Irreversible Thermodynamics

David M Rogers\(^1\) and Susan B Rempe\(^2\)

Center for Biological and Materials Sciences, MS 0895, Sandia National Laboratories, Albuquerque, New Mexico 87185, USA
E-mail: dmroge@sandia.gov, \(^1\) slrempe@sandia.gov\(^2\)

Abstract. We present a maximum-entropy theory of mesoscopic kinetics. The theory gives fully nonlinear nonequilibrium thermodynamic relationships and has no explicit requirement for either microscopic bath variables, an equilibrium energy, or an equilibrium partition function. The entropy maximization process is instead carried out over transition probability distributions with constraints on particle position and velocity updates. The Lagrange multipliers for these constraints express the instantaneous temperature and pressure of external (or microscopic) thermostatic driving systems, with which the distinguished system may or may not eventually reach equilibrium. We show that the analogues of the Gibbs-Maxwell relations and free energy perturbation techniques carry over to fluctuation-dissipation theorems and nonequilibrium ensemble reweighting techniques as should be expected. The result is a fully time-dependent, non-local description of a nonequilibrium ensemble coupled to reservoirs at possibly time-varying thermostatic or mechanical states. We also show that the thermodynamic entropy production extends the generalized fluctuation theorem through the addition of an instantaneous information entropy term for the end-points, leading to a concise statement of the second law of thermodynamics.

1. Far From Non-Equilibrium

Standard treatments of statistical mechanics are heavily influenced by equilibrium, thermostatic, ideas. This reliance is especially evident when one attempts to formulate a kinetic theory without reference to a potential of mean force. As a case-in-point, Eyring’s transition-state theory and Marcus’ electron transfer rate equations are fundamentally based on the ‘free energy change’ in a reversible reaction. Such a reaction has no driving force, does not proceed in the forward direction on average, and dissipates no heat. Although nonequilibrium statistical mechanics is the underlying picture that must be used for analyzing laboratory experiments taking place in open systems, the standard analysis of basic phenomena such as Joule-heating, piezoelectricity, or the Seebeck effect\(^1\) relies on a theory of small, linear, deviations from equilibrium. Further, these deviations assume a spatial dependence that pre-supposes a spatial scale on which the fluid energy function is local. This seems to preclude our discussion of anything but local thermostatic quantities.

On the other hand, a set of fluctuation-dissipation theorems have been derived based on a projection from a fully dynamical system described by a Liouville equation.\(^2\)\(^3\)\(^4\)\(^5\) The projection operation removes the pernicious microscopically reversible information from the phase-space distribution and so allows its information entropy to increase.\(^6\) Driving such a projected system can be modeled by adding a time-dependent Hamiltonian, again with reference to the equilibrium picture. In this case the projection causes dissipation of the added energy into
the unmodeled degrees of freedom. Although this picture gives a complete theory, it introduces unnecessary microscopic variables whose only purpose is to set the rate of heat dissipation from the mesoscopic system. That this rate is linked to the mesoscopically observable fluctuations (the famous fluctuation-dissipation theorem) requires a result from second-order perturbation theory on the system’s equilibrium partition function. This state of affairs therefore seems far from nonequilibrium.

2. Maximum Caliber

In the late 1970s, Jaynes[7] introduced an important idea to solve some of the above difficulties. He conceived of an ‘ensemble’ of trajectories, \( x(t) \), in place of the usual ensemble of microstates, \( x \). The probability of each trajectory, \( P(\chi(t)) \), was then to be determined by maximizing a functional of \( P(\chi(t)) \) subject to some constraints on the average values of functionals of \( \chi(t) \). In order to avoid overloading the term entropy, he named the functional to maximize,

\[
\sigma = \int -\log P(\chi(t)) P(\chi(t)) \, D\chi,
\]

the ‘caliber.’ Pursuing the analogy further, Jaynes showed[7, 8] that the Lagrange multipliers constraining averages of path functionals correspond to thermodynamic forces, and that the first derivatives of a suitable function of the caliber yield those averages. Further, its second derivative (minus the space and time-correlation functions) gives the “first-order” perturbation to the path averages. A more complete set of the analogies which could be anticipated is shown in Table 1.

| Equilibrium | Nonequilibrium |
|-------------|----------------|
| Phase Space | Trajectory Space |
| Free Energy | Path Free Energy |
| Entropy | Caliber |
| Average Value Constraint | Average Flux Constraint |
| Equilibrium Average | Path Average |
| Conditional Free Energy (PMF) | Conditional Path log-Probability |
| Thermodynamic forces (δPMF) | Changes in Path Flux (work) |
| Heat Capacity | Thermal Conductivity |
| (none) | Irreversibility and Entropy Production |

Table 1. Correspondence between single-time and time-dependent path maximum entropy formulations of statistical mechanics.

However, the published articles of Jaynes showed very few calculations using this method. Instead, in Ref. [9], the maximum entropy formalism is given to justify adding a Lagrange multiplier to the Hamiltonian, and the standard perturbation theory used by, e.g. Chapter 4 of Kubo [10] was employed for carrying out the calculation. The maximum caliber approach has seen a resurgence in recent work[11, 12, 13] outlining its connection to the fluctuation and entropy production theorems. [14] Connections to local equilibrium theories have been strengthened in Refs. [15, 16, 17], and applications to discrete master equations have been explored in Refs. [18, 19], and from another perspective in Ref. [20].

The pervasive reference to perturbation theory in these applications is likely because the argument leading to straightforward maximization of Eq. 1 contains a fatal flaw. To see this, first note that the average of a quantity at time \( t \) is dependent on the “control” parameters

1 It is worthwhile to note that Jaynes gave special attention to perturbation terms of all orders here.
throughout the whole trajectory. This is because logical inference does not contain a preferential time direction. Knowledge that the system has a given property at time \( t + \tau \) constrains the state at time \( t \), as it should, but a force \( \lambda_k \) acting at a future time should not be able to influence the past trajectory. Although this could be alleviated by requiring only causal information to enter into the determination of the state at time \( t \), this approach leads to a probability distribution valid only for \( x_t \) and not throughout a trajectory.

Here, we show for the first time the benefits of maximizing the entropy of the transition probability distribution under suitable constraints. We note here that an argument for using this approach in the long-time limit has been given in Ref. [21], and that several authors have also noted the need for another 'entropy' associated with time-dependent processes [22, 23, 24, 25].

Limiting the scope of the maximum trajectory entropy procedure in this way automatically corrects a related re-normalization problem. Suppose phase space were to branch at a future time \( t + \tau \). In this case, a uniform measure on path space would assign points at time \( t \) a different weight depending on future events. A simple example is the Monty Hall problem with a prize assigned to the first door without loss of generality. [26] The contestant’s choice of one of three doors plus Monty opening another (not concealing a prize) constitutes four possible paths, and maximum path entropy would give each path an equal weight – an intuitive, but incorrect, solution. The correct probability assignment is a uniform distribution for each transition, leading to ending weights of \( 1/3 \times 1/2 = 1/6 \) for the two paths following the selection of the first door.

In more abstract terms, the marginal probability at each time should not depend on the future – a concept expressed mathematically by defining a progressively measurable function with respect to the natural filtration of stochastic processes [27].

In practice, this flaw can be avoided by considering only processes where assigning equal a priori weight to all paths is equivalent to assigning equal weight to all transitions. By Liouville’s theorem, this is obviously true for deterministic processes. More generally, the approaches are equivalent when the number of possible transitions does not depend on the starting point. The path entropy approach of Jaynes is therefore valid in the absence of factors re-normalizing for starting-point dependent differences in the number of possible paths \( Z[\lambda, x_i] = Z[\lambda_i] \) in Eq. [1].

3. Maximum Transition Probability Entropy

A realization of a time-dependent process is a history, \( \{x_0, x_1, \ldots, x_S\} \), giving the state, \( x \), of the system at times \( i = 0, \ldots, S \). In a stochastic process, knowledge of the history, \( \{x_0, x_1, \ldots, x_i\} = \Gamma_i \), is not sufficient to determine the future beyond time \( i \). Instead, knowing the process history gives us a probability distribution, \( \mathcal{P}(x_{i+1}|\Gamma_i) \), over the state at time \( i + 1 \) conditional on the observed history. The conditional probability is formally a function of any predictive information we know, such as the past history, \( x_{j\leq i} \). It is required to be normalized, \( \int \mathcal{P}(x_{i+1}|\Gamma_i) \, dx_{i+1} = 1 \). For a Markov process, the conditional probability is fixed and a function of only the previous state, \( x_i \). For a general system with known averages, \( \langle f_k(x_{i+1}; \Gamma_i) | \Gamma_i \rangle \) (where \( f \) is some function, and the notation ‘| \( \Gamma_i \)’ is read as conditional on the history up to time \( i \)), we introduce one (unknown) Lagrange multiplier, \( \lambda_{k,i} \), for each average and carry out the standard constrained maximization of the transition entropy,

\[
\delta \mathcal{H}_{i+1}|\Gamma = \delta \sum_{\{x_{i+1}\}} p(x_{i+1}|\Gamma_i) \left[ - \ln \frac{p(x_{i+1}|\Gamma_i)}{p_0(x_{i+1}|\Gamma_i)} \right]
- (\ln Z[\Gamma_i] - 1) - \sum_{k=1}^m \lambda_{k,i} f_k[x_{i+1}, \Gamma_i]
\]  

\[ (2) \]

\[ (3) \]

\[ ^2 \text{We leave updating the player’s state of knowledge about the prize out of the discussion, and reiterate that the full path probabilities are } 1/6, 1/6, \text{ (contestant’s correct choice plus Monty choosing randomly from the remaining doors), and } 1/3, 1/3 \text{ (contestant’s two incorrect choices plus Monty’s forced choice of an incorrect door).} \]
to yield
\[ p(x_{i+1}|\Gamma_i) = p_0(x_{i+1}|\Gamma_i)e^{-\eta_i[x(t)]}/Z[\lambda_i,\Gamma_i]. \] (4)

The terms in the exponential make a quantity like a Boltzmann weight,
\[ \eta_i[x(t)] \equiv \sum_{k=1}^{m} \lambda_{k,i} f_k[x_{i+1},\Gamma_i], \] (5)
and there is a quantity corresponding to a partition function
\[ Z[\lambda_i,\Gamma_i] = \sum_{\{x_{i+1}\}} p_0(x_{i+1}|\Gamma_i)e^{-\eta_i[x(t)]}, \] (6)
which is in fact a functional of the instantaneous Lagrange multipliers and the system history. The transition probability entropy can be re-written in terms of \( Z \) and the averages as in the equilibrium theory,
\[ H_{i+1}|\Gamma = \ln Z[\lambda_i,\Gamma_i] + \langle \eta_i|\Gamma_i \rangle \] (7)
\[ dH_{i+1}|\Gamma = \sum_{k=1}^{m} \lambda_{k,i} d\langle f_k|\Gamma_i \rangle. \] (8)

These are conditional averages, meaning that they are summed only over the possible values of \( x_{i+1} \) and are still functions of the history, \( \Gamma_i \). We can also take Legendre transforms to yield quantities corresponding to state-functions.
\[ \Psi_{i+1}|\Gamma \equiv \sum_{k=1}^{m'} \lambda_{k,i} \langle f_k|\Gamma_i \rangle - H_{i+1}|\Gamma \] (9)
\[ d\Psi_{i+1}|\Gamma = \sum_{k=1}^{m'} \langle f_k|\Gamma_i \rangle d\lambda_k - \sum_{k=m'+1}^{m} \lambda_{k,i} d\langle f_k \rangle. \] (10)

Of course here the functions \( \Psi \) are functionals of the system history and of \( \lambda_i \).\[25\] We will see that the \( \lambda_i \) control rates of change, e.g. using constraints of the form \( f = x_{i+1} - x_i \), so that the equilibrium notion of a state function is replaced here with a single-time, transition state function. That is, the Lagrange multipliers specify the state of the external, driving system between times \( i \) and \( i + 1 \). In the continuous limit, the multipliers \( \lambda_i \) exist between times \( i \) and \( i + 1 \), and the sum over \( x_{i+1} \) should be viewed as a Stratonovich path integral between these times. This distinction is important when \( \lambda \) are functions of \( \Gamma_i \). We have also found that differentiation of the transition partition function, \( Z \), gives results identical to the stochastic integration of quantities such as instantaneous work and heat.

In the deterministic case, we know that \( x_{i+1} \) can be related to \( x_i, x_{i-1} \) through a discrete action,\[29\] such as
\[ A = \frac{(x_{i+1} - x_i)^2}{2\epsilon^2} + \frac{(x_i - x_{i-1})^2}{2\epsilon^2} - V(x_i) + \text{const.}(\{\ldots, x_{i-1}, x_{i+1}, \ldots\}). \] (11)
Here \( \epsilon \) represents the length of a small time-step. Differentiation of this equation with respect to \( x_i \) must yield zero, so the process is deterministic with the usual Verlet integration,
 whom whose columns correspond to the constraint directions $\partial q$ of the noise is to re-distribute $\hat{\theta}$ system and the bath, then it makes sense to enforce a constraint on the average change $\langle dI(t) \rangle$. If we assume the existence of some “total” invariant between the system on the dynamics. If we assume the existence of a generalized momentum, $\hat{\theta}$, that is nonetheless an invariant of some ‘complete’ system, we can define (using Eq. [12])

$$I(t) - I(\infty) = \int_{-\infty}^{t} -y^T (t') \frac{\delta A}{\delta x(t')} dt'. $$

If the action functional is invariant to this transformation ($A[x(t)] \rightarrow A[x(t) + q(t, a)]$), then the corresponding invariant is a conventional conserved quantity. The substitution $x(t) \rightarrow x(t) + a$ ($y = 1$, the ones-vector) generates the momentum, $x(t) \rightarrow x(t + a) (y = \hat{x})$ generates the energy, and so on.

It is instructive to consider a case where the action is not invariant to the transformation. For example, momentum is not conserved in the simple harmonic oscillator ($L = m\dot{x}^2/2 - kx^2/2 \neq m\ddot{x}^2/2 - k(x + a)^2/2$). However, if we assume the existence of a generalized momentum, $\hat{\rho}$, that is nonetheless an invariant of some ‘complete’ system, we can define (using Eq. [12])

$$\frac{d\hat{\rho}(t)}{dt} = -1^T (F - \hat{\rho}) = m\ddot{x} + kx$$

as a momentum exchange. If the net momentum of the observed system changes by more than its internal force, it implies that an external system has lost exactly this amount of momentum. In the presence of noise, this quantity may not be conserved ($d\hat{\rho}$ is stochastic). Thus one effect of the noise is to re-distribute $\hat{\rho}$ over the system and the bath.

In general, we can define a set of constraints on $\{ (dI_j(t)/dt) \}_{j=0}^{m}$ using an $N \times m + 1$ matrix $Y$, whose columns correspond to the constraint directions $\frac{\partial q_j(t,a)}{\partial a} \bigg|_{a=0}$ and whose leftmost column is reserved for the energy $Y_{*,0} = \hat{x}$. In this case, the vector of exchanges for a given trajectory over a given time interval, $\epsilon$, is

$$dI(t) \equiv -Y(t)^T \frac{\delta A}{\delta x(t)} \epsilon. $$

In this equation, the presence of $\epsilon$ is used to denote the Stratonovich integral implicitly. The selection of constraints must be made based on the allowed interactions of the system with its reservoir.

These considerations have shown a simple method for including the influence of an external system on the dynamics. If we assume the existence of some “total” invariant between the system and the bath, then it makes sense to enforce a constraint on the average change, $\langle dI(t) \rangle$. [12] -- [12]
We should note our fixed sign convention, where $I$ is always taken to be a quantity belonging to the central system under consideration. Before moving on to discuss the obvious connection of these changes to the thermodynamic work, we shall first consider the significance of the Lagrange multipliers in this formalism.

Writing Eq. \ref{eq:lagrange} with $f_0 = -\epsilon \frac{\delta A}{\delta x(t)}^2$, $\lambda_0 = (2C)^{-1}$, and a vector of constraints, $dI$ (Eq. \ref{eq:constraint}), enforced using a vector of Lagrange multipliers, $\lambda = \beta/2$, we find a Gaussian-distributed probability for

$$\frac{\delta A}{\delta x(t)} \epsilon = CY(t)\beta \epsilon/2 + (C\epsilon)^{1/2} z(t),$$

where $z$ denotes a standard normal random variable. To constrain only the energy exchange (the leftmost column of $Y$), we use $f_0 = -\epsilon \dot{x}^T \frac{\delta A}{\delta x(t)} (=dH$, the change in the Hamiltonian) and $\lambda_0 = \beta_0/2$, so

$$\dot{\epsilon} = F(x(t))\epsilon - C(\dot{x}\beta \epsilon/2) + C^{1/2} dW(t).$$

Here the centered (Stratonovich) Wiener process increment, $dW(t) = \epsilon^{1/2} z(t)$, has been substituted for $z(t)$.

This equation is exactly the Langevin equation. Note that our method of derivation generalizes trivially to include imposed time-correlations between $\frac{\delta A}{\delta x(t)}$ and $\frac{\delta A}{\delta x(t')}$ as well as arbitrary external forces. Note especially that the correspondence between a time-correlation function and a time-dependent damping (dissipation) function are automatic in the Gaussian theory.

It is well-known from the Fokker-Planck equation\cite{31},

$$\frac{\partial \rho(x,p)}{\partial t} = -(M^{-1}p)^T \nabla_x \rho - \nabla_p \left[ F\rho - C\dot{x} \beta \rho/2 - C\nabla_p \rho/2 \right],$$

that the solution to Eq. \ref{eq:langevin} is the canonical distribution with temperature $\beta_0^{-1} = k_B T$, where $k_B$ is the Boltzmann constant. See also Ref.\cite{32} for the formulation in the presence of memory terms. This solution is independent of $C$, suggesting that a natural parametrization for the Langevin equation is in terms of the temperature and $C^{1/2}$, which is related to the thermal conductivity or rate of temperature equilibration with the unmodeled external system.

Generally, if an invariant, $I$, can be expressed as a function of $x$ and $p$, then a similar result should hold for other common equilibrium thermodynamic ensembles as well, such as the $N, P, T$ ensemble where $\langle dV \rangle$ constitutes an additional dynamic variable and constraint.\cite{33} This establishes the physical interpretation of the Lagrange multipliers as the thermostatic variables of the bath that dictate the eventual equilibrium of the system. Note that increasing $\beta$ tends to decrease $\langle dI \rangle$, for example increasing pressure will drive the volume downward.

### 4. Irreversible Thermodynamics

Having firmly established the connection between the constrained dynamical equations and the equilibrium distribution, we next construct a thermodynamic view of any process that employs a series of transition probabilities to effect a change in the state of the system. This construction will lead naturally to a view of a (general, nonequilibrium) thermodynamic process in terms of a thermodynamic path transforming one type of energy into another with a concomitant irreversible entropy production. In this section, as in the above, we do not rely on results from the equilibrium theory.

To begin, we exactly define a system state, $A$, as any information known about a system that is sufficient to construct a probability distribution for its variables, $P(x|A)$. The machinery of
statistical mechanics can then be used to propagate this information to system states at other
time points and under alternate possible processes.

The work of Joule and Thomson showed that there exists a series of mechanical operations
that can be performed to effect a transition between any two thermostatic states;[34] however,
this transition can only take place in the direction of increasing entropy. The entropy increase
comes about because of experimental inability to control the detailed motions of all particles, and
is therefore zero in the case of completely controllable mechanical work. Thus, it is important
to define a mechanical, adiabatic process in which all work is completely controlled by letting
\( C \rightarrow 0 \) with constant external force experienced by the system
\( F_{\text{ext},j}^{\text{sys}} = -\frac{\partial x(t)}{\partial I_j} F_{\text{ext},j}(t) \). At a finite
temperature, there is some amount of uncertainty about the exact state of
\( I \) on short time
scales. This leads to a discrepancy between the force exerted by the external system,
\( F_{\text{ext},j}(t) \), and its “long time scale” counterpart experienced by the system,
\( F_{\text{ext},j} = \frac{\partial x(t)}{\partial I_j} F_{\text{ext},j}(t) \). The mechanical, adiabatic process is formally a zero-temperature, continuous-time limit. In general,
only a subset of work values can be controlled. Before proceeding it will be necessary to solidify
the concept of controllable work.

If this work is to be delivered by an external thermostatic system, for example an adiabatically
coupled piston, then the first law of thermostatics gives
\( F_{\text{ext},j} = -dU_{\text{ext}}/dI_j = dU_{\text{ext}}/dI_{\text{ext},j} = -\beta_j/\beta_0 \) if the force can be assumed constant over a sufficiently short time-step. The \( \beta \) variables
are the Lagrange multipliers (2\( \lambda \) in Eqns. 2-10) controlling the rates of change of \( I \) between the
two systems. Mechanically, this force corresponds to the force on a wall exerted by a spring
placed externally to it. The total force on the wall is, of course,
\[
F_{\text{tot},j} = -\frac{dU_{\text{int}} + dU_{\text{ext}}}{dI_j} \equiv F_{\text{int},j} + F_{\text{ext},j}.
\]

This implies that if a two spring system were disconnected after a change
\( d\tilde{I} \equiv [dI_1, \ldots, dI_m]^T \),
their internal potential energies would have changed by an amount
\[
dU_{\text{int}} = -d\tilde{I}^T F_{\text{int}}
dU_{\text{ext}} = -d\tilde{I}^T F_{\text{ext}} = d\tilde{I}^T \tilde{\beta}/\beta_0
\]
\[
\tilde{\beta} \equiv [\beta_1, \ldots, \beta_m]^T.
\]
The sum of these two energy changes is not necessarily zero due to the possibility of momentum
change. Using the known energy change of the system, it should then be possible to solve for
the change in kinetic energy of the constraint. For the system, the total energy change is given by
\[
dE = -\frac{\delta A}{\delta x}^T \dot{x} \epsilon = dI_0
= dW + dQ
\]
\[
dW_j = -\frac{\delta A}{\delta x}^T y_j \dot{a}_j \epsilon
\]
\[
dW = d\tilde{I}^T \dot{a}
\]
\[
dQ = -\left( \ddot{x} - \tilde{Y} \dot{a} \right)^T \frac{\delta A}{\delta x} \epsilon,
\]
where we have used Eq. 12 for \( \frac{\partial A}{\partial a} \), and multiplied by \( da/dt \) to define work values and
a corresponding “non-mechanical” energy transfer, \( dQ \). In the adiabatic limit, all energy
transferred to the system by external forces should be reflected by known mechanical changes
related to \( \{I\} \), which is precisely what \( \dot{a} \) allows us to do. Note also that all quantities refer
to the central system unless specifically denoted ‘ext.’ Thus \(dW\) means the work done on the system.

In the mechanical limit, \(-\frac{\delta A}{\delta x}\) is just \(\dot{p} - F_{\text{int}} = F_{\text{ext}}^\text{sys}\), the external force experienced by the system in the absence of the thermostatting random noise, and \(\dot{y}_{s,0} = \dot{x}\). Substituting this quantity from the Langevin equation (14), an adiabatic, mechanical system must satisfy

\[
dQ = -(\dot{x} - \dot{\tilde{y}} \dot{a})^T \frac{\delta A}{\delta x} \epsilon = -\frac{1}{2} (\dot{x} - \dot{\tilde{Y}} \dot{a})^T C \tilde{Y} \beta \epsilon = 0.
\]

The last section gave some physical insight into the quantities \(\dot{a}\). A mathematical consideration of the previous equation shows that \(\dot{\tilde{y}} \equiv \sum_{j=1}^m y_j \dot{a}_j\) (\(\dot{y}\) being identical to \(\dot{x}\) with the first column removed) can be understood as a projection, removing components of \(\dot{x}\) parallel to \(\delta A/\delta x\). If we therefore define (writing the Moore-Penrose pseudoinverse of \(A\) as \(A^+\))

\[
\dot{a} \equiv (C^{-1/2} \tilde{Y})^+ C^{-1/2} \dot{x},
\]

then \(dQ = 0\) identically. For an example application, \(y = 1\) (ones-vector) is associated with the system momentum, and Eq. (19) generates the average velocity \(\dot{a} = 1^T \dot{x}/N\) when \(C = cI\). Eq. (19) is invariant to multiplication of \(\tilde{C}\) by a constant, and so persists in the deterministic limit. The work done on the system (Eq. 16) is

\[
dW = -\frac{\delta A^T}{\delta x} \tilde{Y} (C^{-1/2} \tilde{Y})^+ C^{-1/2} \epsilon.
\]

Note the information-theoretic quality of the work defined by the equation above. If separate reservoirs existed that were able to influence the motion of each particle in the system independently, then \(\dot{Y}\) would become an identity matrix, and \(dW\) would equal the total energy added to the system, \(dE\), identically. In the presence of noise, the work taken from the thermostatic reservoirs does not necessarily equal the work done on the system. In the \(y = 1\) example above, the momentum work done on the system is

\[
\frac{1}{N} \left( \sum_a \dot{p}_a - F_a \right) \left( \sum_a \dot{x}_a \right) \epsilon.
\]

Here “\(\sum_a\)” denotes a sum over \(N\) (one-dimensional) particles, with obvious extension to multiple dimensions. Because this interaction controls only the total system momentum, the work is computed using the average velocity change. The expectation of this stochastic integral is \(\langle dW_p(\epsilon) \rangle = \frac{1}{N^2} \tilde{Y}^T C (\lambda^1 - \beta_0 \dot{x}) / 2\), while \(dU_{\text{ext,p}}(\epsilon) = \frac{1}{\beta_0} \tilde{Y}^T C (\lambda \mathbf{1} - \beta_0 \dot{x}) / 2\). Similarly, an electric field can couple only to the net dipole moment of a system. This implies the transformation of applied energy to heat if it cannot be manipulated to perfectly match fluctuations in the driven variable.

The kinetic energy change of \(I\), ascribed to the reservoir, can be determined in the mechanical limit from

\[
0 = dU_{\text{ext}} + dK - dW \Rightarrow dK = -dI^T (\dot{\beta}/\beta_0 - \dot{a}),
\]

using Eqns. 15 and 16. The equations above thus completely describe any exchange of mechanical energy between deterministic systems exerting known forces. If the two-spring

\[3\] Note the similarity of this form to \(\int \dot{p} - F \, dt\), describing a possible parallel to the flux connection, or integrated applied potential. [25]
system considered above were disconnected at time $S$, with an outside observer absorbing the kinetic energy of the wall (Eq. 21), the energy change of the reservoir would reduce to the usual thermostatic potential change $\int_0^S dU_{ext}(t)$.

A well known consequence of Liouville’s theorem is that the entropy change is zero in a completely deterministic process. Using adiabatic processes, then, it is possible to propagate a starting state, $A$, to any state with constant entropy. If, however, phase space volume were not preserved, it would amount to discarding information on the state of the system at a given time (for example, by integrating the probability over short time intervals). Then the amount of work that can be recovered from the system will become less than the amount input. In an extreme case, all information about the system may have been lost, flushing the corresponding information content to zero. Starting from this unknown state, $p^0$, the probability of a given frequency distribution, $p$, is approached by $P(p) \propto e^{\mathcal{H}[p]}$, with $\mathcal{H}[p]$ the familiar information entropy functional,

$$\mathcal{H}[p] = -\sum_i p_i \ln p_i / p_i^0.$$  

Further, if exchanges of conserved quantities during some process $A \rightarrow B$ are known, then the set $I(B)$ are also known from $A$. This information can be usefully employed to increase the amount of work that can be recovered, showing the entropy as a measure of “lost work.” Consequently, any adiabatic process should be described by not only the mechanical work values, but also the change in information entropy due to information loss.

Next, consider allowing heat exchange in addition to controllable work. Assuming an infinitely large reservoir (or a short enough time-step), added heat will cause a negligible change in reservoir temperature. Because we have assumed the work done on each reservoir (Eq. 15) can be reversibly stored, these are not associated with an entropy change. We therefore introduce the physical entropy change in the reservoir as due only to exchange of heat, or “non-work” energy, $dS_{ext} = \beta_0 dQ_{ext}$, with $\beta_0^{-1} = k_B T_{ext}$. In this article, different notations are used for the information entropy, $\mathcal{H}$, and the physical entropy, $S$. Both of these quantities are defined to be unitless and have some relation to the information entropy of Eq. 22. The energy rejected to the reservoir as heat, $dQ_{ext}$, originates from heat removed from the system plus a kinetic energy, $dK$, assumed to be recoverable only as heat.

$$dQ_{ext} = -dQ + dK = -dE - dU_{ext} = \frac{\delta A}{\delta x(t)} \dot{x} c - d\tilde{I}^T \frac{\beta}{\beta_0} = -d\tilde{I}^T \frac{\beta}{\beta_0}$$

This heat can alternately be understood as the total energy dumped to the environment minus the energy removed “reversibly” ($dU_{ext}$). This interpretation shows that if some of the changes in the environment were re-classified as irreversibly stored, so that the information $\beta_j dI_j$ becomes lost, then this is equivalent to adding that energy to the total $dQ_{ext}$. The total energy rejected to the environment is recoverable (making a given step mechanical as considered above) if and only if $dQ_{ext} = 0$, implying $dE = dI_0 = -d\tilde{I}^T \beta / \beta_0$. These considerations again highlight the subjective nature inherent in the definition of irreversibility.

Connecting back to the usual thermostatics, $U_{ext}$ plays the role of an energy for the reservoir.

$$dU_{ext} = dE_{ext} - \beta_0^{-1} dS_{ext} = d\tilde{I}^T \beta / \beta_0$$

To use this relation, we must provide an experimental justification for the ability to use or store the energy terms appearing in the last equation.
For any transformation, $A \rightarrow B$, the total entropy change deriving from information loss and interactions with the environment is given by the change in information entropy (Eq. 22) plus the heat exchange term (Eq. 23). This can be integrated over the time in which the transformation takes place as

$$\Delta S_\text{inf,tot} = \mathcal{H}[P(x|B)] - \mathcal{H}[P(x|A)] + \beta_0 Q_{\text{ext}}. \quad (24)$$

It should be noted that this formula is still not complete if there is a change in the phase space between $A$ and $B$, for example if particles are added/removed, or if the state space is uniformly dilated. In this case, we have extra information on the region of phase space occupied after a transition. In general, if the state at time $i$ is known to be $x_i$, then the size of the region of configuration space accessible at time $i+1$ is $Z_i(\beta, x_i)$ (Eq. 6). This reduces the entropy at $i+1$ to $S_{i+1} - \ln Z_i(\beta, x_i)$. Writing this down for each transition,

$$\Delta S_\text{tot} = \left\langle \sum_{i=0}^{S-1} dS_i \right\rangle$$

$$dS_i \equiv -\ln \frac{P_{i+1}(x_{i+1}) Z_i(\beta, x_i)}{P_i(x_i) Z_i(-\beta, x_{i+1})} - dI^T \beta. \quad (25)$$

This equation has been symmetrized by including the corresponding entropy decrease for $i$ given that $x_i$ was inferred from $x_{i+1}$ using the forward step probability but with reversed forces, $-\beta$. Support for this form is given in the next section, where it is also shown that $\Delta S_\text{tot} \geq 0$ using the Gibbs inequality.

Nevertheless, Eq. 24 can already be applied to Langevin and Brownian motion with uniform diffusion constants. As another example, applying Eq. 24 to a process taking any point to the equilibrium distribution shows that the expected entropy change for this process is the usual system entropy difference plus $\beta_0 Q_{\text{ext}} = -\beta_0(\langle E|B \rangle - \langle E|A \rangle) - \Delta \tilde{T}^T \beta = (I(A) - I(B))^T \beta$. Because the end-point entropies are fixed, alternate processes for transforming $A \rightarrow B$ are restricted to varying $\Delta S_{\text{ext}} = \int \beta_0 dQ_{\text{ext}}$. We may give a direct microscopic interpretation for the Clausius form of the second law for such processes.

$$\Delta S_\text{tot} = \Delta \mathcal{H} + \int \beta_0 dQ_{\text{ext}} \geq 0 \quad (26)$$

Here $\Delta \mathcal{H}$ is a function of the end-points, and $\beta_0$ and $dQ_{\text{ext}}$ are fully variable along the path. Choosing a path from a specified distribution, $A$, to a maximum entropy distribution, $I(A)$, and then to the ending maximum entropy distribution $I(B)$, we may employ a quasistatic, “reversible,” path between the two maximum entropy distributions so that $\min \int \beta_0 dQ_{\text{ext}}(A \rightarrow I(B)) = \min \int \beta_0 dQ_{\text{ext}}(A \rightarrow I(A))$. The heat evolved in this best-case process has its origin in the re-classification of information that occurs during the coarsening of $A$, in accordance with conclusions on Maxwell’s demon. 36

The equilibrium theory is therefore contained in the present development in the form of a slow process that quickly relaxes to a canonical distribution at each moment. This perspective shows the intimate connection between coarse-graining that assumes infinitely fast relaxation of the reservoir and the traditional theory of quasi-static processes. The complete theory also permits an investigation of both relaxation processes and entropy production in time and history-dependent systems moved arbitrarily far from equilibrium by coupling to simple thermostatic reservoir systems.

5. Path Perturbation and Connection to Entropy Production Theorems

Free energy perturbation techniques are well-known methods for computationally estimating free energy differences to high accuracy. 37 This method carries directly over to nonequilibrium ensembles and allows us to progress beyond linear perturbation theories.
Any two processes on the same path space, \( \{ \Gamma \} \), can be compared using a likelihood,

\[
e^{\lambda A - B ({x}_i \delta)} = \mathcal{P}_B (\{x\}_i) / \mathcal{P}_A (\{x\}_i) \tag{27}
\]

\[\Rightarrow \langle l \rangle_B = \mathcal{D}[B|A] \geq 0,
\]

where \( \mathcal{D} \) is the Kullback-Leibler divergence (and necessarily positive). The likelihood obeys \( \langle e^{l[\Gamma]} \rangle_A = 1 \), and (by the Gibbs inequality) \( \langle l \rangle_A \leq 0 \leq \langle l \rangle_B \). The distribution of the weight satisfies the large deviation formula [38]

\[
e^{\lambda} \mathcal{P}_A (L = l[\Gamma]) = \int e^{l[\Gamma]} \delta (L - l[\Gamma]) \mathcal{P}_A (\Gamma) \ d\Gamma = \int \delta (L - l[\Gamma]) \mathcal{P}_B (\Gamma) \ d\Gamma = \mathcal{P}_B (L = l[\Gamma]),
\]

where \( d\Gamma \) represents integration over all possible paths.

It is possible to express the path average of a functional, \( a[\Gamma] \), using this weighting factor as

\[\langle a[\Gamma] \rangle_B = \left( a[\Gamma] e^{l[\Gamma]} \right) \right|_A. \tag{28}\]

In the equilibrium theory, however, a constant related to the free energy difference is usually cancelled on the right-hand side of this expression. There is no such constant in the prior equation because we have not identified an appropriate extensive variable.

It has been claimed that entropy production can be gauged by the ratio of forward to reverse path probabilities [4] [11]. For the process derived in Sec. [3], we can define a reverse process by inverting the sign of the generalized forces, \( \beta_i \), and normalizing the distribution separately for each \( x_{i+1} \). This reversal corresponds to an attempt at guessing whether energy has been added or subtracted during a step \( i \to i + 1 \), with the action deviation constraint, \( C \) (Eq. [14]), unchanged. Thus, we can define a ratio

\[
edS_i = \frac{\mathcal{P} (x_{i+1} | x_i, \beta) \mathcal{P} (x_i)}{\mathcal{P} (x_{i+1}, -\beta) \mathcal{P} (x_{i+1})} = \frac{\mathcal{P} (x_i) p_0 (x_{i+1} | x_i) Z_{-\beta} (x_{i+1})}{\mathcal{P} (x_{i+1}) p_0 (x_i | x_{i+1}) Z_{\beta} (x_i)} e^{-d\mathcal{F} \beta}.
\]

Since \( \frac{p_0 (x_{i+1} | x_i) p_0 (x_i | x_{i+1})}{p_0 (x_i) p_0 (x_{i+1})} \) by Bayes’ theorem, this result exactly matches Eq. [25] arrived at through thermodynamic reasoning.

Moreover, several steps can be concatenated to give

\[\sum_{i=0}^{\beta^{-1}} dS_i = \frac{\mathcal{P} (\Gamma | x_0, \beta) \mathcal{P} (x_0)}{\mathcal{P} (\Gamma | x_S, -\beta) \mathcal{P} (x_S) / p_0 (x_S)}. \tag{30}\]

This equation can therefore be viewed as a statistical basis for the second law of thermodynamics. It is physically motivated by observing that the entropy increase is attributed to a combination of environmental entropy changes, \( dQ_{ext} \), and information-like entropy changes,

\[-\ln \frac{\mathcal{P} (x_{i+1}) Z_{\beta} (x_i) / p_0 (x_{i+1})}{\mathcal{P} (x_i) Z_{-\beta} (x_{i+1}) / p_0 (x_i)}.
\]

This term implies an extra contribution to the total entropy change beyond Eq. [24]. To understand this contribution, consider a simple one-state system, to be transformed into a two-state system through the (unbiased) transition probability \( \mathcal{P} (x_{i+1} | x_i) = p_0 (x_{i+1}) = 1/2 \). The
presence of the normalization in Eq. [25] leads to an additional factor of $-\ln Z[x_i]/p_0(x_{i+1}) = -\ln 2$ in the entropy. This term is canceled by the entropy of the resulting state, $\mathcal{P}(x_{i+1}) = 1/2$, so that the total entropy change for this process is zero. If instead we perform a bit-set operation by going in the opposite direction, the entropy will increase if $\mathcal{P}(x_{i+1})$ is any non-uniform state – corresponding to our loss of information on discarding the bit. A term of this form is exactly what we should have expected when writing down Eq. [24]. Applying this equation to a situation where a particle is added to the system, we find that the re-normalization will physically compensate for the expansion of phase space, which leads to a difference in entropy measures. Mathematically, this implies that the “default” measure $p_0(x_{i+1})$ is replaced by the re-normalized measure $p_0(x_{i+1})Z_{st}^* (x_{i+1})/Z_{st} (x_i)$ if we have information about the previous state when calculating the entropy at state $i + 1$.

5.1. Jarzynski’s Equality

Jarzynski equations can be derived in the special case where a stationary distribution (denoted $x \sim ss$) is known during each time-propagation step. The stationary distribution is not required to have a physical temperature and may vary between time-steps. At each step, it can be any single-time distribution that remains a fixed-point of the transition operator. The derivation uses Eq. [27] to re-weight each state between transitions,

$$\langle a(x(T))|ss, T \rangle = \left\langle a(x(T)) \prod_{t=1}^{T} \frac{\mathcal{P}_{ss}(x(t), t)}{\mathcal{P}_{ss}(x(t), t-1)} |x_0 \sim ss, 0 \right\rangle.$$  \hspace{1cm} (31)

The first average represents a single-time average over the steady-state distribution used to propagate the system during the transition $T \rightarrow T + 1$. The second average is a path average over stochastic trajectories beginning in the steady-state at time zero. For the particular choice $a = 1$, the normalization constants cancel between successive $\mathcal{P}_{ss}(x(t), t)$, leaving only the ratio between the end-point partition functions to be expressed as a path average.

The proof is by recursion from the property of the stationary distribution under time-propagation,

$$\mathcal{P}_{ss}(x(t + 1), t) = \int \mathcal{P}_{tr}(x(t + 1)|x(t), t) \mathcal{P}_{ss}(x(t), t) \, dx(t).$$

Because the transition probability, $\mathcal{P}_{tr}(\cdot)$, is not physically constrained, equality of the starting and ending temperatures is not required. Combinations of Eq. [27] can give a very large number of fluctuation theorems.

6. Conclusions

In this article, we have given a generalization of the theory for driven, irreversible processes. Our formulation of maximum transition entropy in Sec. [3] provides several distinct advantages for the derivation of nonequilibrium relations. This theory does not require assuming a Boltzmann-Gibbs form for the starting distribution, nor does it rely on any particular definition of a stationary distribution. Instead, the time-evolution for the system is determined directly from a set of transition probabilities. These are obtained from applying maximum entropy methods, paralleling the canonical form of equilibrium thermostatics (Table [1]). Non-anticipating stochastic trajectories for classical particle and field motion can be cast in this form. The action functional interpretation gives a physical method for defining conserved quantities exchanged between the two during energy conversion processes (Sec. [4]). The energy cost associated with transfers of these quantities from an external environment or experimental apparatus can easily be specialized by an appropriate choice of constraints and computed.

The Lagrangian and Langevin equations are seen seen to form two limits of the theory, with the rate of energy exchange with the bath as the control parameter. Deterministic dynamics are
recovered from the Langevin equation when the deviation of the action functional is strongly constrained to zero (Sec. 3). In this limit, the external forces, which appeared as statistical in the stochastic approach, become the mechanical Lagrange-D’Alembertian forces. Because both limits appear here, and because the Lagrange multipliers can be set to produce constant force or constant flux transitions,\cite{28} the fluctuation-dissipation theorems derived as Gibbs relations from taking second derivatives of Eq. 9 are applicable in the case of both thermal and mechanical driving forces. These equations are completely general because they apply not only arbitrarily far from equilibrium, but also during transient processes that do not possess a steady-state.

A particularly useful aspect of this approach is that it directly connects multiple length and time scales. The formulation of the equations has been in terms of particle motion, but coarse-grained relations are easy to define as appropriate ensemble averages over these motions. Examples of such averages include centers of mass for polymer units or average density and velocity fields. The coarse equations of motion will then lead to polymer coarse-graining models\cite{40} or non-local hydrodynamic models.\cite{41, 42, 43} For the time-evolution of average quantities, we expect the thermodynamic limit argument\cite{44} to apply when the number of averaged degrees of freedom is large, so that the path realized by the system under a given set of constraints will fall arbitrarily close to the maximum entropy solution an overwhelming majority of the time. The present work is therefore a suitable foundation for the theory and analysis of nonequilibrium molecular dynamics.

Our method is free of projection operators and Fourier transforms that can often obscure the physical picture of driven dynamical systems. As a consequence, applications to simplified, standard examples such as circuit theory are easily accomplished.\cite{28} In addition, we find the physical basis of entropy production has become much more apparent. The Joule heating of a resistor, for example, can be seen from Eq. 20 as fundamentally arising from the difference between the velocity added to each ion individually vs. the usable energy in the average ion velocity. Because the energy added to the system in driving the ions is not expressible in terms of the average velocity alone, spreading in the distribution of ion velocities becomes heat. The same remarks follow for driven convective transport, where a spreading in the distribution of forward fluid momentum leads to increases in the local temperature.

Connections of this theory to the formal structure of maximum entropy thermodynamics and Bayesian inference have been elaborated upon in Ref.\cite{28}. These connections allow the definition of thermodynamic cycles expressing differences between driving protocols using the same free energy techniques commonly employed in the equilibrium theory. Some examples have already appeared in the literature for path re-weighting.\cite{45, 46, 47} It is expected that expression in terms of thermodynamic cycles will greatly simplify the derivation and interpretation of these studies.

We have identified a new generalization of the second law for irreversible processes by including the instantaneous information entropy. A traditional analysis shows that the total entropy increase (Eq. 25) is dependent on the details of system dynamics and exchange of conserved quantities with an external system. Connecting this with the fluctuation theorem (Eq. 29) gives a microscopic form for the second law of thermodynamics. The physical device of tracking work performed on individual particles as well as external reversible work sources allows us to track the flow of each type of work (and heat) through the system. Because these changes come directly from the forces on each degree of freedom, this analysis does not depend on arbitrary decompositions of energy functions or definitions of steady-states.

From an informational perspective, entropy increase comes about from discarding information and/or from the information loss associated with coupling to external reservoirs. This is distinguished from the entropy production functional of local equilibrium theory in that the entropy functionals developed here include long-range correlations and are thus not simply functions of locally defined quantities such as composition or temperature.\cite{48} Instead, we have
shown that temperature should not be thought of as a property of the system, but of the external bath. This frees us to define more complicated baths, for example placing constraints on $\partial A/\partial x_a(t)$ only for particles, $a$, in a range of space can be used to model local thermal contacts for imposing thermal or pressure gradients.[49] It should be noted that the analysis in Sec. 4 showed that increasing the number of controllable variables decreases the number of degrees of freedom associated with heat production.

Many questions remain open for attack by this method. Feynman’s path action formulation of quantum mechanics[50] uses the action functional as a rotation operator, whereas the present relations are of the diffusion-type. Is there an appropriate limiting process whereby the latter may be obtained and the multipliers $C$ identified with quantum fluctuations?[51, 52, 53] Ref. [54] gives a Lagrangian related to the Clebsch parametrization for the field equations of Eulerian fluid mechanics. Could application of this derivation method yield novel turbulence models? Addition of electric fields to this picture should result in derivations for the dielectric response in nanoscale fluid mechanics[55, 56]. The present report has not been specialized to cases where near-equilibrium approximations may be applied, and these should also be investigated for the potential of generating new relations in that regime.

Acknowledgment
This work was supported by Sandia’s LDRD program. Sandia National Laboratories is a multi-program laboratory operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-AC04-94AL85000.

References
[1] Kjelstrup S and Bedeaux D 2008 Non-Equilibrium Thermodynamics Of Heterogeneous Systems (Series on Advances in Statistical Mechanics) (World Scientific) ISBN 9812779132
[2] Zwanzig R 1961 Phys. Rev. 124 983–992
[3] Zwanzig R, Nordholm K S J and Mitchell W C 1972 Phys. Rev. A 5 2680–2682 correction to Zwanzig, 1961
[4] Mori H 1965 Prog. Theor. Phys. 33 423–455
[5] Zwanzig R 1965 Ann. Rev. Phys. Chem. 16 67–102
[6] Lebowitz J L 1993 Phys. Today 46 32–38
[7] Jaynes E T 1980 Ann. Rev. Phys. Chem. 31 579–601
[8] Jaynes E T 1985 Complex Systems - Operational Approaches ed Haken H p 254
[9] Jaynes E T 1979 The Maximum Entropy Formalism ed Levine R D and Tribus M (M.I.T Press, Cambridge) p 498 ISBN 0262120801,9780262120807
[10] Kubo R, Toda M and Hashitsume N 2003 Statistical Physics II: Nonequilibrium Statistical Mechanics (Springer) ISBN 354053833X 2nd ed. Translated by M. Toda, R. Kubo, N. Saito, and N. Hashitsume
[11] Maes C 1999 J. Stat. Phys. 95 367–392
[12] Lebowitz J L and Spohn H 1999 J. Stat. Phys. 95(1) 333–365 ISSN 0022-4715
[13] Dewar R 2003 J. Phys. A: Math. Gen. 36 631–641
[14] Crooks G E 1999 Phys. Rev. E 60 2721–2726
[15] Robin W A 1990 J. Phys. A: Math. Gen. 23 2065–2085
[16] Robertson B 1993 Physics and Probability: essays in honor of Edwin T. Jaynes ed Grandy W T (Cambridge Univ. Press) pp 251–260
[17] Luzzi R, Áurea R Vasconcellos and Ramos J G 2002 Predictive statistical mechanics: a nonequilibrium ensemble formalism (Kluwer, Dordrecht) ISBN 1402004826
[18] Stock G, Ghosh K and Dill K A 2008 J. Chem. Phys. 128 194102 ISSN 00219606
[19] Press S, Ghosh K and Dill K A 2011 J. Phys. Chem. B 115 6202–6212
[20] Baiesi M, Maes C and Neton K 2009 Journal of Statistical Physics 135 57–75 ISSN 0022-4715
[21] Filyukov A A and Karpov V Y 1967 J. Eng. Phys. Thermophys. 13(6) 416–419 ISSN 1062-0125
[22] Jou D, Casas-Vázquez J and Lebon G 1988 Rep. Prog. Phys. 51 1105–1179
[23] Jou D, Casas-Vázquez J and Lebon G 1999 Rep. Prog. Phys. 62 1035–1142
[24] Attard P 2006 J. Chem. Phys. 124 024109
[25] Attard P 2009 J. Chem. Phys. 130 194113 (pages 10)
[26] Gill J 2002 Bayesian Methods (Chapman & Hall/CRC) ISBN 1584882883, 978-1584882886
[27] Bobrowski A 2005 Functional Analysis for Probability and Stochastic Processes: An Introduction (Cambridge University Press) ISBN 0521831660
[28] Rogers D M, Beck T L and Rempe S B 2011 J. Stat. Phys. 145 385–409
[29] Marsden J E and West M 2001 Acta Numerica 10 357–514
[30] Noether E 1971 Transp. Theory Stat. Phys. 1 186–207 English translation by M. A. Tavel
[31] Risken H 1996 The Fokker-Planck Equation: methods of solution and applications (Springer) ISBN 354061530X 2nd ed.
[32] Frank T D 2007 Phys. Lett. A 360 552–562 ISSN 0375-9601
[33] Feller S E, Zhang Y, Pastor R W and Brooks B R 1995 J. Chem. Phys. 103 4613–4621
[34] Callen H B 1985 Thermodynamics and an Introduction to Thermostatistics (Wiley, New York) ISBN 0471862568 2nd ed.
[35] Robertson B and Mitchell W C 1971 J. Math. Phys. 12 563–568
[36] Leff H S and Rex A F 2003 Maxwell’s demon 2: entropy, classical and quantum information, computing (IOP Publishing, London) ISBN 0750307595
[37] Chipot C and Pohorille A (eds) 2007 Free Energy Calculations (Springer) ISBN 3540384472
[38] Bennett C H 1976 J. Comput. Phys. 22 245–268 ISSN 0021-9991
[39] Jarzynski C 1997 Phys. Rev. Lett. 78 2690–2693
[40] Akkermans R L C and Briels W J 2000 J. Chem. Phys. 113 6409–6422
[41] Grandy W T 1987 Foundations of Statistical Mechanics (Kluwer, Boston) ISBN 902772489X, 9789027724892
[42] Kubo R 1966 Rep. Prog. Phys. 29 255
[43] Lehoucq R and Von Lilienfeld-Toal A 2010 J. Elasticity 100(1) 5–24 ISSN 0374-3535
[44] Jaynes E T 1957 Phys. Rev. 108 171–190
[45] Zuckerman D M and Woolf T B 1999 J. Chem. Phys. 111 9475–9484
[46] Minh D D L and Chodera J D 2009 The Journal of Chemical Physics 131 134110
[47] Chodera J D, Swope W C, Noe F, Prinz J, Shirts M R and Pande V S 2011 J. Chem. Phys. 134 244107–14
[48] Jaynes E T 1992 Maximum Entropy and Bayesian Methods ed Smith C R, Erickson G J and Neudorfer P O (Kluwer, Dordrecht) pp 1–22 ISBN 978-0792320319
[49] Zhu F, Tajkhorshid E and Schulten K 2004 Biophys. J. 86 50–57 ISSN 0006-3495
[50] Dyson F 2007 Advanced Quantum Mechanics (World Scientific) ISBN 9812706224 transcribed from 1951 Lecture Notes by David Derbes

[51] Christopeit N, Helmes K, Kohlmann M and Yasue K 1986 Stochastic Differential Systems (Lecture Notes in Control and Information Sciences vol 78) (Springer Berlin/Heidelberg) pp 352–361

[52] Davies I M 1989 J. Phys. A: Math. Gen. 22 3199–3203 ISSN 0305-4470

[53] Dammak H, Chalopin Y, Laroche M, Hayoun M and Greffet J J 2009 Phys. Rev. Lett. 103 190601

[54] Mobbs S D 1982 Proc. Royal Soc. London. A. Math. Phys. Sci. 381 457–468

[55] Hubbard J and Onsager L 1977 The Journal of Chemical Physics 67 4850–4857

[56] Ballenegger V, Blaak R and Hansen J P 2006 Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 2 (Lecture Notes in Physics vol 704) ed Ferrario M, Ciccotti G and Binder K (Springer Berlin / Heidelberg) pp 45–63