Stony Brook University

Non-Equilibrium Physics from an Inference Perspective

A Thesis presented

by

Valentin Walther

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Master of Arts

in

Physics

Stony Brook University

August 2014

Stony Brook University

The Graduate School

Valentin Walther

We, the thesis committee for the above candidate for the

Master of Arts degree, hereby recommend

acceptance of this thesis

Ken A. Dill
Professor, Department of Physics and Astronomy, Department of Chemistry

Eden Figueroa
Professor, Department of Physics and Astronomy

Sergei Maslov
Professor, Department of Physics and Astronomy, Brookhaven National Laboratory

This thesis is accepted by the Graduate School

Charles Taber
Dean of the Graduate School

Abstract of the Thesis

Non-Equilibrium Physics from an Inference Perspective

by

Valentin Walther

Master of Arts

in

Physics

Stony Brook University

2014

This thesis describes implications of inference logic and information theory on non-equilibrium physics. Inference logic prescribes a unique way of converting information about mean-values into probability assignments, which are applied to the physics of non-equilibrium systems. This thesis is structured in four chapters:

First, the central theorems of information theory are derived from the basic requirements of any inference procedure. The theory, formulated to apply to physical equilibrium systems with mean-value constraints, is called Maximum Entropy (MaxEnt). This idea bridges the work of Shannon and Boltzmann and sets difficult concepts from statistical physics, such as ergodicity and metric transitivity, into a new context.

Second, MaxEnt is extended to general time-dependent systems, giving a concept named Maximum Caliber (MaxCal). In particular, we develop a theoretical description of steady-state systems in the limit of small driving forces. We recover two important results from linear thermodynamics, namely Onsager's reciprocal relations and Prigogine's theorem. This result is important as it does not rely on the problematic assumption of local equilibrium. Furthermore, higher-order symmetries are revealed, leading to pre-

dictions beyond the linear regime. We show that the same expansion results in corrections to Prigogine's theorem. Both effects are exemplified in a toy model. In this example, we find bifurcations which suggest a spontaneous change in behavior as the system leaves the near-equilibrium regime. These "phase transitions" are reminiscent of physical effects, such as the transition from conductive to convective heat transfer in viscous fluids.

Third, we review the idea and application of Maximum Entropy Production (Max-EP), a theory supposed to explain large-scale organization for complex systems far from equilibrium. We present the theory's success in climate prediction, commenting on the underlying assumptions and problems in the application. Next, we discuss Dewar's hypothesis that MaxEnt and Max-EP could be equivalent. A detailed analysis of Dewar's attempts at proving his conjecture shows the connections to be evasive. We conclude that Max-EP is loaded with ambiguity and that a convincing derivation from inference science has not yet been brought forward.

Fourth, we move beyond the assumption of a steady state in considering a diffusion system, which is arbitrarily far from equilibrium. We simulate the diffusion system (in one dimension) and produce statistical distributions over possible outcomes. Aiming at reproduction of the simulated results, we develop a coarse-grained mathematical model, which can be analyzed using MaxCal. We exploit symmetries to find a computationally efficient representation in terms of a transfer matrix. Our MaxCal analysis reveals a physical constraint which was previously unconsidered. After the implementation of this constraint, MaxCal predicts the simulated distributions at good accuracy. We conclude this chapter by pointing out how MaxCal provides a unique means of including new physical effects.

In summary, the results of this thesis demonstrate the ideas of inference logic to be very useful in physics. Especially our findings about linear thermodynamic steady states pose a novel approach to traditional thermodynamic theorems. While neither MaxEnt nor MaxCal has, by itself, any physical content, they can be used to great effect for the analysis of physical observations as well as in model building. In this sense, MaxEnt and MaxCal provide a rigorous theoretical framework of converting observations into probability assignments and show a remarkable potential for the approach to fundamental questions in statistical physics from an inference perspective.

Contents

List of Figures

Acknowledgments

Foremost, I would like to express my gratitude to my advisor Professor Ken A. Dill for the assignment of this exciting topic and his guidance during my work on the thesis. Ken Dill took care that my time in his group was not just intellectually rewarding, but also introduced me to the many members of his group. Out of all the scientists I had the pleasure to meet at the Laufer Center, I am particularly grateful to have had many inspiring discussions with Michael Hazoglou and Purushottam Dixit.

My entire stay at Stony Brook would not have been possible without the generous support through a Fulbright annual stipend. Additional thanks go to Ken Dill for providing financial support throughout the summer.

Besides my advisor, I thank professors Sergei Maslov and Eden Figueroa who went to the trouble of reading my thesis and of being part of my thesis committee.

Last but not least, I would like to thank Kathrin Kelly for checking the final document for orthographic correctness.

1 Introduction: Inference Theory in Physics

1.1 Inference Problems

One of the fundamental and general problems of quantitative science is the conversion of experimental data into theoretical models. Even the most trivial examples like the rolling of a die pose a conceptual problem: If one assumes a model in which each of the die's sides occurs at probability $p = 1/6$, calculation becomes easy and a precise mathematical theorem about the probability of m successful events in N trials is given by Laplace (**Weak Law of Large Numbers**)

$$P(m|N,p) = \binom{N}{m} p^m (1-p)^{N-m} \tag{1}$$

$$\lim_{N\to\infty} P(p-\epsilon < f = m/N < p+\epsilon) = 1 \quad \forall \epsilon > 0, \tag{2}$$

where f is the observed frequency of successful events. While this question is mathematically easily tractable, a scientific problem usually requires the inverse: Some experimental data is available, and a theoretical model is sought. To start calculation, one needs to make a **prior assumption**. In the case of a die this might be considered trivial: If the probability assignments were not equal, a simple permutation of the sides would give a different probability assignment and thus a contradiction. If, however, not all outcomes are by symmetry equally likely we must still arrive at some assignment of probabilities. Assume, for example, a loaded die, whose known average value is $\bar{x} = 2.9$. Fig. 1 shows some alternative distributions, all of which are compatible with the mean value constraint. Which distribution should be chosen? A related question arises in the case of continuous variables. If no information is available on a measurement, analogy with the discrete case might suggest assigning a uniform prior to all outcomes. But continuous variables pose an additional challenge. A uniform prior on a variable α is non-uniform for equivalent reparametrizations, such as α^5 or $\ln(\alpha)$. Why should the first line in Eq. 3 be preferred over the last two?

$$\begin{aligned} dP/d\alpha &= \text{const.} \\ dP(\alpha^5)/d\alpha &= 5\alpha^4 \cdot dP/d\alpha = \text{const.} \\ dP(\ln\alpha)/d\alpha &= 1/\alpha \cdot dP/d\alpha = \text{const.} \end{aligned} \tag{3}$$

In either case, a rigorous mathematical criterion for the probability assignment is required. The key goal must be that **the assigned probability dis-**

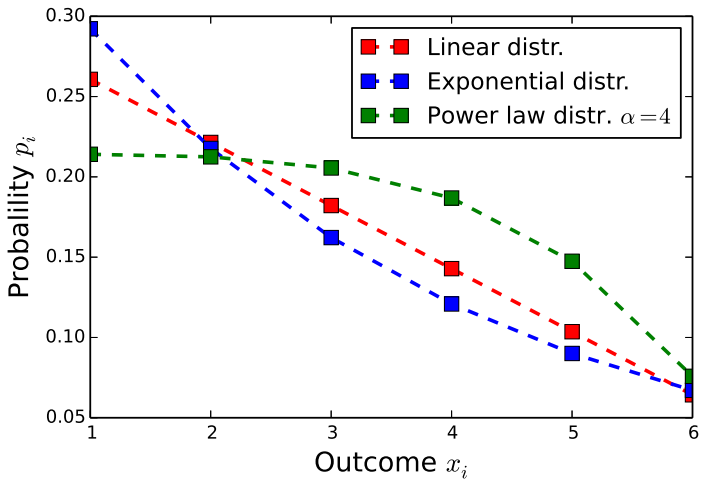


Figure 1: Several discrete probability distributions compatible with $\bar{x} = 2.9$.

tribution represents the information provided by the constraints[1]**, no more, no less**, and that is does so in a mathematically well-defined (unique) way. In the following, this theorem is developed by demonstrating how it naturally emerges from physics and information science. Its great simplicity and explanatory success are illustrated and possible predictive failures of the theorem are interpreted. The chapter concludes with a sketch of a formal proof based on axiomatic requirements to any inference procedure.

1.2 Shannon's Coding Theorem

A key advance in the scientific progress of approaching a unique inference algorithm was made outside physics, by Shannon [49]. He analyzes the potential for data compression when sending messages over physical "channels" and finds that it has a definite upper bound valid for any lossless compression algorithm. In other words, there is a minimum (average) length beyond which loss of information is bound to occur. To evaluate just where this point lies, Shannon uses the following simplified model. A message is a sequence of

[1] Any constraint used in this formalism must be quantitative. Even though complicated constraints in the form of inequalities or functional relations are in theory admissible, the constraints used in this thesis will be (mostly) on mean values.

symbols X_i, all of which are taken from an "alphabet" $A = (a_1, ..., a_N)$. Next, Shannon assumes that the probabilities[2] of each letter are known $(p_1, ..., p_N)$ and supposes that the symbols are statistically uncorrelated. The intuitive requirement that the occurrence of two specific letters should increase "information" linearly while the probabilities multiply leads to

$$I(p_i \cdot p_j) \stackrel{!}{=} I(p_1) + I(p_2) \tag{4}$$

$$\Rightarrow I(p_i) \propto \ln p_i \tag{5}$$

where I **denotes information**. To ensure that information increases with decreasing probabilities, a negative sign is adopted. In the binary world of computers, the typical choice of logarithm is $\log_2$, but to emphasize the relation to the physical form of entropy $\ln = \log_e$ seems a good option here. Averaging over all symbols in the alphabet gives the **Shannon entropy**

$$H(A) \equiv -\sum_{i=1}^{N} p_i \ln p_i. \tag{6}$$

I is an intuitive measure of information for a random source and thus H **can be interpreted as the average information content per symbol.** Since no lossless compression may alter the "information content" of a message, the shortest conceivable compression of a message[3] is $N \cdot H(A)$ symbols long[4]. Shannon entropy assumes its maximum if all N outcomes symbols have the same probability $p = 1/N$. For the binary case, this is illustrated in Fig. 2. A second desirable feature is **additivity**: Coarse-graining a prob-

[2]Actually, only the relative frequency of any letter in an alphabet is known. The identification of relative frequencies with probabilities is not always accurate (Cf. subjective and frequentist interpretation of probabilities).

[3]The messages considered here are infinitely long.

[4]Shorter compressions with prefixes are not considered.

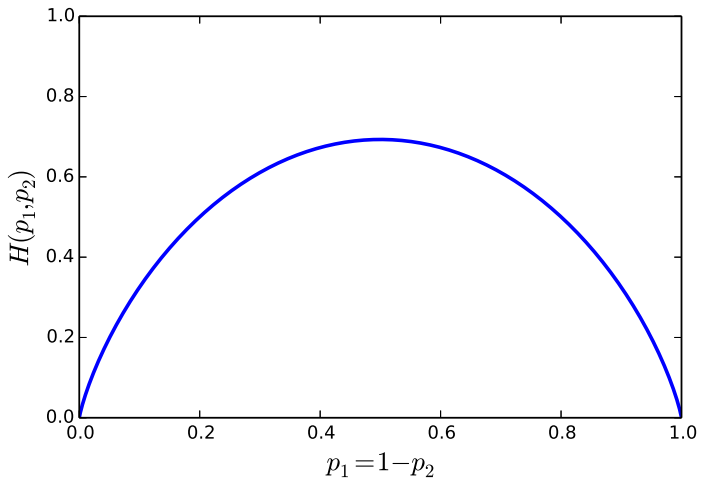


Figure 2: Information entropy for a binary process has a maximum at $p_1 = p_2 = 1/2$.

ability distribution does not change its information content (Fig. 3).

$$
\begin{aligned}
&-\sum_{j=1}^{\alpha} q_j \ln q_q + \sum_{j=1}^{\alpha} q_j H\left(\frac{p_{m_{j-1}+1}}{q_1}, ..., \frac{p_{m_j}}{q_j}\right) \\
&= \sum_{j=1}^{\alpha} -q_j \ln q_i - q_i \left[\frac{p_{m_{j-1}+1}}{q_1} \ln \frac{p_{m_{j-1}+1}}{q_1} + ... + \frac{p_{m_j}}{q_j} \ln \frac{p_{m_j}}{q_j}\right] \\
&= \sum_{j=1}^{\alpha} -q_j \ln q_i + \underbrace{\left(p_{m_{j-1}+1} + ... + p_{m_j}\right)}_{q_j} \ln q_j \\
&\qquad - \left[p_{m_{j-1}+1} \ln p_{m_{j-1}+1} + ... + p_{m_j} \ln p_{m_j}\right] \\
&= H(p_1, ..., p_{m_\alpha})
\end{aligned}
\tag{7}
$$

1.3 Equilibrium Statistical Physics: Boltzmann

The earliest problem of statistical physics was to describe the behavior of gases used in the heat engines of the late 18th century. As these gases only have very weak intermolecular forces, they can be approximated as ideal

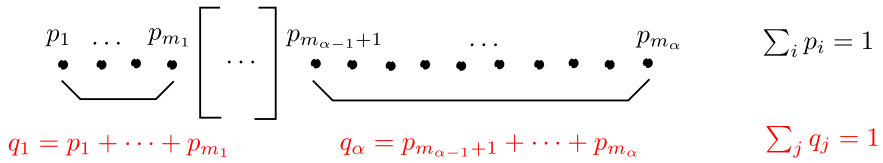


Figure 3: Entropy is additive, i.e. independent of coarse-graining in the probability distribution.

gases, i.e. gases interacting only via elastic collisions. Combined with the idea that gases consist of a great number of indivisible parts, Newton's laws in principle provide the means to evaluate each particle's trajectory for all times as functions of the initial position and velocity.

The number of atoms involved in a typical experiment being of the order $10^{23} - 10^{25}$, direct integration of the equations of motion is unfeasible. Boltzmann was first to consider the problem from an entirely different angle. Using a statistical approach, he triumphantly calculates many important features of ideal gases [2]: In his derivation, he exploits that total energy and particle number in an isolated gas are conserved, without using Newton's equations of motion.

The total energy of N non-interacting[5] particles in an external conservative potential Φ is

$$E = \sum_i^N \left[\frac{1}{2} m_i v_i^2 + \Phi(x_i) \right]. \tag{8}$$

Every particle can be characterized by a vector in position space X and one in momentum space Y, whose product is phase space $R = X \otimes Y$. The system's trajectory in phase space is, however, limited to the subspace in R^N satisfying Eq. 8. To count the number of available states, one can divide phase space into discrete cells of volume h^3 [6]. The boxes are assumed to be sufficiently large to accommodate a great number particles (still $N_k < M_R$), while at the same time being small enough such that the system's energy within a cell is essentially constant. It is then possible to enumerate all cells

[5] i.e. no interaction besides elastic collisions

[6] The uncertainty principle from quantum mechanics provides that h is Planck's constant. In any case, the absolute size of the cells drops out of the final expression for entropy (cf. below).

$R_k, k \in \{1, ..., M_R\}$ in accessible phase space. There is a vast number of possible configurations, but not all sets of occupation numbers are equally likely. For a particular set of $\{N_k\}$ the multiplicity is given as

$$W = \binom{M_R}{N_1, ..., N_{M_R}} = \frac{M_R!}{N_1! \cdots N_{M_R}!} \tag{9}$$

because **particles in one box are indistinguishable** (cf. footnote 6). As M_R and N are macroscopic quantities, W is a very large number, so it makes sense to use logarithms and the Stirling approximation $N! \approx N^N e^{-N} \sqrt{2\pi N}$

$$\begin{aligned}
\ln W &= \ln \left(\frac{M_R!}{N_1! \cdots N_{M_R}!} \right) \\
&\approx \ln \left(M_R^{M_R} e^{-M_R} \sqrt{2\pi M_R} \prod_{k=1}^{M_R} \frac{N_k^{-N_k} e^{N_k}}{\sqrt{2\pi N_k}} \right) \\
&= M_R (\ln M_R - 1) - \sum_k N_k (\ln N_k - 1) + \frac{1}{2} \left(\ln M_R - \sum_k \ln N_k \right) + c \\
&\approx - \sum_k N_k \ln \left(\frac{N_k}{M_R} \right) = -M_R \sum_k \frac{N_k}{M_R} \ln \left(\frac{N_k}{M_R} \right)
\end{aligned} \tag{10}$$

In the second to last equation, only the excellent approximation $N \gg \ln N \gg 1$ for large N was used. In the last expression, one can identify $p_k = \frac{N_k}{M_R}$ as the probability of one particle being in phase space cell R_k. At this point, it also becomes clear that the actual choice of cells is immaterial, since Eq. 10 has the same form as information entropy (Eq. 6) and therefore also satisfies the additivity property. If phase space was re-divided as in Fig. 3, H would remain invariant.

Assuming ergodicity[7], Boltzmann predicts all physical properties to be dominated by the most probable states. The most probable set $\{N_k\}$ occurs when $p_k = p = \frac{1}{M_R}\ \forall i$ because this set maximizes the multiplicity W (cf. Fig. 2). A slightly different situation arises when the system is in contact with a heat bath and only its average energy is conserved. Still, Boltzmann reasons that **the most probable set of occupation numbers $\{N_k\}$ should be**

[7] i.e. one assumes that the system's trajectory in phase space maps out every possible state. Time averages as occurring in a measurement process are then equivalent to statistical averages.

realized since this set outnumbers every other distribution by a very large margin. Mathematically, the distribution can be found with the help of Lagrange multipliers

$$\frac{\delta}{\delta p_k}\left[\ln W - \beta \sum_k E_k p_k\right] \stackrel{!}{=} 0 \tag{11}$$

$$\Rightarrow p_k = \frac{1}{Z}\exp(-\beta E_k), \tag{12}$$

where $Z = \sum_k \exp(\beta E_k)$ is the partition function. The solution of β is determined from

$$E = \sum_k N_k E_k \tag{13}$$

with $N_k = N \cdot p_k$. From this remarkable result many familiar results can be read-off. For example, the earth's gravitational field $E_k = mgz$ directly gives rise to the barometric formula

$$\rho(z) = \rho(0)\exp(-\beta mgz), \tag{14}$$

where $\rho(z)$ is the air density at height z. In this case, thermodynamic equilibrium also requires for the pressure $P = Nk_bT/V = k_bT/m \cdot \rho$

$$-\frac{dP}{dz} = -k_bT/m \cdot \rho'(z) \stackrel{!}{=} g\rho(z) \tag{15}$$

$$\Rightarrow \rho(z) = \rho(0)\exp(-\frac{mgz}{k_bT}). \tag{16}$$

The identification $\beta = \frac{1}{k_bT}$ is immediately obvious. But the distribution even grants indirect statements about dynamics. Because Eq. 12 factorizes into one part containing only space and one containing only momentum components, one sums over one space to find distributions in the marginal space of the other. For example, the distribution of velocities is

$$p(v) = \sum_x p(v,x) \propto \exp(-\beta m v^2/2). \tag{17}$$

It is interesting to see that Eq. 17 (Maxwell distribution) holds irrespective of position. Thus, one can conclude that in any conservative field the Maxwell distribution is maintained automatically without the need for collisions.

1.4 Maximum Entropy (MaxEnt)

1.4.1 Analysis

Poincaré points out that with Boltzmann's method one seems to "get something for nothing". In fact, there is an intuitive gap between the assumptions made and the scope of the results obtained. After all, Boltzmann employs a two-step process to arrive at his predictions:

First, he develops a model for the system at consideration. In the case of the ideal gas, this description includes splitting phase space into elementary cells, followed by a simple counting argument to determine the states with the greatest multiplicities. In a second step, he determines the distribution that maximizes the total number of available states based on the imposed energy constraint. Under the assumption that all states are equally populated, this allows all mentioned results. The derivation shows that Boltzmann's result is very stable against changes of the model. On the one hand, the specific choice of cells in phase space is irrelevant and the assumed size h drops out of the calculation. On the other hand, the used model of fixed cells in phase space implies that phase space available to the system does not change with time. This is the content of Liouville's theorem, which holds for all Hamiltonian systems, and it implements a great deal of the physical nature of Hamiltonian dynamics. Boltzmann's final results are correct and have stood the test of experiment. It is interesting to note, however, that Boltzmann counts the particles in each box as being indistinguishable. While a central postulate of quantum mechanics, Boltzmann must have discovered this a requirement necessary to produce correct results[8]. Boltzmann distilled the central physical content of the gas problem (energy constraint, Liouville dynamics) into his model and inferred an unknown physical constraint, namely the indistinguishability of particles in a small box.

1.4.2 MaxEnt: An Inference Method

Boltzmann's approach to the gas problem is borne from the overwhelming complexity of an enormous many-body system. In practical terms, accurate measurement of the initial conditions is just as impossible as integrating all trajectories. However, it is worth noting that this plethora of microscopic possibilities is equaled by the redundancy in the outcomes. Though predic-

[8]In the case of distinguishable particles the factors $N_k!$ would not appear in Eq. 10

tions concerning the system's microstate are virtually impossible, statements about its macroscopic state are very simple. **Macroscopic reproducibility suggests that a handful of measured macroscopic quantities must be sufficient for a statistical treatment** of the system. Microscopic details evade control in experimentation and can therefore not be relevant to a reproducible result. Jaynes [23, 24, 25] points out that eliminating redundant information at the outset might be one key to overcoming the complexity barrier and revealing the physics underlying some problem. In the following, Boltzmann's method is generalized by using its mathematical relation with Shannon entropy to reach a general inference algorithm.

The first step of Jaynes' algorithm is devising a model. Next, one has to implement all physical constraints known about the system. In a typical situation, the number of constraints will be far smaller than the $(N-1)$ degrees of freedom in the distribution (cf. die example in Sect. 1.1), such that a vast number of distributions would be possible. While all of them satisfy the constraints, some suggest tendencies which are not warranted by the constraints. The idea is to select the distribution which contains no information beyond that enforced by the constraints. At this point, a connection to information entropy becomes apparent. If one identifies the different outcomes of an experiment as the letters of an alphabet, information entropy provides a means to quantify the information contained in probability distribution. In other words, it becomes possible to measure just **how much bias is contained in any distribution** satisfying the constraints. The resulting information-based algorithm can be summarized as follows[9]:

Theorem 1 (*Maximum Entropy (MaxEnt)*)

> *MaxEnt selects the probability distribution that is compatible with all given constraints and maximizes Shannon's entropy. The resulting distribution is "the least biased" given all constraints because it contains the least information.*

Entropy is a positive quantity defined for all possible distributions, which has an upper bound $H_{max} = -\sum_i p_i^* \ln p_i^* = \ln N$. A solution is, therefore, guaranteed to exist on the set of all compatible distributions. In the counting approximations of Eq. 10 and also implicitly contained in Shannon's derivation is the assumption of an infinite number of states and infinitely long

[9]For a derivation of thermodynamic ensembles from MaxEnt cf. [23, 12]. In MaxEnt, strong assumptions such as ergodicity and metric transitivity are unnecessary.

messages, respectively. Because Boltzmann's search for the most probable state only coincides with the MaxEnt result in the limit of large numbers, Sect. 1.5 shows that **MaxEnt is the uniquely correct way of inferring a probability distribution, regardless of the number of states**.

1.5 MaxEnt is a Unique Solution to the Inference Problem

The MaxEnt principle as described above has been criticized by prominent physicists such as Uhlenbeck [25]. The key objection raised against the principle is that it makes predictions that are based on some "state of knowledge" or "ignorance" about the problem at hand. Because different observers may have different information, they are bound to arrive at different predictions based on MaxEnt. Since any system can only be in one state, MaxEnt was dismissed by some as a non-physical theory. The theoretical connection between information and physics is not an entirely established subject. Century-old paradoxes such as Maxwell's Demon are still object of some controversy [34, 13, 14]. Concepts like Landauer's Principle [31] suggest a direct relation between information entropy and thermodynamic entropy that is, for example, supposed to underlie all processes in a computer. Experiments are not yet at the state of settling the issue conclusively.

Regardless of whether or not Landauer's Principle is true, I believe that MaxEnt should be regarded as a sensible tool in physics. On the one hand, physical predictions are, in fact, sometimes dependent on the amount of ignorance an observer has. In the above example of the ideal gas there clearly is a difference between knowing all positions and momenta at one time and knowing only macroscopic (thermodynamic) observables. Concepts like thermodynamic entropy arise from incomplete knowledge. It must be noted, however, that thermodynamic entropy (and all other quantities derivable from it) are still well-defined in a MaxEnt framework. **Thermodynamic entropy is the system's information entropy if only the well-defined thermodynamic constants are known.** Because these macroscopic values are not subjective, the MaxEnt formalism of thermodynamic equilibrium is not, either. On the other hand, **MaxEnt provides the mathematically defined best resulting posterior distribution based on mean-value constraints.** Information entropy can be derived from desired properties of a measure of information [21]. Considering the fundamental role MaxEnt plays

in the results of this thesis, it is worthwhile to review the axiomatic proof from inference principles, reformulated for the considered discrete processes.

1.5.1 Inference Axioms

Consider an experiment with N possible outcome states $\mathbf{x} \in \mathbf{D}$. There exists a space $\mathfrak{D}$ containing all possible probability distributions $\mathbf{q}$, including the "true" distribution $\mathbf{q}^*$. Now, a new piece of information I is added that can reduce the space of compatible distributions to $\mathfrak{R} \subseteq \mathfrak{D}$. I is allowed to take the form of mean value constraints

$$\sum_i \mathbf{q}_\mathbf{i}^* \tilde{a}_{ik} = \bar{a}_k \Leftrightarrow \sum_i \mathbf{q}_\mathbf{i}^* a_{ik} = 0 \quad \text{or} \tag{18}$$

$$\sum_i \mathbf{q}_\mathbf{i}^* \tilde{c}_{ik} \geq \bar{c}_k \Leftrightarrow \sum_i \mathbf{q}_\mathbf{i}^* c_{ik} \geq 0 \tag{19}$$

The inference procedure leading to a "best" distribution (**posterior**) should satisfy a set of axioms suggested in [50].

1. **Uniqueness:** The posterior is unique: $q = (\circ I)$, where $\circ$ denotes the operator incorporating a new piece of information.

2. **Permutation Invariance:** The choice of any particular coordinate system should not alter the result. If $\mathbf{\Gamma}$ is the transformation of coordinates, then $\circ(\mathbf{\Gamma} I) = \mathbf{\Gamma}(\circ I)$.

3. **System Invariance:** Two pieces of information about separate systems should influence probability assignments on the systems separately: $(\circ(I_1 \wedge I_2)) = (\circ I_1)(\circ I_2)$.

4. **Subset Independence:** If the experiment decomposes into disjoint subsets $\mathbf{S_i}$ of outcomes whose union is $\mathbf{D}$, information M that gives the summed probabilities in each subset must satisfy: $(\circ(I_i \wedge ... \wedge I_n \wedge M)) * \mathbf{S_i} = (\circ I_i)$. $*$ denotes the conditional probability on a subspace $\mathbf{S_i}$: $q(\mathbf{x}|\mathbf{x} \in \mathbf{S_i}) = q(\mathbf{x}) / \int_{\mathbf{S_i}} d\mathbf{x}' q(\mathbf{x}')$

H defines the correct probability distribution via

$$H(\mathbf{q}) = \max_{\mathbf{q}' \in \mathfrak{D}} H(\mathbf{q}'). \tag{20}$$

The following proof shows that these axioms determine the form of the entropy functional $H(\mathbf{q})$ entirely.

1.5.2 Sketch of the Proof

Using the axioms, Shore and Johnson [50] prove that the function $H(\mathbf{q})$ can essentially be expressed as[10]

$$H(\mathbf{q}) = \sum_j f(q_j) \tag{21}$$

where f is an arbitrary function. This is already a major reduction in the infinite set of conceivable functional forms. It indicates that the weight of each outcome enters entropy additively. For simplicity, consider a continuous random variable, denoted q. The functional form of Eq. 21 can still be narrowed down. If one imposes the constraint $\int_{\mathbf{D}} d\mathbf{x} q^*(\mathbf{x}) a(\mathbf{x}) = 0$, variational calculus gives

$$\int f(q(\mathbf{x}))d\mathbf{x} + \lambda \left(\int q(\mathbf{x})d\mathbf{x} - 1 \right) + \alpha \left(\int q(\mathbf{x})a(\mathbf{x})d\mathbf{x} - 0 \right) = 0 \tag{22}$$

$$\Rightarrow \lambda + \alpha a(\mathbf{x}) + g(q(\mathbf{x})) = 0, \tag{23}$$

where $g \equiv \frac{\partial f(b)}{\partial b}$. Using Axiom 2, one considers a general coordinate transformation $\Gamma : \mathbf{x} \to \mathbf{y}$. In the new coordinate system (whose quantities are primed) one finds the analogue of Eq. 23

$$\lambda' + \alpha' a'(\mathbf{y}) + g(q'(\mathbf{y})) = 0 \tag{24}$$

However, Axiom 2 requires that probabilities of events are invariant with the change of coordinates: $q'(\mathbf{y}) = \mathbf{J^{-1}} q(\mathbf{x})$.[11]

$$\lambda' + \alpha' a(\mathbf{x}) + g(\mathbf{J^{-1}} q(\mathbf{x})) = 0 \tag{25}$$

$$\Rightarrow g(\mathbf{J^{-1}} q(\mathbf{x})) = g(q(\mathbf{x})) + (\alpha - \alpha')a(\mathbf{x}) + \lambda - \lambda' \tag{26}$$

Consider the special case where it is possible decompose the space of outcomes into a set of disjoint subspaces $\{\mathbf{S_i}\}$ whose union is $\mathbf{D}$, such that that the constraints are constant on each subset, $q(\mathbf{x}) = \text{const.}\ \forall \mathbf{x} \in \mathbf{S_i}$. So far,

[10] Shore and Johnson point out that H is only "equivalent" to the form given in Eq. 21, in the sense that any other function H^* can satisfy Eq. 20 if and only if H is also a solution to Eq. 20.

[11] This essentially means that probabilities are conserved $p = q(\mathbf{x})d\mathbf{x} \stackrel{!}{=} q(\mathbf{y})d\mathbf{y} = p$, $d\mathbf{y} = \mathbf{J}d\mathbf{x}$. For the same reason $a'(\mathbf{y}) = a(\mathbf{x})$.

only the resulting distribution q was considered since MaxEnt always starts with uniform priors. A arbitrary coordinate transformation, however, makes the prior non-uniform. Letting u denote the uniform prior, which must be included in $g(q) \to g(q,u)$ as a second argument, one obtains in place of Eq. 26

$$g(\mathbf{J^{-1}}q(\mathbf{x}), \mathbf{J^{-1}}u(\mathbf{x})) = g(q(\mathbf{x})) + (\alpha - \alpha')a(\mathbf{x}) + \lambda - \lambda' \tag{27}$$

This case lets the right-hand side remain constant, while the left-hand side can vary arbitrarily. The necessity to satisfy Eq. 27 enforces

$$g(b,c) = g(b/c), \tag{28}$$

whose general solution is found from integration

$$f(a,b) = a \cdot h(a/b) + v(b) \tag{29}$$

with arbitrary functions h and v. The variable b accounts for the prior and can be set to unity for the uninformed case considered in MaxEnt, leaving

$$H(q) = \int q(\mathbf{x})h(q(\mathbf{x}))d\mathbf{x}. \tag{30}$$

$$\Rightarrow H(\mathbf{q}) = \sum_i q_i h(q_i) \tag{31}$$

The final major step is to exploit Axiom 3: Two independent systems can be described separately or by a joint distribution. The outcomes should not differ. Imposing constraints separately, one obtains for systems $i \in \{1,2\}$

$$\lambda_1 + \alpha a_i + u(q_i) = 0 \tag{32}$$

$$\lambda_2 + \beta b_i + u(r_k) = 0 \tag{33}$$

where $u(q) \equiv h(q) + q\frac{d}{dq}h(q)$. Describing the systems as one large ensemble gives

$$\lambda' + \alpha' a_i + \beta' b_k + u(q_i r_k) = 0. \tag{34}$$

In Eq. 34, Axiom 3 provides that the last argument factors because of independence $u(q_i, r_k) = u(q_i \cdot r_k)$. In the discrete case, subtracting Eq. 34 from Eq. 32 gives

$$u(q_i r_k) = u(q_i) + u(r_k) + (\alpha - \alpha')a_i + (\beta - \beta')b_k + \lambda_1 + \lambda_2 - \lambda', \tag{35}$$

where $u(x) = f'(x)$. The remaining proof is mathematical manipulation to reveal the final form of H. First, observe that the difference of two values of the function u depends only on the uncommon factors of the arguments

$$\begin{aligned} u(q_i r_k) - u(q_i r_\nu) &= u(q_i) + u(r_k) + (\alpha - \alpha')a_i + (\beta - \beta')b_k + \lambda_1 + \lambda_2 - \lambda' \\ &\quad - [u(q_i) + u(r_\nu) + (\alpha - \alpha')a_i + (\beta - \beta')b_\nu + \lambda_1 + \lambda_2 - \lambda'] \\ &= u(q_u) + u(r_k) + (\alpha - \alpha')a_u + (\beta - \beta')b_k + \lambda_1 + \lambda_2 - \lambda' \\ &\quad - [u(q_u) + u(r_\nu) + (\alpha - \alpha')a_u + (\beta - \beta')b_\nu + \lambda_1 + \lambda_2 - \lambda'] \\ &= u(q_u r_k) - u(q_u r_\nu) \\ &= G(r_k, r_\nu) \end{aligned} \tag{36}$$

From Eq. 36 one can immediately read off that G must satisfy

$$G(x, y) = s(x) - s(y) \tag{37}$$

for an arbitrary functions s. Next, one observes

$$\begin{aligned} G(x,z) - G(w,y) &= u(\boxed{w} \cdot x) - u(\boxed{w} \cdot z) \\ &\quad - [u(\boxed{x} \cdot w) - u(\boxed{x} \cdot y)] = u(xy) - u(wz) \\ \Leftrightarrow u(wz) + s(x) - s(z) &= u(xy) + s(w) - s(y) \\ \Leftrightarrow u(wz) - s(w) - s(z) &= u(xy) - s(x) - s(y) \end{aligned} \tag{38}$$

The right-hand side only depends on x, y, the left-hand side only on w, z. Both sides must therefore equal a constant, proving that

$$u(xy) = g(x) + g(y). \tag{39}$$

The general solution to this equation is [54]

$$f'(x) = u(x) = A \ln(x) + B \tag{40}$$

$$\Rightarrow f(x) = Ax \ln(x) + Bx + C. \tag{41}$$

Evaluating the functional H from Eq. 21 at the N possible outcomes gives

$$H(\mathbf{q}) = A \sum_i^N q_i \ln(q_i) - N \cdot A + B. \tag{42}$$

Maximizing this function requires choosing a negative sign since $q_i < 0$. The prefactor's magnitude is irrelevant.

$$H(\mathbf{q}) = -\sum_{i=1}^{N} q_i \ln(q_i). \tag{43}$$

This concludes the proof that maximizing $H(\mathbf{q})$ is the uniquely correct way of making inferences about a prior based on the average constraints used in this thesis [12] [13].

1.6 What it Means When MaxEnt is Wrong

In summary, the above proof shows that MaxEnt is the uniquely defined principle that can be used to infer probability distributions from mean value constraints. Reminiscent of Boltzmann's solution of the gas problem, the algorithm consists of building a model and assigning a probability distribution to all states based on information entropy. The proof demonstrates that MaxEnt gives the "best" distribution based on both the model and the constraints used. MaxEnt does not, however, provide a means of predicting new physics. By itself, MaxEnt contains no physical information. This implies that MaxEnt cannot be falsified and for this reason does not meet the criteria of a **scientific theory** [25, 9]. In any case, MaxEnt is useful in two distinct ways:

1. If its predictions are correct, there is strong indication that the used constraints represent the "essential physics".

2. If its predictions are wrong, it is sure that the system's physical behavior is not exhaustively described by the used constrains and/or enumeration of states.

Thus, MaxEnt provides a unique means of evaluating the validity of physical assumptions one believes to lie at the heart of some problem. This thesis will use MaxEnt and generalizations thereof to test the statistical content of well-known results from non-equilibrium physics.

[12] So far, it has been shown that any functional other that $H(\mathbf{q})$ is incompatible with the axioms. Technically, it remains to show that $H(\mathbf{q})$ actually satisfies them, which is omitted for brevity.

[13] The result can be generalized to continuous distributions with non-uniform priors: $H(q,p) = \int dx q(x) \ln(q(x)/p(x))$.

2 Linear Non-Equilibrium Physics: Derivations from MaxCal

Entropy is the most fundamental quantity in thermodynamics. While conservation of energy only limits the processes possible in a system, entropy defines the Second Law and gives a direction to any physical process

$$S(t_f) - S(t_t) \geq 0. \tag{44}$$

Because Eq. 44 is the only law that provides an **arrow of time**, entropy is intrinsically associated with non-equilibrium. However, most thermodynamic quantities of equilibrium physics, in particular entropy, lack a clear thermodynamic definition outside equilibrium. As indicated in Eq. 44, one typically compares initial and final states of an isolated process, and can predict that the system will evolve from a state with low entropy to a state with high entropy. This perspective is unequivocal, but it does not grant description of experiments without equilibrium states, such as non-equilibrium steady-state systems.

A standard approach is to approximate a thermodynamic system as being in **"local equilibrium"** [30]: If the system is perturbed slightly away from equilibrium, one splits it up into parts which can be assigned equilibrium quantities, such as temperature. While molecular dynamics simulations offer evidence that the approach to local equilibrium is rapid [33], assuming it has weaknesses. Because the theory is formulated entirely in terms of local equilibrium variables, it excludes effects due to (strong) gradients. More gravely, a straight-forward microscopic interpretation of total entropy is rendered impossible.

2.1 MaxCal: MaxEnt for Time-Dependent Systems

Chapter 1 introduces MaxEnt as the fundamental inference algorithm used in this thesis. Compared to traditional thermodynamics, it offers the advantage that it is mathematically rigorously defined, both in and and outside equilibrium. So far, the method was demonstrated only on state distributions, but expanding into non-equilibrium is straight-forward. The key point criticized about local equilibrium is that a non-equilibrium system cannot be described in terms of local states.

Instead, **MaxEnt can be formulated in terms of microtrajectories** $\{\Gamma\}$ **(MaxCal)**. A microtrajectory is any path in phase space which the

system can take between t_i and t_f. Unlike thermodynamic theories, MaxCal defines each path without appealing to the notion of equilibrium. After the enumeration of all possible paths, they are given weights $\{p_\Gamma\}$ which maximize information entropy

$$H = -\sum_\Gamma p_\Gamma \ln p_\Gamma \tag{45}$$

based on a set of mean value constraints. Naturally, these constraints take the form of dynamical quantities, such as mean values of microcurrents[14]

$$J = \sum_\Gamma j_\Gamma p_\Gamma. \tag{46}$$

2.2 Linear Thermodynamics According to Onsager & Prigogine

2.2.1 Onsager's Reciprocal Relations

It is a long-standing experimental result that there is cross-influence between different thermodynamic forces and flows. For example, consider a system reacting to both an electric and a gravitational force. The electric field will drive an electric current, while the gravitational field will force a flow of matter. In experiments, it is surprising that the electric field will also influence the flow of matter and gravitation will affect the electric current. Similar effects have been discovered in thermocouples (Peltier and Seebeck [30]), diffusion [12] and many more [36]. The striking feature in these very different systems is that the coupling is always symmetrical. In the example given above, this means that the electric field's effect on the flow of matter equals the gravitational influence on the electric current. This principle was derived from the assumption of microscopic reversibility by Onsager [38, 39]. His derivation from fluctuation theory is reviewed and commented in Appendix 5.1. Onsager's result can be summarized mathematically as

$$\vec{J} = \mathbf{L}\vec{F} \tag{47}$$

where $\mathbf{L}$ is a symmetric matrix.

[14] Again, these constraints to not require thermodynamic interpretation.

2.2.2 Entropy Production

Pushing concerns about its conceptual validity aside, Prigogine derives entropy production[15] based on the assumption of local equilibrium [30]. Define the local entropy s and consider (for simplicity) a system where no kinetic energy is dissipated due to convection and diffusion and where there is no external field. Invoke a conservation law with source term

$$\frac{\partial s}{\partial t} + \nabla J_s = \sigma \tag{48}$$

Using the Gibb's relation $Tds = du - \sum_k \mu_k dn_k$ (the term pdV does not appear as ds is a local quantity), the term $\frac{\partial s}{\partial t}$ can be re-written as

$$\frac{\partial s}{\partial t} = \frac{1}{T}\frac{\partial u}{\partial t} - \sum_k \frac{\mu_k}{T}\frac{\partial n_k}{\partial t} \tag{49}$$

Next, the following conservation laws for energy and mole number n_k for each chemical species k are inserted

$$\frac{\partial n_k}{\partial t} = -\nabla J_k + \sum_j \nu_{jk} v_j, \quad \frac{\partial u}{\partial t} + \nabla J_u = 0, \tag{50}$$

where the term $\sum_j \nu_{jk} v_j$ accounts for the chemical reactions (source term in the mole number equation) with v_j being the reaction velocity and ν_{jk} the stoichiometric coefficients.

$$\frac{\partial s}{\partial t} = -\frac{1}{T}\nabla J_u + \sum_k \frac{\mu_k}{T}\nabla J_k - \sum_{k,j} \frac{\mu_k}{T}\nu_{jk} v_j \tag{51}$$

Using an expression for the reaction affinities $A_j = -\sum_k \mu_k \nu_{jk}$[16] and exploiting the product rule from vector calculus $\nabla \cdot (g\vec{J}) = \vec{J} \cdot (\nabla g) + g(\nabla \cdot \vec{J})$, Eq. 51 is transformed into

$$\frac{\partial s}{\partial t} + \nabla \cdot \left(\frac{J_u}{T} - \sum_k \frac{\mu_k J_k}{T} \right) = J_u \cdot \nabla \frac{1}{T} - \sum_k J_k \cdot \nabla \frac{\mu_k}{T} + \sum_j v_j \cdot \frac{A_j}{T} \tag{52}$$

[15] Entropy production σ is defined as a creation of entropy. Even though systems under consideration need not be isolated, entropy production is always positive. Export to other systems is considered separately.

[16] This is the definition of A_j.

Comparing Eqs. 52 and 48 leads to the identification of thermodynamic expressions for J_s and σ

$$\sigma = J_u \cdot \nabla \frac{1}{T} - \sum_k J_k \cdot \nabla \frac{\mu_k}{T} + \sum_j v_j \cdot \frac{A_j}{T} \geq 0. \tag{53}$$

Prigogine splits entropy into an 'exchange component' and an 'irreversible component' $ds = d_e s + d_i s$. As defined in the continuity equation Eq. 48, σ is the entropy production (i.e. it represents the irreversible part), and by the second law $\sigma \geq 0$. Forces and fluxes can be read off Eq. 53 directly and give pairs of the form

$$\dot{s} \equiv \frac{d_i s}{dt} \equiv \sigma = \sum_\alpha J_\alpha X_\alpha \tag{54}$$

Conservative force fields (e.g. a static electric field or a gravitational field) with potential ψ can now be easily incorporated: Eq. 49 is extended into

$$\frac{\partial s}{\partial t} = \frac{1}{T}\frac{\partial u}{\partial t} - \sum_k \frac{\mu_k + \tau_k \psi}{T} \frac{\partial n_k}{\partial t} \tag{55}$$

By the very same reasoning, one finds the extension to Eq. 52

$$\begin{aligned} &\frac{\partial s}{\partial t} + \nabla \cdot \left(\frac{J_u}{T} - \sum_k \frac{(\mu_k + \tau_k \psi) J_k}{T} \right) \\ &= J_u \cdot \nabla \frac{1}{T} - \sum_k J_k \cdot \nabla \frac{\mu_k}{T} + \frac{I \cdot (-\nabla \psi)}{T} + \sum_j v_j \cdot \frac{A_j}{T} \end{aligned} \tag{56}$$

One may find this derivation straight-forward and based on very safe grounds (conservation laws and local equilibrium). It must be noted, however, that there is some arbitrariness in Eq. 52. For example, the second term on the right-hand side could be split into two parts

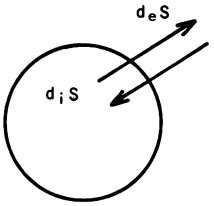
$$\sum_k J_k \cdot \nabla \frac{\mu_k}{T} = \sum_k J_k \cdot \frac{\nabla \mu_k}{T} + \sum_k J_k \cdot \mu_k \nabla \frac{1}{T}. \tag{57}$$

This would lead to the identification of different flows and currents from a changed expression of σ

$$\Rightarrow \sigma = J'_u \cdot \nabla \frac{1}{T} - \sum_k J_k \cdot \frac{\nabla \mu_k}{T} + \sum_j v_j \cdot \frac{A_j}{T} \geq 0, \text{ with} \tag{58}$$

$$J'_u = J_u - \sum_k J_k \cdot \mu_k. \tag{59}$$

This observation indicates that even in Prigogine's theory, a fixed set of 'thermodynamic forces' is non-existent. In the linear regime, there are allowed transformations changing the basis vectors of the force vector space.



	Force F_α	Flow (Current) J_α
Heat conduction	$\nabla\frac{1}{T}$	Energy flow $\mathbf{J}_u$
Diffusion	$-\nabla\frac{\mu_k}{T}$	Diffusion current $\mathbf{J}_k$
Electrical conduction	$\frac{-\nabla\phi}{T}=\frac{\mathbf{E}}{T}$	Ion current densities $\mathbf{I}_k$
Chemical reactions	$\frac{A_j}{T}$	Velocity of reaction $v_j=\frac{1}{V}\frac{d\xi_j}{dt}$

Figure 4: Left: Exchange of entropy between the outside and the inside [44], Right: Flows and forces which are thermodynamically coupled [30]

In summary, Prigogine finds certain couples of flows and forces. Also, Prigogine can define a quantity $\dot{s} = \sigma = \sum_i F_i J_i$ (Eq. 54) which is strictly positive. Combined with the linear response theory assumed by Onsager (Eq. 47), this gives

$$P = \int \dot{s} dV, \tag{60}$$

$$\dot{s} = \sum_{ij} F_i \mathbf{L_{ij}} F_j \geq 0. \tag{61}$$

This implies that $\mathbf{L}$ needs to be positive definite. Mathematically, the Onsager and Prigogine theory boils down to a linear response ansatz with a symmetric and positive definite coefficient matrix.

2.2.3 Prigogine's Principle

So far, the theory provides notation but has no predictive power beyond the assumption of linear response. In many situations, a selection principle would be very desirable. Consider, as a simple example of a non-equilibrium steady state, the system coupling electric and gravitational fields from above. Assume that the gravitational force is fixed but the experiment exerts no control over the electric field. This is not to say that the applied voltage has a fixed but unknown value. Rather, the system can autonomously evolve towards a voltage through internal processes. Which value will the electrical

field assume? In the tradition of physical extremal principles[17], Prigogine postulates as a selection principle:

Theorem 2 (*Prigogine's Principle, I*)

In a steady-state system satisfying Eq. 47, $\dot{s}$ will assume its minimal possible value under the imposed constraints.

As explained above, $\dot{s}$ is a conceptually evasive quantity without strictly defined meaning. Prigogine's principle can, however, be reformulated into a completely unambiguous form.

Theorem 3 (*Prigogine's Principle, II*)

In a steady-state system satisfying Eq. 47, all currents whose corresponding forces are unconstrained will vanish.

Proof 1 ($N = 2$)

$$P = \frac{d_i S}{dt} = \int (F_1 J_1 + F_2 J_2) dV \rightarrow \textit{Minimum} \tag{62}$$

Use linear response laws

$$J_1 = L_{11} F_1 + L_{12} F_2, \quad J_2 = L_{21} F_1 + L_{22} F_2 \tag{63}$$

Insert into Eq. 62, use $L_{12} = L_{21}$ (Onsager) and minimize

$$P = \int (L_{11} F_1^2 + 2 L_{12} F_1 F_2 + L_{22} F_2^2) dV \tag{64}$$

$$\frac{\partial P}{\partial F_2} = \int 2(L_{12} F_1 + L_{22} F_2) \stackrel{!}{=} 0 \tag{65}$$

$$\Leftrightarrow J_2 = L_{21} F_1 + L_{22} F_2 = 0 \tag{66}$$

This cannot be a maximum since Eq. 62 is unbounded above as a function of J_2.

[17] Extremal principles, such as the Hamilton-Lagrange Principles, have enjoyed enormous success in all fields of physics.

Proof 2 (*general N*)

Let l be an unconstrained force

$$P = \frac{d_i S}{dt} = \int \vec{F}^t \vec{J} dV = \sum_{ij} \int F_i L_{ij} F_j dV \tag{67}$$

$$\Rightarrow \frac{\partial P}{\partial F_l} = \sum_{lj} \int 2 L_{lj} F_j dV \stackrel{!}{=} 0 \tag{68}$$

$$\Rightarrow J_l = \int \sum_j L_{lj} F_j = 0 \tag{69}$$

This is a minimum since L is positive definite. In the following, the second version of Prigogine's theorem will be used.

2.3 Statistical Linear Thermodynamics

2.3.1 Linear Response and the Onsager Relations

Case $N = 2$: Starting from MaxCal, we want to obtain a linear response theory and observe how it relates to Onsager's reciprocal relations. Consider a system with $N = 2$ **average currents**, which are implemented in the maximization of entropy as constraints. The partition function then reads (cf. Appendix 5.2)[18]

$$Z(\lambda_1, \lambda_2) = \sum_{\Gamma} \exp(\lambda_1 j_{1,\Gamma} + \lambda_2 j_{2,\Gamma}). \tag{70}$$

The macroscopic currents are averages over all paths and therefore they are functions only of λ_1, λ_2. In general, a constraint's effect on the partition function as well as on any physical quantities derived from it shrinks as $|\lambda| \to 0$. In the limit $\lambda_i = 0$, the constraint loses its effect altogether, i.e. the PDF is independent of j_i. Thus, it makes sense to expand the currents J_1, J_2 about $(\lambda_1 = 0, \lambda_2 = 0)$ for weak constraints. At $(\lambda_1 = 0, \lambda_2 = 0)$, the system has no constraints. In other words, we expect the system to have no average currents $J(0, 0) = 0$. One can justify this because $(\lambda_1 = 0, \lambda_2 = 0)$ is the **equilibrium point**, where zero currents are guaranteed by **detailed**

[18] Note the mathematical correspondence to $Z = \sum_i \exp(-\beta E_i)$, the MaxEnt solution to Boltzmann's equilibrium gas problem.

balance. Another nice effect is that the theory is an **expansion around equilibrium**, where all thermodynamic quantities are well-defined. Note that as Eq. 70 is an analytic function by definition, all expansions are well-defined and convergent.

$$\begin{aligned} J_1(\lambda_1,\lambda_2) &= \underbrace{J_1(\lambda_1,\lambda_2)|_{(0,0)}}_{=0} + \left.\frac{\partial J_1(\lambda_1,\lambda_2)}{\partial\lambda_1}\right|_{(0,0)}\lambda_1 + \left.\frac{\partial J_1(\lambda_1,\lambda_2)}{\partial\lambda_2}\right|_{(0,0)}\lambda_2 + \mathcal{O}(\lambda^2) \\ J_2(\lambda_1,\lambda_2) &= \underbrace{J_2(\lambda_1,\lambda_2)|_{(0,0)}}_{=0} + \left.\frac{\partial J_2(\lambda_1,\lambda_2)}{\partial\lambda_1}\right|_{(0,0)}\lambda_1 + \left.\frac{\partial J_2(\lambda_1,\lambda_2)}{\partial\lambda_2}\right|_{(0,0)}\lambda_2 + \mathcal{O}(\lambda^2) \end{aligned} \tag{71}$$

This already gives a form that is reminiscent of **linear response**. To see that the Onsager relations are satisfied, one needs to use an alternative expression of the currents (from Eqs. 46 and 70)

$$\frac{\partial \ln Z}{\partial\lambda} = \sum_\Gamma j_\Gamma \frac{\exp(\lambda j_\Gamma)}{Z} = \sum_\Gamma j_\Gamma p_\Gamma = J. \tag{72}$$

In the special case, this gives

$$J_1 = \frac{\partial \ln(Z)}{\partial\lambda_1} \qquad J_2 = \frac{\partial \ln(Z)}{\partial\lambda_2}. \tag{73}$$

Now, the expansion simplifies formally into

$$\begin{aligned} J_1(\lambda_1,\lambda_2) &= \left.\frac{\partial^2 \ln(Z(\lambda_1,\lambda_2))}{\partial\lambda_1\partial\lambda_1}\right|_{(0,0)}\lambda_1 + \left.\frac{\partial^2 \ln(Z(\lambda_1,\lambda_2))}{\partial\lambda_1\partial\lambda_2}\right|_{(0,0)}\lambda_2 + \mathcal{O}(\lambda^2) \\ J_2(\lambda_1,\lambda_2) &= \left.\frac{\partial^2 \ln(Z(\lambda_1,\lambda_2))}{\partial\lambda_1\partial\lambda_2}\right|_{(0,0)}\lambda_1 + \left.\frac{\partial^2 \ln(Z(\lambda_1,\lambda_2))}{\partial\lambda_2\partial\lambda_2}\right|_{(0,0)}\lambda_2 + \mathcal{O}(\lambda^2), \end{aligned} \tag{74}$$

which is nothing but the Hessian matrix $\mathbf{H}$ evaluated at the origin

$$\vec{J}(\lambda_1,\lambda_2) = \mathbf{H}(\ln Z(\lambda_1,\lambda_2))|_{(0,0)} \cdot \vec{\lambda} + \mathcal{O}(\lambda^2). \tag{75}$$

Proof 3 (*Case N arbitrary*)

The above reasoning can immediately be generalized to an arbitrary number of constraints/currents

$$\vec{J}(\vec{\lambda}) = \mathbf{H}(\ln Z(\vec{\lambda}))\Big|_{\vec{\lambda}=0} \cdot \vec{\lambda} + \mathcal{O}(\lambda^2). \tag{76}$$

This directly implies the Onsager relations in their full generality.

2.3.2 Higher-order Expansions

The currents being analytic functions of $\vec{\lambda}$, they can be expressed exactly as a Taylor series

$$\begin{aligned} J_i(\vec{\lambda}) &= \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty}\cdots\sum_{n_d=0}^{\infty}\frac{\lambda_1^{n_1}\cdots\lambda_{n_d}^{n_d}}{n_1!\cdots n_d!}\left(\frac{\partial^{n_1+\cdots+n_d}J_i(\vec{\lambda})}{\partial\lambda_1^{n_1}\cdots\partial\lambda_d^{n_d}}\right)\Bigg|_{\vec{\lambda}=0} \\ &= \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty}\cdots\sum_{n_d=0}^{\infty}\frac{\lambda_1^{n_1}\cdots\lambda_{n_d}^{n_d}}{n_1!\cdots n_d!}\underbrace{\left(\frac{\partial^{n_1+\cdots+(n_i+1)+\cdots+n_d}\ln Z(\vec{\lambda})}{\partial\lambda_1^{n_1}\cdots\partial\lambda_i^{n_i+1}\cdots\partial\lambda_d^{n_d}}\right)\Bigg|_{\vec{\lambda}=0}}_{\equiv\mathbf{L_i}^{\mathbf{n_1 n_2\cdots n_d}}}. \end{aligned} \tag{77}$$

The d-dimensional tensors $\mathbf{L_i}$ (there are d such tensors, one for each force λ_i) have certain symmetries because of the interchangeability of second derivatives (Schwarz' theorem)

$$\mathbf{L_i}^{n_1\cdots n_i\cdots(n_j+1)\cdots n_d} = \mathbf{L_j}^{n_1\cdots(n_i+1)\cdots n_j\cdots n_d}. \tag{78}$$

2.3.3 What's Special About This?

Eq. 76 is a concise mathematical description of the expected currents as a function of small $\vec{\lambda}$ according to MaxCal. It contains the Onsager relations because the Hessian matrix $\mathbf{H}$ is symmetric. For the same reason, $\mathbf{H}$ is diagonizable and invertible[19], two very convenient properties.
Here are the essential conditions that led to this result

1. Z is an analytic function $\Rightarrow \vec{J}(\vec{\lambda})$ has a convergent Taylor series.

[19] if none of its eigenvalues is zero

2. Detailed balance at equilibrium ⇒ the zeroth-order term in the expansion vanishes, linear response.

3. The partition function acts as a "potential" to the current $\vec{J}(\vec{\lambda}) = \nabla \ln Z(\vec{\lambda}) \Rightarrow$ Onsager relations.

Which of these points is specific to MaxCal? To answer this question, consider an arbitrary theory that predicts currents as functions of some set of parameters $\vec{\kappa}$. Points 1 and 2 can safely be postulated for any physical theory close to equilibrium. This alone gives a linear response theory. In my opinion, it makes sense to **define the parameters $\vec{\lambda}$ as forces** (cf. Sect. 2.3.4).

If Points 1 and 2 are given, the potential from Point 3 can be constructed for an arbitrary theory in the linear approximation

$$V(\vec{(}\kappa)) = \vec{J}_c \cdot \vec{\kappa}, \tag{79}$$

where $\vec{J}_c$ is a vector of arbitrary coefficients. Eq. 73 holds exactly, though, and this is specific to MaxCal (cf. Helmholtz theorem and [53]). Thus, **the Onsager relations are not a universal result but originate specifically from the MaxCal theory.**

Furthermore, the Hessian matrix $\mathbf{H_{ij}} \equiv \mathbf{H}(\ln Z(\vec{\lambda}))\Big|_{\vec{\lambda}=0}$ is **positive definite**!

Proof 4

$\mathbf{H_{ij}}$ *is the correlation matrix (evaluation at* $\vec{\lambda} = 0$ *implied, though the statement is general)*

$$\begin{aligned} \mathbf{H_{ij}} &= \frac{\partial^2 \ln Z}{\partial\lambda_i \partial\lambda_j} = \frac{\partial}{\partial\lambda_i}\left(\frac{1}{Z}\frac{\partial Z}{\partial\lambda_j}\right) = \left(\frac{1}{Z}\frac{\partial^2 Z}{\partial\lambda_i\partial\lambda_j} - \frac{1}{Z^2}\frac{\partial Z}{\partial\lambda_i}\frac{\partial Z}{\partial\lambda_j}\right) \\ &= \langle j_i j_j\rangle - \langle j_i\rangle \langle j_j\rangle \\ &= \mathrm{Cov}(j_i, j_j) \end{aligned} \tag{80}$$

Correlation matrices are positive (semi-) definite

$$\begin{aligned} \sum_i \sum_j \lambda_i \mathbf{H_{ij}} \lambda_j &= \sum_i \sum_j \lambda_i \lambda_j \,\mathrm{Cov}(j_i, j_j) \\ &= \sum_i \sum_j \mathrm{Cov}(\lambda_i j_i, \lambda_j j_j) = \mathrm{Var}(\sum_i \lambda_i j_i) \geq 0. \end{aligned} \tag{81}$$

Here, the response matrix's positive definiteness is a mathematical necessity, not a consequence of the (phenomenological) Second Law.

2.3.4 MaxCal Implies Prigogine's Principle

It seems that we are now in the same position as Prigogine before postulating what is today known as **Prigogine's Principle**[20]. In fact, we could proceed completely analogously by defining $\dot{\tilde{S}}$

$$\dot{\tilde{S}} \equiv \vec{\lambda}^T \vec{J} = \vec{\lambda}^T \mathbf{H} \vec{\lambda} + \mathcal{O}(\lambda^3) \tag{82}$$

$$\Rightarrow \dot{\tilde{S}} \geq 0, \quad |\vec{\lambda}| \ll 1 \tag{83}$$

because $\mathbf{H}$ is positive semi-definite. Like Prigogine, we might go on to assert a minimum principle

$$\delta_{\{\lambda_i\}} \dot{\tilde{S}} = 0 \ \forall i \text{ unconstrained} \tag{84}$$

This gives rise to the prediction that currents corresponding to unconstrained forces vanish (Box II in Sect. 2.2.3). **Amazingly, MaxCal provides an intrinsic selection principle that renders an additional postulate obsolete (Proof 5)!**

[20] In the literature, Prigogine's Principle is also referred to as the "Minimum Entropy Production Principle".

Proof 5

Assume that there are N forces/flows in the linear system, out of which w.l.o.g the first $(k-1)$ are fixed. Eq. 76 makes no prediction about the values of the unconstrained variable. ***As the guiding principle, use the maximization of the caliber****. This amounts to maximizing the information entropy S_I with respect to the the unconstrained forces.*

$$\begin{aligned} H(\{\lambda_k, ..., \lambda_N\}) &= -\sum_{\Gamma} p_\Gamma \ln(p_\Gamma) \\ &= -\sum_{\Gamma} \frac{e^{\lambda_1 j_{1,\Gamma} + ... + \lambda_N j_{N,\Gamma}}}{Z} \cdot \ln\left(\frac{e^{\lambda_1 j_{1,\Gamma} + ... + \lambda_N j_{N,\Gamma}}}{Z}\right) \\ &= \frac{1}{Z} \cdot \ln(Z) \underbrace{\sum_{\Gamma} e^{\vec{\lambda}\cdot\vec{j_\Gamma}}}_{=Z} - \frac{1}{Z} \sum_{\Gamma} e^{\vec{\lambda}\cdot\vec{j_\Gamma}} \cdot (\vec{\lambda} \cdot \vec{j_\Gamma}) \\ &= \ln Z - \vec{\lambda} \cdot \vec{J} \\ &= \ln Z(\{\lambda_k, ..., \lambda_N\}) - \vec{\lambda} \cdot \vec{J}(\vec{\lambda}) \end{aligned} \tag{85}$$

As explained above, entropy inference methods demand that

$$\frac{\partial H}{\partial \lambda_l} = 0 \ \forall l \in \{k, (k+1), ..., N\} \tag{86}$$

Using the linear expansion in Eqs. 76

$$\begin{aligned} &\Leftrightarrow J_l - \underbrace{\sum_i \mathbf{H_{il}} \lambda_i}_{J_l} - \underbrace{\sum_i \mathbf{H_{li}} \lambda_i}_{J_l} = 0 \\ &\Leftrightarrow J_l = 0 \end{aligned} \tag{87}$$

This completes the proof since

$$J_l = 0 \Leftrightarrow \frac{\partial \dot{\tilde{S}}}{\partial \lambda_l} = 0 \tag{88}$$

Note: *The solutions to Eq. 86 are general.* ***They only give Prigogine's principle in linear approximation****.*

2.3.5 Interpretation of λ

In all of statistical mechanics there is some difficulty in assigning thermodynamic meaning to statistical quantities. Here, I will argue that this issue is resolved for equilibrium systems by one key property of (almost all) forms of matter and suggest by example that a similar interpretation might be viable for non-equilibrium steady-state systems.

In equilibrium statistical mechanics, the Lagrange multiplier corresponding to the energy is β. In entropy maximization procedures, the Lagrange multiplier is always given as

$$\beta = \frac{1}{k_B} \frac{\partial S^*}{\partial \langle E \rangle}, \tag{89}$$

where S^* is the maximum entropy under the energy constraints. Temperature is then defined via $\beta = \frac{1}{k_B T}$. This way, temperature is expressed as $T = \frac{\partial E}{\partial S}$, but this is not a useful relation since absolute entropy cannot be measured directly. It is an explicit formula (Sackur-Tetrode) which relates T with an experimentally accessible quantity, namely the energy E:

$$S(E) = k_B N \left(\frac{3}{2} \ln(E/N) + \ln(V/N) + \ln(c) \right) \tag{90}$$

$$\Rightarrow T \equiv \frac{\partial E}{\partial S} = \frac{2}{3} \frac{E}{k_B N}. \tag{91}$$

I think it is important to realize that **this relation is specific to the ideal gas**. Most other materials behave similarly, i.e. they also obey $S(E) \propto \ln(E)$, but exotic systems such as finite spins in a magnetic field do not satisfy this relation [3]. In these systems, temperature is not proportional to the average energy. **In some cases temperature can even be negative!**
In non-equilibrium statistical mechanics, the Lagrange multipliers have the same function and mathematical definition. While in equilibrium the standard relation is $T \propto \frac{1}{E}$, in non-equilibrium it is $\lambda \propto J$. Since this is the case, the expansion around $\lambda = 0$ makes sense. For illustration, consider the simple dog-flea model with N fleas on each dog. Within one time unit $\Delta\tau$ each flea can either jump or stay on the dog. The net number of fleas jumping in one direction constitutes the current J. Initially devised by the Ehrenfests [29], this model is described in detail in [43, 18]. In this case one

finds analytically

$$Z = (1 + e^{\lambda})^N \cdot (1 + e^{-\lambda}) \tag{92}$$

$$J = \frac{\partial \ln Z}{\partial \lambda} = N \left[\frac{e^{\lambda}}{1 + e^{\lambda}} - \frac{e^{-\lambda}}{1 + e^{-\lambda}} \right] = N \tanh(\lambda/2) \approx N\lambda/2, \quad |\lambda| < 1 \tag{93}$$

$$\lambda = 2 \operatorname{arctanh}(J/N) \approx 2J/N. \tag{94}$$

Thus, the experimental meaning is established: As β is determined by measuring energy, $\vec{\lambda}$ is determined from measured currents.

2.4 MaxCal Predictions of a General System

In the following, I will demonstrate the analytical results for a simple and still quite general system. Working with the Lagrange multipliers as independent parameters, the striking relations to Onsager's reciprocal relations and Prigogine's principle will be shown.

2.4.1 A Generic Model

The proofs reveal that MaxCal predictions share two fundamental properties of linear non-equilibrium steady-state systems: They satisfy the Onsager relations and Prigogine's principle. In the following, I will refer to these principles in terms of the system's Lagrange multipliers, knowing that they are really defined thermodynamically. The proofs from above are formulated

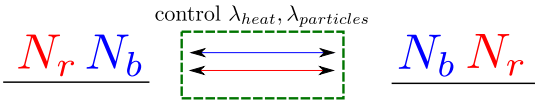


Figure 5: A very general model allowing MaxCal predictions

for any number of species M. For illustration I will use $M = 2$ as depicted in Fig. 5. Assume the system is initially in equilibrium

$$J_i(\lambda_p = 0, \lambda_h = 0) = 0, \tag{95}$$

but two forces drive it into a non-equilibrium steady state. This system can be interpreted in various ways; I will choose for concreteness (and in the tradition of the hot-dog model [18]):

$$J_{particles} = \text{Net particle flow to the right} = J_{red} + J_{blue} \tag{96}$$
$$J_{heat} = \text{Net heat flow to the right} = J_{red} + a \cdot J_{blue}, \tag{97}$$

where the parameter a (set $a = 0.3$) determines to which extent the blue particles contribute to heat transport. The same model can describe different physics if the coupling is done differently.

2.4.2 Onsager's Reciprocal Relations

For the described model, all quantities can be given analytically. The partition function reads

$$\begin{aligned} Z =&(1 + e^{\lambda_p + a\cdot\lambda_h})^{N_b} \cdot (1 + e^{\lambda_p + \lambda_h})^{N_r} \\ &\cdot(1 + e^{-\lambda_p - a\cdot\lambda_h})^{N_b} \cdot (1 + e^{-\lambda_p - \lambda_h})^{N_r}. \end{aligned} \tag{98}$$

The average currents can be found by taking derivatives

$$J_h = N_b a \tanh\left(\frac{\lambda_p + a\lambda_h}{2}\right) + N_r \tanh\left(\frac{\lambda_p + \lambda_h}{2}\right) \tag{99}$$
$$J_p = N_b \tanh\left(\frac{\lambda_p + a\lambda_h}{2}\right) + N_r \tanh\left(\frac{\lambda_p + \lambda_h}{2}\right). \tag{100}$$

To lowest order the currents can be expanded into

$$\begin{aligned} \begin{pmatrix} J_h \\ J_p \end{pmatrix} &\approx \begin{pmatrix} \frac{\partial^2 \ln Z}{\partial\lambda_h \partial\lambda_h} & \frac{\partial^2 \ln Z}{\partial\lambda_h \partial\lambda_p} \\ \frac{\partial^2 \ln Z}{\partial\lambda_h \partial\lambda_p} & \frac{\partial^2 \ln Z}{\partial\lambda_p \partial\lambda_p} \end{pmatrix}_{\lambda_h=\lambda_p=0} \begin{pmatrix} \lambda_h \\ \lambda_p \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} N_b a^2 + N_r & N_b a + N_r \\ N_b a + N_r & N_b + N_r \end{pmatrix} \begin{pmatrix} \lambda_h \\ \lambda_p \end{pmatrix}. \end{aligned} \tag{101}$$

Note that this matrix is **symmetric** and **positive definite**. This corresponds to the **Onsager relations**. Higher derivative can be easily calculated and will satisfy the generalized symmetry relations from Eq. 78. For example, we find in the next non-vanishing order (using the notation from Eq. 78):

$$\mathbf{L_1}^{21} = \mathbf{L_2}^{30} = \left.\frac{\partial^4 \ln Z}{\partial\lambda_h^3 \partial\lambda_p}\right|_{\lambda_h=\lambda_p=0} = -\frac{1}{4}\left(a^3 N_b + N_r\right) \tag{102}$$

2.4.3 Deviations from the Linear Regime

Fig. 6 shows a comparison between the currents as calculated exactly according to Eqs. 99-100 with the linear approximation from Eq. 101. At first

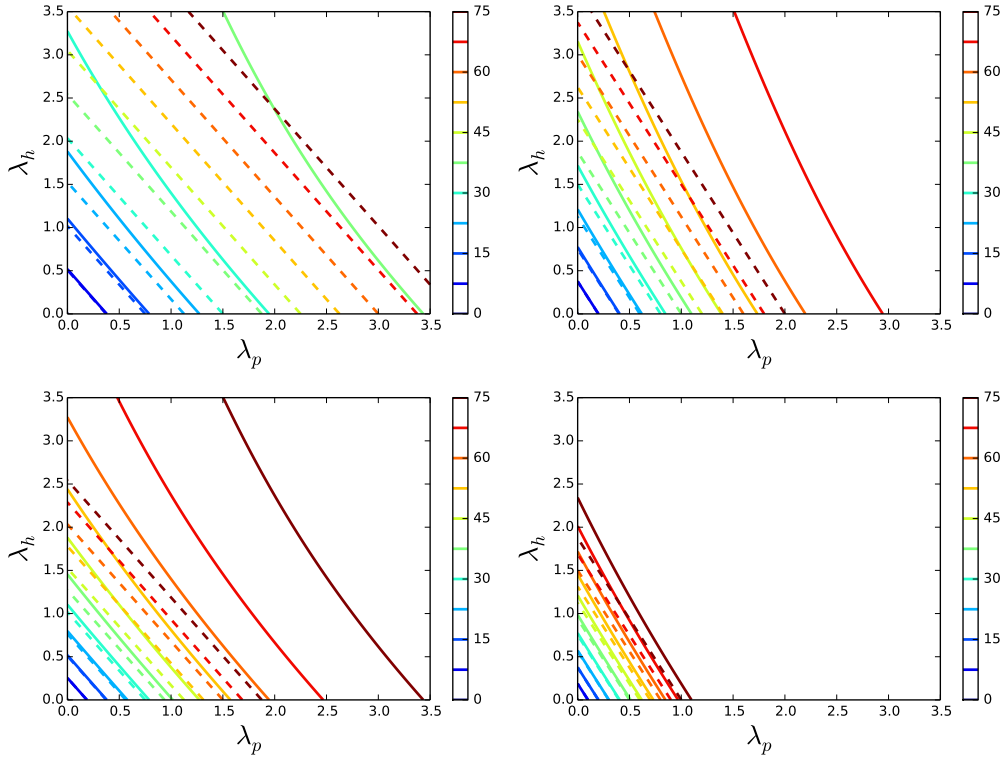


Figure 6: Comparison of the exact currents (solid lines) with the linear approximation (dashed lines), Top: $N_{blue} = 50, N_{red} = 25$, Bottom: $N_{blue} = 100, N_{red} = 50$. Left: heat current, Right: particle current.

glance, the linear approximation in the top row seems reasonably accurate for $\lambda < 1$, for $\lambda > 1$ it quickly diverges from the true solution. These calculations were performed setting $N_{blue} = 50, N_{red} = 25$. As can be anticipated from a similar setup in Sect. 2.3.5, increasing the number of particles shifts the linear approximation's range of validity. For $N_{blue} = 100, N_{red} = 50$ (bottom row) the linear approximation is feasible for a much greater range of the λ's.

2.4.4 Prigogine's Principle

A general question to ask is: **Given some set of Lagrange parameters is fixed (say λ_p), what will the other Lagrange parameters (in this case λ_h) adjust to?** In MaxCal, this is determined by maximizing the information entropy with respect to the "free" Lagrange multipliers. As was shown in Eq. 86, this amounts to

$$\begin{aligned} H &= \ln Z - \sum_i \lambda_i J_i \\ \Rightarrow \frac{\partial H}{\partial \lambda_l} &= J_l - \sum_i \frac{\partial}{\partial \lambda_l}(\lambda_i J_i) = -\sum_i \lambda_i \frac{\partial J_i}{\partial \lambda_l} \stackrel{!}{=} 0. \end{aligned} \tag{103}$$

I find it very interesting to observe that Eq. 103 in the linear approximation gives $J_l = 0$. **This is Prigogine's principle!** The true solution, however, is only zero for small λ. This is depicted in Fig. 7, while Fig. 8 shows the predicted values of the uncontrolled Lagrange multiplier. The larger λ_p gets, the further will J_h deviate from zero. Thus, MaxCal gives Prigogine's principle as a first approximation and provides corrections outside the linear regime. It is interesting to observe that the solutions to Eq. 103 only depend on the ratio in the particle numbers (Fig. 8), while the predicted current J_h deviates further from Prigogine's linear behavior as there are more particles (Fig. 7). This can be easily seen by writing out the explicit expression for Eq. 103.

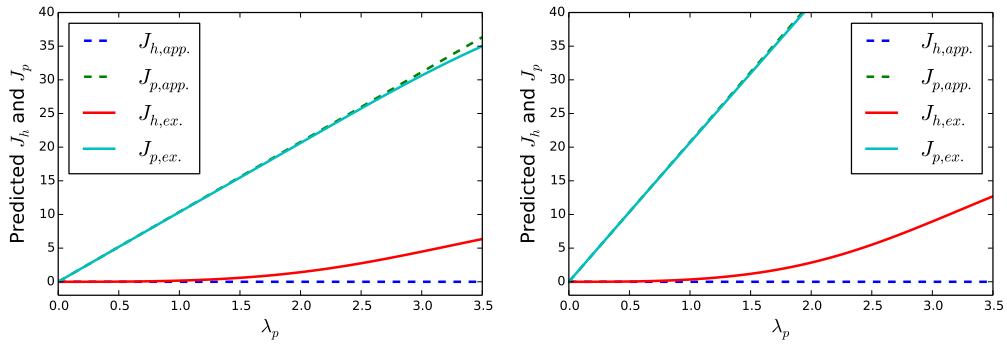


Figure 7: Current comparison between linear approximation (Prigogine's principle) and the MaxCal prediction, Left: $N_{blue} = 50, N_{red} = 25$*, Right:* $N_{blue} = 100, N_{red} = 50$

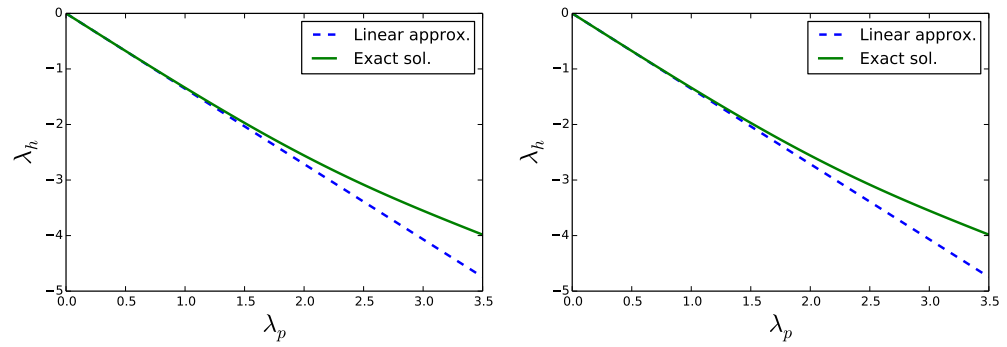


Figure 8: Prediction for λ_h *in linear approximation (Prigogine) and according to MaxCal, Left:* $N_{blue} = 50, N_{red} = 25$*, Right:* $N_{blue} = 100, N_{red} = 50$

2.4.5 Bifurcations

Eq. 103 is the **necessary condition** for a local extreme point. From the partition function the limiting behavior of H is clear for (very) small and very large numbers of λ_h at fixed λ_p: H is diminished. Physically, this is due to the fact that a great value of $|\lambda_h|$ asks for great currents, which reduces the number of possible microtrajectories available to the system. Since H is a continuous function, we can conclude that **there is an odd number of maximum points and an even number of minimal points**. For relatively small λ_h, there is a sharply peaked maximum (the one that is also represented in Fig. 8), as can be seen in Fig. 10. As λ_p gets bigger this

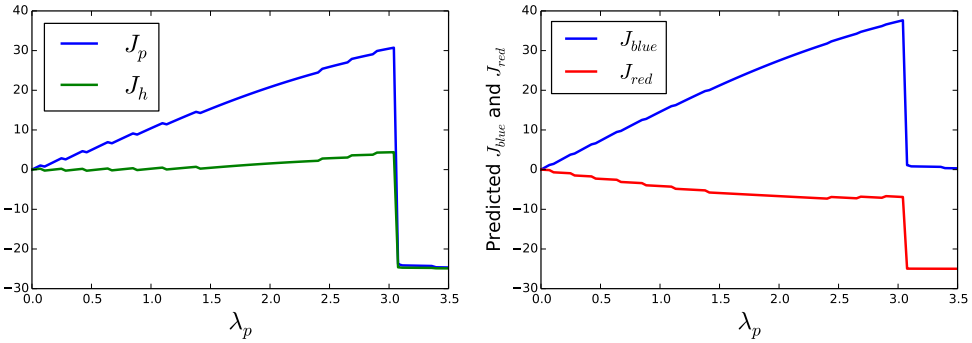


Figure 9: At some point the currents jump: indication of a phase transition. Left: Representation of macroscopic currents, right: currents of the two particle species (The wobbling line is due to the maximum finding algorithm), $N_{blue} = 50, N_{red} = 25$

changes drastically: A second maximum develops, which eventually becomes the global maximum. Assuming that the system can freely adjust the second Lagrange multipliers, it would jump to the value where the new global maximum emerges, leading to what in dynamic systems theory is called a **first-order phase transition**! One can get an idea about this observation by looking at the two particle species, rather than the coupled flows: In this new "phase" all hot (red) particles flow from the right to the left, while there is hardly any (net) current of cold (blue) particles. I suspect this phenomenon is due to the rapid decline in the number of permutations as the blue (cold) current increases. With blue outnumbering red particles by a factor of 2, this overcompensates for the reduction in red pathways.

Instabilities and semi-stable steady states are essential to the physics beyond the linear regime. Prigogine developed a theoretical framework based on bifurcations to explain failures the quantitative changes observed when a system leaves the linear regime [30, 44]. Evidently, the given toy model is a tremendous simplification of any natural system. However, we have demonstrated that MaxCal provides a theoretical description that allows to bridge the linear world of Prigogine's theorem with his ad-hoc theory of bifurcations. Applying MaxCal to more realistic systems will reveal if the formalism is able to reproduce realistic phase transitions.

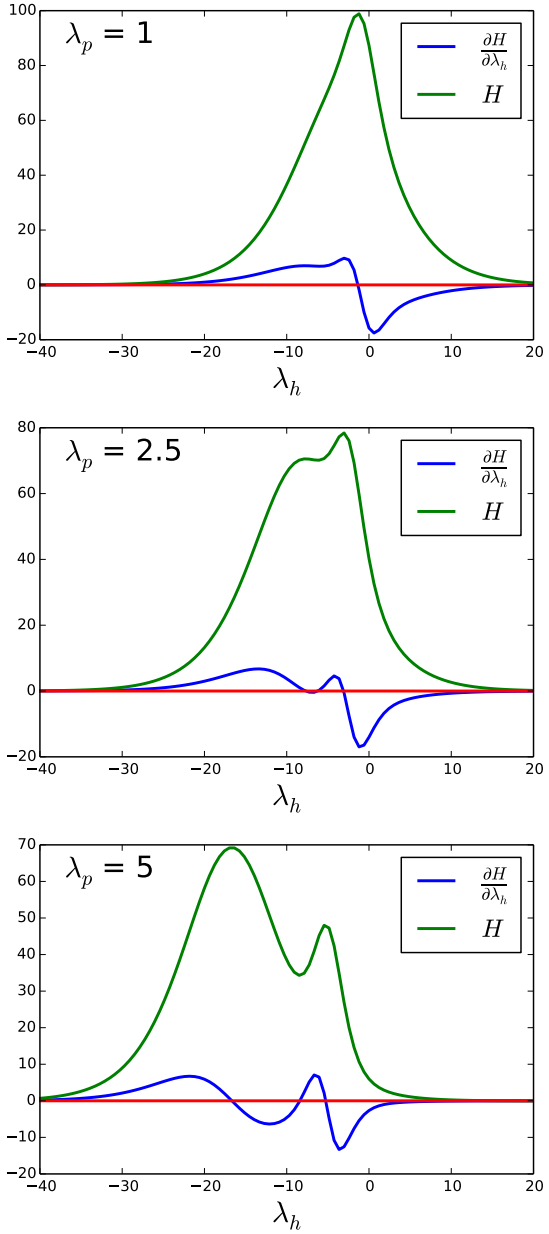


Figure 10: Transitions in the predicted current, $N_{blue} = 50, N_{red} = 25$

3 Far from Equilibrium: A Survey of Maximum Entropy Production

In the previous chapter, theorems from near-equilibrium thermodynamics are analyzed from a unique inference perspective. Two of the central results are demonstrated to appear as special cases from a MaxCal expansion in small parameters $\{\lambda_i\}$, which measure distance from equilibrium. For the case of a simple model system, Sect. 2.4.5 reveals the emergence of several peaks in the entropy curve, each corresponding to a different macroscopic state of the system. This already hints at the fact that non-equilibrium thermodynamics becomes both much more complex and far less understood as one leaves the direct vicinity of equilibrium.

One central and still unresolved conundrum from non-equilibrium thermodynamics is nature's tendency to form low entropy structures. Some small-scale thermodynamic systems (e.g. the Rayleigh-Bénard convection cell), global climate and all forms of life assume highly-ordered structures which are associated with low entropy[21]. The Second Law predicts that a closed thermodynamic system will increase its entropy when transitioning from one equilibrium state to another. How and how fast this process occurs remains unanswered. Still, the ubiquitous appearance of (open) ordered structures in the real world has challenged researchers to come up with theories explaining why the observed structures seem to be favored in nature.

Prigogine termed this observation **dissipative structures** because the production of thermodynamic entropy in these systems increases rapidly, while their internal entropy content drops. Trying to expand his theory of minimum entropy production, Prigogine develops a phenomenological theory based on bifurcations to explain the observed qualitative phase transitions [44]. He also presented the beginnings of a non-unitary microscopic theory incorporating a **microscopic meaning of entropy**. Contrasting Prigogine's attempts to find a *microscopic* explanation for systems far from equilibrium, a *macroscopic* analogue of Prigogine's principle was brought forward: the Principle of Maximum Entropy Production (Max-EP).

[21] Schrödinger even claims that the essence of life is exploiting free energy by exporting entropy to the environment [47].

3.1 Statement of Max-EP

Max-EP has been proposed by several researchers, both independently and in various scientific fields, including formulations by Ziegler, Swenson, Dewar and Kleidon. The latter two are the most general and appear to form a kind of mainstream within the school of Max-EP. What is more, Dewar is accepted as having come closest to a derivation of his Max-EP principle from information theory. Alex Kleidon summarizes this principle as:

Theorem 4 (*Maximum Entropy Production [28]*)

> *"The Max-EP principle states that non-equilibrium thermodynamic systems are organized in a steady state such that the rate of entropy production is maximized."*

Thus, Max-EP is an extremal principle[22]. A simple argument shows, however, that **finding a universal extremum principle for non-equilibrium thermodynamics is problematic**. At first glance, there is an overt contraction between Max-EP and Prigogine's Principle discussed in the last chapter. Why should the same entropy production be minimized and maximized? The answer appears to be that Max-EP is only applicable in systems that have many steady states compatible with the boundary conditions, a situation usually associated with great "distance" from equilibrium. Because some articles on Max-EP do not emphasize the limited range of the principle, consider a steady-state electrical circuit [26] as an illustrative example: Because of charge conservation, no charge can accumulate

$$\nabla \cdot (\underbrace{\sigma \nabla \Phi}_{-\vec{J}}) = 0. \tag{104}$$

Eq. 104, however, is equivalent to the minimum point of the functional $\mathcal{L}$ as can been seen by using Euler-Lagrange equations:

$$\begin{aligned} \mathcal{L} &= \int_V \sigma \left(\nabla \Phi\right)^2 dV \\ \Rightarrow \frac{\delta \mathcal{L}}{\delta \Phi(x)} &= 0 \Leftrightarrow \nabla \frac{\partial \mathcal{L}}{\partial(\nabla \Phi)} - \frac{\partial \mathcal{L}}{\partial \Phi} = 0 \\ \Leftrightarrow \nabla \cdot (\sigma \nabla \Phi) &= 0. \end{aligned} \tag{105}$$

[22] Physical theories based on the assumption that some quantity assumes an extreme value have been most successful in all fields of physics. Classical mechanics, quantum mechanics, quantum field theory and equilibrium statistical mechanics are all described by finding the extreme points of some quantity.

Physically, one finds $\mathcal{L} = \int_V \vec{j}\vec{E}dV = \int_V \dot{q}dV = \dot{Q}$, the heat production. This means that Kirchhoff's law is equivalent to a law of minimum heat dissipation. In isothermal systems, the minimum in heat production corresponds to a minimum in entropy production

$$\dot{Q} = T\frac{dS}{dt} = 0 \Leftrightarrow \frac{dS}{dt} = \frac{1}{T}\dot{Q} = 0. \tag{106}$$

Thus, one can conclude that Max-EP cannot be valid near equilibrium, which is the domain of validity of Prigogine's Principle[23]. This example shows that Max-EP can at best be applied to systems with a **sufficiently large number of degrees of freedom**. In other words, Max-EP is only applicable to systems that are under-constrained with respect to basic physical laws, such as energy, charge and (in classical systems) mass conservation. The notion of "sufficient complexity" has been the object of some controversy [10, 40]. This is probably the reason why Max-EP is primarily used in climate and vegetation models, which have a myriad of steady states compatible with basic physical conservation laws.

3.2 Applications

3.2.1 The Simplest Example

In [27] Kleidon uses an example to illuminate the transition from applicability to non-applicability of Max-EP. He uses the notation convention

$$\frac{dS}{dt} = \sigma - NEE, \tag{107}$$

where the net entropy exchange (NEE) is positive if entropy is exported to the environment. By definition, the stationary state is then given as $\sigma = NEE$. In Fig. 11, Kleidon compares two setups: In the first, heat is transferred gradually through a closed system, permitting only heat conduction. In the second, heat can additionally be lost from the first bath. A

[23] Jaynes [25] points out the following subtlety: If the circuit is not at one temperature, minimum entropy production is no longer equivalent to minimum heat production (cf. Eq. 106). Maes and Netočný [32] explain that the additional thermal gradient does not couple to the electric current, but still enters the entropy production formula. According to their article, this effect is non-linear and therefore beyond the scope of Prigogine's Principle.

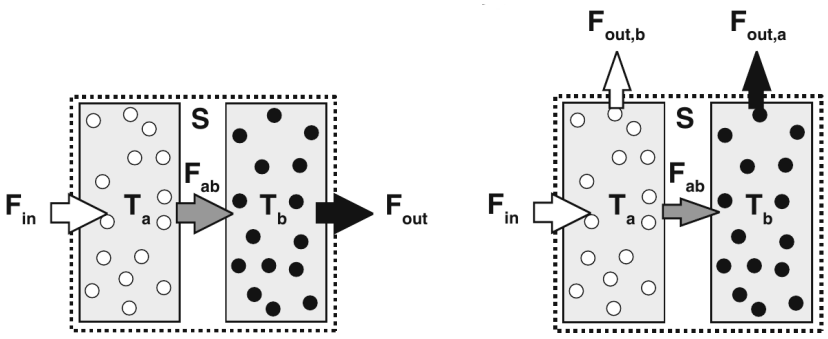


Figure 11: Left: heat can only enter the system in compartment a and leave it in compartment b, Right: heat can be dissipated in many ways (convection, friction, radiation, etc.) [27].

state of maximum entropy production exists only in the second case, where the dynamics are not fully dictated by heat conduction. Kleidon argues that the introduction of the extra flows $F_{out,a}$ and $F_{out,b}$ is equivalent to having a variable conductivity between the boxes. The system's additional complexity is thus absorbed into the effective conductivity, turning the constant κ into a variable. The maximum of the entropy production rate σ_{ab} coincides with

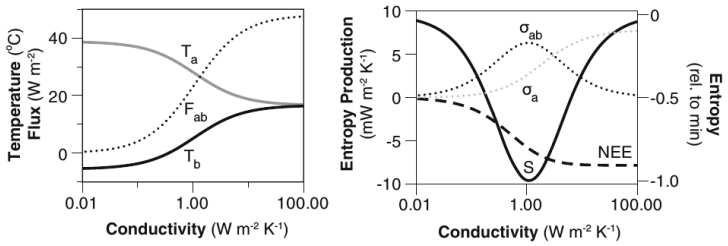


Figure 12: A maximum in entropy production appears. [27]

a state of minimum entropy relative to all other steady states. This might not be the result of a fundamental general principle, even though Kleidon ventures speculations along these lines.

Next, consider Kleidon's derivation in more detail: Eq. 107 says that a system's "entropy budget" is influenced by its internal entropy production as well as its exchange of entropy with the environment. The key aspect about

his example is the existence of an uncontrolled parameter (the conductivity, Fig. 11). Physically, this could account for convection between the boxes A and B. The energy flow and corresponding temperature change are modeled by (notation from Fig. 11 (right)):

$$\begin{aligned} c \cdot \frac{dT_a}{dt} &= F_{in} - F_{ab} - F_{out,a} \\ c \cdot \frac{dT_b}{dt} &= F_{ab} - F_{out,b}, \end{aligned} \tag{108}$$

where $F_{out,a/b}$ are **fixed functions** of the respective temperatures. Kleidon assumes a linear dependency with constants taken from atmospheric measurements

$$F_{out} = a + bT, \quad a = 12 \left[\frac{W}{m^2}\right], \quad b = 2.17 \left[\frac{W}{m^2 K}\right] \tag{109}$$

The system's exchange of entropy with the environment (NEE) is then given by

$$NEE = \frac{F_{out,a}}{T_a} + \frac{F_{out,b}}{T_b} - \frac{F_{in}}{T_{in}}, \tag{110}$$

where T_{in} is the assumed temperature of a heat bath from which the heat flow F_{in} originates. Internally, there are two sources of entropy production: a) heat exchange between A and B

$$\sigma_{ab} = F_{ab} \cdot \left(\frac{1}{T_b} - \frac{1}{T_a}\right) \tag{111}$$

and b) the entropy of mixing the incoming heat flow with bath A

$$\sigma_a = F_{in} \cdot \left(\frac{1}{T_a} - \frac{1}{T_{in}}\right) \tag{112}$$

Finally, **it is the plain assertion of Max-EP that lets the system choose the (existing) state of maximum entropy production**. In other words, Kleidon provides neither physical nor mathematical backing for the theory. His example simply illustrates the key feature necessary, such that a Max-EP state exists and the principle may be applied. However, Kleidon does not address why σ_{ab} is the entropy entropy production to which the Max-EP principle is to be applied. Another (possibly more plausible) choice would be $\sigma_{ab} + \sigma_a = NEE$, the system's total entropy production. This ambiguity of the system's boundary conditions is a well-known issue [52, 28, 40].

3.2.2 Climate Prediction

Given the many problems and misconceptions connected to Max-EP, why is any effort devoted to its application and theoretical development? It appears to be tantalizing predictive success that has won Max-EP its many supporters. The principle was initiated by climatologist Garth Paltridge who sought a thermodynamic principle showing the general features of earth's climate [41]. The example given above is a strongly simplified version of the model he developed and the maximization principle he applied. I will present a more detailed overview of this theory as given by Paltridge and Ozawa [41, 40], which seems to be Max-EP's first and most striking triumph.

In essence, earth's global climate is characterized by turbulent flows which are associated with some production of entropy. The exact derivation of entropy production in turbulent flows is given in Landau and Lifshitz

$$\dot{S}_{turb} = \int_V \frac{1}{T}\left[\frac{\partial(\rho c T)}{\partial t} + \nabla(\rho c T \vec{v}) + p\nabla\vec{v}\right] dV + \int_A \frac{F}{A} dA, \tag{113}$$

where the first term describes the rate of internal entropy change, while the second accounts for entropy transfer across the system's boundaries. In a steady state, only the boundary term is non-zero, giving Max-EP the specific form

$$\dot{S} = \int_A \frac{F}{T} dA = \text{maximum}. \tag{114}$$

Paltridge divides the earth up into ten boxes, each of which is characterized by three variables: surface temperature T, partial cloud cover θ and meridional (i.e. north-south) heat flow F_m. T and θ determine the average vertical transport of heat by radiation into space via empirical functions. Of course, this system of 30 unknowns is highly under-determined. Paltridge finds remarkably good agreement with experimental data (cf. Fig. 13) by maximizing

$$\sum_{i\in\{\text{boxes}\}} \frac{F^{atm}_{long,i} - F^{atm}_{short,i}}{T^{atm}_i}, \tag{115}$$

where F_i are the emission (absorption) rate of long (short) wavelength radiation at the top of the atmosphere and T^{atm}_i denote the average atmospheric temperature of box i.

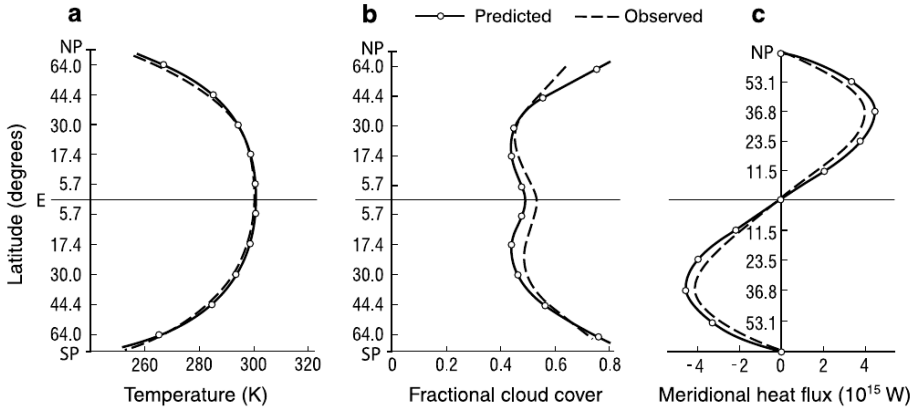


Figure 13: Three important quantities from earth's climate, comparison Max-EP prediction (solid lines), observed values (dashed lines) [41]

Technically, there should be an atmospheric as well as a ground/oceanic contribution to entropy production. The atmosphere receives energy flows both from outer space as well as from the ground/ocean system, such that the entire entropy production reads

$$\begin{aligned}\dot{S} &= \dot{S}^{atm} + \dot{S}^{go} \\ &= \int_A \frac{F^{atm}_{long} - F^{atm}_{short} - F^{go}_{long} + F^{go}_{short}}{T^{atm}} dA + \int_A \frac{F^{go}_{long} - F^{go}_{short}}{T^{go}} dA \\ &= \int_A \frac{F^{atm}_{long} - F^{atm}_{short}}{T^{atm}} dA + \int_A \left[F^{go}_{short} - F^{go}_{long}\right] \left(\frac{1}{T^{atm}} - \frac{1}{T^{go}}\right) dA.\end{aligned} \tag{116}$$

Eq. 115 contains only the first term from this last equation. It represents the entropy production due to horizontal heat transfer[24]. The second term represents vertical heat transport and is non-zero if $T^{atm} \neq T^{go}$. Noda and Tokioka [37], however, derive quite similar results just from $\dot{S} = \max$. In other words, the second term is negligible compared to the first.

A central objection to Paltridge's approach raised by Essex [16] is that the predominant part of total entropy production in the earth's system is due to the down-conversion of solar radiation in the atmosphere as well as

[24]The term describes the atmosphere's change in entropy due to emission and absorption. It only accounts for horizontal heat transfer indirectly, because horizontal fluxes maintain temperature differences on earth.

on the ground. The thermodynamic relationship $dQ = TdS$ used above is not valid for the entropy contained in a radiation field. Even if the same amount of heat is absorbed and emitted by the earth, the particular spectral composition determines an additional entropy contribution, which is too large to be neglected in Paltridge's climate model. This gives rise to a general question. Which of the following forms of entropy should be considered: entropy produced in turbulent flows, entropy changes of the radiation fields or both? Quite similar to Kleidon's simple example above, there is some vagueness in the system's boundaries. Paltridge only considers the earth, but offers no argument why the sun should be excluded[25] (cf. Sect. 3.2.1 for references).

3.3 Dewar's Derivations

Many attempts have been made to show the validity of Max-EP from a general principle. The approach I most often encountered is to derive Max-EP from Jaynes' MaxEnt and MaxCal. Interpreting Jaynes, Virgo [52] simply claims that entropy should be maximized and that this practically is equivalent to maximizing entropy production. To this end, he equates thermodynamic and information entropy, without offering any proof for this step. As Virgo is primarily concerned with climate prediction, a lot of attention is devoted to the system boundary question, i.e. which (sub-)system the principle must to be applied to. Whereas correct predictions are obtained if the sun is excluded, it is hard to argue that the climate system is in any sense closed. I do not find Virgo's argument convincing that the source of heat is arbitrary and can therefore be excluded.

The most promising attempts to derive Max-EP from MaxEnt have - according to [52] and [35] - been published by Dewar [9, 10, 6, 8, 7]. I will outline his two fundamental approaches, providing comments and references to criticizing articles along the way.

3.3.1 Heuristic Derivations

In two early papers [7, 10], Dewar attempts to portray non-equilibrium physics as a search for the most probable state. Dewar argues heuristically that this state is given by Max-EP. In this context, he even describes "dissipative structures" (life etc.) as the most probable states.

[25] Including the sun would give a more closed system.

Fig. 14 illustrates the relation between macroscopic reversibility and microscopic phase space conserving time evolution (Liouville equation). Each box encompasses all microscopic configurations in phase space representative of one macroscopic state A. If the transition $A \to B$ is reproducible, one can conclude that the phase space volume $TA \subseteq B$[26]. Let $W(X)$ denote the phase space volume of X. Because of Liouville's Theorem one knows

$$W(X) = W(TX). \tag{117}$$

The second law can then be described intuitively from $S(X) = k_b \ln |W(X)|$

$$\begin{aligned} & TA \subseteq B, \quad W(TA) = W(A) \\ \Rightarrow\ & |W(A)| \leq |W(B)| \\ \Rightarrow\ & S(A) \leq S(B) \end{aligned} \tag{118}$$

In this formalism, the (time-averaged) entropy production connected with one path Γ is immediately given as

$$\sigma_\Gamma = \frac{1}{\tau}(S(\tau) - S(0)) = \frac{k_b}{\tau} \ln \frac{W(\Gamma(\tau))}{W(\Gamma(0))} \tag{119}$$

where the notion of "phase space volume" is extended to all points along the trajectory Γ. Dewar now considers a direct decomposition of a system into

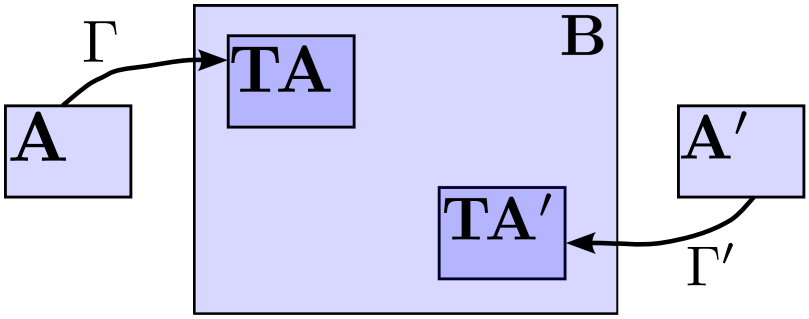


Figure 14: Macroscopic reproducibility in phase space

an open system and the environment. The environment is vaguely defined as "that part of the entire system detailed knowledge of which is assumed to be

[26] if the extremely small minority of non-reproducible phenomena is excluded

irrelevant to the macroscopic state of the open system". Consider now the evolution of one entire macrostate A (system $X(A)$ and environment $E(A)$) over an interval τ. Dewar focuses solely on steady-state systems, such that $X(A)$ cannot change over time. As a consequence, all states evolving into B have the same open system X and the same entropy production rates σ_X; only their environments can change. Because the time evolution operator is phase space conserving[27], the number of possible "environmental histories" is given by

$$N_X = \frac{W(B)}{W(A)} = \exp\left(\frac{\tau\sigma_X}{k_b}\right), \tag{120}$$

where A is one of the possible states defined above and the last equality is a direct result of Eq. 119.

In the end, Max-EP aims at selecting one state B. To find the most probable state, compare any two states X and Y. Without further knowledge about the environment, it seems reasonable to select the state with the greater number of compatible histories. In the language of dynamical systems theory, this would be the state representing the larger basin of attraction:

$$\frac{N_X}{N_Y} = \exp\left(\frac{\tau(\sigma_X - \sigma_Y)}{k_b}\right). \tag{121}$$

The distribution function of all possible macrostates is - in the limit of large τ - very sharply peaked at the state of maximum entropy production.

There are two problems with this argumentation: First and foremost, entropy production in Eq. 119 is defined between two *equilibrium* states, but Dewar uses the expression for non-equilibrium systems. Outside equilibrium, thermodynamic entropy lacks a clear definition and the elegant derivation of Eq. 119 breaks down. Secondly, Dewar uses Liouville's theorem and the stationarity constraint to claim that two systems A and A' must have the same phase space volumes. However, only the open systems X are steady states, while the environments necessarily evolve in time. There are two logical alternatives: $W(A) \neq W(A')$ or Liouville's theorem is applied to an **open** system, each leading to a contradiction in Dewar's derivation.

[27] i.e. it represents an injective map $\forall(b, b') : T(b) = T(b') \Rightarrow b = b'$

3.3.2 Quantitative Derivations

In contrast to the very general derivation in Sect. 3.3.1, Dewar also attempts to formulate a precise mathematical framework for his theory. For that purpose, he employs two very different ideas in his papers [6, 11], which are sketched below.

A Derivation Based on the Entropy Production Along a Path [6]: Consider a general open system ($\{V, \Omega\}$) in the grand-canonical ensemble, i.e one that can exchange mass and energy with its surroundings. If the system consists of different chemical constituents, it is useful to define

$$\begin{aligned} \mathbf{d} &= (u, \{\rho_i\}) : \text{ energy density and mass densities} \\ \mathbf{F} &= (\mathbf{f_u}, \{\mathbf{f_i}\}) : \text{ corresponding volume flows} \\ \mathbf{F}^n &= (f_u^n, \{f_i^n\}) : \text{ corresponding flows, normal to boundary } \Omega \end{aligned} \tag{122}$$

Let $\langle X \rangle$ denote an ensemble average and $\bar{X}$ a time average. In a first step, Max-Ent is applied to incorporate some fixed average values $\langle d(x,0) \rangle$ and $\left\langle \bar{\mathbf{F}}^n(x) \right\rangle$. In physical terms, Dewar fixes the system's initial state and the average flows across its boundaries. By the usual method, this gives a Gibbs-like distribution with spatially varying Lagrange multipliers

$$p_\Gamma = \frac{1}{Z} \exp\left(\underbrace{\int_V \lambda(x) \mathbf{d}(x,0)_\Gamma + \int_\Omega \eta(x) \bar{\mathbf{F}}^n(x)_\Gamma}_{A_\Gamma} \right) \tag{123}$$

Imposing spatially inhomogeneous constraints generalizes the method from taking derivatives to taking variations

$$\begin{aligned} \frac{\delta \ln Z}{\delta \lambda(x)} &= \langle \mathbf{d}(x,0) \rangle \\ \frac{\delta \ln Z}{\delta \eta(x)} &= \left\langle \bar{\mathbf{F}}^n(x) \right\rangle . \end{aligned} \tag{124}$$

Dewar allows both for the interconversion of particle species as well as including sources $\mathbf{Q}$ of particles and energy. Mathematically, this can easily be implemented by a conservation (balance) law

$$\frac{\partial \mathbf{d}(x,t)_\Gamma}{\partial t} = -\nabla \mathbf{F}_\Gamma + \mathbf{Q}. \tag{125}$$

To incorporate these terms, Dewar uses the fundamental theorem of calculus

$$\begin{aligned}\mathbf{d}(x,0)_\Gamma &= \frac{1}{2}\left[(\mathbf{d}(x,0)_\Gamma + \mathbf{d}(x,\tau)_\Gamma) - (\mathbf{d}(x,\tau)_\Gamma - \mathbf{d}(x,0)_\Gamma)\right] \\ &= \frac{1}{2}\left[(\mathbf{d}(x,0)_\Gamma + \mathbf{d}(x,\tau)_\Gamma) - \tau\cdot\frac{1}{\tau}\int_0^\tau dt\frac{\partial \mathbf{d}(x,\tau)_\Gamma}{\partial t}\right] \\ &= \frac{1}{2}\left(\mathbf{d}(x,0)_\Gamma + \mathbf{d}(x,\tau)_\Gamma\right) - \frac{\tau}{2}\cdot\overline{\frac{\partial \mathbf{d}(x,t)_\Gamma}{\partial t}} \\ &= \frac{1}{2}\left[\mathbf{d}(x,0)_\Gamma + \mathbf{d}(x,\tau)_\Gamma\right] - \frac{\tau}{2}\left(-\nabla\overline{\mathbf{F}}_\Gamma + \overline{\mathbf{Q}}\right).\end{aligned} \tag{126}$$

Substituting this into A_Γ from Eq. 123 gives

$$\begin{aligned}A_\Gamma = \frac{1}{2}\int_V \lambda(x)\left[\mathbf{d}(x,0)_\Gamma + \mathbf{d}(x,\tau)_\Gamma\right] - \frac{\tau}{2}\int_V \lambda(x)\left(-\nabla\overline{\mathbf{F}}_\Gamma + \overline{\mathbf{Q}}\right) \\ + \int_\Omega \eta(x)\overline{\mathbf{F}^n}(x)_\Gamma.\end{aligned} \tag{127}$$

Integration by parts shifts derivatives in the second term, which can then (by virtue of the divergence theorem) be absorbed into the boundary integral

$$\begin{aligned}A_\Gamma = \frac{1}{2}\int_V \lambda(x)\left[\mathbf{d}(x,0)_\Gamma + \mathbf{d}(x,\tau)_\Gamma\right] - \frac{\tau}{2}\int_V \left(\overline{\mathbf{F}}_\Gamma\nabla\lambda(x) + \lambda(x)\overline{\mathbf{Q}}\right) \\ + \underbrace{\int_\Omega \left(\frac{\lambda\tau}{2} + \eta(x)\right)\overline{\mathbf{F}^n}(x)_\Gamma}_{\text{neglected}}\end{aligned} \tag{128}$$

Dewar continues by dropping the last term, arguing that the necessary "information" was also included in the second term. I cannot quite agree with this, as this step completely throws η out of the derivation. In fact, neglecting the term is equivalent to demanding $\frac{\lambda\tau}{2} = -\eta(x)$, which would eliminate one of the Lagrange multipliers. These multipliers must be independent to account for two constraints. The key step in Dewar's derivation is to **define the remaining Lagrange multiplier λ as**

$$\lambda(x) = \frac{1}{k_b T(x)}(-1, \{\mu_i(x)\}). \tag{129}$$

This assumption is well-defined in the equilibrium limit $\mathbf{F}_\Gamma = \mathbf{Q} = 0$. Outside equilibrium, defining local temperatures and chemical potentials relies on

the **assumption of local equilibrium**, which is feasible, but restricts the claimed universality of the results. Using this definition in Eq. 128, one finds

$$A_\Gamma = \frac{1}{2}\int_V \frac{\sum_i \mu_i \left(\rho_i(0)+\rho_i(\tau)\right) - \left(u(0)+u(\tau)\right)}{k_b T} + \overbrace{\frac{\tau}{2k_b}\sigma_\Gamma}^{A^{irr}}, \tag{130}$$

$$\sigma_\Gamma = \int_V \left[\overline{\mathbf{f_u}}\nabla\frac{1}{T} - \frac{1}{T}\overline{Q}_u - \sum_i \overline{\mathbf{f_i}}\nabla\frac{\mu_i}{T} + \frac{\mu_i}{T}\overline{Q}_i\right]. \tag{131}$$

The expression σ_Γ is equivalent to Prigogine's derivation of entropy production. In some sense, Dewar's derivation can be considered a paraphrase of Prigogine's work as both are based on conservation laws. The new aspect is that Dewar takes a Gibbs-like distribution provided by Jaynes' algorithm as his starting point.

Dewar's final step is to note that in Eq. 130, the first part is time-reversal symmetric, while the second is anti-symmetric. Every symmetric contribution having an inverted contribution canceling it, only the anti-symmetric part is assumed to contribute

$$Z = \sum_\Gamma \exp(A_\Gamma) \approx W(A^{irr}) \exp A_\Gamma \tag{132}$$

$$S = -\sum_\Gamma p_\Gamma \ln p_\Gamma = \ln Z - \langle A \rangle \tag{133}$$

$$\Rightarrow S \approx \ln W(\langle A^{irr} \rangle), \tag{134}$$

where W denotes the multiplicity of all paths with action A_Γ and the second line is demonstrated in detail in Chapter 2[28]. Thus, the maximum of S coincides with a maximum of σ **if** W **is an monotonically increasing function in** A^{irr}. A basis for this assumption is not given. In fact, Bruers et. al [4] find a counter example to this assumption. In summary, the above derivation is founded upon the problematic assertion of local equilibrium, which is required to postulate Eq. 129. The same issue flaws a second and very similar paper by Dewar [8]. Again, Dewar does not admit that proximity to equilibrium is a premise to his derivation, but this deficiency is addressed in [4, 20].

[28]Dewar makes no clear distinction between thermodynamic entropy S and information entropy H. While $S = k_b H$ at equilibrium, the existence of a correspondence in non-equilibrium is, in fact, not clear.

A Derivation Based on Kullback-Leibler Divergence [11]: In this derivation, Dewar assumes a system which is a) in a steady state, b) far from equilibrium and c) has several states compatible with the constraints ("instability condition"). Naturally, this derivation applies the general strategy of finding the extreme point of the information entropy functional. Interestingly, Dewar makes a statement about the distance from equilibrium his central constraint. Distance in the space of probability distributions is typically measured by the Kullback-Leibler divergence[29]

$$I = \int p(\mathbf{f}) \ln \frac{p(\mathbf{f})}{p(-\mathbf{f})} d\mathbf{f}, \tag{135}$$

where the notation is adopted from above. I is positive and vanishes only if $p(\mathbf{f}) = p(-\mathbf{f})$. Because this is the case only in a state of equilibrium, I measures the distance from equilibrium. With this definition, Dewar finds it intuitive to demand that the distance from equilibrium should remain within certain bounds

$$I_{min}(C) < I \leq I_0, \tag{136}$$

where $I_{min}(C)$ is the (close) distance from equilibrium, at which only one compatible steady state exists. I_0 is an arbitrary distance in the range $I_{min} < I_0 \leq I_{max}$, introduced as an auxiliary variable in the proof. To highlight this new constraint, I will leave out standard mean-value constraints as they are used above.

In maximization problems, inequalities are implemented as equalities, with the exception that the corresponding Lagrange multiplier is set to zero if the constraint is satisfied as an inequality. In other words, the constraint only applies if the solution is on the boundary of the "area" prescribed by the inequality:

$$\begin{aligned} \int d\mathbf{f} \frac{\delta}{\delta p(\mathbf{f}')} \left[-p(\mathbf{f}) \ln p(\mathbf{f}) + \mu \left(p(\mathbf{f}) \ln \frac{p\mathbf{f}}{p(-\mathbf{f})} - I_0 \right) \right] = 0 \\ \Rightarrow \ln p(\mathbf{f}) \propto \mu \cdot \left(\ln p(\mathbf{f}) - \ln p(-\mathbf{f}) - p(\mathbf{f}') \frac{1}{p(-\mathbf{f}')} \delta(\mathbf{f}' - (-\mathbf{f})) \right) = 0 \\ \Rightarrow \ln p(\mathbf{f}) \propto \mu \left(\ln d - \exp(-d) \right), \end{aligned} \tag{137}$$

[29] Actually, the Kullback-Leibler divergence satisfies only one (semi-definiteness) out of three properties (semi-definiteness, symmetry, triangle inequality) required of a norm. It is still the generally accepted measure in the space of probability distributions.

where $d = \ln p(\mathbf{f})/p(-\mathbf{f})$. For the constraint to take effect $N \neq 0$ and it is reasonable to assume $\mu = \partial S/\partial I_0 > 0$[30]. The strict inequality chosen as the constraint would otherwise be inactive. One can directly infer

$$\mu = 0 \Rightarrow I = I_0. \tag{138}$$

The last step is to take the limit $I_0 \to I_{max}$. The hypothetical limiting case $I_0 > I_{max}$ leads to $\mu = 0$. Assuming that μ is continuous, $\mu \to 0$ as $I_0 \to I_{max}$ from below. As a result, μ drops from the MaxEnt solution, which only contains the (omitted) standard terms originating from mean value constraints. In a physical interpretation, this means that MaxEnt predicts a system to "choose" the state corresponding to $I_0 = I_{max}$. However, the result is more or less directly contained in the assumptions that were made:

- Distance from equilibrium is measured in terms of the Kullback-Leibler divergence, i.e. it **has no direct thermodynamic interpretation**.
- $\mu > 0$: Entropy grows as the system is moved further from equilibrium.
- $I \leq I_{max}$: There is an upper bound to the distance from equilibrium.

Essentially, Dewar proves that MaxEnt selects the state with a maximum Kullback-Leibler divergence. This state must exist as he postulates an upper postulates an upper bound I_{max}. In conclusion, the derivation is self-contained, but it fails to connect to thermodynamic entropy production. Therefore, this derivation cannot be considered a complete proof of the Max-EP principle.

3.4 What Is To Be Thought About This?

The Principle of Maximum Entropy Production is a current research topic, which, at this point, faces a number of significant issues that prevent it from becoming an established theory.

First of all, the exact statement of Max-EP is not generally agreed upon. Paltridge provides the most prominent example: His interpretation of the theory can be used to successfully make predictions about the earth's climate system. Paltridge uses his conjecture of Max-EP to infer a realized

[30]Greater distance from equilibrium is assumed to result in a larger amount of accessible states and, thus, greater entropy.

steady state from the set of all stationary states compatible with the constraints. However, several conceptual problems arise, which are discussed above and in the references. Ambiguities include the arbitrariness of the system's boundaries, the considered part ofthermodynamic entropy production (including or not including entropies of the radiation fields) and the neglect of virtually all physical factors central to earth's climate[31].

Secondly, Dewar picks up on Paltridge's central ideas in his heuristic attempts to derive Max-EP from MaxEnt. These derivations are founded on reversibility as the key property and try to infer Max-EP from very general considerations in phase space. Though heuristics cannot be considered sufficient in exact science, these approaches appear to capture some of Max-EP's very general nature. Dewar's quantitative derivations, however, fall short of that generality and contain technical weaknesses. Without stating it explicitly, the first derivation assumes local equilibrium and is therefore restricted to the near-equilibrium regime. In this range, however, Max-EP cannot be valid as Prigogine's Principle is in effect. The second derivation fails to establish a sound connection to thermodynamic entropy production and also appears to require very strong assumptions. Most gravely, Dewar derives MaxEnt from an information theoretical quantity, thus presenting a kind of circular reasoning failing to connect to the statement of Max-EP.

Dewar's stated hope is to demonstrate a similarity in spirit between Max-EP and Max-Ent in the sense that "[Max-EP] is the messenger, not the message". Assuming this conjecture holds, Max-EP would, on the one hand, gain justification for its failures in some systems and its glorious success in others, defying the huge amount of unknown information. MaxEnt, on the other hand, would be equipped with the possibility of a useful application to thermodynamic systems.

Dewar's and other researchers' many attempts have led to skepticism: Dewar sees that Max-EP's "apparent successes remain something of a curiosity, while the interpretation of its apparent failures is fraught with ambiguity" [11], while even Max-EP's initiator Paltridge no longer places much trust in it [42]. Despite the great challenges of an exact formulation and a derivation from a fundamental principle, Max-EP heralds great physical versatility. If applicable, it is a strong principle which claims to be valid for very complex systems and as well as being sufficiently general to cover an enormously broad

[31]This specific point is pointedly illustrated in [9], where Dewar discusses the predictions' implausible invariance if "the seas were made of vinegar".

spectrum of phenomena. Future research will have to clarify if Max-EP is a physical theory, an inference algorithm equivalent to MaxEnt or a lucky coincidence.

4 MaxCal: Evolution of a Model

4.1 Application of MaxCal to Diffusion

In the previous chapters, we discuss the connection between the inference theories MaxEnt and MaxCal and non-equilibrium steady-state theories. The relations are derived from MaxCal's general mathematical structure as well as an appropriate choice of constraints. Next, we move beyond these restrictions, by considering a problem that is a) far from equilibrium and b) not in a steady state. In the following, the idea, application and particular strengths of MaxCal are demonstrated for an example of a more complex system, namely diffusion of particles in a tube.

First, the diffusion system under consideration is implemented in a computational model. Results from this simulation will serve as "experimental data" to be utilized as constraints in a MaxCal framework. Second, we develop a coarse-grained mathematical model adequate for a MaxCal calculation. Combining model and constraints, MaxCal is capable of making predictions about other dynamical quantities or higher moments of probability distributions. Finally, predictions are compared with the simulated results. Thus, we demonstrate the whole cycle of a typical application of MaxCal (Fig. 15).

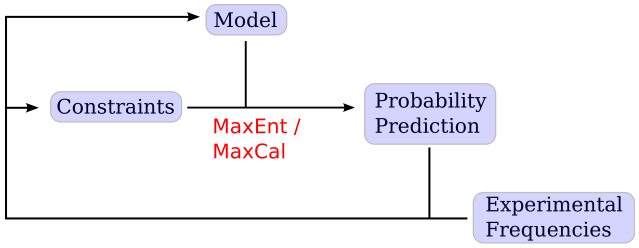


Figure 15: MaxCal is a theory that faithfully converts constraints into a probability assignment.

4.2 Simulation of Diffusion

Consider a very general diffusive system, consisting of a one-dimensional tube of length L filled with some solvent. Colloidal particles are inserted at the left end of the tube at $t = 0$. Due to Brownian (thermal) motion, the colloids

will diffuse across the tube. After some time, the particles have given rise to currents J

$$J(x) \equiv N(x' > x, t_f) - N(x' > x, t_i). \tag{139}$$

where we only consider $x \in \{1/4L, 1/2L, 3/4L\}$. A very similar setup is studied experimentally by Seitaridou et al. [48]. Instead of performing experiments in the laboratory, we simulate them on the computer. The main advantage lies in great speed as well as flexibility regarding the parameters under consideration. A standard mathematical approximation for colloidal dynamics is the **Langevin equation**

$$ma(t) = F(t) - \gamma m v(t) + R(t), \tag{140}$$

where $F(t)$ is the external (or inter-molecular) force on one molecule, γ is a friction parameter and $R(t)$ represents the stochastic motion of the water molecules. Assuming that the thermal motion is uncorrelated, it is given by a Gaussian distribution with $\mu = 0$ and standard deviation [46]

$$\sigma = \sqrt{2k_B T \gamma m \delta(t - t')} \approx \sqrt{\frac{2k_B T \gamma m}{\Delta t}}. \tag{141}$$

Table 1 gives the set of parameters used in the simulations[32].

Eq. 140 is solved numerically using the BKK integrator for stochastic differential equations [5]. A detailed description of the implementation is given in Appendix 5.3.1. Fig. 16 shows a typical set of trajectories in the system. Because of the fixed boundaries, particles are reflected off the tube's ends.

4.3 MaxCal Model of Diffusion

Instead of exploiting the entire information about all trajectories directly, assume that the experiment (here replaced by a simulation) only provides very limited information. Hypothetically speaking, the system might be too small to make accurate measurements on $J(1/4L)$ and $J(3/4L)$, but give an

[32]Depending on the interpretation of these computational units, they can represent various real-world experiments.

Parameter	Value
No. particles N	10
m_A	1
Tube length l	100
Simulation time t_f	300
k_b	1
Temperature T	1
γ	1
Boundary conditions	fixed
No. repetitions	1000

Table 1: List of parameters used in the simulations

average value $\langle J(1/2L)\rangle$. In the following, a model is formulated to treat the physical problem of diffusion in mathematical terms, which can be solved by MaxCal. As a first approximation, the system can be divided into two parts, corresponding to the tube's right and left half. Because the particles are distinguishable[33], a trajectory is defined as any combination of particles "hopping" from left to right. This model is called the "dog-flea model" [43, 18]. As demonstrated in Appendix 5.2, the partition function reads

$$Z = (1 + e^{\lambda})^N. \tag{142}$$

Due to the binomial coefficients occurring in the expansion of the partition function the distribution underlying this partition function is identified as the binomial distribution (cf. Appendix 5.2)

$$\Rightarrow J = \mathrm{Bin}(J, p, N) = \binom{J}{N} p^J (1-p)^{N-J}, \quad p = \frac{\langle J \rangle}{N}, \tag{143}$$

where the last equation converts the measured average current into a "hopping probability" on the single-particle level. Based only on the mean value constraint, this model provides a distribution function, i.e. it gives all higher moments of the distribution. At this point, note that the model has no means of accessing other valuable insights. For example, currents $J(1/4L)$ and $J(3/4L)$, both non-existent in the model, may be physically relevant and information about them could be desirable. Also, currents at other positions or times as well as other physical observables could be asked from the model.

[33] in contrast to the notion of indistinguishability from quantum mechanics

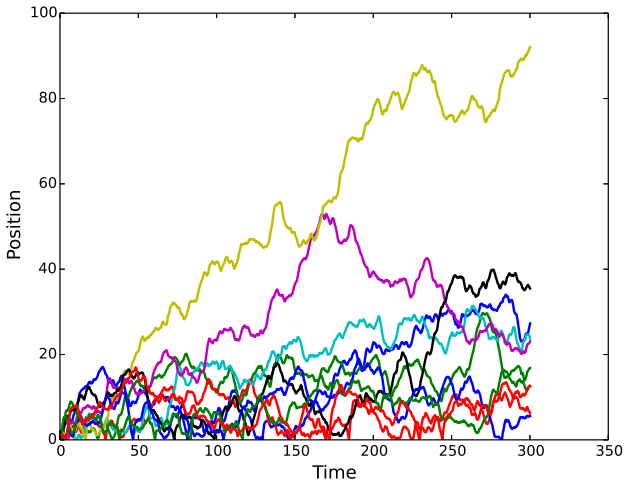


Figure 16: A typical trajectory of particles diffusing in the tube

Therefore, it is necessary to develop a new model which is sufficiently comprehensive to predict currents $J(1/4L)$ and $J(3/4L)$. It will turn out that this model also predicts other quantities, including some rather remote ones.
The usual MaxCal procedure of enumerating all micro-trajectories and assigning weights to them is implemented in several steps:

1. List all states that the system can be in. Even though particles are assumed to be independent, states which are identical upon permutation of the particles are not listed separately to reduce computational effort.

2. For any two states s_i and s_j calculate the number of transitions between them. This accounts for the distinguishability of particles.

3. Assign MaxCal weights to each path. This is accomplished by inserting factors $\exp(\lambda)$ for each mean-value constraint into the formalism.

4. Calculate the transfer matrix.

5. Determine the Lagrange multipliers by inversion of the mean-value constraint equations.

6. Evaluate all desired quantities. A partition function is useful for collective properties, such as higher moments of the probability distribution function. For other purposes, one can project out certain states.

To illustrate the procedure as well as to assure that the algorithm works correctly, consider first the case of $N_b = 3$ boxes and $N_p = 2$ particles. This case in non-trivial and can, unlike the case of the simulation $N_b = 4, N_p = 10$, still be easily counted by hand.

Step 1: The particles from the simulation are distinguishable. They represent large particles, such as colloids, whose wave functions do not overlap. For N_p particles and N_b boxes, there are $N_b^{N_p}$ states of the system. Even though the particles are in principle distinguishable, all constrained quantities (currents, numbers of particles per box, etc.) do not depend on the identity of each particle. In the given example, there are only six such states. They are represented pictorially in Fig. 17.

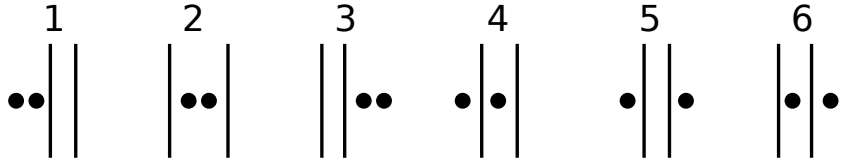


Figure 17: State classes of the system

Each symbol represents a class of

$$\frac{N!}{\prod_i n_{p_i}!} \tag{144}$$

states, where n_{p_i} is the number of particles in box i. In general, the number of different classes is

$$N_C = \frac{(N_b - 1 + N_p)!}{(N_b - 1)! N_p!}. \tag{145}$$

The division into classes of states brings about an enormous simplification in computational terms. Taking $N_p = 20$ particles and $N_b = 10$, each class represents on average

$$\frac{N_b^{N_p}}{(N_b - 1 + N_p)!/((N_b - 1)! N_p!)} \approx 10^{13} \tag{146}$$

states. Remember that the simplified notation is only possible due to the kinds of physical constraints imposed, it is not a fundamental consequence of quantum mechanics.
Step 2: The fundamental quantities are the distinguishable states, and the trajectories connecting them. To see how the trajectories can be counted from classes, observe the following:

- $s_k \in C_i$: Each state s_k belongs to one class C_i.
- $i \neq j : C_i \cap C_j = \emptyset$: The classes are disjoint. One state belongs to only one class.

Also, **transitions between states are reversible**, i.e. if there is a transition $s_k \to s_l$, the reverse $s_l \to s_k$ exists, too. Any permutation $P \in S_{N_p}$ can act on two elements s_k, s_f from C_i and C_j, respectively. It follows that $P(s_k) \leftrightarrow P(s_l) \Leftrightarrow s_k \leftrightarrow s_l$. A direct consequence of this is that the degree of every element in a class is identical. In summary

- $s_k \to s_l \Leftrightarrow s_l \to s_k$: reversibility
- $P(s_k) \leftrightarrow P(s_l) \Leftrightarrow s_k \leftrightarrow s_l$: permutation symmetry of classes
- $\deg(s_i) = \deg(s_j) \;\; \forall s_i, s_j \in C_k$: All elements in one class are essentially identical.

All these properties can be visualized. As the number of states grows rapidly, consider for illustration the overly simple case $N_p = 2$, $N_b = 2$ (Fig. 18). While the graph[34] of this small example is fully connected, this is usually not the case. Some states cannot be reached within one time interval τ. Each arrow in Fig. 18 represents one microtrajectory that must be counted. The colored circles indicate the classes. To count every arrow, one can consider every pair of two classes C_i, C_j. Because of reversibility one concludes $|C_i \to C_j| = |C_i \to C_j|$. The computer algorithm f described in Appendix 5.3.2 takes any two classes C_i, C_j and returns

$$f(C_i, C_j) = |\{s_j \in C_j : s_i \leftrightarrow s_j\}|. \tag{147}$$

f can be regarded as the partial degree of each node in C_i, which is not symmetric (cf. Fig. 18). For the more complex example above ($N_p = 2$,

[34]Technically speaking, it is a pseudo-graph as it contains loops.

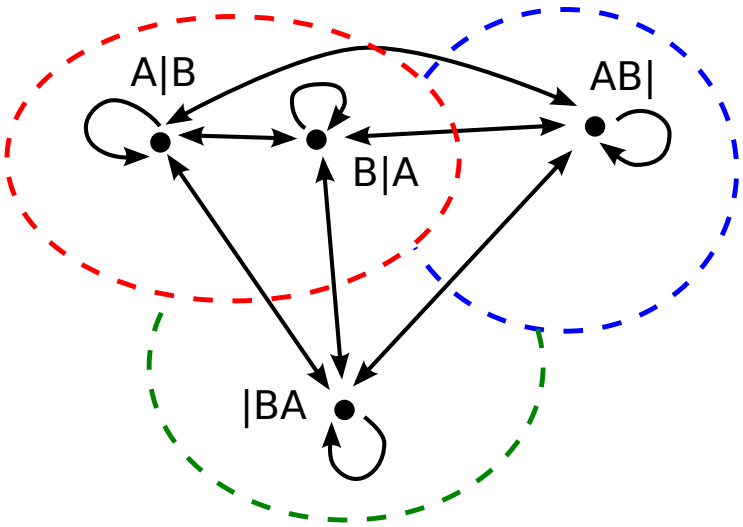


Figure 18: States, classes and trajectories for $N_p = 2$, $N_b = 2$

$N_b = 3$), the matrix for the classes reads

$$D_p = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 2 & 2 & 0 & 2 & 1 & 1 \\ 0 & 2 & 0 & 1 & 1 & 1 \\ 0 & 2 & 2 & 1 & 1 & 2 \end{pmatrix} \end{array}. \tag{148}$$

The total number of trajectories (arrows) M_{ij} between classes C_i and C_j can now be simply calculated:

$$M_{ij} = f(C_i, C_j) \cdot |C_i| = f(C_j, C_i) \cdot |C_j|. \tag{149}$$

In the example system one finds

$$M = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \begin{pmatrix} 1 & & & & & \\ 1 & 1 & & & & \\ 0 & 1 & 1 & & & \\ 2 & 2 & 0 & 4 & & \\ 0 & 2 & 0 & 2 & 2 & \\ 0 & 2 & 2 & 2 & 2 & 4 \end{pmatrix} \end{array}, \tag{150}$$

where the blank spots are symmetric to the upper triangle, $M_{ij} = M_{ji}$ (cf. Eq. 149). As far as computational implementation is concerned, Step 2 turns out to be the most expensive. Because the matrix has to be evaluated only once, this does not pose a problem.

Step 3: Assigning weights to the paths is the second important incorporation of physics (after forming a model). For simplicity, assume that only an average current across the border between boxes 1 and 2 (defined according to Eq. 139) is imposed. Each transition from state i to state j is associated with a microcurrent j_{kl}

$$j = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \begin{pmatrix} 0 & & & & & \\ 2 & 0 & & & & \\ 2 & 0 & 0 & & & \\ 1 & -1 & -1 & 0 & & \\ 1 & -1 & -1 & 0 & 0 & \\ 2 & 0 & 0 & 1 & 1 & 0 \end{pmatrix} \end{array}, \tag{151}$$

the blank spaces being anti-symmetric to the lower triangle $j_{ij} = -j_{ji}$. The matrix of path weights is then given by $W_{kl} = \exp(\lambda j_{kl})$.

Step 4: Now, the transfer matrix can be computed by componentwise multiplication

$$T_{kl} = M_{kl} \cdot W_{kl}. \tag{152}$$

Step 5: Fig. 19 shows the functional relation between the current J and the Lagrange multiplier λ. For reasonably small λ, this function can clearly be

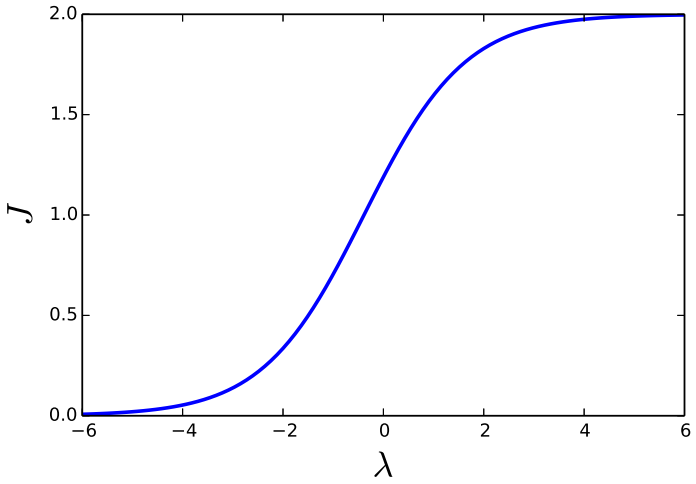


Figure 19: $J(\lambda)$ for the case of the system starting in $|C_i\rangle = 1$, after $t = 2$ time units. The current is clearly bounded: $0 \leq J \leq 2$.

inverted. For simplicity, use λ as the independent variable in the following.
Step 6: All relevant quantities can now be calculated from the transfer matrix, which maps the probabilities at time t_i onto state probabilities at $t_i + \tau$. Matrix powers reveal the system's complete temporal evolution. For t time steps

$$p_f = \langle C_f | \, T^t \, | C_i \rangle \,, \tag{153}$$

where $|C_i\rangle \,, |C_f\rangle$ are projectors of the initial and final classes. Fig. 20 shows the probability distribution for $t = 2$, $|C_i\rangle = 1$ and $\lambda = 1$.

4.4 First Results

Having introduced the model for a simple case, it will now be used to treat the simulation. Out of the 286 classes, the class representing all particles on the tube's left end is chosen as the initial configuration. As explained above, the matrix power t gives the number of transitions in the MaxCal model. It is assumed that the particles can only travel once across the tube, such that $t = 3$. This corresponds well with the distribution of particles

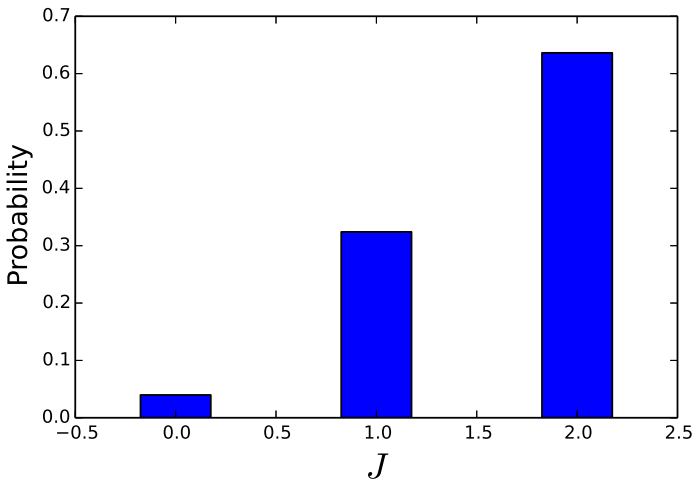


Figure 20: Current distribution for $t = 2$, $|C_i\rangle = 1$ *and* $\lambda = 1$.

after the simulations (Fig. 21)[35]. In the following, the short-hand notation $J_1 = J(1/4L), J_2 = J(1/2L), J_3 = J(3/4L)$ is used.

Again, one can calculate the function $J_2(\lambda)$ directly from the model (Fig. 22). As expected, the function increases from a minimum current $J_2 = 0$ from no particles jumping to all ten particles crossing the central line. From simulations one obtains $\langle J_2 \rangle = 0.547$ after $R = 1000$ repetitions of the experiment. Evidently, $J_2(\lambda) = \langle J_2 \rangle$ can be solved numerically and the calculation yields $\lambda_0 = -2.24$.

Next, one can compare the simulated distribution with MaxCal's prediction (Fig. 23). Numerical comparison between two probability distributions is challenging and can be done in various ways. From an information theoretical perspective, the decisive measure is the Kullback-Leibler divergence

$$D_{KL} = R \sum_i g_i \ln \frac{g_i}{p_i}, \tag{154}$$

where i enumerates all possible outcomes, g_i and p_i are the relative frequencies from the simulation and the MaxCal probabilities, respectively. In statistics, a more common choice is the χ^2-test, which is used to test if a

[35] If this information is not available, t is an additional (effective) parameter in the optimization problem.

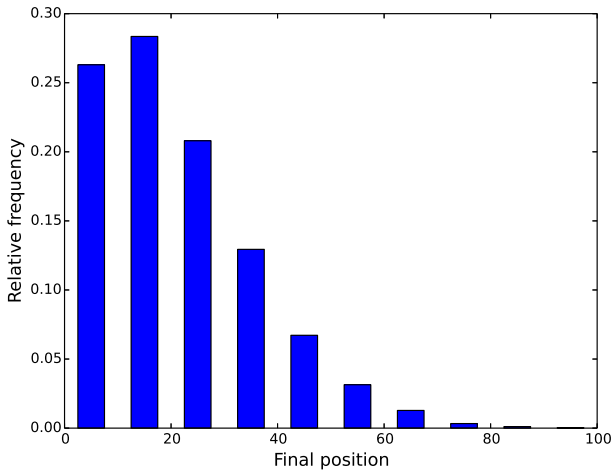


Figure 21: Simulated particle distribution after the full simulation

hypothetical distribution is approximated by experiment. The test statistic

$$\chi^2 = R \sum_i \frac{(g_i - p_i)^2}{p_i}. \tag{155}$$

is compared to a critical value[36] depending on the number of degrees of freedom f and the considered statistical significance α. In the problem of diffusion, $f = 11 - 1 - 1 = 9$ since there are eleven possible outcomes, reduced by the normalization constraint and one fit parameter λ. The χ^2 test is applied here as a heuristic, disregarding details such as a minimum number of events per bin as discussed in [22]. For completeness, the first moments of the distributions, mean $\langle\cdot\rangle$ and variance $\langle(\Delta\cdot)^2\rangle$, are also given. While only a few central values are discussed in the text, Table 2 contains all results. All measures show that MaxCal's prediction is excellent. Most intuitively, $\chi^2 = 3.3$ supports this conclusion.

Next, one calculates a prediction for J_1 (Fig. 24). All observables indicate that MaxCal's prediction is wrong. In particular, predictions of $\langle J_1\rangle$ and $\langle J_3\rangle$ exceed simulations by a factor of up to 5! Does this mean that the model is incorrect?

[36]The critical value can be obtained from standard tables.

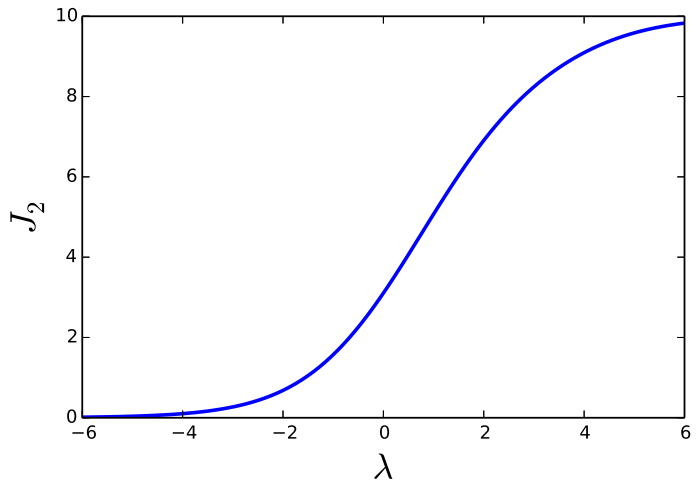


Figure 22: J_2 increases as a function of λ within the bounds $0 \leq J_2 \leq N$

Referring back to the conceptual nature of MaxCal as depicted in Fig. 15, one recognizes that there is the possibility of a missed constraint leading to incorrect predictions. In fact, the overly large MaxCal values for the unconstrained currents suggest that MaxCal does not take friction into sufficient account. So far, a trajectory is weighted based on how many particles cross the central separation. To emphasize this point, consider a tube filled with only one particle. Any trajectory of 4 steps, for example, that starts and ends in the same box would have the same weight, regardless of whether the particles remained in the box four times or jumped four times. In other words, friction is not considered in the model. In terms of simulations, the limit $\gamma \to 0$ reduces the Langevin dynamics to a random walk in one dimension. The effect of over-prediction is then eliminated. Jaynes phrased the discovery of new constraints in the following words [25]: "Comparing experimental observations with some existing theory, or calculation, one will never find perfect agreement. Are the discrepancies so small that they might reasonably be attributed to measurement errors, or are they so large that they indicate, with high probability, the existence of some new systematic cause?"

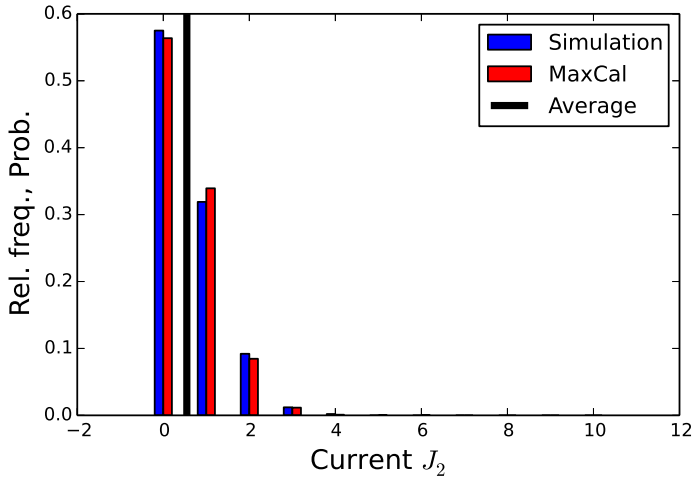


Figure 23: Without friction: Comparison of simulated frequencies and MaxCal's prediction for J_2

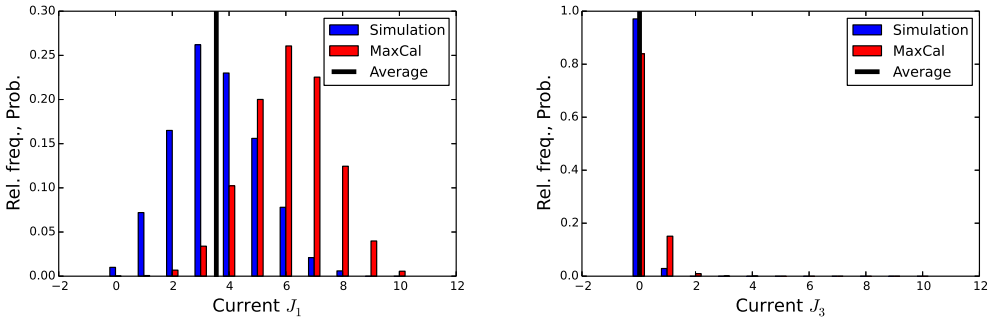


Figure 24: Without friction: Comparison of simulated frequencies and MaxCal's prediction for J_1 and J_3

4.5 Physical Constraint: Friction

Clearly, the next step is to include drag as a physical constraint into the model. MaxCal prescribes a factor e^{γ} for every jumping particle, where the name of the Lagrange multipliers is chosen to remind of the friction term in Eq. 140. However, the number of jumping particles can be *different* for two transitions between the *same* classes C_i and C_j. To some extent, this breaks the permutation symmetry of the groups: While the degree of each element in a class must still be identical, the weights of the paths connecting them need not be. In the calculation, the new constraint is implemented by changing the matrix of multiplicities M to include weights $e^{\gamma \cdot n}$ if n is the number of particles jumping during the transition. Both Lagrange parameters can be found by simultaneously solving the constraint equations

$$\frac{\partial \ln Z}{\partial \lambda} = \langle J_2 \rangle \quad \frac{\partial \ln Z}{\partial \gamma} = \langle n \rangle \,. \tag{156}$$

Instead of using the number of jumping particles as an immediate constraint, one can work directly with the parameter γ[37]. A trial and error procedure is now used to find as a first approximation $\gamma = -3$. This negative value suppresses trajectories with many jumps compared to those with few jumps by a factor $e^{\gamma \Delta n}$. Evaluating the partition function for $\gamma = -3$ and re-evaluating λ by local optimization leads to good agreement with the simulation: Due to this first estimate, prediction and simulation in J_1 are very close as far as the mean value is concerned. The predicted value of J_3 is now of the correct order of magnitude. In this case, the χ^2 heuristic even indicates that the predicted distribution is correct.

It may be objected that the MaxCal probability distribution for J_2 becomes less accurate after the introduction of friction. Intuitively, this can be understood by considering that the effective[38] number of trajectories available in the entropy maximization for J_2 is reduced, imposing an effective additional constraint upon J_2. In fact, the increased value of χ^2_2 is still within a plausible regime.

Having worked with the raw Lagrange multiplier γ, the essential consistency requirement is to check if Eq. 156 holds. The derivative can be

[37] The cross check is calculated below.

[38] While, of course, no trajectory is excluded, suppressing those with many jumps has a similar effect.

approximated as

$$\left.\frac{\partial \ln Z}{\partial \gamma}\right|_{\lambda',\gamma=-3} \approx \frac{\ln Z(-2.9,\lambda') - \ln Z(-3.1,\lambda')}{0.2} = 4.3. \tag{157}$$

To cross-check this result, one can simulate diffusion without boundaries: Extending the tube to a great length, particles are placed randomly into a section of a box's size $L = 25$ located very far from the boundaries. After $t = t_f/3 = 100$, the number of particles diffused out of the box is counted. One thousand repetitions result in $\langle n \rangle = 4.3$. This equivalence confirms that the friction constraint is physically justified and that the numerical value chosen is a good approximation. It might occur closer at hand to check the jumping statistics in the original simulation, including boundary effects. However, the jumping rate in the first third of the evolution is significantly lower than that in the next two thirds, because all particles start out on the left end of the tube. Related to this subtlety regarding the initial conditions, one can also understand why predictions improve along the tube: After all, the model of four boxes is too crude to resolve the particles' initial positions on the far left. This influences the statistics of J_1 most gravely, while fluctuations weaken the effect in J_2 and J_3.

4.6 Discussion and Outlook

In summary, the model developed above is capable of supporting a full MaxCal calculation. Through the MaxCal approach, the additional constraint of friction was "discovered", i.e. realized as relevant and then included in the partition function. After the consideration of friction, predictions are in reasonably good agreement with simulations.

One must take into account two aspects when assessing the quality of the method. On the one hand, the physical system is quite simple, lacking complicated interaction patterns which would have to be captured by the model. On the other hand, only two constraints were used to make predictions about **all moments of three probability distributions**. Also, the model is very coarse, consisting of just four boxes. As mentioned above, details about the particles' initial configurations (particularly position, but also velocity) are not accounted for. Introducing more boxes would clearly improve the results. Furthermore, the two Lagrange multipliers λ and γ should be solved simultaneously. One could either use a numerical scheme or seek

an iterative solution, by fixing one parameter and maximizing the other. For well-behaved equations, this procedure will converge to the correct solution if appropriate first values are chosen. Any systemic solution strategy would surely improve the numerical calculation beyond the first approximation.

It might seem that the mathematical formalism needed to arrive at theses results is rather cumbersome in the light of the model's simplicity. Why could one not assume a simple model, such as a fixed hopping probability of each particle, to make predictions? The conceptual issue with simplistic models is that they make ad hoc assumptions, which require a comprehensive understanding of the underlying physics. For example, in a model of a constant hopping probability p, one would have to assume a priori that the particles move independently. Then, the parameter p can, in principle, be calculated from the recursive relation

$$P(J_2, t') = \sum_{t<t'} \sum_{s_2, s_3} \left[s_2(t) P(s_2, t) - s_3(t) P(s_3, t) \right] p. \tag{158}$$

where s_i is the occupation of box i at time t and $P(s_i, t)$ is the corresponding probability function. Hopping probability being the only parameter, the model would be solved and allow predictions[39].

However, there is a fundamental difficulty if predictions do not match the simulation: There is no mathematically defined way of modifying the model to incorporate new physics. Consider, for example, an interaction between the particles in the form of repulsion. Eq. 158 loses it's validity together with the entire model. A next ad hoc model could be formulated, solved and compared to experiment. The example of friction above already demonstrates that MaxCal provides a means of improving a model in a mathematically unique way. To add evidence that MaxCal is generally able to incorporate new physics, here is how particle interaction can be implemented: Interactions can be understood as a change in potential energy for a number of particles in one region or box. As such, static interactions are best defined as properties of the states and their corresponding classes, rather than trajectories. Although MaxCal is founded upon the idea of trajectories, states retain a clearly defined meaning: states are the configurations of the system whose sequence forms a trajectory. If one assumes repulsion, one would, for example, introduce a Lagrange multiplier κ corresponding to the number of particles in one box. Mathematically, this would be easily implemented via

[39] Even this calculation is non-trivial without the use of computer methods.

	J_1	J_2	J_3	$\langle(\Delta J_1)^2\rangle$	$\langle(\Delta J_2)^2\rangle$	$\langle(\Delta J_3)^2\rangle$
Simulation	3.55	0.55	0.03	2.28	0.53	0.03
MaxCal, no fric.	6.08	(0.55)	0.17	2.01	0.5	0.16
MaxCal, fric.	3.4	(0.55)	0.02	1.24	0.44	0.02

	λ	γ	χ_1^2	D_1^{KL}	χ_2^2	D_2^{KL}	χ_3^2	D_3^{KL}
MaxCal, no fric.	-2.24	0	16455	1435	3.3	1.5	129	93
MaxCal, fric.	0.35	-3	479	120	28.9	11.2	7.95	3.35

Table 2: MaxCal indicates that friction is a relevant parameter in the system.

matrix multiplication

$$\begin{pmatrix} e^{\kappa V(C_1)} & & \\ & \ddots & \\ & & e^{\kappa V(C_{N_C})} \end{pmatrix} \cdot T^{\text{no interaction}}, \tag{159}$$

where $V(C_i)$ denotes the potential energy of each state in class C_i. This automatically assigns each class a correct weight based on its occupation number.

In conclusion, we demonstrate the generality of MaxCal by developing a mathematical model for the example of diffusion. The model is sufficiently general to describe the system's most relevant features. Most importantly, the model can be extended to include new physics in a well-defined way. The essence of an application of MaxCal is to detect the relevant physical constraints at work and to distill them into a probability assignment in a mathematically unique fashion.

Bibliography

[1] Benjamin Andrae, Jonas Cremer, Tobias Reichenbach, and Erwin Frey. Entropy production of cyclic population dynamics. *Physical Review Letters*, 104(21):218102, 2010.

[2] Ludwig Boltzmann. *Vorlesungen über Gastheorie*. Verlag von Johann Ambrosius Barth, 1896.

[3] Simon Braun, Jens Philipp Ronzheimer, Michael Schreiber, Sean S Hodgman, Tim Rom, Immanuel Bloch, and Ulrich Schneider. Negative absolute temperature for motional degrees of freedom. *Science*, 339(6115):52–55, 2013.

[4] Stijn Bruers. A discussion on maximum entropy production and information theory. *Journal of Physics A: Mathematical and Theoretical*, 40(27):7441, 2007.

[5] Axel Brünger, Charles L Brooks, and Martin Karplus. Stochastic boundary conditions for molecular dynamics simulations of ST2 water. *Chemical Physics Letters*, 105(5):495–500, 1984.

[6] Roderick Dewar. Information theory explanation of the fluctuation theorem, maximum entropy production and self-organized criticality in non-equilibrium stationary states. *Journal of Physics A: Mathematical and General*, 36(3):631, 2003.

[7] Roderick C Dewar. Maximum entropy production and non-equilibrium statistical mechanics. In *Non-equilibrium Thermodynamics and the Production of Entropy*, pages 41–55. Springer, 2005.

[8] Roderick C Dewar. Maximum entropy production and the fluctuation theorem. *Journal of Physics A: Mathematical and General*, 38(21):L371, 2005.

[9] Roderick C Dewar. Maximum entropy production as an inference algorithm that translates physical assumptions into macroscopic predictions: don't shoot the messenger. *Entropy*, 11(4):931–944, 2009.

[10] Roderick C Dewar and Amos Maritan. The second law, maximum entropy production and Liouville's theorem. *arXiv preprint arXiv:1107.1088*, 2011.

[11] Roderick C Dewar and Amos Maritan. A theoretical basis for maximum entropy production. In *Beyond the Second Law*, pages 49–71. Springer, 2014.

[12] Ken A Dill and Sarina Bromberg. *Molecular driving forces: statistical thermodynamics in chemistry and biology*. Garland Science, 2003.

[13] John Earman and John D Norton. Exorcist XIV: the wrath of Maxwell's demon. Part I. From Maxwell to Szilard. *Studies in History and Philosophy of Modern Physics*, 29(4):435–471, 1998.

[14] John Earman and John D Norton. Exorcist XIV: The wrath of Maxwell's demon. Part II. From Szilard to Landauer and beyond. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics*, 30(1):1–40, 1999.

[15] Albert Einstein. Zur allgemeinen molekularen Theorie der Wärme. *Annalen der Physik*, 319(7):354–362, 1904.

[16] Christopher Essex. Radiation and the irreversible thermodynamics of climate. *Journal of the Atmospheric Sciences*, 41(12):1985–1991, 1984.

[17] Pierre Gaspard. Time-reversed dynamical entropy and irreversibility in Markovian random processes. *Journal of Statistical Physics*, 117(3-4):599–615, 2004.

[18] Kingshuk Ghosh, Ken A Dill, Mandar M Inamdar, Effrosyni Seitaridou, and Rob Phillips. Teaching the principles of statistical dynamics. *American Journal of Physics*, 74(2):123–133, 2006.

[19] Josiah W Gibbs. *Elementary Principles in Statistical Mechanics: Developed with Especial Reference to the Rational Foundation of Thermodynamics*. Cambridge University Press, 2010.

[20] G Grinstein and R Linsker. Comments on: A derivation and application of the maximum entropy production principle. *Journal of Physics A: Mathematical and Theoretical*, 40(31):9717–9720, 2007.

[21] Claudius Gros. *Complex and Adaptive Dynamical Systems*. Springer, 2008.

[22] Frederick James. *Statistical Methods in Experimental Physics*. World Scientific Publishing Company Incorporated, 2006.

[23] Edwin T Jaynes. Information theory and statistical mechanics. *Physical Review*, 106(4):620, 1957.

[24] Edwin T Jaynes. Information theory and statistical mechanics II. *Physical Review*, 108(2):171, 1957.

[25] Edwin T Jaynes. Where do we stand on maximum entropy? In *The Maximum Entropy Formalism*, pages 15–118. MIT Press, 1978.

[26] Edwin T Jaynes. The minimum entropy production principle. *Annual Review of Physical Chemistry*, 31(1):579–601, 1980.

[27] Axel Kleidon. Nonequilibrium thermodynamics and maximum entropy production in the earth system. *Naturwissenschaften*, 96(6):1–25, 2009.

[28] Axel Kleidon, Yadvinder Malhi, and Peter M Cox. Maximum entropy production in environmental and ecological systems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1545):1297–1302, 2010.

[29] Martin J Klein. Entropy and the Ehrenfest urn model. *Physica*, 22(6):569–575, 1956.

[30] Dilip Kondepudi and Ilya Prigogine. *From Heat Engines to Dissipative Structures*. John Wiley & Son, 1998.

[31] Rolf Landauer. Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, 5(3):183–191, 1961.

[32] Christian Maes and Karel Netočný. Minimum entropy production principle. *Scholarpedia*, 8(7):9664, 2013.

[33] M Malek Mansour and Florence Baras. Microscopic simulation of chemical systems. *Physica A: Statistical Mechanics and its Applications*, 188(1):253–276, 1992.

[34] Dibyendu Mandal, HT Quan, and Christopher Jarzynski. Maxwell's refrigerator: An exactly solvable model. *Physical Review Letters*, 111(3):030602, 2013.

[35] Leonid M Martyushev and Vladimir D Seleznev. Maximum entropy production principle in physics, chemistry and biology. *Physics Reports*, 426(1):1–45, 2006.

[36] Donald G Miller. Thermodynamics of irreversible processes: The experimental verification of the Onsager reciprocal relations. *Chemical Reviews*, 60(1):15–37, 1960.

[37] Akira Noda and Tatsushi Tokioka. Climates at minima of the entropy exchange rate. *Journal of the Meteorological Society of Japan*, 61(6):894–908, 1983.

[38] Lars Onsager. Reciprocal relations in irreversible processes I. *Physical Review*, 37(4):405, 1931.

[39] Lars Onsager. Reciprocal relations in irreversible processes II. *Physical Review*, 38(12):2265, 1931.

[40] Hisashi Ozawa, Atsumu Ohmura, Ralph D Lorenz, and Toni Pujol. The second law of thermodynamics and the global climate system: A review of the maximum entropy production principle. *Reviews of Geophysics*, 41(4), 2003.

[41] Garth W Paltridge. Global dynamics and climate – a system of minimum entropy exchange. *Quarterly Journal of the Royal Meteorological Society*, 101(429):475–484, 1975.

[42] Garth W Paltridge. A story and a recommendation about the principle of maximum entropy production. *Entropy*, 11(4):945–948, 2009.

[43] Steve Pressé, Kingshuk Ghosh, Julian Lee, and Ken A Dill. Principles of maximum entropy and maximum caliber in statistical physics. *Reviews of Modern Physics*, 85(3):1115, 2013.

[44] Ilya Prigogine. Nobel lecture – chemistry. *Nobel Foundation*, 1977.

[45] Ilya Prigogine and Erwin N Hiebert. From being to becoming: Time and complexity in the physical sciences. *Physics Today*, 35:69, 1982.

[46] Tamar Schlick. *Molecular Modeling and Simulation: An Interdisciplinary Guide*. Springer, 2010.

[47] Erwin Schrödinger. *What is life?: With mind and matter and autobiographical sketches*. Cambridge University Press, 1944.

[48] Effrosyni Seitaridou, Mandar M Inamdar, Rob Phillips, Kingshuk Ghosh, and Ken Dill. Measuring flux distributions for diffusion in the small-numbers limit. *The Journal of Physical Chemistry B*, 111(9):2288–2292, 2007.

[49] Claude E Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 1948.

[50] John Shore and Rodney Johnson. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Transactions on Information Theory*, 26(1):26–37, 1980.

[51] Gerhard Stock, Kingshuk Ghosh, and Ken A Dill. Maximum caliber: A variational approach applied to two-state dynamics. *The Journal of Chemical Physics*, 128:194102, 2008.

[52] Nathaniel Virgo. From maximum entropy to maximum entropy production: A new approach. *Entropy*, 12(1):107–126, 2010.

[53] Jin Wang, Li Xu, and Erkang Wang. Potential landscape and flux framework of nonequilibrium networks: Robustness, dissipation, and coherence of biochemical oscillations. *Proceedings of the National Academy of Sciences*, 105(34):12271–12276, 2008.

[54] Wikipedia. Functional equation, 2014. [Online; accessed 30-June-2014].

5 Appendix

5.1 Onsager's Derivation of the Reciprocal Relations

The following presentation of Onsager's reciprocal relations is based on Onsager's two articles on the topic [38, 39] and Prigogine's book [30]. Onsager was awarded the 1968 Nobel Prize in Chemistry for his proof of the reciprocal relations.

Theorem 5 (*Onsager's Reciprocal Relations*)

In thermodynamic transport processes, currents $\vec{J}$ can be expressed as linear functions of the thermodynamic forces $\vec{X}$

$$\vec{J} = \mathbf{L}\vec{X} \tag{160}$$

The matrix $\mathbf{L}$ is positive semi-definite and symmetric for all t-even systems.

5.1.1 Example: Detailed Balance and Microscopic Reversibility Imply Onsager's Relations

Consider a system in which three different conformations A, B and C of the same substance coexists. The transformation rates between the species

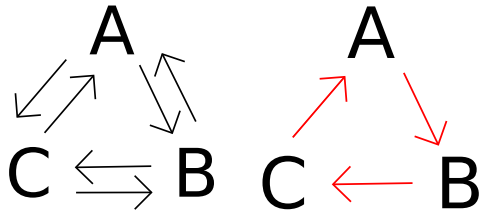


Figure 25: Left: Diagram of possible transitions, Right: Diagram of forbidden (t-odd) transitions

can be described by a first-order differential equation, since the number of transformed particles is proportional to the total number of particles of a

species n_i

$$\begin{aligned}\frac{dn_a}{dt} &= -(k_{ba}+k_{ca})n_a + k_{ab}n_b + k_{ac}n_c \\ \frac{dn_b}{dt} &= k_{ba}n_a - (k_{ab}+k_{cb})n_b + k_{bc}n_c \\ \frac{dn_c}{dt} &= k_{ca}n_a + k_{cb}n_b - (k_{ac}+k_{bc})n_c.\end{aligned} \tag{161}$$

While knowledge of all constants k_{ij} defines a unique equilibrium solution via $\frac{d\vec{n}}{dt} = 0$, the converse is not true. In fact, measuring the relative populations $n_a : n_b : n_c$ together with a constraint arising from particle number conservation $N = \sum_i n_i$ gives only 3 out of necessary 6 constraints.

It is possible (and common) to close the system by invoking **detailed balance**. This is the assertion that after a sufficiently long time (typically equilibrium) every elementary conversion occurs just as often as its reverse

$$\begin{aligned}k_{ba}\bar{n}_a &= k_{ab}\bar{n}_b \\ k_{cb}\bar{n}_b &= k_{bc}\bar{n}_c \\ k_{ac}\bar{n}_c &= k_{ca}\bar{n}_a.\end{aligned} \tag{162}$$

This constraint is not necessary to achieve equilibrium, but it is sufficient to find a unique solution for each k_{ij}. In the given example, detailed balance rules out cyclic transitions such as the one displayed in Fig. 25 (right). Onsager argues that **detailed balance is a direct consequence of time-reversal invariance of all known (relevant) laws of nature**. This fundamental feature of nature thus dictates that at equilibrium every elementary transition must be balanced by its reverse transition as **they are equivalent if the direction of time is reversed**.

Some further manipulation is necessary to reveal that the assumption of detailed balance in Eq. 162 gives rise to Onsager's reciprocal relations. As a first step, define $x_i = n_i - \bar{n}_i$[40] and subtract the equilibrium condition from Eq. 161

$$\begin{aligned}\frac{dx_a}{dt} &= -(k_{ba}+k_{ca})x_a + k_{ab}x_b + k_{ac}x_c \\ \frac{dx_b}{dt} &= k_{ba}x_a - (k_{ab}+k_{cb})x_b + k_{bc}x_c \\ \frac{dx_c}{dt} &= k_{ca}x_a + k_{cb}x_b - (k_{ac}+k_{bc})x_c\end{aligned} \tag{163}$$

[40] x_i measures distance from equilibrium.

To make the response matrix symmetric it is clear that the forces must be re-defined to include a factor $1/\bar{n}_i$. This appears naturally when considering the free energy F and its expansion in concentrations

$$\begin{aligned} F &= F_{\min} + RT \sum_{i=a,b,c} n_i \ln(n_i/\bar{n}_i) \\ \delta F &= RT \sum_{i=a,b,c} \ln(n_i/\bar{n}_i)\delta n_i, \end{aligned} \tag{164}$$

where the logarithm does not need to be expanded since its argument at equilibrium is close to unity. Converting to a notation measuring small deviations from equilibrium ($x_i \ll \bar{n}_i$), one finds

$$\begin{aligned} -RT\ln(n_i/\bar{n}_i) &= -RT\ln\left[(x_a+\bar{n}_i)/\bar{n}_i\right] \\ &= -RT\ln(1+x_i/\bar{n}_i) \approx -\frac{RT}{\bar{n}_i}x_i \equiv X_i. \end{aligned} \tag{165}$$

With this definition of the forces, the final equations show the announced symmetry relations if and only if Eq. 162 is true

$$\begin{aligned} \frac{dx_a}{dt} &= (k_{ba}+k_{ca})\frac{\bar{n_a}}{RT}X_a - \frac{k_{ab}\bar{n}_b}{RT}X_b - \frac{k_{ac}\bar{n}_c}{RT}x_c \\ \frac{dx_b}{dt} &= -\frac{k_{ba}\bar{n}_a}{RT}X_a + (k_{ab}+k_{cb})\frac{\bar{n}_b}{RT}X_b - \frac{k_{bc}\bar{n}_c}{RT}X_c \\ \frac{dx_c}{dt} &= -\frac{k_{ca}\bar{n}_a}{RT}X_a - \frac{k_{cb}\bar{n}_b}{RT}X_b + (k_{ac}+k_{bc})\frac{\bar{n}_c}{RT}X_c. \end{aligned} \tag{166}$$

5.1.2 General Proof

Onsager proves the reciprocal relations' general validity by considering fluctuations around equilibrium. He defines the "asymmetry" α_i (comparable to a center of gravity) of the thermodynamic quantity i as

$$\alpha_i = \int \epsilon \cdot x_l dV, \quad l \in \{1,2,3\}, \tag{167}$$

where ϵ is the local density of i. For example, ϵ could be the local energy density in the crystal. For each of the three coordinate directions l one would obtain one value α. Commonly, the origin can be chosen such that

$$\bar{\alpha}_k = 0 \quad \forall k \tag{168}$$

Detailed balance allows a very general statement about these fluctuations: If a perturbation $\alpha_j = \alpha_j'$ is on average followed by a perturbation $\alpha_i = \alpha_i'$ τ seconds later, then on average the perturbation in $\alpha_i = \alpha_i'$ causes a perturbation $\alpha_j = \alpha_j'$ the same time τ later:

$$A_{ji}(\tau) \equiv \overline{\alpha_j(t)\alpha_i(t+\tau)} = \overline{\alpha_i(t)\alpha_j(t+\tau)} \equiv A_{ij}(\tau) \tag{169}$$

Onsager assumes that deviations α_i decay linearly according to

$$\frac{d\bar{\alpha}_i}{dt} = \dot{\alpha}_i = \sum_{r=1}^{n} L_{ir}F_r, \text{ where} \tag{170}$$

$$F_r = \frac{\partial S}{\partial \alpha_r} \tag{171}$$

is defined to be the thermodynamic force. The temporal evolution of a decay can thus be expressed as

$$\bar{\alpha}_i(\Delta t, \alpha_i') \approx \bar{\alpha}_i(0, \alpha_i') + \dot{\alpha}_i \Delta t = \bar{\alpha}_i(0, \alpha_i') + \sum_{r=1}^{n} L_{ir}F_r \Delta t \tag{172}$$

This leads to the conclusion

$$A_{ji}(\Delta t) = \overline{\alpha_j(t)\alpha_i(t+\tau)} = \underbrace{\overline{\alpha_j'\bar{\alpha}_i(0, \alpha_i')}}_{A_{ji}(0)} + L_{ij}\Delta t \overline{\alpha_j F_j} \tag{173}$$

The first identification is true since $\alpha_j(t) = \alpha_j^{'}$ is a constant value, which can be absorbed into the average. The correlation function can be evaluated by exploiting the fluctuation theorem by Einstein [15]. It looks like an inversion of Boltzmann's formula, but the physical interpretation is different: Here, entropy is the fundamental quantity that gives rise to a certain probability distribution. In Boltzmann's formula entropy is defined by the probabilities.

$$\begin{aligned} P(\alpha_1, ..., \alpha_n) &\propto e^{\Delta S/k_b} \\ \Rightarrow k_b \ln P &= S + \text{const.} \\ \Rightarrow k_b \frac{dP}{d\alpha_j} &= P \cdot \frac{dS}{d\alpha_j} = P \cdot F_j \end{aligned} \tag{174}$$

Evaluate the correlation function using Eq. 174

$$\begin{aligned}\overline{\alpha_j F_j} = \overline{\alpha_j \cdot dS/d\alpha_j} = \int_{-\infty}^{\infty} \alpha_j \frac{dS}{d\alpha_j} P d\alpha_j \; &= k_b \int_{-\infty}^{\infty} \alpha_j \frac{dP}{d\alpha_j} d\alpha_j \\ = k_b \underbrace{[\alpha_j P]_{-\infty}^{\infty}}_{=0 \text{ (boundedness)}} &- k_b \underbrace{\int_{-\infty}^{\infty} P d\alpha_j}_{=1 \text{ (norm.)}},\end{aligned} \tag{175}$$

where integration by parts is used to obtain the second line. Integration over $\alpha_l, l \neq j$ gives unity and is omitted for brevity. Evidently, one finds by the same calculation $\overline{\alpha_i F_j} = 0, i \neq j$. Thus one is left with

$$\begin{aligned}A_{ji}(\Delta t) &= A_{ji}(0) - k_b \Delta t L_{ij} \\ A_{ij}(\Delta t) &= A_{ij}(0) - k_b \Delta t L_{ji}.\end{aligned} \tag{176}$$

Because of the symmetry in A_{ij} (Eq. 169) this directly implies Onsager's reciprocal relations

$$L_{ij} = L_{ji} \tag{177}$$

5.1.3 Summary

In my opinion, Onsager justifies on safe grounds that detailed balance has to hold near equilibrium. The key assumption, therefore, remains Eq. 170. Onsager considers fluctuations in a system, but it is plausible that the fluctuations' origin should not change the physics of a problem (memoryless process). As a result, reciprocal relations should be encountered close to equilibrium for most physical systems, in particular in transport phenomena (electrical, heat, diffusion). The Onsager relations cannot hold if microscopic reversibility is not given (presence of magnetic fields, other velocity-dependent t-odd force fields).

5.2 Explicit MaxCal Solutions

Consider first the simple system from Sect. 4.3 with N particles. Which is the MaxCal solution for the jump dynamics? Following MaxCal, one maximizes information entropy over possible microtrajectories with respect

to the average (known) currents. His can be implemented with the help of Lagrange multipliers

$$\begin{aligned} \mathfrak{C} &= -\sum_{\Gamma} p_\Gamma \ln p_\Gamma + \alpha \left(\sum_{\Gamma} p_\Gamma - 1 \right) + \lambda \left(\sum_{\Gamma} p_\Gamma j_\Gamma - J \right) \\ \frac{\delta C}{\delta p_{\Gamma'}} &= -1 - \ln p_{\Gamma'} + \alpha + \lambda j_{\Gamma'} \stackrel{!}{=} 0 \\ \Rightarrow p_{\Gamma'} &= \exp(\alpha - 1 + \lambda j_{\Gamma'}) = 1/Z \exp(\lambda j_{\Gamma'}). \end{aligned} \tag{178}$$

Since factor $\exp(\alpha - 1)$ is constant, it can be absorbed into the normalization constant Z, the partition function. The next important step is to determine the set of all trajectories $\{\Gamma\}$. The particles being distinguishable, a current involving j particles can be realized in $\binom{N}{j}$ ways, such that the collective weight of all microtrajectories with j is

$$W(j) = \sum_{\Gamma : j_\Gamma = j} \exp(\lambda j_{\Gamma'}) = \binom{N}{j} \exp(\lambda j) \tag{179}$$

This form can be simplified using the binomial expansion

$$Z = \sum_{\Gamma} p_\Gamma = \sum_{j=0}^{N} W(j) = \sum_{j=0}^{N} \binom{N}{j} \exp(\lambda j) = (1 + e^{\lambda})^N. \tag{180}$$

The generalization to the system from Sect. 2.4 consisting of N_r red and N_b blue particles in two reservoirs[41] is straightforward: Just as above both particles species contribute to particle currents. For the particle current each jumping particle counts ± 1, depending on the direction it jumps. Red and blue particle might, however, carry different specific energies. In the evaluated example, $J_{heat} = J_r + a \cdot J_{blue}$ with $a = 0.3 < 1$. This forms a new constraint, which mathematically very similar to the one above. The very same maximization and counting of trajectories with certain currents results in

$$\begin{aligned} Z = &(1 + e^{\lambda_p + a \cdot \lambda_h})^{N_b} \cdot (1 + e^{\lambda_p + \lambda_h})^{N_r} \\ &\cdot (1 + e^{-\lambda_p - a \cdot \lambda_h})^{N_b} \cdot (1 + e^{-\lambda_p - \lambda_h})^{N_r}. \end{aligned} \tag{181}$$

[41] i.e. these numbers are constant in time

5.3 Computer Algorithms

5.3.1 Integrating Langevin's Stochastic Differential Equation

Simple diffusion systems can be modeled by Langevin's differential equation

$$ma(t) = F(t) - \gamma m v(t) + R(t). \tag{182}$$

Compared to Newton's equation of motion, it contains a friction term and a random force term. The friction term accounts for the slowing down of fast colloids in a solvent due to drag by the solvent molecules. The very same solvent molecules, however, are subject to thermal motion, resulting in occasional collisions with the colloids. Langevin assumes white noise, i.e.

$$\begin{aligned} \langle R(t) \rangle &= 0 \\ \langle R(t) R(t') \rangle &= 2 k_b T m \delta(t - t'). \end{aligned} \tag{183}$$

In a computer implementation, white noise is commonly represented by a Gaussian distribution with zero mean $\mu = 0$ and finite variance $\sigma^2 = 2k_bTm\delta(t-t')$. The delta-distribution is approximated by $\frac{1}{\Delta t}$, where Δt is one time step in the integration algorithm [46].

There are several ways of discretizing the differential equation. I choose the standard Verlet procedure. First, calculate a Taylor expansion of the particle's position in time

$$r(t + \Delta t) = r(t) + v(t)\Delta t + a(t)\Delta t^2/2 + \dot{a}(t)\Delta t^3/6 + \mathcal{O}(\Delta t^4). \tag{184}$$

The Taylor expansion of $r(t - \Delta t)$ gives the same result, with all t-odd terms having switched signs. As a result, adding gives

$$\begin{aligned} r(t + \Delta t) = 2r(t) - r(t - \Delta t) \\ + \Delta t^2/m \left[F(t) - \gamma m v(t) + R(t) \right] + \mathcal{O}(\Delta t^4). \end{aligned} \tag{185}$$

Next, the function $v(t)$ is replaced by a finite difference approximation

$$v(t) = \frac{r(t + \Delta t) - r(t - \Delta t)}{2} + \mathcal{O}(\Delta t^3). \tag{186}$$

Finally, solving the equation for $r(t + \Delta t)$ gives

$$\begin{aligned} r(t + \Delta t) = \frac{1}{1 + \gamma \frac{\Delta t}{2}} \Bigg[2r(t) - r(t - \Delta t) + \gamma \frac{\Delta t}{2} r(t - \Delta t) \\ + \frac{\Delta t^2}{m} \left[F(t) + R(t) \right] + \mathcal{O}(\Delta t^4) \Bigg]. \end{aligned} \tag{187}$$

The resulting integrator is known as the BKK integrator [5].

5.3.2 Calculating All States

The general procedure all of constructing the transfer matrix for arbitrary N_p and N_b is straightforward. However, some parts of the algorithm can consume very different computational time, depending on the specific implementation. For example, the enumeration of all states can be done very efficiently using a recursive function. The following function in pseudo-code iterates only up to the required depth.

```
Data: Number of boxes N_b, number of particles N_p
Result: All states compatible with N_b and N_p
RecursiveFunction(level)
    if sum_{j=0}^{N_b} c_j equals N_p then
        save {c_j} as a state
    else
        if sum_{j=0}^{level} c_j ≤ N_p then
            for c_i = 0 to N_p do
                Set c_level = c_i
                Call RecursiveFunction at level + 1
            end
        end
    end
end
```

Algorithm 1: Enumeration of all states with N_b boxes and N_p particles

The remaining steps necessary to construct the transfer matrix are implemented following the steps outlined in Sect. 4.3 directly.