Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Work in Quantum thermodynamics: Discussion of an axiomatic approach

Master Thesis

Frederik Lohof

April 27, 2016

Supervisors: Prof. Dr. R. Renner, P. Kammerlander
Department of Theoretical Physics, ETH Zürich

Abstract

In quantum thermodynamics there is no straight forward definition of work and different solutions were put forward in the past. Recently a new axiomatic approach to the topic was presented by Gallego, Eisert and Wilming [arXiv:1504.05056] and was claimed to formally give a way of finding definitions of work. This approach, however, explicitly rules out definitions of work previously used in the so called single shot framework and criticizes it for violating the proposed axioms. In order to determine whether this criticism is justified we present an extensive discussion of the two mentioned frameworks and compare their underlying assumptions as well as their applicability. We provide arguments why the single shot framework gives a reasonable model for the regime of individual quantum systems that is operationally and physically justified. Furthermore we find that the new approach fails to incorporate the single shot framework and is in fact not as general as claimed by the authors. The axioms pose a priori restrictions on the form in which thermodynamic work can be defined and therefore are too strong in order to provide a good description of this regime.

Contents

Chapter 1

Quantum Thermodynamics

1.1 Introduction

The correspondence between statistical mechanics and classical thermodynamics is only valid in the limit of large numbers of particles in the system. It is the regime of "the big and the many". However, in the light of recent developments of nanoscale devices [2–4] it becomes increasingly important that we also understand the thermodynamics of those systems which are at the opposite end of the thermodynamic limit. In this cases, special attention is given to the question, how to define work [5–10]. Classically, work, unlike heat, is defined as an ordered and predictable form of energy. In classical thermodynamics one question that is often considered is how much (mechanical) work can be extracted from a given system, e.g. a steam engine or a piston containing a gas. However, in the micro regime the clear distinction between work and heat vanishes and a straight forward definition of work is not possible any more.

As an illustrative example, consider a classical thermodynamic engine that performs work by lifting a stone. The difference of the stone's height is a direct measure for the performed work. It can be stored and, at a later time, used in another process. Note that we could easily bring the stone into contact with a heat bath of higher temperature. This would transfer heat to the stone but would not change the amount of work that is stored in the system – the stone's height stays the same. In general, the macroscopic degrees of freedom, used to measure work (like the height of a stone), are not notably affected by thermal fluctuations due to the increased temperature of the system. Therefore, there are distinct notions of work and heat, respectively. Imagine shrinking the whole setup to the nano regime. Intuitively it is understood that, as the size of the stones becomes smaller, fluctuations become increasingly important. If one considers a "stone" only consisting of, say, a few atoms, it becomes unclear if a change in its horizontal position is due to an ordered motion or just a displacement due to thermal (or even quantum) fluctuations.

In classical thermodynamics, the optimal work that can be extracted from a system ρ in the presence of a single heat bath at temperature T is governed by the system's free energy [11]

$$F(\rho) = U(\rho) - TS(\rho) \tag{1.1}$$

where $U(\rho)$ is the system's internal energy and S its entropy. Work is an average value, which is a good description since it has vanishingly small fluctuations. However, this is not true in the microscopic regime as indicated by the above example of a lifted stone. For small systems the extractable work can easily have fluctuations similar to the energy scale of the system [6], which makes the average work more heat like than work like.
To address this issue there is ongoing research in formulating a consistent description of thermodynamics in the quantum regime [12]. This also comprises a generalized notion of work. Important results in this field of research were obtained by applying methods from quantum information theory to the thermodynamic setting [13]. In particular, the formulation of quantum thermodynamics in terms of a resource theory [14, 15] provides a fruitful description that allows a reformulation of the laws of thermodynamics [16] and recovers classical results in the thermodynamic limit [15]. It will be the main language in which this our discussion is formulated.

In this work we present an extensive analysis of recent results by R. Gallego et. al. (2015) [1] concerning the definition of work in quantum thermodynamics. The authors present a novel approach towards the topic by imposing a set of operational axioms that any possible notion of work is supposed to satisfy. The axioms are inspired by the second law of thermodynamics. Formulated in the language of resource theories, the authors derive strong restrictions on the form in which thermodynamic work is quantified. Notably, these restrictions partly seem to contradict previously introduced definitions of work, in particular those formulated in the framework of single shot thermodynamics [5–7, 10].
The objective of our discussion is to provide a context in which the results of [1] must be interpreted and how one consolidates this framework and those of single shot thermodynamics. We find that the framework put forward in [1] is not as general as claimed by the authors, but impose a priori restrictions which lead to a particular model of thermodynamic work. Furthermore, we discuss the single shot approach, provide arguments for its usefulness and highlight the cases where the model in [1] produces the same results and where the models are incompatible.

The second part of this chapter gives an introduction to quantum thermodynamics as a resource theory, since it is the language used to formulate the approaches towards defining work, that are considered here. Chapter 2 discusses the framework, introduced in [1] and summarizes its results. Chapter 3 gives motivation for the single shot approach to quantum thermodynamics and discusses the criticism of it put forward in [1]. The final chapter compares the different frameworks in its applicability and gives a general context in which they can be set into context.

1.2 Quantum thermodynamics as a resource theory

Quantum thermodynamics can be formulated as a resource theory [5,15]. In this approach it is described by a restriction on the quantum operation that can be implemented to manipulate a system. The allowed operations in a resource theory are called free operations and will be specified below for the case of thermodynamics. Under this restrictions one examines which states can be transformed into one another. Importantly, the states that can be created using only free operations are called free states. All others represent a resource. Possessing a resource state allows, in general, the creation of states that could not be created using only free states and operations.

Objects and transitions between objects In the following we will adapt the notation from [1] by describing quantum systems as objects p, which are defined as pairs (ρ, H) of a quantum state ρ with a Hamiltonian H on an arbitrary Hilbert space. In comparison to normal quantum information theory, we also consider the Hamiltonian since thermodynamics accounts for all changes in energy of a system. Some examples of objects include

- a single qubit in an arbitrary state with a trivial Hamiltonian $(\rho, \mathbb{1}_2)$.
- a single qubit in an arbitrary state with a specific Hamiltonian $(\rho, E|1\rangle\langle 1|)$.
- thermal objects $w_H = (\omega_{H,\beta}, H)$ with $\omega_{H,\beta}$ being the Gibbs state corresponding to the Hamiltonian H (see next paragraph).

If we consider two systems described by the objects $p = (\rho, H_1)$ and $q = (\sigma, H_2)$ these can be combined using a generalized tensor product

$$p \otimes q := (\rho \otimes \sigma, H_1 + H_2), \tag{1.2}$$

where $H_A + H_B$ is a short hand notation for $H_A \otimes \mathbb{1}_B + \mathbb{1}_A \otimes H_B$. The interaction between subsystems will be introduced manually in the next paragraph. Furthermore, we will consider transitions between different states of a system. If we transform $p_1 = (\rho_1, H_1)$ into $p_2 = (\rho_2, H_2)$ we write $p_1 \to p_2$. It is important to note that not only the state but also the Hamiltonian of the system is allowed to change.

In the following we will often drop the distinction between state and objects by referring to p or q as the state of the system. Implicitly we also include the Hamiltonian in this notion. In general the real quantum state (without Hamiltonian) will be denoted by Greek letters ρ, σ, etc. while the objects describing the current pair $p = (\rho, H)$ will always be denoted with Roman letters.

Thermal operations Following the frameworks of [1, 5, 15] we introduce catalytic thermal operations which cast quantum thermodynamics as a resource theory [17]. Note that thermal operations are only one possible choice of quantum operations for a resource theory of thermodynamics and that there are alternative models (e.g Gibbs preserving operations [18], thermal contact). Thermal operations are considered free (i.e executing them costs no work) and aim at modelling the thermodynamic setting in which we have access to arbitrary heat baths at fixed inverse temperature β, described by thermal states. Furthermore, thermal operations assume arbitrary control over internal degrees of freedom of the heat bath as well as the system itself. It models the interaction between the system and the heat bath. Explicitly, catalytic free operations are a combination of the following operations:

- To our system $p = (\rho, H)$ we can tensor a heat bath described by a thermal state with an arbitrary Hamiltonian H_B and inverse Temperature β (which is fixed throughout the whole process). It is written as a thermal object $w_{H_B,\beta} = (\omega_{H_B,\beta}, H_B)$ with the thermal state (Gibbs state)

$$\omega_{H_B,\beta} = \frac{\exp(-\beta H_B)}{Z_{H_B}} \tag{1.3}$$

  and the partition function

$$Z_{H_B} = \mathrm{Tr}(\exp(-\beta H_B)). \tag{1.4}$$

- We can add any ancillary system $p_C = (\sigma, H_C)$ (referred to as a catalyst) which is, at the end of the process, returned in exactly the same state and furthermore is returned uncorrelated to the rest of the system.

- We can apply any energy conserving unitary operation U on the combined system $p \otimes w_{H_B,\beta} \otimes p_C$ that conserves the total energy of the system, i.e. it commutes with the total Hamiltonian of the combined system of heat bath, catalyst and the original system,

$$[U, H + H_B + H_C] = 0. \tag{1.5}$$

  The state is transformed as

$$\rho \otimes \omega_{H_B,\beta} \otimes \sigma \mapsto U(\rho \otimes \omega_{H_B,\beta} \otimes \sigma)U^\dagger = \tilde{\rho} \otimes \sigma \tag{1.6}$$

  with σ being the state of the catalyst system which is returned in its initial state and uncorrelated to the rest of the system.

- We can trace out any subsystem S after applying the unitary. This corresponds to discarding parts of the total system. Explicitly,

$$\mathrm{Tr}_S(p_S \otimes p_{S'}) = p_{S'}, \tag{1.7}$$

where here the partial trace is defined on the level of objects instead states.

Thermal operation are defined as a combination of these operations. They are maps on the system $p = (\rho, H)$

$$\mathcal{E}(\rho) = \mathrm{Tr}_{SC}\left(\left(U(\rho \otimes \omega_{H_B,\beta} \otimes \sigma)U^\dagger, H + H_B + H_C\right)\right) \tag{1.8}$$

The resource theory of thermodynamics now asks the question whether or not certain states (or in our case object) can be transformed into others. Importantly, non thermal states can be used as a resource: A state p can act as a resource by assisting transformations of other states in the sense that it enables the transformation $q_1 \to q_2$, if the transition $p \otimes q_1 \to q_2$ is possible using catalytic thermal operations only. In general the resource is consumed in such a process. Since only non thermal states can be used in such a way, the theory is also called the resource theory of a-thermality.

Chapter 2

Definition of work from operational principles

In the following I will present the framework introduced in [1]. The framework is formulated in the language of a resource theory and takes a novel approach towards defining work in quantum thermodynamics. The authors postulate axioms that any function quantifying work should satisfy and derive restrictions on the form of those quantifiers. It is important to note that previous approaches [5–7] where opposite to that. They a priori proposed notions of work and derived properties, that these satisfied.
The framework considers two agents who interact with each other according to a certain set of rules. They possess different quantum systems and are able to manipulate them.

2.1 Two agents playing a game

The first player, Arthur, possesses work storage devices which are quantum systems described as objects p_A. They are restricted to be from a set $\mathcal{P}$ which contains all valid objects describing work storage devices. The importance of the set $\mathcal{P}$ will become clear later. The second agent, Merlin, holds a system which we refer to as fuel. They are described by another object p_M. In comparison to Arthur's system, p_M is not restricted to have a certain form. In the game that the two agents play Arthur gives his system to Merlin, who transforms it using his system as a resource before giving it back to Arthur. Explicitly, Merlin performs a free catalytic operation on the joint system $p_A \otimes p_M$ and returns the final state back to Arthur, which again has to be from the set $\mathcal{P}$. Classically, this can be interpreted as the expenditure of fuel (p_M) in order to lift a suspended weight (p_A) where the catalyst (p_C) employed in the catalytic thermal operation can be thought of as an engine that, undergoing a cyclic transition, ends up in its initial state.

Every transformation that Merlin performs on Arthur's system comes with a price that the two player must agree on. It is determined by some function $\mathcal{W}(p_A \to q_A, \beta)$, where p_A and q_A are the initial and final state[1] of a transformation respectively. The function value is interpreted as the exchanged amount of work and the function is referred to as a work quantifier. We use the convention, that $\mathcal{W}(p_A \to q_A, \beta) \geq 0$ means that Arthur pays to Merlin and the price that Arthur pays is interpreted as the positive amount of work, stored in Arthur's system.
Note, that the systems of Arthur and Merlin are not on an equal footing. The reason for this is that Merlin can possibly hold any quantum system while Arthur's systems are restricted to the set $\mathcal{P}$, which is of crucial importance because of the following consideration.

Work is an ordered type of energy in the sense that it comes in a well defined form and can be stored in and extracted from a work storage device in a deterministic and predictable way. The set $\mathcal{P}$ of valid work storage devices accounts for that notion. In classical analogy, if we see the task of performing work as lifting a suspended weight, any work storage device should be described as having a continuous state variable (the height of the weight under a conservative force) while any other system is not considered a valid work storage system. Similarly, in the framework that we introduce here, work storage systems p_A are considered to be from a set $\mathcal{P}$ which defines what form of energy is considered as work. It is important to note, that any definition of work quantifier $\mathcal{W}$ will in general depend on the form of the work storage systems given by the set $\mathcal{P}$.

Although not mentioned explicitly in [1], Merlin does not have access to arbitrary systems p_M at any time in a given process. Instead one has to consider the individual situation where Merlin starts with well defined resources and transforms them according to the rules of the game. In any process, the set $\mathcal{P}$ can be interpreted as a work reservoir held by Arthur. One can obtain two different points of view. Arthur can transform systems $p_A \in \mathcal{P}$ by giving them to Merlin, who uses his system as a resource (or a catalyst) if necessary. Arthur does not know how Merlin performs the transformations which is reflected by the fact that the work quantifier $\mathcal{W}$ does not depend on the transition on Merlin's system. From this point of view, work is a process independent quantity.
On the other hand we can take the point of view that Merlin uses Arthur's systems as a resource in order to transform his system p_M. To capture this

[1] For convenience, we refer to p_A, q_A, etc. as states although they actually describe objects containing also the Hamiltonian: $p_A = (\rho_A, H_A)$. The distinction becomes clear from the use of Roman letters (objects) and Greek letters (quantum state without Hamiltonian).

we define the work of transition on Merlin's system as

$$W_{\text{trans}}(p_M^{(1)} \to p_M^{(2)}, \beta) := \sup_{\substack{p_A^{(1)}, p_A^{(2)} \in \mathcal{P} \\ p_M^{(2)} \otimes p_A^{(2)} \in \mathcal{F}_{\mathcal{C}}(p_M^{(1)} \otimes p_A^{(1)})}} \mathcal{W}(p_A^{(1)} \to p_A^{(2)}, \beta), \tag{2.1}$$

where $\mathcal{F}_{\mathcal{C}}$ is referred to as the free catalytic image and $\mathcal{F}_{\mathcal{C}}(p_M^{(1)} \otimes p_A^{(1)})$ is the set of all objects that can be reached from $p_M^{(1)} \otimes p_A^{(1)}$ using free catalytic operations only. The value W_{trans} is the optimal work transfer that Merlin can achieve in the sense that it maximizes the stored work in the reservoir held by Arthur when performing the transition $p_M^{(1)} \to p_M^{(2)}$. Importantly, W_{trans} is also a function of $\mathcal{P}$, i.e. the form of the work storage systems that we consider. Note that Merlin is by no means forced to perform the optimal process when transforming his system. From this point of view, work can be seen as dependent on the process that Merlin applies. It will be a function of Arthur's system as well. From an operational point of view however, stored work, as a predictable form of energy, is valuable and Merlin aims at minimizing its expenditure, i.e. he aims at performing the optimal process with work cost $W_{\text{trans}}(p_M^{(1)} \to p_M^{(2)}, \beta)$.

The framework also introduces the extractable work and the work of formation of a system p_M. They are defined as

$$W_{\text{ext}}(p_M, \beta) := W_{\text{trans}}(p_M \to \omega_\beta, \beta), \tag{2.2}$$

$$W_{\text{form}}(p_M, \beta) := -W_{\text{trans}}(\omega_\beta \to p_M, \beta). \tag{2.3}$$

The extractable work is defined as the maximal energy that can be stored in a work storage system while thermalizing the state p_M. Analogously the work of formation is the minimum amount of energy that must be extracted from a work storage system in order to create the state p_M from a thermal state. The functions are defined as optimal transitions to and from the thermal state ω_β because thermal states are considered free and, as a basic assumption, no work can be extracted from them. Since W_{ext} and W_{form} are defined via W_{trans} they are also a function of the form of the work storage systems given by $\mathcal{P}$.

The extractable work is defined as a function on Merlin's system only and likewise for the work of formation. On Arthur's system, such functions are not introduced since it is assumed that extractable work and work of formation are the same for any work storage device in order to account for the deterministic nature of work. That means, the same amount of energy that we deposit in a work storage device must be extractable at some later time without losses. This issue will be discussed further in Section 4.2. In fact one has to note that the definition of extractable work and work of formation as given here results in an inherent irreversibility emerging in the model in the sense that

$$W_{\text{ext}}(p_M, \beta) \leq W_{\text{form}}(p_M, \beta). \tag{2.4}$$

This irreversibility results from the evaluation of the supremum in (2.1) and is only true for transformations on Merlin's system, quantified by W_{trans}. It is a manifestation of an inherent second law arising in the model of [1]. As we will discuss in Section 2.3 this irreversibility does not arise for transformations on Arthur's system which are quantified by $\mathcal{W}$.

2.2 Axioms

The foundation of the paper are two basic axioms, that any work quantifier is supposed to satisfy. They are inspired by the second law of thermodynamics. Especially, they prevent the construction of a perpetuum mobile. In the language of the paper, the axioms assure that neither of the two players can become arbitrarily rich by playing the game:

- **Axiom 1** Consider a cyclic sequence of transitions of Arthur's system where he repeatedly gives his system to Merlin who performs the transitions
$$\{p_A^{(1)} \to p_A^{(2)} \to \ldots \to p_A^{(n)} = p_A^{(1)}\}. \tag{2.5}$$
Then the sum of the work of each step satisfies
$$\sum_{k=1}^{n-1} \mathcal{W}(p_A^{(k)} \to p_A^{(k+1)}, \beta) \geq 0. \tag{2.6}$$
This means that Arthur can not become arbitrarily rich by repeating a cyclic process many times. Operationally it assures, that Merlin can not extract arbitrary amounts of energy from a work storage system without changing it.

- **Axiom 2** Consider a cyclic sequence of transitions of Merlin's system where he repeatedly takes Arthur's system and uses it as an assisting system to perform the transitions
$$\{p_M^{(1)} \to p_M^{(2)} \to \ldots \to p_M^{(n)} = p_M^{(1)}\}. \tag{2.7}$$
Then the total work of transition satisfies
$$\sum_{k=1}^{n-1} W_{\text{trans}}(p_M^{(k)} \to p_M^{(k+1)}, \beta) \leq 0 \tag{2.8}$$
This means that Merlin also can not become arbitrarily rich by repeating such a process. Operationally the axiom assures that Merlin can not extract arbitrary amounts of of energy from his system without spending any resources.

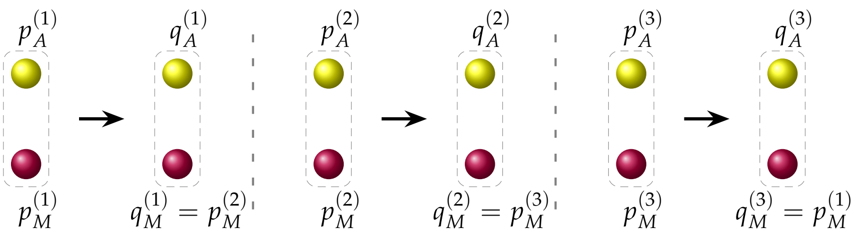

Figure 2.1: A set of transitions in the spirit of Axiom 2. Merlin obtains $p_A^{(1)}$ from Arthur and uses his system $p_M^{(1)}$ as an assistant to transform the joint state into the joint state $q_A^{(1)} \otimes q_M^{(1)}$. After that Arthur gives Merlin a new system, so that he now holds $p_A^{(2)} \otimes q_M^{(1)}$. Merlin performs several transitions that way where after the last transition his system is back to the initial state $q_M^{(1)}$. Note that during the whole process the two states remain uncorrelated. The axiom states, that Merlin should not have a net benefit from performing such a cyclic process on his system.

There is an equivalent formulation of this axiom, that is intuitively less appealing but mathematically easier since it does not involve W_{trans}, which might be difficult to evaluate. The equivalence of the formulations is shown in the appendix of [1].

- **Axiom 2 (reformulated)** Let $\{p_A^{(k)} \to q_A^{(k)}\}_{k=1}^{n-1}$ be a collection of transitions of Arthur's system, assisted by $(p_M^{(k)}, q_M^{(k)})$ with $p_M^{(n)} = p_M^{(1)}$ and $p_M^{(k+1)} = q_M^{(k)}$ (see Figure 2.1). Then the total work satisfies

$$\sum_{k=1}^{n-1} \mathcal{W}(p_A^{(k)} \to q_A^{(k)}, \beta) \leq 0. \tag{2.9}$$

2.3 Consequences from the axioms

In [1] the authors derive very stringent conditions for valid work quantifiers from the axioms. First they show, that the axioms 1 and 2 are equivalent to the following properties.

1. For all $p, q, r \in \mathcal{P}$

$$\mathcal{W}(p \to q) = -\mathcal{W}(q \to p), \tag{2.10}$$

$$\mathcal{W}(p \to q) + \mathcal{W}(q \to r) = \mathcal{W}(p \to r). \tag{2.11}$$

2. For all $p^{(1)}, \dots, p^{(m)}$ and $q^{(1)}, \dots, q^{(m)}$ in $\mathcal{P}$ such that $\bigotimes_{i=1}^m q^{(i)} \in \mathcal{F}_C\left(\bigotimes_{i=1}^m p^{(i)}\right)$,

$$\sum_{i=1}^{m} \mathcal{W}(p^{(i)} \to q^{(i)}) \leq 0. \tag{2.12}$$

Although the proof of the equality is mathematically correct, one can argue that it lacks operational plausibility. For example, in order to prove the equality with property 1, the authors show that every cyclic sequence of transformations on Arthur's system can be performed by Merlin without spending resources. In the light of corollary 23 in [1] this shows that every such sequence on Arthur's system would have total work cost $\mathcal{W} = 0$. However, the proof assumes that Merlin already possesses all the intermediate states of the sequence as a catalyst. These are then subsequentially swapped with the work storage system, which is assumed to be a free operation as the swapping is defined on the level of objects, i.e. the swap $p \otimes q \to q \otimes p$ also involves swapping of the Hamiltonians. In a nutshell, if we perform a sequence of transitions and assume that we already possess all the intermediate states of the sequence it seems obvious that we can perform the sequence of transformation without spending resources.

One important consequence of the properties 1 and 2 is, that there is some function M acting on the working system, such that the work quantifier can be written in the form

$$\mathcal{W}(p \to q) = M(q) - M(p), \tag{2.13}$$

where M satisfies the property

- Additive monotonicity: For all $p^{(1)}, \ldots, p^{(m)}$ and $q^{(1)}, \ldots, q^{(m)}$ in $\mathcal{P}$ such that $\bigotimes_{i=1}^{m} q^{(i)} \in \mathcal{F}_C\left(\bigotimes_{i=1}^{m} p^{(i)}\right)$,

$$\sum_{i=1}^{m} M(q^{(i)}) \leq \sum_{i=1}^{m} M(p^{(i)}). \tag{2.14}$$

This means that the work storage device acts analogous to a lifted weight under a conservative force, in the sense that the transferred work is the difference of a function of the system's state before and after the transition. Furthermore, the additive monotonicity implies

$$M(q) \leq M(p) \qquad \forall q \in \mathcal{F}_C(p), \tag{2.15}$$

i.e. the function M acts as a (decreasing) monotone under catalytic thermal operations.

Using this results, the authors identify the Rényi divergences as possible monotones M, i.e.

$$M_\alpha(p, \beta) = \frac{1}{\beta} S_\alpha(\rho \parallel \omega_{H,\beta}), \tag{2.16}$$

where S_α is the quantum Rényi divergence. It is a valid monotone in (2.13) for any $\alpha > 0$.

Chapter 3

Single shot thermodynamics

3.1 Introduction

Classical thermodynamics is valid in the thermodynamic limit, where the system is described by macroscopic variables. Statistical mechanics tells us that these variables represent average values with negligible fluctuations of the underlying random variable. However, for small systems these fluctuations can be of the same size as their averages [6] and there is need of an extended theory of thermodynamics away from the thermodynamic limit. One important approach towards such a generalized theory is the single shot thermodynamics [5–7, 10, 16]. It considers the laws, analogous to the ones in classical thermodynamics, for individual runs of a process and turn away from describing average quantities, which do not anymore govern processes like work extraction and state transitions. Importantly, single shot thermodynamics relies on methods from quantum information theory. The connection between the two topics can be made via Maxwell's Demon and the Landauer principle, for instance [19–21].

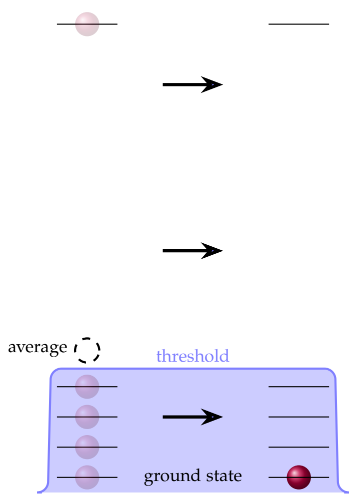

Figure 3.1: Left: A quantum system is in an equal superposition of different energy states. One eigenstate is much higher than the other ones. The average energy lies above some energy threshold. Right: Extracting energy from the system will on average yield work, that is higher than the threshold. In a single run however, the extracted energy is most likely to be lower than the threshold.

As an example, that highlights the importance of the single shot approach, consider the case where we want to use a quantum state as a source of energy to trigger some other process (say, a chemical reaction) that requires some threshold or activation energy. Assume that the system is in an equal superposition of many energy eigenstates where one energy level lies high above the threshold while all the others lie beneath it. If we repeated the process many times, we might very well be able to extract an amount of energy above the threshold on average (see Figure 3.1). However, in a single run of the experiment the energy of the system is likely to be found below the threshold. This example shows that the description of quantum thermodynamics in terms of average quantities can be inappropriate in processes involving only small systems. Furthermore, the example also highlights the probabilistic nature of energy in the microscopic regime. The single shot approach instead aims at finding laws that govern individual processes as the one described above.

In single shot thermodynamics, the definition of work is of central interest [5–7]. It aims at a notion of work as a well-ordered form of energy, i.e. predictable and stored in a specific form, that allows a deterministic use at later times. In the following we introduce the single shot approach to a definition of work and give arguments for its usefulness.

3.2 ε-deterministic work

In single shot thermodynamics extracting work can be defined as deterministically bringing a two level system from its ground state to the excited state [5]. Explicitly we consider a system in state ρ with Hamiltonian H and a single qubit with Hamiltonian $H_W = \Delta|1\rangle\langle 1|$. We ask the question what is the largest Δ such that the transition

$$\rho \otimes |0\rangle\langle 0| \overset{\text{free cat.}}{\longrightarrow} |1\rangle\langle 1| \tag{3.1}$$

is possible using only catalytic thermal operations. The single qubit acts as a work storage device (analogous to a lifted weight) and is referred to as a *wit* (work qubit). Similarly the work of formation of the system described by ρ and H is the minimal Δ such that the transition

$$|1\rangle\langle 1| \overset{\text{free cat.}}{\longrightarrow} \rho \otimes |0\rangle\langle 0|. \tag{3.2}$$

is possible with catalytic thermal operations.

Importantly, the framework allows for a failure probability ε when performing the above tasks. This notion accounts for the probabilistic behaviour of quantum systems and introduces a risk-reward trade-off into the framework [7]: In general it is possible to extract more energy from a given state when accepting a higher probability of failure. The notion of ε-deterministic work can be modelled by demanding that the state of the *wit* is not necessarily in an eigenstate of its Hamiltonian, but can be in a state ε-close to it, i.e. the *wit* is in a state ρ_W s.t.

$$\|\rho_W - |E\rangle\langle E|\|_1 < 2\varepsilon \tag{3.3}$$

where $|E\rangle\langle E|$ is an eigenstate of the *wit*-Hamiltonian H_W, i.e. $|0\rangle\langle 0|$ or $|1\rangle\langle 1|$ and the norm being the trace norm

$$\|\rho - \sigma\|_1 = \mathrm{Tr}(\sqrt{(\rho - \sigma)^2}). \tag{3.4}$$

Operationally this means that upon performing a projective measurement on the *wit* in the energy eigenbasis after work extraction (3.1), we will find with probability ε that the *wit* is in the 0-state and thus the work extraction failed.

The optimal values for Δ in (3.1) for the process of extracting work from the state ρ while allowing for a failure probability ε is given by[1]

$$W^{\varepsilon}_{\mathrm{ext}}(\rho) = F^{\varepsilon}_{\mathrm{min}}(\rho) - F_{\mathrm{min}}(\tau), \tag{3.5}$$

and similar for the work of formation of a state ρ the optimal Δ in (3.2) is given by

$$W^{\varepsilon}_{\mathrm{form}}(\rho) = F^{\varepsilon}_{\mathrm{max}}(\rho) - F_{\mathrm{max}}(\tau). \tag{3.6}$$

where τ is the thermal state corresponding to the Hamiltonian of the system in state ρ. The functions F_{min} and F_{max} are the min and max free energies of the state. The functions $F^{\varepsilon}_{\mathrm{min}}$ and $F^{\varepsilon}_{\mathrm{max}}$ are the smoothed min and max energies given by [5]

$$F^{\varepsilon}_{\mathrm{min}}(\rho) = \sup_{\|\rho-\rho'\|_1<\varepsilon} F_{\mathrm{min}}(\rho), \tag{3.7}$$

and

$$F^{\varepsilon}_{\mathrm{max}}(\rho) = \inf_{\|\rho-\rho'\|_1<\varepsilon} F_{\mathrm{max}}(\rho). \tag{3.8}$$

They are related to smoothed max and min entropies [22]. It is important to note that the work of formation given in (3.6) is only valid for states that are diagonal in the energy eigenbasis. Coherent states (in the energy eigenbasis) can not be reached using catalytic thermal operations. The extractable work (3.5) is optimal for all states block diagonal in the energy eigenbasis. For coherent states the optimal extractable work is not known [5].

The expressions for the extractable work and work of formation were independently derived in [5] and [6] and they are a generalization of the previous

[1] we write $W^{\varepsilon}_{\mathrm{ext}}$ and $W^{\varepsilon}_{\mathrm{form}}$ instead of W_{ext} and W_{form} in order to distinguish the expressions here from the ones in (2.2) and (2.3). The relation of these expressions and how they correspond to one another will be discussed further in Section 4.1.

results in [7], which only considers states with fully degenerate Hamiltonian $H = \lambda\mathbb{1}$.

Furthermore, the framework recovers classical results in the thermodynamic limit, in the sense that

$$\lim_{n\to\infty} \frac{1}{n} F^{\varepsilon}_{\text{max}}(\rho^{\otimes n}) = \lim_{n\to\infty} \frac{1}{n} F^{\varepsilon}_{\text{min}}(\rho^{\otimes n}) = F(\rho). \tag{3.9}$$

This is an asymptotic equipartition theorem derived for min and max entropies [23] which can directly be applied to the min and max free energies as well. Equation (3.9) means that in the thermodynamic limit the extractable work is equal to the work of formation

$$W^{\text{s.shot}}_{\text{form}}(\rho) = W^{\text{s.shot}}_{\text{ext}}(\rho) = F(\rho) - F(\tau) \tag{3.10}$$

which is the classical result with $F(\rho) = \text{Tr}(\rho H) - \frac{1}{\beta}S(\rho)$ being the von Neumann free energy.

The framework in [1] proposes properties that possible quantifiers for thermodynamic work should satisfy. The authors argue that ε-deterministic work does not satisfy the axioms as stated in Section 2.2 and thus does not qualify as valid work quantifier. In the following I will discuss in detail why the criticism of single shot definitions of work, as stated in [1], is not valid and how it should be interpreted. The aim is to provide context to what extent the frameworks of the paper and the one of single shot thermodynamics are compatible with each other and in which aspects they provide different models for quantum thermodynamics.

3.3 Criticism of the single shot definitions of work

3.3.1 Extracting work from a single heat bath

In [1] it is stated that the notion for extractable work derived from the single shot framework as discussed above is not valid in the sense that it does not satisfy axioms 1 and 2. The authors argue that in this framework it is possible to obtain a positive amount of work using a heat bath and thermal operations only. As an example we consider the *wit* from the previous section. It is shown that it can be brought to a state with high excitation probability using only catalytic thermal operations. This corresponds to extracting work from a single heat bath which violates the second law as incorporated in the axioms 1 and 2 in [1]. Here we will demonstrate how this apparent violation is performed and in the next section give arguments why it actually does not imply a violation of physical laws.

Using the language of [1] the violation becomes apparent when we consider the following work quantifier that captures the notion of ε-deterministic work

$$\mathcal{W}(p_A^{(1)} \to p_A^{(2)}) = f(\rho_A^{(2)}, H_A^{(2)}) - f(\rho_A^{(1)}, H_A^{(1)}), \tag{3.11}$$

with

$$f(\rho, H) = \begin{cases} \Delta & \text{if } \| \rho - |1\rangle\langle 1| \|_1 < \frac{1}{2} \\ 0 & \text{if } \| \rho - |0\rangle\langle 0| \|_1 < \frac{1}{2}. \end{cases} \tag{3.12}$$

The work storage systems are *wit*s which are two level systems with an energy gap Δ and in a state that is at least ε-close to an energy eigenstate of the system. The function (3.12) simply measures if the state of the *wit* is close to the excited or close to the ground state. Therefore, if we bring the *wit* from the ground into the excited state, we say that work Δ was extracted.

In this framework the following two statements are equivalent [17]:

1. For two (block diagonal) states ρ_0 and ρ_1, corresponding to probability vectors v_0 and v_1 there exists a catalytic thermal operation (see Section 1.2)
$$\rho_1 = \mathcal{E}(\rho_0) = \mathrm{Tr}_{S'}\left[U(\rho_0 \otimes \omega_{H,\beta})U^\dagger\right] \tag{3.13}$$
transforming ρ_0 into ρ_1.
2. There exists a stochastic map A with
$$Av_0 = v_1 \qquad \text{and} \qquad A\omega_{H,\beta} = \omega_{H,\beta} \tag{3.14}$$
with the Gibbs state
$$\omega_{H,\beta} = \frac{e^{-\beta H}}{Z_H} \tag{3.15}$$

Consider now the states

$$\begin{aligned} \rho_0 &= r_0|1\rangle\langle 1| + (1-r_0)|0\rangle\langle 0|, \\ \rho_1 &= (1-r_1)|1\rangle\langle 1| + r_1|0\rangle\langle 0|, \end{aligned} \tag{3.16}$$

with $r_0,\ r_1 < \varepsilon$. Define the stochastic map

$$\mathcal{G}_\Delta^0 = \begin{pmatrix} 0 & e^{-\beta\Delta} \\ 1 & 1-e^{-\beta\Delta} \end{pmatrix}. \tag{3.17}$$

It is easy to see that $\mathcal{G}_\Delta^0$ indeed maps the probability vector corresponding to ρ_0 into the one of ρ_1, i.e

$$\begin{pmatrix} (1-r_1) \\ r_1 \end{pmatrix} = \mathcal{G}_\Delta^0 \begin{pmatrix} r_0 \\ (1-r_0) \end{pmatrix}, \tag{3.18}$$

with $r_1 = 1-(1-r_0)e^{-\beta\Delta}$. We find that $r_1 < \varepsilon$ can only be satisfied if

$$\Delta < \frac{1}{\beta}\log\left(\frac{1-r_0}{1-\varepsilon}\right) \approx \frac{1}{\beta}(\varepsilon - r_0), \tag{3.19}$$

where the approximation holds for small r_0 and ε. We conclude, that for every $\varepsilon > 0$ we can find a $\Delta > 0$ such that the transition $\rho_0 \to \rho_1$ can be performed without spending additional resources and using only catalytic thermal operations and $f(\rho_0, H_\Delta) - f(\rho_1, H_\Delta) = \Delta > 0$ which violates axioms 1 and 2 in the paper. It is important to note that for $\varepsilon = 0$ this process is not possible any more. In this case we find $\Delta < 0$ which means that no positive work can be extracted from a heat bath.

3.3.2 Discussion: Work extraction from single heat bath

The above example shows an apparent violation of the second axiom as stated in [1]. Explicitly, we found a work quantifier $\mathcal{W}$ and a process with

$$\mathcal{W}(p \to q) > 0 \qquad \text{with} \qquad q \in \mathcal{F}_C(p), \tag{3.20}$$

which contradicts (2.12) and therefore also the axioms proposed in [1]. Here we will discuss this issue and explain why there is no fundamental violation of physical laws in single shot thermodynamics, which is crucial for its applicability. In particular we argue why it is not possible to construct a perpetuum mobile using work extraction from a single heat bath. We will conclude that, since the axioms rule out the ε-deterministic definition of work, the axioms pose too strong restrictions on work quantifiers and thus do not represent a good model for describing thermodynamic work.

The possibility of extracting work from a single heat bath is well known in single shot thermodynamics [6] (implicitly also [5]). As indicated in equation (3.19) the success probability of extracting work from heat bath is exponentially suppressed in the amount of extracted work, i.e.

$$e^{-\beta W_{\text{ext}}(\omega_{H,\beta})} \geq 1 - \varepsilon, \tag{3.21}$$

with $1 - \varepsilon$ being the success probability. Furthermore, in a single process as given above, it is very likely to extract a positive amount of work from a heat bath. However, the average extractable work of a state is still given by the difference of its von Neumann free energy and the von Neumann free energy of the thermal state, i.e.

$$\langle W_{\text{ext}}(\rho) \rangle = F(\rho) - F(\tau) \tag{3.22}$$

Thus the average extractable work is zero for any thermal state [6]. This means that in the improbable event of a failure one looses a large amount of energy and thus compensates for the energy extracted in the previous runs of the process. In order to explain the event of a failure in which a large amount of energy is lost, we have to consider the catalysts employed in the catalytic thermal operation. In a successful process the catalyst will be returned unchanged. However, with small probability the experimenter will find the catalyst in a state of much lower energy and thus will have lost energy.

The situation can also be pictured by considering the following betting game. Suppose the bet is a million dollar and one has the chance of winning 1 dollar with a very high probability of $1 - 10^{-6}$. The question is, whether one would risk that kind of game a large number of times. Certainly there is the chance to win some amount but it gets compensated by the risk of loosing a very large amount, when one fails. In particular, one can not get arbitrarily rich by playing this game. This example highlights, that it is not possible to construct a perpetuum mobile using the process of extracting work from a heat bath. Doing so would require an arbitrary repetition of the process. In this case one would consider the average extractable work which does not violate the second laws as captured by the axioms in [1].

The reason that the framework in [1] observes an apparent violation of the second law is due to the fact it defines work a priori as a deterministic function on quantum states (or more general: objects). It neglects the probabilistic nature of measurement outcomes and looks only at state transformations which are assumed to be performable with unit probability. It does not consider the notion of different possible outcomes (e.g. success, failure) of an individual process. Single shot thermodynamics on the other hand considers work as the outcome of some energy measurement and as such defines work a priory as a random variable [6]. From the probability distribution of this variable it then derives bounds on work that can be extracted (or used otherwise) with high probability while the underlying notion is still the uncertain outcome in each single run of a process.

We conclude that the framework of ε-deterministic work is an operationally well motivated model for thermodynamic work. In the thermodynamic limit it recovers the classical notion of work, it does not violate any fundamental physical laws and is a generalization of the classical scenario while simultaneously incorporating the probabilistic nature of quantum mechanics. Since the framework in [1] explicitly excludes the ε-deterministic definition of work, we argue that its axioms are too strong in order to describe a good model for thermodynamic work. In particular the axioms demand that a strict second law holds for individual processes in the micro regime – a notion that is also not motivated by statistical mechanics where the entropy of a closed system increases only on average while the entropy can momentarily decrease in the system due to fluctuations.

Chapter 4

Synthesis

In the previous chapters we discussed two different approaches towards definitions of work in quantum thermodynamics. The one from [1] claims that any notion of work has to satisfy certain properties (the axioms) and derives restrictions on the form of valid definitions of work from that. The demanded properties enforce that the second law is strictly satisfied for any definition of work. The single shot framework takes an opposite approach. It defines a priory notions of work and determines, in turn, properties of these.
The goal of this chapter is to put the definition of work in [1] and those of single shot thermodynamics into perspective. We point out to what extent the two different frameworks can be incorporated into one another and what the distinguishing features of the frameworks are. Furthermore, we show that the framework in [1] presents a self consistent model with the restriction of a strictly satisfied second law in the sense that it defines work in a consistent way that is distinct from the notion of heat like energy. Finally, we further compare the two frameworks with respect to its underlying assumptions and applicabilities.

4.1 Single shot work model for $\varepsilon = 0$

Here we show, that the framework in [1] can reproduce a special case of the single shot framework which is the definition of ε-deterministic work for $\varepsilon = 0$. To formulate this definition of work in the language of [1] we consider work storage devices given by objects in the set

$$\mathcal{P}_{\varepsilon=0} = \{(\rho, H_\Delta) \mid H_\Delta = \Delta|1\rangle\langle 1|,\ \rho = |E\rangle\langle E|,\ \Delta \geq 0\}, \tag{4.1}$$

where $|E\rangle$ is either $|0\rangle$ or $|1\rangle$, i.e. an eigenstate of the Hamiltonian. Furthermore we consider the work quantifier

$$\mathcal{W}(p_A^{(1)} \to p_A^{(2)}, \beta) = f(p_A^{(2)}) - f(p_A^{(1)}) \tag{4.2}$$

where

$$f(\rho, H) = \begin{cases} \Delta & \text{if } \rho = |1\rangle\langle 1| \\ 0 & \text{if } \rho = |0\rangle\langle 0|. \end{cases} \tag{4.3}$$

as in (3.12) but with deterministic states ρ.

The set $\mathcal{P}_{\varepsilon=0}$ together with the work quantifier given here satisfy the operational axioms in [1]. Thus it can be seen as a particular instance of a definition of work in the framework. We now look at he extractable work as given in (2.2). For any resource state (object) $p_M = (\rho_M, H_M) \in \mathcal{P}_{\varepsilon=0}$ held by Merlin the extractable work is then defined by

$$W_{\text{ext}}(p_M, \beta) = \\ W_{\text{trans}}(p_M \to \omega_\beta, \beta) = \sup_{(|1\rangle\langle 1|, H_\Delta) \in \mathcal{F}_{\mathcal{C}}(p_M \otimes (|0\rangle\langle 0|, H_\Delta))} \Delta. \tag{4.4}$$

where w.l.o.g. we assume that $p_A^{(1)} = (|0\rangle\langle 0|, \Delta|1\rangle\langle 1|)$ and $p_A^{(2)} = (|1\rangle\langle 1|, \Delta|1\rangle\langle 1|)$ from the definition (2.1). In words, the extractable work is the optimal (maximal) value of Δ such that the transition

$$p_M \otimes (|0\rangle\langle 0|, H_\Delta) \to (|1\rangle\langle 1|, H_\Delta) \tag{4.5}$$

is possible using only catalytic thermal operations. This is just the process described in (3.1) and it was shown in [5] that the optimal value is given by

$$W_{\text{ext}}^{\varepsilon=0}(\rho) = F_{\min}(\rho) - F_{\min}(\tau) \tag{4.6}$$

as in (3.5). A very similar consideration leads to the work of formation that also recovers the same expression as in Chapter 3, i.e.

$$W_{\text{form}}^{\varepsilon=0}(\rho) = F_{\max}(\rho) - F_{\max}(\tau) \tag{4.7}$$

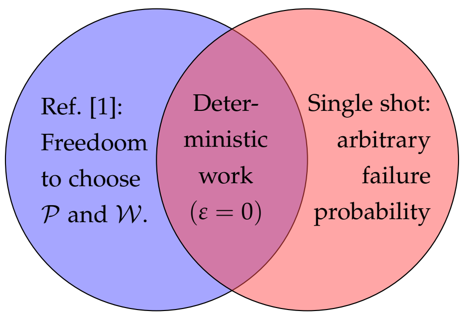

Figure 4.1: The domains of the frameworks presented in [1] and the one of single shot thermodynamics. Blue: The framework in [1] does in principle allow for arbitrary work storage devices given by the set $\mathcal{P}$. The work quantifier $\mathcal{W}$ is restricted to satisfy the two axioms. Red: The ε-deterministic work allows (in principle) for arbitrary failure probabilities ε. For $\varepsilon > 0$ the work quantifier put forward in [5, 6] does not satisfy the axioms given in [1]. Overlap: The only case in which the to frameworks are compatible is the case of deterministic work corresponding to the case of $\varepsilon = 0$.

We therefore see that the single shot definition of work for the case $\varepsilon = 0$ can easily be reproduced in the framework of [1]. However, for failure probabilities $\varepsilon > 0$ the framework of [1] does not include the corresponding notion of work as discussed in the previous chapter, since it allows for work extraction from a single heat bath and hence violates the axioms.
While failing to incorporate single shot definitions of work, the framework in [1] includes the freedom to choose arbitrary sets $\mathcal{P}$ describing valid work storage systems and thus the form in which work is stored. It therefore includes (in principle) a variety of different notions of work, but makes no claim about their operational interpretation, see also Figure 4.1. In particular, all notions of work are bound to satisfy the two imposed axioms. These also enforce a strict second law to be satisfied for all possible work definitions (cf. also Section 2.3).

4.2 Self-consistency of the model

Work in quantum thermodynamics should, as in the classical case, be defined to be an ordered and predictable form of energy. In the language of [1] this means that the energy which was deposited in a work storage system p_A can deterministically be extracted at a later time. Thus we would expect for the extractable work and the work of formation given in (2.2) and (2.3)

$$W_{\text{ext}}(p, \beta) = W_{\text{form}}(p, \beta) \qquad \forall p \in \mathcal{P}, \tag{4.8}$$

which is a special case of (2.4) where equality always holds. As discussed in Section 2.1 in general we have the inequality

$$W_{\text{ext}}(p_M, \beta) \leq W_{\text{form}}(p_M, \beta) \tag{4.9}$$

which is a manifestation of a second law in the sense that it introduces an inherent irreversibility for processes when creating and using resources. The special case in (4.8) can be interpreted as the defining feature of work that distinguishes it from heat. Work can be transferred reversibly without dissipation of energy into heat.

As we recall, in [1] the two expressions (referred to as work value and work cost of a state) are defined as

$$W_{\text{ext}}(p_M, \beta) := W_{\text{trans}}(p_M \to \omega_\beta, \beta), \tag{4.10}$$

$$W_{\text{form}}(p_M, \beta) := -W_{\text{trans}}(\omega_\beta \to p_M, \beta). \tag{4.11}$$

Suppose we consider any given definition of work in this framework determined by a set of work storage systems $\mathcal{P}$ and a work quantifier $\mathcal{W}$. If we assume $p_M \in \mathcal{P}$ we can lower bound the extractable work as

$$W_{\text{trans}}(p_M \to \omega_\beta, \beta) = \sup_{\substack{p_A^{(1)}, p_A^{(2)} \in \mathcal{P} \\ \omega_\beta \otimes p_A^{(2)} \in \mathcal{F}_\mathcal{C}(p_M \otimes p_A^{(1)})}} \mathcal{W}(p_A^{(1)} \to p_A^{(2)}, \beta) \tag{4.12}$$

$$\geq \mathcal{W}(\varnothing \to p_M, \beta), \tag{4.13}$$

where we set $p_A^{(1)} = \varnothing$ (the empty object corresponding to the absence of a work storage system) and $p_A^{(2)} = p_M$. This is feasible since $p_M \in \mathcal{P}$ and $\omega_\beta \otimes p_M \in \mathcal{F}_\mathcal{C}(p_M \otimes \varnothing)$. Equally we upper bound the work of formation

$$-W_{\text{trans}}(\omega_\beta \to p_M, \beta) = \inf_{\substack{p_A^{(1)}, p_A^{(2)} \in \mathcal{P} \\ p_M \otimes p_A^{(2)} \in \mathcal{F}_\mathcal{C}(\omega_\beta \otimes p_A^{(1)})}} -\mathcal{W}(p_A^{(1)} \to p_A^{(2)}, \beta) \tag{4.14}$$

$$\leq -\mathcal{W}(p_M \to \varnothing, \beta) = \mathcal{W}(\varnothing \to p_M, \beta), \tag{4.15}$$

where this time we set $p_A^{(1)} = p_M$ and $p_A^{(2)} = \varnothing$. Combining the two inequalities yields

$$\mathcal{W}(\varnothing \to p_M, \beta) \leq W_{\text{ext}}(p_M, \beta) \leq W_{\text{form}}(p_M, \beta) \leq \mathcal{W}(\varnothing \to p_M, \beta), \tag{4.16}$$

which shows, that within the framework of the paper the condition (4.2) is satisfied for all $\mathcal{P}$.

However, the proof of this self-consistency relies on the fact that swapping of objects is supposed to be a catalytic thermal operation in the framework of [1], i.e.

$$p_1 \otimes p_2 \in \mathcal{F}_\mathcal{C}(p_2 \otimes p_1). \tag{4.17}$$

This corresponds to exchanging all parts of a system between the two players of the game. It is clear, that if Merlin possesses a system in the form of a work storage device, the best strategy he can employ is to simply give all parts of the system to Arthur since he will accept it as a valid form of work. It is important to note that here swapping is not interpreted as interchanging the states (without the Hamiltonian) of the systems. This would in general not be energy conserving and thus not be considered a free operation.
The formation of a state would analogously correspond to Arthur giving the system back to Merlin. Operationally it is questionable if the physical exchange of systems would be a feasible operation in all circumstances. For example, if (Arthur's) work storage systems and (Merlin's) working systems are stationary and spatially separated, physical exchange of the systems is not an option for the optimal procedure.

4.3 Comparison of the models

In the previous chapters we discussed a novel approach towards a consistent definition of work in quantum thermodynamics as introduced in [1]. It started from the premise that work has to be defined as a function on the state of a work storage device, i.e. work is considered a priori deterministic. It neglects the probabilistic notion of work incorporated e.g. in [5–7]. Importantly, this premise was justified by and tied to the condition, that any notion of work should strictly satisfy a second law of thermodynamics implicitly enforced by two axioms. From the axioms, in turn, emerged a condition of the form

$$W_{\text{ext}}(p, \beta) \leq W_{\text{form}}(p, \beta), \tag{4.18}$$

which can be interpreted as an explicit manifestation of the second law already implicitly introduced by the axioms. The authors of the framework claim to take up the most general point of view that any reasonable definition of work should satisfy the given axioms, enforcing the strict second law. We, on the other hand, argue that this condition in fact poses a restriction on definitions of work and only enforces a restricted model of it that does not account for the probabilistic properties of work. In comparison, as discussed in the previous chapter, ε-deterministic work incorporates this probabilistic notion and furthermore also defines a reasonable framework for defining work. The fact that the framework in [1] is unable to incorporate the ε-deterministic work definition shows that it can not be the most general approach towards this issue.

Instead of relating work to a deterministic function on quantum states the single shot framework considers work as the outcome of an energy measurement on systems which are, in general, inherently random. Thus, work is treated as a random variable and single shot thermodynamics makes predictions about its probability distribution. This point of view is supported by statistical mechanics, especially with regard to the apparent violation of the second law: In statistical mechanics it is possible that in any instance the second law is violated due to fluctuation. This also is related to the possibility to extract work from a single heat bath when we consider individual processes in the micro regime (cf. Section 3.3.2). Furthermore, as indicated by the previous section, the framework for ε-deterministic work definitions reduce to the one from [1] in the case of $\varepsilon = 0$.

We conclude that the notion of work put forward in [1] is not reasonable when one wants to define work as an energy that is transferred in individual processes. This is related to the fact that the axioms are imposing a strict second law also for individual processes, a notion that is not justified by results from statistical or quantum mechanics. As discussed in the previous chapter the regime governing single processes is well described in the single shot framework. However, the framework in [1] can be useful when one aims at defining work as an average quantity which always obeys the second law of thermodynamics. Under this premise it presents a self consistent framework that mathematically generalizes some of previous work definitions in a new approach that starts from general properties of work definitions and derives restrictions on the form in which work can be described.

Chapter 5

Conclusion

In order to find a definition of work in quantum thermodynamics, several approaches can be put forward. They all aim at defining work as an ordered and controllable form of energy as opposed to heat which can not be transformed and transported between different systems in a deterministic way. However, it is not clear which approach provides the best description of thermodynamic work when transforming individual quantum systems – a regime where the thermodynamic limit can not be applied. Thus, a comprehensive analysis of different approaches and a comparison of their applicability is vital for finding a good model for the physics of those systems.
The recent approach presented in [1] puts forward an axiomatic framework where work is defined by assigning values to state transitions of some work storage system analogous to a lifted weight or battery. The approach leads to strong restrictions on what can be considered a valid way of quantifying work for individual quantum systems. In particular, these restrictions explicitly rule out previous definitions of work that were established using the single shot approach as laid out in [5–7]. Since the single shot approach has a well established operational interpretation and a mathematically rigorous foundation it is interesting to know why exactly the single shot approach is excluded and if the new axiomatic approach can withstand a rigorous analysis with respect to its operational plausibility.

In this work we presented an extensive analysis of the mentioned axiomatic framework for defining work [1]. We clarified the operational interpretation that the two agents play and discussed the axioms that lead to the restrictions on how to quantify work. In addition, we also discussed results from the single shot approach and provided arguments why it is reasonable from a physical point of view to consider this framework. Furthermore we argued why its exclusions by the results in [1] are operationally not justified. Finally we compared the underlying assumptions of both frameworks as well as their capability for providing a model describing the physics of quantum thermodynamics.

In our analysis we found that the two compared frameworks make very different underlying assumptions about what kind of quantities can represent thermodynamic work. The single shot framework considers work to be a random variable associated to some energy measurements on quantum systems before and after a transition. It then discovers properties of the underlying probability distributions and thus respects the probabilistic nature of microscopic systems as it is suggested by statistical and quantum mechanics. In particular, it allows a "gambling for work" in the sense that better work values can be obtained in a thermodynamic process when one allows a finite probability for the process to fail. This gambling in turn permits processes that are not possible in the classical setting, like spontaneously violating the second law of thermodynamics. At the same time classical thermodynamics is recovered in the thermodynamic limit or when average quantities are considered.

In comparison to the single shot approach, the framework in [1] does not take into account the probabilistic behaviour of quantum systems. It a priori considers work to be a deterministic quantity associated to transformations of a quantum system. Axioms are stated that implicitly enforce a strict second law of thermodynamics for any kind of process. The idea behind this approach is to find a description of physics in the micro regime that resembles the one from classical thermodynamics and therefore takes the second law as its foundation. Since in the single shot approach the second law is not strictly satisfied, it gets a priori excluded by the new approach. Only for the case where we assume a zero probability for a process to fail the framework in [1] recovers the results from the single shot approach. In this case, however, we are again in the regime of deterministic transformations which complies with the new framework but is not operationally justifiable. We provided arguments for the point of view that the single shot approach in fact represents the more useful framework for describing thermodynamics in the micro regime and that the inability of incorporating it is a major weakness of the framework put forward in [1]. The latter is only capable of describing work quantities that behave like averages as those strictly satisfy the second law. It therefore only describes a special case, although claimed to be the most general approach towards the topic. Its axioms enforcing the strict second law are too strong in order to provide a sensible model for quantum thermodynamics.

Although the language of the framework [1] is quite abstract, the novel approach towards defining the properties of work quantifiers in the first place is not unreasonable at all. A topic for future research might be to replace the axioms, which proved to be too strong, with a modified version that is able to incorporate more general definitions like the ones from the single shot approach. This might lead to a generalization of work definitions supported by an axiomatic basis giving new insight to the general structure of an extensive theory of quantum thermodynamics.

Appendix A

The role of correlations for extracting work

A.1 Extension of axiom 2 to incorporate correlation

In [1] also the role of correlations between Merlin's and Arthur's systems is considered. How does the picture change, if after a transition of Merlin's system only the marginal state of the whole system must coincide with the final state, i.e. $p_M^{(f)} = \mathrm{Tr}\left(p_{MA}^{(f)}\right)$? In order to discuss this case, first a different version of the work of transition (2.1) is defined. It incorporates the case where at the end of a process correlations exist between the systems and we refer to it as work of transition with correlations:

$$W_{\text{trans}}^{\text{corr}}(p_M^{(1)} \to p_M^{(2)}, \beta) := \sup_{\substack{p_A^{(1)}, p_A^{(2)} \in \mathcal{P} \\ p_{MA}^{(2)} \in \mathcal{F}_C(p_M^{(1)} \otimes p_A^{(1)})}} \mathcal{W}(p_A^{(1)} \to p_A^{(2)}, \beta), \tag{A.1}$$

with $p_{MA}^{(2)}$ having marginal states corresponding to the ones in the definition of the work of transition in (2.1).

Having established this definition an extended version of Axiom 2 is stated:

- **Axiom 3** Consider a cyclic sequence of transitions of Merlin's system where he repeatedly takes systems from Arthur and uses them as assisting systems to perform the transitions

$$\{p_M^{(1)} \to p_M^{(2)} \to \ldots \to p_M^{(n)} = p_M^{(1)}\}. \tag{A.2}$$

Then the total work of transition with correlations satisfies

$$\sum_{k=1}^{n-1} W_{\text{trans}}^{\text{corr}}(p_M^{(k)} \to p_M^{(k+1)}, \beta) \leq 0 \tag{A.3}$$

This means that Merlin also can not become arbitrarily rich by repeating such a process even if he builds up correlations between his and Arthur's system. Note that there is an alternative formulation of this axiom, analogous to the case in Axiom 2. The only difference is, that in each single transition the catalyst starts uncorrelated with Arthur's system $p_A^{(k)}$ but might be correlated with the final state $q_A^{(k)}$(see Figure A.1). The equality of these formulations is shown in Section A.2.

As stated in the paper, one can take two different points of view. The first one just forbids any correlation between work storage system held by Arthur and Merlin's system in the beginning and at the end of each transition in the spirit of axiom two. One then needs to assert this condition since it was shown that otherwise, the "absence of correlation" could be used as a resource in order to extract work from a system using only free operations and thus violating axiom 2. The other point of view just incorporates correlations in the notion of free operation and demands axiom 3 to be satisfied instead of axiom 2. It is important to note, that axiom 3 is strictly stronger than axiom 2 and thus, the consequences of the axioms in the following section depend on which axioms are applied.

A.2 Reformulation of axiom 3

Here we derive an alternative formulation of axiom 3 in analogy to the alternative formulation of axiom 2.

- **Axiom 3 (reformulated)** Let $\{p_A^{(k)} \to q_A^{(k)}\}_{k=1}^{n-1}$ be a collection of transitions of Arthur's system, assisted by $(p_M^{(k)}, p_M^{(k+1)})$ with $p_M^{(n)} = p_M^{(1)}$ (see Figure A.1) where after each transition the systems of Arthur and Merlin can end up correlated. Then the total work satisfies

$$\sum_{k=1}^{n-1} \mathcal{W}(p_A^{(k)} \to q_A^{(k)}, \beta) \leq 0. \tag{A.4}$$

This formulation is similar to the alternative formulation of axiom 2. The only difference is, that for each transition the work-storage device starts uncorrelated with the assisting system but can (possibly) end up correlated with it. We will show now that the two formulations of axiom 3 are indeed equivalent. Let us assume the version in the paper to be true: For any cyclic sequence of Merlin's system $p_M^{(1)} \to \ldots \to p_M^{(n)} = p_M^{(1)}$, we have

$$\sum_{k=1}^{n-1} W_{\text{trans}}^{\text{corr}}(p_M^{(k)} \to p_M^{(k+1)}, \beta) \leq 0. \tag{A.5}$$

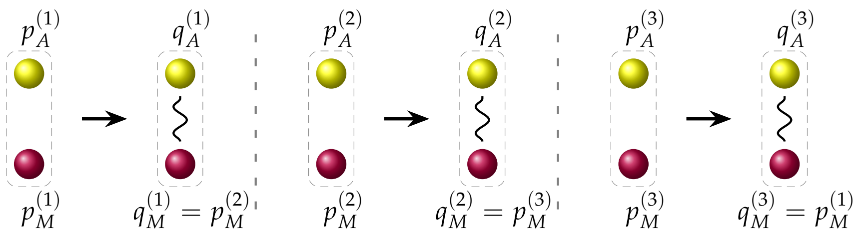

Figure A.1: A set of transitions in the spirit of Axiom 2. Merlin obtains $p_A^{(1)}$ from Arthur and uses his system $p_M^{(1)}$ as an assistant to transform the joint state into the (possibly) correated joint state $q_{AM}^{(1)}$. After that Arthur gives Merlin a new system, so that he now holds $p_A^{(2)} \otimes q_M^{(1)}$. Merlin performs several transitions that way where after the last transition his system is back to the initial state $q_M^{(1)}$. Note that during the whole process the systems can be correlated with each other indicated by wiggly lines (cf. Figure 2.1 for absence of correlations). Nevertheless in each transition the systems must start off uncorrelated. The axiom states, that Merlin should not have a net benefit from performing such a cyclic process on his system.

Now consider a set of assisted correlated transitions $\{p_A^{(k)} \to q_A^{(k)}\}_{k=1}^{n-1}$ of the work-storage device, assisted with correlations by $(p_M^{(k)}, p_M^{(k+1)})$ respectively[1], with $p_M^{(n)} = p_M^{(1)}$, as stated in axiom 3. Using the definition in (A.1), we see that

$$\mathcal{W}(p_A^{(k)} \to q_A^{(k)}, \beta) \leq W_{\text{trans}}^{\text{corr}}(p_M^{(k)} \to p_M^{(k+1)}, \beta), \tag{A.6}$$

for all $k \in \{1, \dots, n-1\}$. From that we immediately obtain (A.4).

Now we show the converse direction. We have to show, that for a given sequence $p_M^{(1)} \to \dots \to p_M^{(n)} = p_M^{(1)}$, Eq. (A.4) implies Eq. (A.5). Each transition $p_M^{(k)} \to p_M^{(k+1)}$ will also induce a transition on the marginal of the work storage system, given by $p_A^{(k)} \to q_A^{(k)}$ with

$$q_{AM}^{(k)} \in \mathcal{F}_C(p_A^{(k)} \otimes p_M^{(k)}, \beta) \tag{A.7}$$

and

$$\mathrm{Tr}_A\left(q_{AM}^{(k)}\right) = p_M^{(k+1)} \qquad \mathrm{Tr}_M\left(q_{AM}^{(k)}\right) = q_A^{(k)}. \tag{A.8}$$

That said, all possible marginal transitions on Arthur's system $p_A^{(k)} \to q_A^{(k)}$ form an assisted transitions with correlations, assisted by $(p_M^{(k)}, p_M^{(k+1)})$. By our assumption (Eq. (A.4)) this implies that the total work-value satisfies

$$\sum_{k=1}^{n-1} \mathcal{W}(p_A^{(k)} \to q_A^{(k)}, \beta) \leq 0 \tag{A.9}$$

[1]The upper indices seem a bit confusing at first but they simply reflect the process of the form depicted in Figure A.1

for all $p_A^{(k)} \to q_A^{(k)}$ as given in Eq. (A.7) and (A.8). Then, this trivially implies

$$\sum_{k=1}^{n-1} \sup_{\substack{p_A^{(1)}, p_A^{(2)} \in \mathcal{P} \\ p_{MA}^{(2)} \in \mathcal{F}_{\mathcal{C}}(p_M^{(1)} \otimes p_A^{(1)})}} \left(\mathcal{W}(p_A^{(k)} \to q_A^{(k)}, \beta) \right) \leq 0. \tag{A.10}$$

Now notice that the supremum in Eq. (A.10) is the same as the one in Def. 6 of $W_{\text{trans}}^{\text{corr}}$ in the paper, which concludes the proof.

Bibliography

[1] R. Gallego, J. Eisert, H. Wilming, arXiv:1504.05056v3.

[2] P. Hänggi, F. Marchesoni, *Rev. Mod. Phys.* **81**, 387–442 (2009).

[3] G. Cerefolini, *Nanoscale Devices* (Springer, 2009).

[4] A. Dechant, N. Kiesel, E. Lutz, *Phys. Rev. Lett.* **114**, 183602 (2015).

[5] M. Horodecki and J. Oppenheim, *Nature Comm.* **4**, 2059 (2013).

[6] J. Åberg, *Nature Comm.* **4**, 1925 (2013).

[7] O. Dahlsten, R. Renner, E. Rieper, V. Vedral, *New J. Phys.* **13** (2011).

[8] D. Reeb and M. M. Wolf, *New J. Phys.* **16**, 103011 (2014).

[9] R. Alicki, M. Horodecki, P. Horodecki, R. Horodecki, *Open Systems & Information Dynamics (OSID)* **11**, 205 (2004).

[10] D. Egloff, O. Dahlsten, R. Renner, V. Vedral, *New J. Phys.* **17**, 073001 (2015).

[11] P. Skrzypczyk, A. J. Short, and S. Popescu, *Nat. Commun.* **5**, 4185 (2014).

[12] J. Millen, A. Xuereb, *New J. Phys.* **18**, 011002 (2016).

[13] J. Goold, M. Huber, A. Riera, L. del Rio, P. Skrzypczyk, arXiv:1505.07835 (2015).

[14] M. Horodecki and J. Oppenheim, *Int. J. Mod. Phys. B* **27**, 1345019 (2013).

[15] F. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, R. W. Spekkens, *Phys. Rev. Lett.* **111**, 250404 (2013).

[16] F. Brandão, M. Horodecki, N. Ng, J. Oppenheim, S. Wehner, *Proc. Nat. Acad. Sci. USA* **112**, 3275 (2015).

[17] D. Janzing, P. Wocjan, R. Zeier, R. Geiss, T. Beth, *Int. J. Th. Phys.* **39**, 2717 (2000).

[18] P. Faist, J. Oppenheim, and R. Renner, arXiv:1406.3618.

[19] K. Maruyama, F. Nori, V. Vedral, *Rev. Mod. Phys.* **81**, 1 (2009).

[20] R. Landauer, *Phys. Today* **44** (5), 23 (1991).

[21] L. Del Rio, J. Åberg, R. Renner, O. Dahlsten, V. Vedral, *Nature* **474**, 61–63 (2011).

[22] R. Renner, arXiv:quant-ph/0512258.

[23] M. Tomamichel, arXiv:1203.2142

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Declaration of originality

The signed declaration of originality is a component of every semester paper, Bachelor's thesis, Master's thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor.

Title of work (in block letters):

Authored by (in block letters):
For papers written by groups the names of all authors are required.

Name(s): **First name(s):**

With my signature I confirm that
- I have committed none of the forms of plagiarism described in the 'Citation etiquette' information sheet.
- I have documented all methods, data and processes truthfully.
- I have not manipulated any data.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for plagiarism.

Place, date **Signature(s)**

For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.