET}} = U_{\mathrm{T}} - U_{\mathrm{E}} = -k_B T\,\ln\left(\frac{[\mathrm{ATP}]^*\,\kappa_{\mathrm{ET}}}{\kappa_{\mathrm{TE}}}\right), \tag{7.9}$$

$$\Delta U_{\mathrm{T\Theta}} = U_{\Theta} - U_{\mathrm{T}} = -k_B T\,\ln(\kappa_{\mathrm{T\Theta}}/\kappa_{\mathrm{\Theta T}}), \tag{7.10}$$

$$\Delta U_{\mathrm{\Theta D}} = U_{\mathrm{D}} - U_{\Theta} = -k_B T\,\ln\left(\frac{\kappa_{\mathrm{\Theta D}}}{\kappa_{\mathrm{D\Theta}}\,[\mathrm{P}]^*}\right), \tag{7.11}$$

and

$$\Delta U_{\mathrm{DE}} = U_{\mathrm{E}} - U_{\mathrm{D}} = -k_B T\,\ln\left(\frac{\kappa_{\mathrm{DE}}}{\kappa_{\mathrm{ED}}\,[\mathrm{ADP}]^*}\right). \tag{7.12}$$

These expressions hold for any value of $\Delta\mu$. For the equilibrium case with $\Delta\mu = 0$, these energy differences determine the equilibrium distribution P_i^{eq} via the detailed balance condition $P_i^{\text{eq}}\,\omega_{ij} = P_j^{\text{eq}}\,\omega_{ji}$ as in (5.8). In this latter case, we can then replace the transition rate ratios ω_{ij}/ω_{ji} in the four relations (7.5) – (7.8) by the probability ratios $P_j^{\text{eq}}/P_i^{\text{eq}}$. As a result, we find the equilibrium distribution $\{P_i^{\text{eq}}\}$ as given by

$$P_i^{\text{eq}} \equiv \frac{1}{Z_{4s}}\hat{P}_i^{\text{eq}} \tag{7.13}$$

with the unnormalized probabilities or Gibbs factors

$$\hat{P}_{\text{E}}^{\text{eq}} \equiv e^{-U_{\text{E}}/k_B T}, \tag{7.14}$$

$$\hat{P}_{\text{T}}^{\text{eq}} \equiv e^{-(U_{\text{T}}-\mu_{\text{ATP}})/k_B T}, \tag{7.15}$$

$$\hat{P}_{\Theta}^{\text{eq}} \equiv e^{-(U_{\Theta}-\mu_{\text{ADP}}-\mu_{\text{P}})/k_B T}, \tag{7.16}$$

and

$$\hat{P}_{\text{D}}^{\text{eq}} \equiv e^{-(U_{\text{D}}-\mu_{\text{ADP}})/k_B T} \tag{7.17}$$

as well as the partition function

$$Z_{4s} \equiv \hat{P}_{\text{E}}^{\text{eq}} + \hat{P}_{\text{T}}^{\text{eq}} + \hat{P}_{\Theta}^{\text{eq}} + \hat{P}_{\text{D}}^{\text{eq}}. \tag{7.18}$$

It is important to note that we could have obtained the same distribution from a (restricted) grand canonical ensemble for the motor system which provides another consistency check of the theory.

7.2. *Balance conditions for the 3-state model.* The 4-state model in Fig. 4(a) can be reduced by combining the hydrolysis transition $|T\Theta\rangle$ with the P release transition $|\Theta D\rangle$ into the new, effective transition $|TD\rangle$. As a result, one obtains the 3-state model as shown in Fig. 4(b) which now contains the single cycle $\mathcal{H}_{3s} \equiv \langle ETDE\rangle$. The steady state balance condition for the dicycle $\mathcal{H}_{3s}^+$ is then given by

$$\Xi(\mathcal{H}_{3s}^+) = k_B T\,\ln\left(\frac{\omega_{\text{ET}}\,\omega_{\text{TD}}\,\omega_{\text{DE}}}{\omega_{\text{TE}}\,\omega_{\text{DT}}\,\omega_{\text{ET}}}\right) = \Delta\mu \tag{7.19}$$

We now require that the 3-state network and the 4-state network describe the same energy balance for the same value of $\Delta\mu$. Comparison of the two steady state balance conditions (7.2) and (7.19) then implies the relation

$$\frac{\omega_{\text{TD}}}{\omega_{\text{DT}}} = \frac{\omega_{\text{T}\Theta}\,\omega_{\Theta\text{D}}}{\omega_{\Theta\text{T}}\,\omega_{\text{D}\Theta}} \tag{7.20}$$

between the transition rates $\omega_{\text{T}\Theta}$, $\omega_{\Theta\text{T}}$, $\omega_{\Theta\text{D}}$ and $\omega_{\text{D}\Theta}$ of the 4-state network and the transition rates ω_{TD} and ω_{DT} of the 3-state network. This relation may also be expressed in terms of the transition entropies $\Delta S_{ij} = k_B\ln(\omega_{ij}/\omega_{ji})$ which leads to

$$\Delta S_{\text{TD}} = \Delta S_{\text{T}\Theta} + \Delta S_{\Theta\text{D}}. \tag{7.21}$$

Furthermore, using the local balance conditions (6.23), we also find that the relation (7.20) implies

$$\Delta U_{\text{TD}} = -\mu_{\text{P}} - k_B T\,\ln\left(\frac{\omega_{\text{TD}}}{\omega_{\text{DT}}}\right) = \Delta U_{\text{T}\Theta} + \Delta U_{\Theta\text{D}} \tag{7.22}$$

with the energy differences $\Delta U_{T\Theta}$ and $\Delta U_{\Theta D}$ as given by (7.6) and (7.7) for the 4-state model.

In equilibrium with $\Delta\mu = 0$, we may again express the probabilities P_i^{eq} in terms of the internal energies U_i and the chemical potentials via the detailed balance condition $P_i^{\mathrm{eq}} \omega_{ij} = P_j^{\mathrm{eq}} \omega_{ji}$. We then recover the Gibbs factors (7.14), (7.15), and (7.17) for the unnormalized probabilities $\hat{P}_{\mathrm{E}}$, $\hat{P}_{\mathrm{T}}$, and $\hat{P}_{\mathrm{D}}$ and the equilibrium probabilities

$$P_i^{\mathrm{eq}} = \frac{1}{Z_{3s}} \, \hat{P}_i \quad \text{with} \quad Z_{3s} \equiv \hat{P}_{\mathrm{E}} + \hat{P}_{\mathrm{T}} + \hat{P}_{\mathrm{D}}. \tag{7.23}$$

Therefore, the relation (7.20) between the transition rates $\omega_{T\Theta}$, $\omega_{\Theta T}$, $\omega_{\Theta D}$ and $\omega_{D\Theta}$ of the 4-state network and the transition rates ω_{TD} and ω_{DT} of the 3-state network implies that the 3-state network is governed by the same transition entropies, internal energies, and Gibbs factors as the 4-state network. In this way, the relation (7.20) illustrates the fact that the transition rates of the coarse-grained model must satisfy certain relations with the transition rates of the original model, if we want to ensure that both models describe the same process of energy transduction. In particular, let us assume that the transition entropies $\Delta S_{T\Theta}$ and $\Delta S_{\Theta D}$ provide good estimates for the heat released by the motor during the transitions $|T\Theta\rangle$ and $|\Theta D\rangle$ in the 4-state model. We may then also identify the transition entropy ΔS_{TD} with the heat released during the transition $|TD\rangle$ in the 3-state model *provided* the transition entropies satisfy the relation (7.21) which is equivalent to (7.20).

8. Kinesin: A case study for chemomechanical coupling.

We have recently applied the general network theory as described in the previous sections to the cytoskeletal motor kinesin [3]. Since kinesin is a two-headed motor, we start from the network description as discussed in subsection 4.2, see Fig. 5. As mentioned, this description is based (i) on the distinct chemical states that the two-headed motor can attain and (ii) on the assumption that the mechanical steps are fast compared to the chemical transitions. Recent experimental observations by Carter and Cross [21] provide strong evidence that such a separation of time scales applies to kinesin.

In Fig. 5, we have illustrated the reduction of the 9-state network to the 6-state network, in which all motor states are characterized by two motor heads that differ in their chemical composition. A slightly different representation of this 6-state network is displayed in Fig. 6(a). As before, the network consists of 6 vertices representing the 6 possible motor states, of solid edges corresponding to the chemical forward and backward transitions, and of one broken edge for the forward and backward mechanical steps. In addition, we have now indicated the direction of the forward mechanical step by a black arrow and the direction of the ATP hydrolysis by a white double-arrow.

The 6-state network contains the two chemomechanical cycles $\mathcal{F} = \langle 12561 \rangle$ and $\mathcal{B} = \langle 23452 \rangle$ where the motor states are labeled as in Fig. 6(a). For small ADP concentrations, the competition of these cycles is predicted to determine the stall force of the motor at which the motor velocity vanishes. In the steady state, the motor velocity v is proportional to the local excess flux $\Delta J_{25}^{\mathrm{st}}$ through the mechanical transition $|25\rangle$ and is given by

$$v = \ell \, \Delta J_{25}^{\mathrm{st}} = \ell \left[P_2^{\mathrm{st}} \, \omega_{25} - P_5^{\mathrm{st}} \, \omega_{52} \right] \tag{8.1}$$

with the step size $\ell = 8$ nm. Using the cycle decomposition (5.21) for the local flux $\Delta J_{25}^{\mathrm{st}}$, one obtains

$$v/\ell = J^{\mathrm{st}}(\mathcal{F}) - J^{\mathrm{st}}(\mathcal{B}) \tag{8.2}$$
$$= J^{\mathrm{st}}(\mathcal{F}^+) - J^{\mathrm{st}}(\mathcal{F}^-) - [J^{\mathrm{st}}(\mathcal{B}^+) - J^{\mathrm{st}}(\mathcal{B}^-)].$$

Therefore, in the 6-state model, this velocity is governed by the competition between the two motor cycles $\mathcal{F}$ and $\mathcal{B}$, and vanishes for $J^{\mathrm{st}}(\mathcal{F}) = J^{\mathrm{st}}(\mathcal{B})$ which provides an implicit equation for the stall force F_{s}. This latter equation can be rewritten in the simple form

$$\frac{\omega_{25}}{\omega_{52}} = \frac{\omega_{23}}{\omega_{56}} \quad \text{for} \quad F = F_{\mathrm{s}}. \tag{8.3}$$

This equation remains unchanged in the limit of small [ADP], in which the transition rates ω_{65} and ω_{32} and the dicycle fluxes $J^{\mathrm{st}}(\mathcal{F}^-)$ and $J^{\mathrm{st}}(\mathcal{B}^-)$ vanish. Thus, for small [ADP], the motor velocity is given by $v \approx \ell[J^{\mathrm{st}}(\mathcal{F}^+) - J^{\mathrm{st}}(\mathcal{B}^+)]$, and the stall force is determined by the two dicycles $\mathcal{F}^+$ and $\mathcal{B}^+$ which both involve ATP hydrolysis.

This prediction of our network model differs qualitatively from those of unicycle models, see, e.g., [37, 38]. Indeed, in motor models with a single motor cycle, the stall force is necessarily determined by the flux balance between ATP hydrolysis and ATP synthesis [38]. Since the ATP synthesis rate is proportional to ADP concentration, the latter balance would imply that the ATP hydrolysis rate vanishes at the stall force in the limit of small ADP concentration. The experiments by Carter and Cross [21] show, however, that the ATP hydrolysis rate is finite in this limit in agreement with our network models. As shown in [3], the 6-state network in Fig. 6(a) agrees even quantitatively with the experimental data in Ref. [21] for an appropriate choice of the transition rates.

Another quantity which can be easily calculated is the ratio, q, of forward to backward mechanical steps which is given by

$$q = \frac{P_2\,\omega_{25}}{P_5\,\omega_{52}} = 1 + \frac{v/\ell}{P_5\omega_{52}}. \tag{8.4}$$

In the 6-state network, an explicit calculation leads to

$$q = \frac{1 + \omega_{56}/\omega_{52}}{1 + \omega_{23}/\omega_{25}}. \tag{8.5}$$

If the load force F is equal to the stall force F_{s}, one has $\omega_{25}/\omega_{52} = \omega_{23}/\omega_{56}$ as in (8.3) which implies $q = 1$. For the transition rates as determined in [3], the step ratio q as calculated for the 6-state model agrees very well with the ratio as observed experimentally in [21]. In addition, the 6-state model also agrees with the data from the Block group [16, 17] and with the data on the P-dependence of the motor velocity as observed by the Howard group in [18]. Thus, the 6-state network provides a good description for almost all measured properties of kinesin's processive motion. There is one exception, however: the decrease of motor velocity with increasing ADP concentration, as calculated for the 6-state model, is not as fast as observed experimentally in [18]. In order to match the observed velocity/ADP relationship, we found it necessary to include another motor state in which both heads contain bound ADP. The resulting 7-state network is shown in Fig. 6(b).
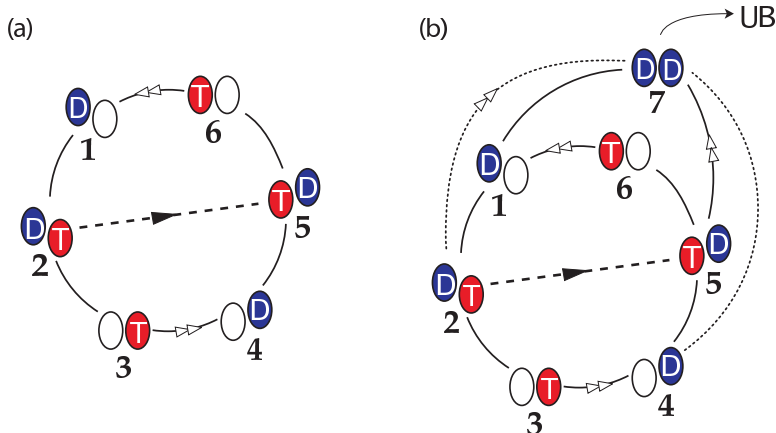
Fig. 6. Network models for kinesin: (a) 6–state network with two chemomechanical cycles, the forward cycle $\mathcal{F} = \langle 12561 \rangle$ and the backward cycle $\mathcal{B} = \langle 23452 \rangle$; and (b) 7-state network with two additional chemomechanical cycles, the additional forward cycle $\mathcal{F}_{DD} = \langle 12712 \rangle$ and the additional backward cycle $\mathcal{B}_{DD} = \langle 27452 \rangle$. In the limit of small ADP concentration, the observed dependence of the motor properties on the load force arises from the competion of the forward cycle $\mathcal{F}$ and the backward cycle $\mathcal{B}$. In the limit of small load force, on the other hand, the available experimental data on the ADP dependence of the motor velocity can be understood from the concerted action of the two forward cycles $\mathcal{F}$ and $\mathcal{F}_{DD}$ [3]. The unbinding of the motor from the filament is most likely to occur from the weakly bound state (D,D).

Inspection of Fig. 6(b) shows that the 7-state network contains two additional chemomechanical cycles, the additional forward cycle $\mathcal{F}_{DD} = \langle 12712 \rangle$ and the additional backward cycle $\mathcal{B}_{DD} = \langle 27452 \rangle$. For small load forces, the motor dynamics is found to be governed by the two forward cycles $\mathcal{F}$ and $\mathcal{F}_{DD}$. The observed reduction of the motor velocity with increasing ADP concentration can be understood as follows. First, the increase of [ADP] leads to an increase in the transition rates ω_{17} and ω_{65}, which acts to reduce the local excess fluxes $\Delta J_{12}^{\mathrm{st}} = P_1 \omega_{12} - P_2 \omega_{21}$ and $\Delta J_{56}^{\mathrm{st}} = P_5 \omega_{56} - P_6 \omega_{65}$, respectively. As a consequence, the increase of [ADP] reduces the flux $J^{\mathrm{st}}(\mathcal{F}^+)$ through the forward dicycle $\mathcal{F}^+$. This reduction of $J^{\mathrm{st}}(\mathcal{F}^+)$ is only partially compensated by the new dicycle flux $J^{\mathrm{st}}(\mathcal{F}_{DD}^+)$, since this latter flux is also impeded by the more frequent transitions $|17\rangle$, which belong to the reverse dicycle $\mathcal{F}_{DD}^-$. As a result, the flux through the di-edge $|25\rangle$, which is proportional to the motor velocity, decreases substantially with increasing [ADP] as observed in experiment [3].

In summary, we have developed a network theory for the chemomechanical coupling of the molecular motor kinesin [3] that is based on the chemical state space of this motor and on the recent experimental observation [21] that its mechanical steps do not exhibit substeps. Our theory provides a quantitative description of the functional dependencies of motor velocity, ratio of forward to backward steps, and randomness parameter on the external control parameters as observed in single molecule experiments [16, 17, 18, 21].

One important aspect of our theory is the presence and competition of several chemomechanical motor cycles. For low [ADP], the motor properties are determined by the two

cycles $\mathcal{F}$ and $\mathcal{B}$, compare Fig. 6, which balance each other at the stall force. Therefore, one has to go beyond the usual unicycle models in order to understand the motor properties of kinesin. For large [ADP], yet another cycle becomes relevant that involves the weakly bound DD state. As will be described elsewhere [60], our theory can also be used to calculate additional motor properties such as the hydrolysis rate and the efficiency of the motor, which are difficult to measure experimentally.

References

[1] R. Lipowsky and S. Klumpp, *Life is motion—multiscale motility of molecular motors*, Physica A 352 (2005), 53–112.

[2] S. Liepelt and R. Lipowsky, *Steady-state balance conditions for molecular motor cycles and stochastic nonequilibrium processes*, EPL 77 (2007), 50002.

[3] S. Liepelt and R. Lipowsky. *Kinesin's network of chemomechanical motor cycles*, Phys. Rev. Lett. 98 (2007), 258102.

[4] R. Lipowsky and S. Liepelt, *Chemomechanical coupling of molecular motors: Thermodynamics, Network representations, and balance conditions*, J. Stat. Phys. 130 (2008), 39–67.

[5] T. L. Hill, *Free Energy Transduction and Biochemical Cycle Kinetics*, Springer Verlag, New York, 1989.

[6] U. Seifert, *Entropy production along a stochastic trajectory and an integral fluctuation theorem*, Phys. Rev. Lett. 95 (2005), 040602/1–040602/4.

[7] B. Alberts, D. Bray, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Essential Cell Biology: An Introduction to the Molecular Biology of the Cell*, Garland, New York, 1998.

[8] J. Howard, *Mechanics of Motor Proteins and the Cytoskeleton*, Sinauer, New York, 2001.

[9] M. Schliwa and G. Woehlke, *Molecular motors*, Nature 422 (2003), 759–765.

[10] R. D. Vale, *The molecular motor toolbox for intracellular transport*, Cell 112 (2003), 467–480.

[11] S. M. Block, L. S. B. Goldstein, and B. J. Schnapp, *Bead movement by single kinesin molecules studied with optical tweezers*, Nature 348 (1990), 348–352.

[12] K. Svoboda, Ch. F. Schmidt, B. J. Schnapp, and St. M. Block, *Direct observation of kinesin stepping by optical trapping interferometry*, Nature 365 (1993), 721–727.

[13] L. Romberg and R. D. Vale, *Chemomechanical cycle of kinesin differs from that of myosin*, Nature 361 (1993), 168–170.

[14] M. J. Schnitzer and St. M. Block, *Kinesin hydrolyses one ATP per 8-nm step*, Nature 388 (1997), 386–390.

[15] S. P. Gilbert, M. L. Moyer, and K. A. Johnson, *Alternating site mechanism of the kinesin ATPase*, Biochemistry 37 (1998), 792–799.

[16] K. Visscher, M. J. Schnitzer, and S. M. Block, *Single kinesin molecules studied with a molecular force clamp*, Nature 400 (1999), 184–189.

[17] M. J. Schnitzer, K. Visscher, and S. M. Block, *Force production by single kinesin motors*, Nature Cell Biology 2 (2000), 718–723.

[18]   W. R. Schief, R. H. Clark, A. H. Crevenna, and J. Howard, *Inhibition of kinesin motility by ADP and phosphate supports a hand-over-hand mechanism*, PNAS 101 (2004), 1183–1188.

[19]   A. Yildiz, M. Tomishige, R. D. Vale, and P. R. Selvin, *Kinesin walks hand-over-hand*, Science 303 (2004), 676–678.

[20]   D. D. Hackney, *The tethered motor domain of a kinesin–microtubule complex catalyzes reversible synthesis of bound ATP*, PNAS 102 (2005), 18338–18343.

[21]   N. J. Carter and R. A. Cross, *Mechanics of the kinesin step*, Nature 435 (2005), 308–312.

[22]   N. R. Guydosh and S. M. Block, *Backsteps induced by nucleotide analogs suggest the front head of kinesin is gated by strain*, Proc. Nat. Acad. Sci. USA 103 (2006), 8054–8059.

[23]   S. Klumpp and R. Lipowsky, *Cooperative cargo transport by several molecular motors*, PNAS 102 (2005), 17284–17289.

[24]   J. Beeg, S. Klumpp, R. Dimova, R. S. Gracia, E. Unger, and R. Lipowsky, *Transport of beads by several kinesin motors*, Biophys. J., submitted.

[25]   R. Lipowsky, S. Klumpp, and T. M. Nieuwenhuizen, *Random walks of cytoskeletal motors in open and closed compartments*, Phys. Rev. Lett. 87 (2001), 108101/1–108101/4.

[26]   R. Lipowsky, Y. Chai, S. Klumpp, S. Liepelt, and M. J. I. Müller, *Molecular motor traffic: From biological nanomachines to macroscopic transport*, Physica A 372 (2006), 34–51.

[27]   S. P. Gross, *Hither and yon: a review of bi-directional microtubule-based transport*, Phys. Biol. 1 (2004), R1–R11.

[28]   M. A. Welte, *Bidirectional transport along microtubules*, Current Biology 14 (2004), R525–R537.

[29]   M. J. I. Müller, S. Klumpp, and R. Lipowsky, *Tug-of-war of molecular motors*, in preparation.

[30]   P. Kraikivski, R. Lipowsky, and J. Kierfeld, *Enhanced ordering of interacting filaments by molecular motors*, Phys. Rev. Lett. 96 (2006), 258103.

[31]   J. Kierfeld, P. Gutjahr, T. Kühne, P. Kraikivski, and R. Lipowsky, *Buckling, bundling, and pattern formation: From semi-flexible polymers to assemblies of interacting filaments*, J. Comp. Theo. Nanosci. 3 (2006), 898.

[32]   If one dissolved the chemical species X in an ideal gas instead of water, the activity scale $[X]^*$ would be equal to $[X]^* = 1/V_X N_{Av}$ with the quantum volume V_X of the species X.

[33]   R. A. Alberty, *Thermodynamics of the hydrolysis of ATP as a function of temperature, pH, pMg, and ionic strength*, J. Phys. Chem. B 107 (2003), 12324–12330.

[34]   R. D. Vale, D. R. Soll, and I. R. Gibbons, *One-dimensional diffusion of microtubules bound to flagellar dynein*, Cell 59 (1989), 915–925.

[35]   M. Guthold, X. Zhu, C. Rivetti, G. Yang, N. G. Thomson, S. Kasas, H. G. Hansma, B. Smith, P. Hansma, and C. Bustamante, *Direct observation on one-dimensional diffusion and transcription by Escherichia coli RNA polymerase*, Biophys. J. 77 (1999), 2284–2294.

[36]   Note that the symbol T is used for the motor head with bound ATP whereas the symbol T is used for the temperature.

[37]   M. E. Fisher and A. Kolomeisky, *Simple mechanochemistry describes the dynamics of kinesin molecules*, Proc. Natl. Acad. Sci. USA 98 (2001), 7748–7753.

[38]   M. E. Fisher and Y. C. Kim, *Kinesin crouches to sprint but resists pushing*, Proc. Natl. Acad. Sci. USA 102 (2005), 16209–16214.

[39]   N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, Elsevier, Amsterdam, 1992.

[40]   J. R. Norris, *Markov Chains*, Cambridge University Press, Cambrigde, 1997.

[41]   R. Lipowsky, *Universal aspects of the chemo-mechanical coupling for molecular motors*, Phys. Rev. Lett. 85 (2000), 4401–4404.

[42] R. Lipowsky and N. Jaster, *Molecular motor cycles: from ratchets to networks*, J. Stat. Phys. 110 (2003), 1141–1167.

[43] G. Kirchhoff, *Über die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Verteilung galvanischer Ströme geführt wird*, Ann. Phys. Chem. 72 (1847), 497–508.

[44] W.T. Tutte, *Graph Theory*, Cambridge University Press, Cambridge, 2001.

[45] T. L. Hill, *Studies in irreversible thermodynamics, IV, Diagrammatic representation of steady state fluxes for unimolecular systems*, J. Theoret. Biol. 10 (1966), 442–459.

[46] J. Schnakenberg, *Network theory of microscopic and macroscopic behavior of master equation systems*, Rev. Mod. Phys. 48 (1976), 571–585.

[47] The relation (5.11) follows from (5.8) by subsequently applying the latter relation to all di-edges of a dicycle. In addition, the relation (5.11) also implies (5.8) as follows from the flux decomposition in (5.21) and the dicycle fluxes as given by (5.12).

[48] T. A. Bak, *Contributions to the Theory of Chemical Kinetics*, Benjamin, New York, 1963.

[49] T. L. Hill and Y.-D. Chen, *Stochastics of cycle completions (fluxes) in biochemical kinetic diagrams*, Proc. Nat. Acad. Sci. USA 72 (1975), 1291–1295.

[50] H.-H. Kohler and E. Vollmerhaus, *The frequency of cyclic processes in biological multistate systems*, J. Math. Biology 9 (1980), 275–290.

[51] J. L. Lebowitz and H. Spohn, *A Gallavotti–Cohen-type symmetry in the large deviation functional for stochastic dynamics*, J. Stat. Phys. 95 (1999), 333–365.

[52] D. J. Evans, E. G. D. Cohen, and G. P. Morris, *Probability of second law violations in shearing steady states*, Phys. Rev. Lett. 71 (1993), 2401–2404.

[53] G. E. Crooks, *Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences*, Phys. Rev. E 60 (1999), 2721–2726.

[54] C. Jarzynski, *Hamiltonian derivation of a detailed fluctuation theorem*, J. Stat. Phys. 98 (2000), 77–102.

[55] In this study, the transition rates ω_{ij} are taken to be time-independent model parameters. More complex situations would arise if we allowed the transition rates to vary with time or considered ensembles of stochastic networks for which the transition rates represent fluctuating variables.

[56] J. L. Luo, C. van den Broeck, and G. Nicolis, *Stability criteria and fluctuations around non-equilibrium states*, Z. Physik B 56 (1984), 165–170.

[57] C. Maes and M. H. van Wieren, *A Markov model for kinesin*, J. Stat. Phys. 112 (2003), 329–355.

[58] U. Seifert, *Fluctuation theorem for a single enzym or molecular motor*, Europhys. Lett. 70 (2005), 36–41.

[59] A. Valleriani, S. Liepelt, and R. Lipowsky, *Waiting time distributions for kinesin's mechanical steps*, EPL (in press).

[60] S. Liepelt and R. Lipowsky, *Hydrolysis rate and efficiency of the molecular motor kinesin*, in preparation.