The Thermodynamic Limit in the Resource-Theoretic Framework

Richard David Petrak East

University of Oxford

A thesis submitted for the degree of

Master of Science

August 2017

Acknowledgements

I want to thank Professor Bob Coecke for setting me up on this project and directing me to an area of study which complements my past understanding and also has allowed me to see categorical methods applied to a technically and philosophically interesting area of fundamental physics.

Thanks is also due to Professor Michal Horodecki for his help in constructing a protocol to correct the proof in earlier literature - without his assistance no doubt much time would have been squandered on the issue.

I also want to thank Carlo Maria Scandolo for his weekly discussion of my work and patience in peering over proofs and ideas of highly variable quality in pursuit of the results presented here. With his assistance in grasping the larger concepts and themes of this area this I have been able to ground the study of rates in a larger practical context in a way that would not normally be possible given the limited time allotted for this research.

In addition I would like to thank Hannah whose scale of lack of contribution to this thesis is equalled only by the scale of her desire to be acknowledged in it: well here we are I suppose.

Finally I want to thank my supportive family and friends who have made this year one to remember. In particular I wish to single out my Grandparents whose generosity has allowed my being here at all; I am concious that in having your support I am fortunate where many are not.

Abstract

In this thesis the abstract resource theoretic approach to thermodynamics is used to derive a number of results regarding the transition between resources of different *values* assigned by monotone functions. In particular the focus is on the behaviour of these systems as one approaches the thermodynamic limit. Transition rates are presented in a novel manner in relation to resource theories and new results are presented alongside generalisations and corrections of work found within quantum information theory literature.

Explicitly these new results are as follows. It is shown that the value of a regularised monotone on the input resource of a transition is bounded from below by the value of the regularised monotone of the output resource times the rate of transition. Following this it is shown that for reversible resource theories with a currency all regularisable monotonic functions are unique up to a constant. The study of rates is then extended to asymptotically exact transitions which allows for the inclusion of transitions which only exist in the limit into the existing framework. The key result shows that rates of transition between two resources in asymptotically reversible theories are related to the ratio of their value as dictated by the regularised monotones.

Finally the resource theory of purity is outlined as a proving ground for the usefulness of these results and final new result is presented where sharp theories with purification are shown to be asymptotically reversible as far as the resource theory of purity is concerned.

Contents

1

Introduction

If one takes the leap past solipsism then there are few things as certain as thermodynamics. Solitons may be sophistry, neurons nonsense, and mitochondria misguided - but hot things get cold. Thermodynamics is the most phenomenologically accessible physical theory, as such it is the most well tested . When one derives results that contravene thermodynamics, one has obtained solid evidence they are wrong. Phenomenological conviction however, does not indicate understanding. As the years have passed thermodynamics has been reinterpreted through the lens of each scientific paradigm. In particular the advent of statistical mechanics provided fertile soil to give justification to many of the macroscopic rules and properties discovered. Statistical mechanics uses probability to understand large scale properties of mechanical systems[23], for instance if one considers an ideal gas, though one doesn't know the exact position and momentum of every particle, the temperature of the system can arrived at by using using the average speed of all its particles. Into the 20^{th} century quantum mechanics lead to new analysis of the underlying structure and extensions of previous results, this however, was still focussed on statistical methods with numerous particles barely interacting. This traditional approach, or philosophy, to thermodynamics was the standard for well over 100 years. In more specific terms, in both the classical or quantum setting, the analysis focussed on systems *in the thermodynamic limit*. This means one has a large number of weakly interacting particles; the prototypical example being the kinetic theory of gasses. In this work the focus will be on a

broader approach known as *single-shot thermodynamics* which is capable of analysing small and strongly correlated (quantum) systems immersed in heat baths.

Single shot thermodynamics is distinct from many physical theories in that it takes an "agent-centric" approach. One models thermodynamics as a resource theory, where an agent has access to certain states, and based on their *value*, certain transitions may or may not be possible. Specifically for thermodynamics one is interested in states of knowledge and the energy of a system[1]. The focus here will be on systems of fixed energy which is to say the thermodynamics of an isolated system. At fixed energy the deciding factor is the aforementioned states of knowledge, the certainty regarding a system, which will be denoted its *purity* - a term which takes its name from the idea of pure states in quantum mechanics, where one is certain of what quantum state one has, as opposed to mixed states where one is unsure. This discussion of purity as a resource motivates a different name for single-shot thermodynamics - *The resource theory of purity.*

This resource approach to thermodynamics is pertinent to nano-scale systems [19] as thermodynamics on this scale is anything but in the thermodynamic limit. It is interesting to note that much of the impetus behind thermodynamics was originally motivated by the industrial revolution in the united kingdom during the 18th century[26]. The capability for nano-scale engineering is likely to provide similar levels of motivation to understanding systems far from the limit. Resource theories allow one to deduce quantifiable boundaries on behaviours of systems in this region. Past success in this area include Landauer's work on the physical bounds on information processing, he showed that the necessary physicality of information - due to the fact information must fundamentally be represented by the state of a system- imposes a fundamental limit on the energy it takes to 'forget' [2] something in a system. This result also brought to a close a long-standing problem in the philosophy of physics[3] by showing that Maxwell's demon cannot break the second law of thermodynamics [3]. To understand this better consider memory in the most basic binary terms, a one for true a zero for false, this is stored by a system being in one of two states which correspond to these binary options. Suppose further one has a

[1]The work of Emmy Noether famously links energy to the existence of time symmetry - in fact the resource theory of symmetry in combination with the, to be discussed, theory of purity encapsulates thermodynamics proper[18].

[2]Namely $K_B T ln2$ where k_B is the Boltzmann constant and T is the temperature

[3]Actually there was still a bit of philosophical debate about this centring on the Ladyman and Norton controversy [22], however the introduction of further rigour by Abramsky in [1] brought this to an end, Maxwell's Demon, for all it knows, is as much a slave to entropy as the rest of us.

"*forgetting*" procedure which takes the current state of the system and always leaves the system in the zero state. In entropic terms one has gone from possibly being in two states to one state - this corresponds to an entropy change of $(-ln2)$, for a system at a temperature T, with the Boltzmann constant K_BT, there is an associated energy change K_BTln2[3]. One can approach this fact from another perspective, it implies that the minimum energy separation between a certain state and a state where the previous state is completely unknown is this very energy. In other words the minimum energy required to go from a mixed state to a pure state is K_BTln2. Extrapolating from this one can view the cost of forming any state, of some arbitrary level of uncertainty, in terms of the amount it costs to form a pure state multiplied by the amount of the state one can, in a sense, extract from said pure state. Before discussing this idea of using pure states to reach other states, there is slightly more to be said on the relationship between pure states and energy - namely that while one can spend energy to go from a mixed state to the pure, the converse allows the extraction of energy. One has a physical manifestation physical manifestation of the old adage that "*knowledge is power*".

To understand this better it is helpful to consider the work of Szilard and his eponymous engine shows that knowledge of the state of a system can be directly linked to energy extraction. To understand this, Szilard proposes the classical situation of a particle bouncing around in a box that has a removable partition down the centre and two pistons as sides, see the below figure for a clearer idea.

The idea is that one slides in the partition, and then if one knows what side the particle is in, one can push the piston in on the other side. When the partition is removed the piston will be pushed out extracting useful work. In this way a pure state, the state where one knows the position of the particle, equates to work! The application of this principle for quantum systems has been studied in the bosonic and fermionic setting [20].

The question remains as to what one does without a pure state. This motivates the study of rates. In quantum mechanics one can use a collection of mixed states to get a smaller number of pure states. Essentially by knowing a quite a few things with a degree of uncertainty one can extract a smaller number of certainties. This motivates one to consider exactly how many pure states can be *distilled* from a collection of other states. It also motivates the contrary question of how few pure states one can get away with spending to get some mixed states should one require them - this is a question of *cost*. Rates of transition between states can be studied

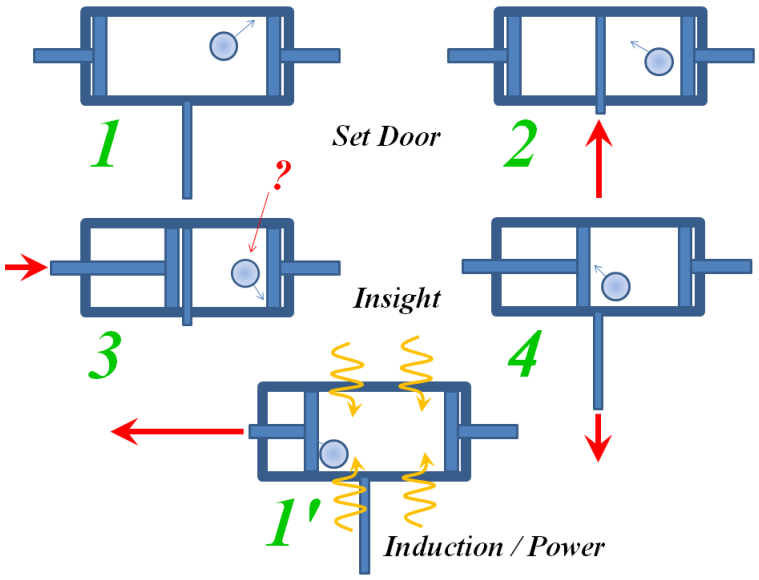


Figure 1.1: In 1 the particle is bouncing around the box unimpeded. In 2 the partition is imposed, a process that can occur arbitrarily slowly which implies there is no lower bound on the energy this requires. The partition effectively reduced the number of micro-states the particle can inhabit in half, which is an entropy change of $-\ln 2$. In 3 one uses the knowledge of which side the particle is on and slides the piston on the other side of the partition forward, again a process with no lower energy bound. In 4 the the partition is removed. In 5 the particle bounces into the piston pushing it back, at which point energy can be extracted to the tune of $K_B T \ln 2$ joules. Note the particle maintains its energy because the system is in a heat-bath; *this is not energy creation.*

abstractly in terms of resource theories. In this way one can focus on the underlying structure and avoid technicalities induced by working with a particular example of a resource theory. Once the results are obtained one can take a particular case, like the resource theory of purity, and directly deduce bounds relevant to that system .

The informational approach is not limited to thermodynamics, this can be seen as part of the encroaching informational tide washing over physics. Another field in which this approach is important is in the reformulation of quantum mechanics. From the outset there has been dissatisfaction with the mathematical underpinning and philosophical interpretation of quantum mechanics. Consider the visceral foundations of other physical theories and principles: Newtons laws, the way objects interact when they collide; Special Relativity, the maximum speed of something is the speed of light; General relativity, how the mass of something affects how something moves around it; Fermat's Principle, light goes in straight lines; Even thermodynamics, traditionally one of the murkier frameworks is essentially the fact energy is constant, and entropy, a measure of how many ways things can be arranged to get the same general result, is low when its cold and increases as things happen in the universe. While these explanations may not be the

most erudite or even clear description of these theories, what they do outline is they are founded on physical principles, reasoning about physical things. The traditional foundations of quantum mechanics on the other hand rely on abstract statements about Hilbert spaces and projectors. In the words of Asher Peres "*...quantum phenomena do not occur in a Hilbert space, they occur in a laboratory*"[29]; this notion motivates alternative phenomenological foundational axioms. The key principles for this philosophy is that one should have physical axioms underpinning a physical theory. In quantum mechanics the physical elements are the observations and what an observer can know - therefore the axioms should talk of states of knowledge.

One approach to achieve this is through what are called generalised probability theories. This framework was created to describe a *landscape of theories*, one of which is quantum mechanics, where each different theory possesses different combinations of structure inherent to quantum mechanics. The aim of this is to pin point exactly what the structural source of various quantum mechanical behaviours are. It also gives researchers flexibility when analysing candidate replacement theories to quantum mechanics motivated by research such as quantum gravity[15].

As current research stands one prominent reformulation of quantum mechanics, motivated by the work done in generalised probability theories, is based on what are termed *sharp theories with purification*. The axioms of these theories attempt to provide the informational foundation desired for the reasons discussed above.

The study of thermodynamics and the resource theoretic questions it motivates, applied to the resource theory of purity, considered from within the framework of generalised probability theories, can be considered the general theme of the thesis. Specifically, resource exchange rates will be analysed and the results will be applied to studying the resource theory of purity. By describing (quantum) thermodynamics via sharp theories with purification as the resource theory of purity, the relationship between the single-shot behaviour of a system and the behaviour of the system in the asymptotic or *thermodynamic* limit, will be presented.

The specifics of the the sections ahead are as follows. Firstly the framework of OPTs is introduced (chapter 2) as the variant of generalised probabilistic theories used in this thesis. This structure is then formalised in terms of category theory (chapter 3). With the category theory framework

outlined, resource theories are defined as a specific structure where the categories maps are partitioned into free and costly processes (section 3.2).

With resource theories defined the next section will focus on rates (chapter 4) as they relate to resources in the abstract. Definitions and properties in a wide variety of situations will be presented. This section constitutes the vast majority of the novel results provided by this author. A number of properties are discussed and propositions and theorems are presented. In addition generalisations of work specific to entanglement literature is also presented, including a theorem which generalises a major result which was discovered by this author to be incorrectly proven - a correct proof is herein presented (Theorem 4.1.7). In addition to the novel results, this section also brings a certain degree of order to what is currently a jumbled area[11][12][21] of study where numerous, ostensibly contradictory, notations exist. The work here has tried to present, from first principles, a clear and reasoned approach.

Finally the specifics of purity theory are addressed (chapter 5) and this author presents result that was previously proven specifically for quantum theories and is here generalised to sharp theories with purification (Theorem 5.7.3). This result shows that in such theories there is but one entropy in the thermodynamic limit and as such shows coherence between single shot thermodynamics at a fixed temperature as approached through the framework presented here and macroscopic thermodynamics as its understood in the thermodynamic limit.

2

Operational Probability Theories

The general approach in this section is to create a framework in which one is focussed on the implementation and combination of measurements in the abstract. The motivation behind this is to try and outline a *landscape of theories* focussed on describing experiments, one of which is to be quantum theory. In this way one can begin to identify the logical origins of the properties of quantum mechanics relative to the only phenomenological experience one ever actually has of quantum systems, the measurement and extraction of information from experiments.

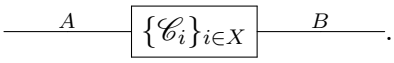
2.1 Foundational Concepts [4]

The initial step is to consider in rough terms what defines a set of experiments, in a sense one aims to define a meta-laboratory. The basis of such a thing would surely be an abstract *device* which would have an input and an output that in some way yet to be defined "performs an experiment". Next attach a label to the input and the output in this way one can tell what goes in and what comes out. Now suppose one can create a new device by putting two old ones in series, but one may only do this if the output label of the first matches the input label of the second. From a computer science perspective one has a typing. Consider also the parallel composition of devices, creating a new device which acts on the combined inputs of the old devices. Its output is then the combination of the composite outputs. In this way one begins to form a non-specific framework for experimental observation.

With composition outlined the natural next question is: what are these devices? The answer is that a device corresponds to a test, which is a process that produces an outcome $i \in X$ where X is the set of possible outcomes that this test can produce. The outcome of a test indicates an event has occurred. In a physical theory this situation corresponds to a measuring device acting on system A and returning a system B; these systems can be anything from the contents of a test tube to the click of a Geiger counter. The specific elements of B returned by the device are referred to as the measurement. With these ideas in mind it is appropriate to begin to formalise this non-specific experimental framework.

Suppose there exists some arbitrary physical systems denoted with letters such as $A, B, C, ...$ reserving I for the trivial system which simply represents nothing. These are the systems one moves between when using a device

Definition 1 (Test)**.** *A test with input system A and output system B is a collection of events $\{\mathscr{C}_i\}_{i\in X}$ labelled by outcomes in some outcome set X. Diagrammatically the test $\{\mathscr{C}_i\}_{i\in X}$ is represented as*



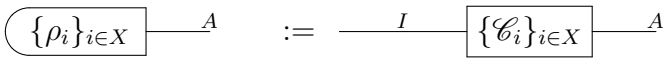
(2.1)

with each event represented as

(2.2)

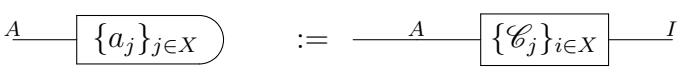
Note the wires for the trivial system are omitted. A test can be thought of as the collection of events that constitute the possible outcomes of using some device.

The definition above allows for a test to have trivial input, that is to say a corresponding device would act on nothing and just provides an output, such a thing is called a *preparation test*; conversely a test with a trivial output is called an *observation event* which only takes an input and produces nothing. Diagrammaticality these are written



(2.3)

and



(2.4)

respectively.

It is practical to write $\mathcal{E}(A, B)$ to denote the set of all events occurring in all tests taking one from A to B; written $\mathcal{E}(A)$ if $A = B$. In this theme, the set $\mathcal{P}(A) := \mathcal{E}(I, A)$ is the set of all preparation events. It can be helpful to write this formulaically, a natural choice is $|\rho)_A$, where $\rho \in \mathcal{P}(A)$ - note it has a deliberate '*Dirac-like*' structure. One can similarly define the set of all observation events $\mathcal{O}(A) := \mathcal{E}(A, I)$ with the formulaic representation $(a|_A$ for $a \in \mathcal{O}(A)$.

Remark 1. *[27] In quantum theory an observation-test is a quantum measurement and is represented by a positive operator valued measurement (POVM) which is a collection of positive operators $\{P_i\}_{i\in X}$ satisfying $\sum_{i\in X} P_i = I_A$ where I_A is the identity on the space of a system A.*

An important case of tests are those that return a guaranteed result.

Definition 2 (Deterministic Test)**.** *A test with a single outcome is deterministic. In this case* $|X| = 1$ *as the set of outcomes has a single element.*

With tests outlined the next step is to consider their sequential composition - the act of applying one device after another. The idea that the outputs and inputs must match immediately asserts itself - one cannot reasonably apply the technique of NMR spectroscopy to the resonance of a crystal: the systems are not compatible. For contrast note that one could apply spectroscopy to the result of the introduction of a catalyst to a substrate. Experiments are inherently typed.

Definition 3 (Sequential composition of tests)**.** *If* $\{\mathscr{C}_i\}_{i\in X}$ *is a test from* A *to* B *and* $\{\mathscr{D}_j\}_{j\in Y}$ *is a test from* B *to* C*, then their sequential composition is a test from* A *to* C*, with outcomes* $(i,j) \in X \times Y$*, and events* $\{\mathscr{D}_j \circ \mathscr{C}_i\}_{(i,j)\in X\times Y}$*. This is written diagrammatically as*

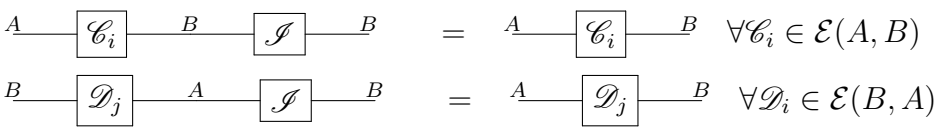
(2.5)

One says that test $\{\mathscr{C}_i\}$ *precedes* $\{\mathscr{D}_i\}$*, or conversely, that* $\{\mathscr{D}_i\}$ *follows* $\{\mathscr{C}_i\}$*.*

This ordering is not necessarily a temporal relation - it refers only to whatever the notion of composition happens to mean in the particular instance in which these tests are combined.

This sequential composition naturally suggests an identity.

Definition 4 (Identity Test)**.** *The identity test for system* A *is a test with a single event* $\mathscr{I}_A$ *such that for every system* B



(2.6)

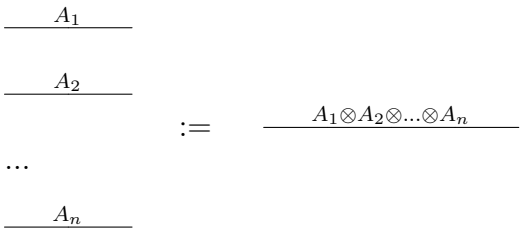
The other form of composition was based on taking two devices and combining them in parallel. To define this formally one requires a definition for their combined inputs and outputs.

Definition 5 (Composite System)**.** *If A and B are systems, the corresponding composite system is $A \otimes B$, often shortened to AB. Composition of systems are such that*

1. $A = I \otimes A = A \otimes I$
2. $A \otimes B \cong B \otimes A$
3. $A \otimes (B \otimes C) \cong (A \otimes B) \otimes C := A \otimes B \otimes C$

The first property outlines composition with nothing is simply what you already have. The second and third demonstrate that the orderings and groupings of the systems compositions are irrelevant.

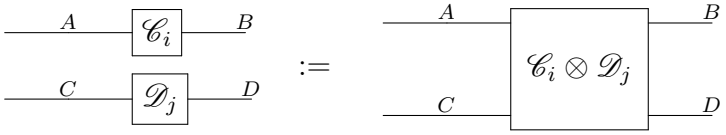
In general any N-partite composite system $A_1 \otimes A_2 \otimes ... \otimes A_N$ with N wires will be represented as



(2.7)

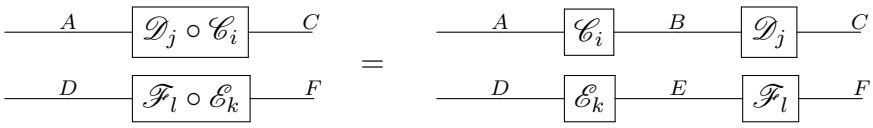
where the trivial systems wires are omitted.

Definition 6 (Parallel composition of tests)**.** *If $\{\mathscr{C}_i\}_{i \in X}$ is a test from A to B and $\{\mathscr{D}_i\}_{i \in X}$ is a test from C to D, then their parallel composition is a test from $A \otimes C$ to $B \otimes D$ with outcomes $(i, j) \in X \times Y$ and events $\{\mathscr{C}_i \otimes \mathscr{D}_j\}_{(i,j) \in X \times Y}$ is represented as*



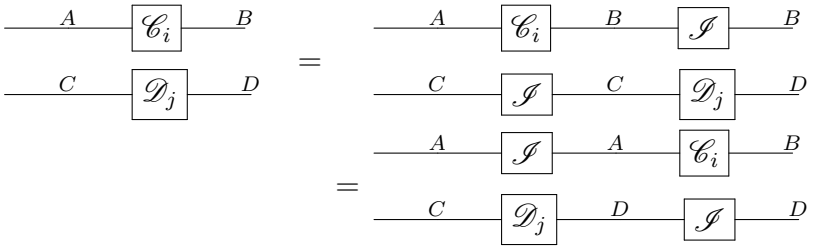
(2.8)

If $\mathscr{C}_i, \mathscr{D}_j, \mathscr{E}_k, \mathscr{F}_l$ are events form , A to B, B to C, D to E, E to F, respectively, their parallel composition is such that



(2.9)

Note that as the ordering of parallel composition is irrelevant as the above implies that tests on different systems commute. Indeed for any pair of events $\mathscr{C}_i$ and $\mathscr{D}_j$ one has that



from here on the identity test will not be explicitly stated but left as a wire.

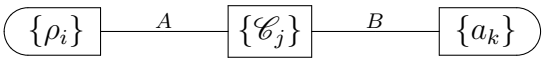
Given the diagrammatic compositions presented above one might ask how to represent this formulaically. For $\mathscr{C} \in \mathcal{E}(A, B)$ and $\rho \in \mathcal{P}(A \otimes C)$ one writes $\mathscr{C} \circ |\, \rho)_{A\otimes B}$ in the place of $(\mathscr{C} \otimes \mathscr{I}_B)|\, \rho)_{A\otimes B}$. On the subject of contractions, the following shorthand is also introduced: $\mathscr{C} \circ |\, \rho)_{A\otimes B}$ will be written $\mathscr{C}|\, \rho)_{A\otimes B}$, and $|\, \rho)_A \otimes |\, \rho)_B$ written as $|\, \rho)_A\, |\, \rho)_B$.

With systems and tests adequately described and their compositions studied, one is in a position to describe *theories* of such systems via their closure under said compositions - these are called *operational theories*.

Definition 7 (Operational theory)**.** *An operational theory is specified by a collection of systems, closed under composition, and by a collection of tests, closed under parallel and sequential composition.*

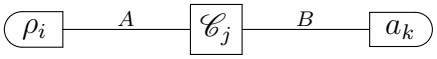
As will be discussed later, the informational structure of tests and systems that has emerged from the idea of composing measuring devices bears a strong resemblance to categories.

In an operational theory one draws *circuits* that represent the connection between experimental apparatus such as:



(2.10)

one can also represent a specific sequence of events



(2.11)

here one can see this circuit depicts the preparation event ρ_i followed by event $\mathscr{C}_j$ from the system A to the system B followed by the observation event a_k. One can consider the whole thing as a single event writing $P_{kji} := (a_k|_B \mathscr{C}_j|\rho_i)_A \in \mathcal{E}(I, I)$ an event from the trivial system to itself.

An operational theory is a framework in which one can describe experiments in the abstract both in their totality and for some specific outcome. A physical theory however, requires more - it requires prediction - one seeks to give the probability of any possible outcome. Technically one requires the assignment of a probability to every event from the trivial system to itself. In this way any complete sequence of preparations, tests, and observations results in a number between zero and unity denoting its likelihood. Furthermore, summing over the possibility of every possible sequence of events gives unity, in short one guarantees that one of the possible outcomes must actually occur.

Definition 8 (Operational-probabilistic theory)**.** *An operational theory is probabilistic if for every test* $\{P_i\}_{i\in X}$ *from the trivial system* I *to itself one has* $p_i \in [0,1]$ *and* $\sum_{i\in X} p_i = 1$*, and the composition of two events from the trivial system to itself is given by the product of probabilities:* $p_i \otimes q_j = p_i \circ q_j = p_i q_j$*. These theories are often abbreviated to OPTs which is how they will be referred to hereafter.*

From a preparation-event ρ_i for a system A one can define a function $\hat{\rho}_i$ sending observation-events of A to probabilities:

$$\hat{\rho}_i : \mathcal{P}(A) \to [0,1], \quad (a_j \mid \mapsto (a_j|\rho_i). \tag{2.12}$$

Similarly, for an observation-event a_j defines a function $\hat{a}_j$ from preparation-events to probabilities

$$\hat{a}_j : \mathcal{O}(A) \to [0,1], \quad \mid a_j) \mapsto (a_j|\rho_i). \tag{2.13}$$

Observe that in this context two preparation-events or observation-events corresponding to the same function are indistinguishable. From this one can move towards defining states and effects as the distinguishable preparation and observation events.

Definition 9 (States and effects)**.** *Equivalence classes of indistinguishable preparation-events are called states. Equivalence classes of indistinguishable observation-events are called effects.*

From here on indistinguishable preparation-events will be defined as states which are elements of $St(A) := \mathcal{P}(A)/_\sim$, the quotient space of preparation events over the equivalence class defined by indistinguishable elements. In a similar way $Eff(A) := \mathcal{O}(A)/_\sim$ where the equivalence class here is indistinguishable observation events. The distinction between an event $\rho_i(a_j)$ and the corresponding function $\hat{\rho}_i(\hat{a}_j)$ will also be dropped.

One can see that a consequence of states and effects being equivalence classes of indistinguishable events is that states are *separating* for effects, and the effects are *separating* for states

$$|\ \rho_0)_A = |\ \rho_1)_A \ \ iff \ \ (a|\rho_0)_A = (a|\rho_1)_A \ \ \forall a \in Eff(A)$$
$$(a_0\ |_A = (a_1\ |_A \ \ iff \ \ (a_0|\rho)_A = (a_1|\rho)_A \ \ \forall a \in St(A)$$

As states are functions from effects to probabilities, and vice versa for effects acting on states, one can take linear combinations of them. This defines two real vector spaces $St_{\mathbb{R}}(A)$ and $Eff_{\mathbb{R}}(A)$, one dual to the other. The focus in this work is on states spanning finite dimensional vector spaces, due to its relevance for quantum mechanics and generalisations thereof. This, by construction implies $St_{\mathbb{R}}(A)$ and $Eff_{\mathbb{R}}(A)$ are of the same dimension, as vector spaces have the same dimension as their dual spaces.

One should note that every event $\mathscr{C}_k$ from A to B induces a linear map $\hat{\mathscr{C}_k} : St_{\mathbb{R}}(A) \to \hat{St}_{\mathbb{R}}(B)$, uniquely defined by

$$\hat{\mathscr{C}_k} : |\rho) \in St(A) \mapsto \mathscr{C}_k\ |\rho)_A \in St(B)$$

And in a similar manner for every system C the event $\mathscr{C}_i \otimes \mathscr{I}_C$ induces a linear map from $St_{\mathbb{R}}(AC)$ to $St_{\mathbb{R}}(BC)$. From the statistical perspective if two events $\mathscr{C}_i$ and $\mathscr{C}_i'$ induce the same maps for every possible system C, then they are indistinguishable. In a similar manner to states and effects one then collapses indistinguishable events to single transformations.

Definition 10 (Transformations)**.** *Equivalence classes of indistinguishable events from A to B are called transformations from A to B.*

Remark 2. *One should be aware that two transformations $\mathscr{C}, \mathscr{D} \in \mathcal{E}(A, B)$ can be different even if $\mathscr{C}\ |\rho)_A = \mathscr{D}\ |\rho)_A$ for every $\rho \in St(A)$: to make $\mathscr{C}$ different from $\mathscr{D}$ all that is required is that there exists an ancillary system C and a joint state $|\rho)_{AC}$ such that $(\mathscr{C} \otimes \mathscr{I}_C)|\ \rho)_{AC} \neq (\mathscr{D} \otimes \mathscr{I}_C)|\ \rho)_{AC}$. This relates to the notion of local discriminability[4].*

Remark 3. *[27] In traditional quantum theory systems are associated with Hilbert spaces. The deterministic states of a system A are density matrices on the corresponding Hilbert space: a deterministic state ρ is a matrix such that $\rho \geq 0$ and $Tr[\rho] = 1$.*

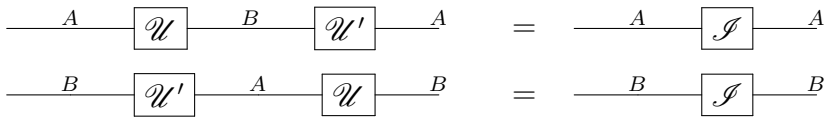
A non-deterministic preparation test $\{\rho_i\}_{i\in X}$ - a quantum information source - is a collection of positive operators where one finds that $\sum_{i\in X} Tr[\rho_i] = 1$. Accordingly the set $St(A)$ of all states of system A in the collection is an un-normalised density matrices ρ with $Tr[\rho_i] \leq 1$. An effect here is a positive operator P with $P \leq I_A$ where I_A is the identity operator.

The probability associated with the occurrence of the state ρ followed the effect P is given by the Born rule: $(P|\rho)_A = Tr[P\rho]$. A transformation in quantum mechanics is referred to as a quantum operation*. In technical terms a quantum operation from A to B is a linear, completely positive, trace non-increasing map sending density matrices of system A to un-normalised density matrices of the system B. One calls a test $\{\mathscr{C}_i\}_{i\in X}$ from A to B as a quantum instrument which is a collection of quantum operations with the property that $\sum_{i\in X} \mathscr{C}_i$ is trace-preserving, namely $\sum_{i\in X} Tr[\mathscr{C}_i(\rho)] = Tr[\rho]$ for every state ρ.*

An important case is when the transformation is deterministic.

Definition 11 (Channel)**.** *A deterministic transformation $\mathscr{C} \in \mathcal{E}(A, B)$ is called a channel.*

Definition 12 (Reversible Channel)**.** *A channel $\mathscr{U} \in \mathcal{E}(A, B)$ is called reversible if there exists a second channel $\mathscr{U}' \in \mathcal{E}(B, A)$ such that*



(2.14)

that is they compose in either possible sequence to the identity.

The reversible channels from A to itself actually forms a group denoted G_A. With a group defined ones mind turns to the question of those states that are invariant under its action.

Definition 13 (Invariant states)**.** *A state $\rho \in St(A)$ is invariant under the action of the group G_A if*

$\forall \mathscr{U} \in G_A$ (2.15)

2.2 Refinement and Coarse Graining [4]

It is a common that one considers several outcomes of some experiment to indicate the same thing. Conversely one may wish to refine a detected outcome - perhaps in order to attain a more specific understanding of some process. These ideas are contained within the notions of test *coarse-graining* and *refinement*.

Coarse-graining allows one to define tests that have outputs that are groupings of the outputs of some other test.

Definition 14 (coarse-graining)**.** *A test* $\{\mathscr{C}_i\}_{i\in X}$ *is a coarse-graining of the test* $\{\mathscr{D}_j\}_{j\in Y}$ *if there is a partition of* Y *into disjoint sets* Y_i *such that* $\mathscr{C}_{i\in X} = \sum_{j\in Y_i} \mathscr{D}$ *for every* $i \in X$*.*

One can define refinements in a simple manner by considering them as the tests that are coarse grained.

Definition 15 (Test refinement)**.** *If* $\{\mathscr{C}_i\}_{i\in X}$ *is a coarse-graining of* $\{\mathscr{D}_j\}_{j\in Y}$ *one says that,* $\{\mathscr{D}_j\}_{j\in Y}$ *is a refinement of* $\{\mathscr{C}_i\}_{i\in X}$*.*

This idea becomes clearer if one considers what it is to refine the individual events that compose a test. In essence a refinement of an event is the assignment of an entire test to that event, in this way it becomes a more nuanced description of the original event.

Definition 16 (Refinement of an event)**.** *A refinement of the event* $\mathscr{C}$ *is given by a test* $\{\mathscr{D}_j\}_{j\in Y}$ *and a subset* $Y_0 \subseteq Y$ *such that* $\mathscr{C} = \{\mathscr{D}_j\}_{j\in Y_0}$*.*

Once can then define the *act* of refining as follows:

Definition 17. *An event* $\mathscr{D} \in \mathcal{E}(A, B)$ *refines* $\mathscr{C} \in \mathcal{E}(A, B)$*, and write* $\mathscr{D} \prec \mathscr{C}$*, if there exists a refinement of* $\mathscr{C}$ *such that* $\mathscr{D} \in \{\mathscr{D}\}_{j\in Y_0}$*.*

One is then in a position to describe the set of all refinements of an event:

Definition 18 (refinement set)**.** *The refinement set* $D_{\mathscr{C}}$ *of an event* $\mathscr{C} \in \mathcal{E}(A, B)$ *is the set of all events* $\mathscr{D}$ *that refine* $\mathscr{C}$*, which corresponds to* $D_{\mathscr{C}} := \{\mathscr{D} \in \mathcal{E}(A, B) | \mathscr{D} \prec \mathscr{C}\}$

The trivial case is of particular note as it gives one the notion of the primal or atomic event, which cannot be subdivided, and on which all other coarse-grainings stand.

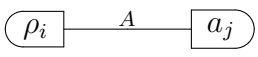
Definition 19 (Atomic and refinable events)**.** *An event $\mathscr{C}$ is called atomic if it admits only trivial refinements which is to say that $\mathscr{D} \prec \mathscr{C}$ implies that $\mathscr{D} = \lambda\mathscr{C}$ for some $\lambda \in [0, 1]$. An event is refinable if it is* not *atomic.*

The idea of these refinements for preparation-events gives us a notion of pure and mixed states

Definition 20 (Pure and Mixed States)**.** *An atomic preparation event $\rho \in St(A)$ is called a pure state. A refinable preparation state is called a mixed state.*

2.3 Metrised Operational Probability Theories [4]

Through experiment one can attempt to discriminate between different devices. Imagine one has a set-up that prepares one of two deterministic states $\rho_0, \rho_1 \in St(A)$, and that one wants to find out which one. To discriminate between two states one can perform a binary observation-test a_0, a_1. The probabilities of the outcomes are given by

$$p(a_j|\rho_i) =: \qquad \text{  } \qquad i,j = 0,1. \tag{2.16}$$

Assuming prior probabilities of π_0 and π_1 for returning the states ρ_0, ρ_1, respectively, one attempts to maximise the probability discriminating correctly, defined as $p_{succ} := \pi_0 p\,(0|0) + \pi_1 p\,(1|1)$. Substituting the expression for the probabilities equation 2.16 recalling that the probabilities sum to unity, one obtains

$$\begin{aligned} p_{succ} =:& \pi_0 + (a_1|\pi_1\rho_1 - \pi_0\rho_0)_A \\ =& \pi_1 + (a_0|\pi_0\rho_0 - \pi_1\rho_1)_A \end{aligned} \tag{2.17}$$

if one then takes the supremum over all such binary tests one finds that

$$\begin{aligned} p^{sup}_{succ} =:& \pi_0 + \sup_{a_1 \in Eff(A)} (a_1|\pi_1\rho_1 - \pi_0\rho_0)_A \\ =& \pi_1 + \sup_{a_0 \in Eff(A)} (a_0|\pi_0\rho_0 - \pi_1\rho_1)_A \end{aligned} \tag{2.18}$$

Summing these expressions gives us

$$p^{sup}_{succ} = \frac{1 + ||\pi_1\rho_1 - \pi_0\rho_0||_A}{2} \tag{2.19}$$

where $||-||_A$ is the operational norm defined by

$$||\delta||_A = \sup_{a_1 \in Eff(A)} (a_1|\delta) - \inf_{a_1 \in Eff(A)} (a_0|\delta)_A \qquad \delta \in Eff_{\mathbb{R}}(A). \tag{2.20}$$

The norm $||\pi_1\rho_1 - \pi_0\rho_0||_A$ ranges from 0 representing when the two states and probabilities are equal, and 1, when the two states are perfectly discriminable. For real numbers $x \in St_{\mathbb{R}}(I) \equiv \mathbb{R}$ one has $||x||_I = |x|$.

Remark 4. *In quantum theory the operational norm is the trace norm.*

For some generic $\rho \in Eff(A)$ the operational norm given in equation 2.20 becomes

$$||\rho||_A = \sup_{e \in Eff(A)} (e|\rho_i) \tag{2.21}$$

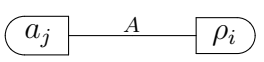
where the supremum is over deterministic effects.

One can now define states of unit size in terms of their operational norm.

Definition 21 (Normalised States)**.** *A state $\rho \in St(A)$ is normalised if $||\rho||_A = 1$. We will denote such states by $St_1(A)$*

If ρ is deterministic then it corresponds to a single outcome preparation test and as e is a single outcome observation test the probability of the former then the latter occurring is unity; this implies that ρ is normalised.

Definition 22 (Distinguishable States)**.** *The states $\{\rho_i\}_{i \in X}$ are perfectly distinguishable if there exists a test $\{a_i\}_{i \in X}$ such that*



$$= ||\rho_i||_A \delta_{ij}. \tag{2.22}$$

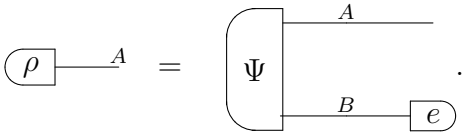
The test $\{a_i\}_{i \in X}$ is called a discriminating test.

Remark 5. *[27] In quantum theory a set of distinguishable states $\{\rho_i\}_{i=1}^n$ is a set of density matrices with orthogonal support. An example of a discriminating test for this set is the collection of orthogonal projectors $\{P_i\}_{i=1}^n$, where P_i is the projector on the support of ρ_i for all $i < n$, while $P_n = I - \sum_{i=1}^{n-1} P_i$. The maximum number of distinguishable states available for a certain system is the dimension d of the corresponding Hilbert space. The distinguishable states are rank-one projectors on an orthonormal basis, and the corresponding discrimination test is the projective measurement on the same basis.*

2.4 Purification and Marginal States [8]

In almost any practical situation one is actually considering only a part of a larger composite system. It is therefore essential, for any framework capable of representing real physical system, to understand the act of restricting larger states to smaller ones.

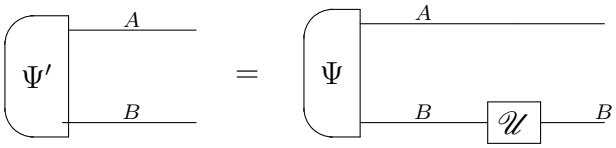
Definition 23 (Purification)**.** *A pure state $\Psi \in St_1(A)$ is a purification of some arbitrary state $\rho \in St_1(A)$ if $|\rho) = (e|_B|\Phi)_{AB}$, where e is some Deterministic effect. In diagram form*



The state ρ is called the marginal *of Ψ.*

Definition 24 (Purification System)**.** *If system $A \otimes B$ contains a purification of $\rho \in St_1(A)$, we call system B a purifying system for ρ.*

Definition 25 (Essentially Unique Purification)**.** *We say that a pure state $\Psi \in PurSt(A \otimes B)$ is an essentially unique purification of its marginal state ρ_A if the every other pure state $\Psi' \in PurSt(A \otimes B)$ satisfying condition equation 3.1 is of the form*



for $\mathscr{U}$, some reversible transformation.

What has been accomplished here is an abstract framework for measurement has been defined, where one can compose measurements in parallel and series, should the typings match. Moreover, by allowing for the assignment of probabilities to outcomes the framework provides an abstract framework for *prediction*. Finally through purification the framework is capable of reflecting the fact that from the human perspective only partial measurement of a much larger system is ever possible. In the following section the material here will be given rigorous mathematical grounding.

3

Resource Theories

The topic of OPTs provided a framework on which one can reason about measurement in its most abstract setting. In this section there are two aims, to give a formal mathematical grounding to the OPT framework, and to develop a theory of resources which can be studied in OPTs that models the idea of vale and exchange. This is something which is at the heart of thermodynamics and quantum theory where one can view certain states as more valuable than others. In thermodynamics for example, states that are hotter than average are more valuable than a cooler ones: energy can be extracted. In quantum mechanics entangled states are valuable to those only able to perform local operations[10] - one cannot create physically separated entangled states but one can use them. The ideas behind this will be developed as the chapter progresses.

3.1 OPTs as categories

The mathematical formalism behind the intuition of an OPT is the symmetric monoidal category. It, by its construction, provides the scalable compositional structure required. In the language of categories, in an OPT, a state is an object of a category and the tests are morphisms, and one has certain extra isomorphisms between compositions of these things not usually in a general category. To formally define this we must however begin with the broader object of the *category*:

Definition 26. *A category D consists of*

1. *A class $ob(D)$ of objects, also denoted $|D|$.*
2. *A class $hom(D)$ of morphisms or arrows between objects. Each morphism $f \in hom(D)$ has a source object a and target object b where a and b are in $ob(C)$. One writes $f : a \to b$ to represent the mapping of a to b and $D(a, b)$ to denote the class of all morphisms that perform this mapping.*

   For every three objects a, b and c, a binary operation $hom(a, b) \otimes hom(b, c) \to hom(a, c)$ called composition of morphisms; the composition of $f : a \to b$ and $g : b \to c$ is written as $g \circ f$.

   One also requires that the following axioms hold. They morphisms are associative:

   If $f : a \to b, g : b \to c$ and $h : c \to d$ then $h \circ (g \circ f) = (h \circ g) \circ f$.

   There exists an identity:

   For every object x, there exists a morphism $id_x : x \to x$ called the identity morphism for x, such that for every morphism $f : a \to x$ and every morphism $g : x \to b$, we have $id_x \circ f = f$ and $g \circ id_x = g$.

Category theory is an expansive field in its own right, for the purposes of this thesis the interest rests on the composition of morphisms which provides the sequential composition. To fully capture OPTs, more is required, namely something that can represent parallel composition. The needed structure is a symmetric monoidal category. To formally define this further definitions are required. In particular the idea of parallel composition implies that one takes two categories and somehow maps to a single category - one requires mappings of categories - such things are known as functors.

Definition 27. *Let C and D be categories. A functor F from C to D is a mapping that*

1. *Associates to each object $X \in C$ an object $F(X) \in D$*
2. *Associates to each morphism $f : X \to Y \in hom(C)$ $X \in C$ a morphism $F(f) : F(X) \to F(Y)$ in D such that:*

   $F(id_X) = id_{F(X)}$

   $F(g \circ f) = F(g) \circ F(f)$ *for all morphisms $f : X \to Y$ and $g : Y \to Z$ in C*

The attentive reader will note that in the previous paragraph the ad-hoc requirement was for a mapping of two categories to a single one. Functors however, provide a mapping between categories. To understand why functors are useful in defining parallel composition one must look at the idea of a product category, a category which is in a sense composed of other categories.

Definition 28. *A product category $C \times D$ is category containing:*

1. *As objects:*

   pairs of objects (A, B) wjere A is an object of C and B of D.

2. *As arrows from (A_1, B_1) to (A_2, B_2): pairs of arrows (f, g) where $f : A_1 \to A_2$ is an arrow of C and $g : B_1 \to B_2$ is an arrow of D.*

3. *Where composition, component wise composition from the contributing categories $(f_2, g_2) \circ (f_1, g_1) = (f_2 \circ f_1, g_2 \circ g_1)$*

4. *Where the identities are pairs of identities from the contributing categories $id_{(A,B)} = (id_A, id_B)$.*

One now has access to a way to compose two categories in way that one gets a single category as a result. Taking this notion and then applying a functor allows us to get the structure we desire. A functor of this sort is defined as follows:

Definition 29 (Bi-functor)**.** *A bi-functor is a functor whose domain is a product category.*

The structure in place now allows one to create a strict monoidal category which has almost all the properties required.

Definition 30. *A strict monoidal category (SMC) denoted $(D, \otimes, I)$ is a category D equipped with a bi-functor $\otimes : D \times D \to D$ called the tensor product and an identity I such that*

1. $(A \otimes B) \otimes C = A \otimes (B \otimes C)$
2. $A \otimes I = A = I \otimes A$.
3. $(A \otimes I) \otimes B = A \otimes B = A \otimes (I \otimes B)$.
4. $(A \otimes B) \otimes (C \otimes E) = ((A \otimes B) \otimes C) \otimes E = (A \otimes (B \otimes C)) \otimes E = A \otimes (B \otimes (C \otimes E)) = A \otimes ((B \otimes C) \otimes E)$

Where $|D|$ denotes the set of objects in D, $A, B, C, E \in |D|$ and the hom-set $D(A, B)$ represents the set of morphisms in D with A as a domain and B as a codomain.

One can check that this structure allows for many of the properties of an OPT, however it is lacks a certain symmetry - in OPTs one can compose parallel tests in any order one desires as this detail is functionally irrelevant - this is lacking in the above $A \otimes B \neq B \otimes A$. This fact must be imposed which results in the following:

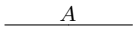
Definition 31 (Symmetric Monoidal Category)**.** *A strict symmetric monoidal category(SMC hereafter) is a strict monoidal category where $A \otimes B = B \otimes A$*

Such a category admits a graphical calculus [11]. This calculus is identical to that of the OPT diagrams above, in fact an alternative definition for OPTS is the following:

Definition 32. *[4] An OPT is an SMC where the systems are the objects and events the are morphisms.*

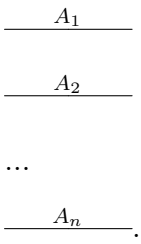
The distinction between the OPT and the SMC is not mathematical but a matter of perspective[11]. For the sake of clarifying this link between SMCs and OPTs the diagrammatic calculus is expressed in categoric terms - with more emphasis on the calculus' nature than with the OPT structure as before - the reader is invited to compare this with the last section.

Firstly objects are represented by wires



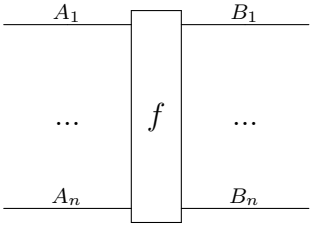
(3.1)

with the n-composite $A_1 \otimes A_2 \otimes ... \otimes A_n$ given by



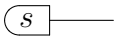
(3.2)

A general morphism $f : A_1 \otimes ... \otimes A_n \to B_1 \otimes ... \otimes B_m$ can depicted as follows



(3.3)

where the trivial object is represented by the absence of a wire. The absence of a wire for the trivial object leads us to the notion of a morphism $s : I \to A$ with no inputs, which can be depicted symbolically as follows



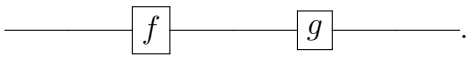
(3.4)

where we have omitted the wire labelling as there is no ambiguity, which will be done from here on where appropriate. Note this is a preparation state from the OPT perspective. The 'symmetry' in a symmetric monoidal category arises from the fact wires may cross:

.

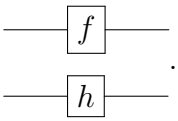
This symmetry is exactly the structure in OPTs that allows one to disregard ordering of systems composed in parallel; the existence of this symmetry enforces that swapping them around is equivalent to the original order.

On the topic of composition one has that sequential composition of $f : A \to B$ and $g : B \to C$ is represented by attaching the relevant matching wires together:



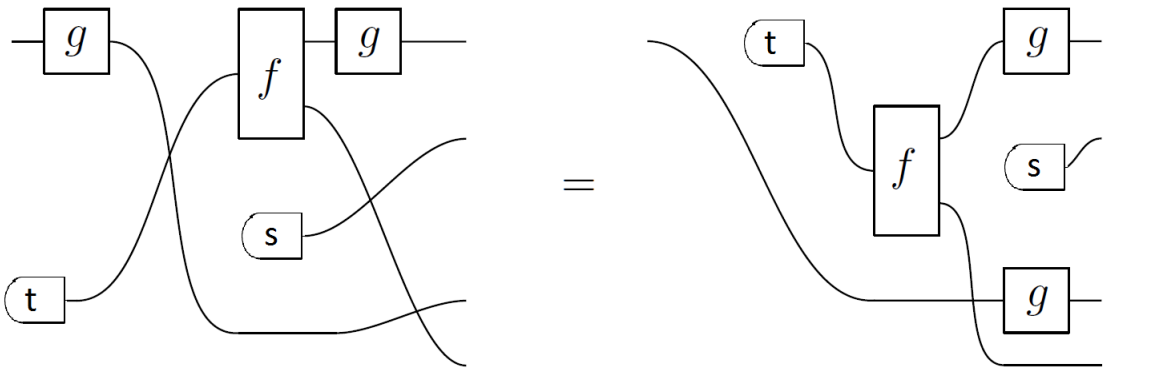
. (3.5)

Parallel composition of f with $h : C \to D$ is accomplished by adjacent box placement



. (3.6)

Again the relationship with between this and the circuit structure of OPTs is clear. A simple summary of this diagrammatic framework is that it allows equational reasoning to be performed by deforming the representative diagram without altering its topology[11].



.

One will notice the use of strict in some of the above definitions which at the time went unexplained. It is implicit in the diagrammatic structures presented that the grouping of objects is irrelevant, that is to say, any bracketed groundings $A(BC)$ or $(AB)C$ are equivalent. Technically, in a general SMC they are only required to be isomorphic. Strictness strengthens this to equality[24].

One may worry that the fact the diagrammatic calculus does not obviously respect the unit and associator isomorphisms, which are required to describe SMCs in general, limits the scope

of diagrams. This is not the case however, as MC Lane's strictification theorem[24] shows that any non strict monoidal category is equivalent to a strict one, thus all SMCs can be represented.

Notice that the enforced symmetry isomorphisms of an SMC are not equivalent to identities under this theorem. That said, they already exist within the diagrammatic calculus as the wire crossings which represent the fact $A \otimes B$ is equivalent to $B \otimes A$ - or equivalently in any SMC there exists a map between such objects. In summary one can be certain that the diagrammatic structure can describe all the things one would hope.

3.2 Resource Theories [11]

At this point a framework for discussing physical systems and the measurements upon them has been outlined and made rigorous through categorical considerations. The aim of this work however is to discuss quantum thermodynamics through an abstract understanding of the properties that define it. The framework used to achieve this is resource theories. Thermodynamics is at its core a set of rules about the allowed transition between states. In extremely broad terms one can move from hot to cold - and one can extract work from this. The converse direction requires work. In this way the hotter state is more valuable - it is a resource. It is at first surprising how far this observation can take one.

The intuitive notion of a resource theory is that one has different kinds of resources, denoted by $A, B, ...$, and one can transform between them via conversions denoted $f : A \to B$, a transformation or process from A to B, or $g : C \to D, ...$, a transformation from C to D, etc. The labelling is necessary as the transition itself is not enough to characterise it; there may be more than one transition between the same resources. These transformations can be composed sequentially if the resources match: if $f : A \to B$ and $h : B \to C$ then the composition $h \circ f : A \to C$. Resources A and B can also be combined together forming a composite resource $A \otimes B$[1], which lifts to composite maps: for maps f and g, as defined above, one has a composite transformation $f \otimes g : A \otimes C \to B \otimes D$. In addition to all this one can presume there exists a 'void' resource I which when composed with any resource is equivalent to the resource on its own.

It takes little imagination to see this description coincides with that of an SMC. In fact the separation between such a category and a resource theory is merely a matter of perspective.

[1]The notation $\otimes$ is not necessarily the traditional tensor product just a kind of parallel composition -tellingly just as in monoidal categories - that said, in suitable resource theories it can be the tensor product.

The resource perspective has more to it than the compositional requirements seen in the OPTs for example. It contains an implicit notion value. The unit object I as a void resource is in essence $'valueless'$ as it can be added endlessly or removed from a system without actually altering anything. Extending this, any resource that can be generated from the void resource is a free resource that can be attained without any requirement, or 'cost', so any A such that the hom-set $D(I, A)$ is non-empty, also has no cost. The set of all such objects are the *free resources*; by contrast its complement is the set of *costly resources*

Definition 33. *The collection of free resources in D is* $|D_{free}| := \{\, A \in |D| \;\; |D(I, A) \neq \varnothing \}$.

It is important to note that this collection is defined to be closed - if one can approximate something to arbitrary accuracy with a free resource - it too is presumed to be free. Broadening this idea one can ask how does one represent a theory where one chooses certain transitions to be $'free'$. One is in essence asking for a category D_{free} which contains all the same objects, or *systems*, but where one is restricted to certain morphisms, or *processes* considered *free*. This is the subject of the next section.

3.3 Free operations and partitioned theories

The 'free operations' of D_{free} are defined as those which one can execute at *'no cost'* - a phrase which finds meaning in the context of a particular application. As an example consider LoCC quantum information, here one quantum system is physically separated, one has tensored quantum states and only local operations and classical communication are for free - an excellent model for information communication in quantum systems[10]. Observe that in this situation one cannot create non-local entangled states that is to say entangled between the separated components.

There is a natural line of reasoning from having a set of free processes in that if f and g are free processes then $f \otimes g$ should be also; In addition the sequential $f \circ g$, where such composition is defined, should also be free. The identity map will be a free process, as it alters nothing. In this way we see that D_{free} is a sub-SMC, that is to say it is a symmetric monoidal category itself[2] that includes all the objects C but only this limited set of free transformations. This motives the following definition:

Definition 34. *A partitioned theory consists of a some system modelled by an SMC denoted* $(D, \circ, \otimes, I)$ *and a sub-theory of free morphisms, described by an SMC called* D_{free} *that contains all objects of* D*. The partitioned theory is written as* (D, D_{free})*.*

Again the free processes are closed in the limit as what can be attained to arbitrary accuracy from free processes is also free.

Remark 6. *The free morphisms themselves induce a preorder structure on the objects of the category, the states, where* ρ *is more* valuable *than* σ *if there is a free operation going from* ρ *to* σ*.*

The free morphisms themselves induce a preorder structure on the objects of the category, the states, where ρ is more *valuable* than σ if there is a free operation going from ρ to σ.

In practice D_{free} will not be given in its totality but in term of a generating set of operations one defines as free; it will be the smallest SMC closed under this set through parallel and sequential composition which contains all objects of the super-category.

[2]Its symmetry is inherited directly from the super-category.

3.4 The Resource Theory of States [11]

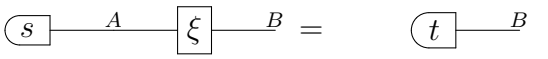
There are many examples of resource theories but the one of practical interest here is the resource theory of states. Hereafter the resource theories of states will be referred to simply as resource theories.

To begin one considers a state as the transition whose input is the trivial system. The set of these states is given by $\cup_{A\in|D|}\ D(I,A)$. To obtain a resource theory one considers when one can transition between these states by a free process, that is, the structure of $\cup_{A\in|D|}D(I,A)$ under $\cup_{A,B\in|D|}D_{free}(A,B)$.

Formally, a resource theory of states, in terms of a partitioned resource theory (D, D_{free}), will be denoted $S(D, D_{free})$. The objects of $S(D, D_{free})$ are taken to be states of D,

$$|S(D,D_{free})| := \bigcup_{A\in|D|} D(I,A). \tag{3.7}$$

A state $s : I \to A$ is said to be converted into another state $t : I \to B$ by a free transformation $\xi : A \to B$ if one has



. (3.8)

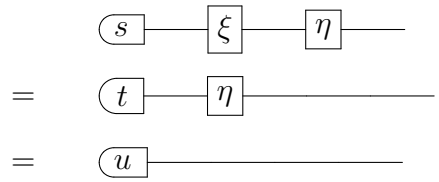
One then defines the hom-set $S(D, D_{free})(s, t)$ for $s, t \in |S(D, D_{free})|$ to be the set of such free transformations where $s \to t$.

One can equate two transformations if they have the same effect on all states, which includes when they act only on one part of a composite system. In the resource theory of states the equivalence classes of free transformations are the morphisms.

The hom-set $S(D, D_{free})(s, t)$ for $s, t \in |S(D, D_{free})|$ is the set of equivalence classes of free processes such that $s \to t$, which is to say

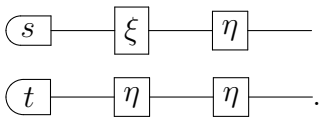
$$S(D, D_{free})(s, t) := \{\xi \in D_{free}(A, B) : \xi \circ s = t\}/\sim . \tag{3.9}$$

Regarding sequential composition if $\xi \in D_{free}(A, B)$ is a free process turning s into t. and $\eta \in D_{free}(B, C)$ is another free process turning t into a third state $u : I \to D$, then $\eta \circ \xi \in D_{free}(A, C)$ is a free process turning s into u,



(3.10)

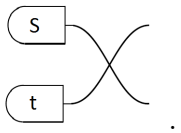
as this respects the structure of equivalence classes one can see that $S(D, D_{free})$ is a category. To be an SMC one requires parallel composition of objects and morphisms. On objects in $S(D, D_{free})$, which are states in D, parallel composition is inherited from D. Morphisms in $S(D, D_{free})$, which are equivalence classes of transformations between states in D using free processes are composed in the natural manner



(3.11)

where one should recall D_{free} is closed under parallel composition. Again this is compatible with the equivalence structure.

One should note that as D is a strict monoidal category then $S(D, D_{free})$ is a strict monoidal category as well. The symmetry isomorphism $(s \otimes t) \to (t \otimes s)$ on objects in $S(D, D_{free})$ is given by symmetry in D.



.

This symmetry is guaranteed to be a free process since D_{free} was assumed to contain all the objects of the super-category of D and thereof it inherits the symmetries from D. The identity object I for the SMC $S(D, D_{free})$ is the tensor unit of D, id_I, the result of all this is that one has proven the following:

Theorem 3.4.1. *[11] For any partitioned theory (D, D_{free}), the procedure outlined above allows one to define a symmetric monoidal category $S(D, D_{free})$ that can be interpreted as a resource theory .*

To summarise what has been done here, firstly OPTs were mathematically formalised as examples of SMCs. Following this the categorical framework was elaborated on and this was expanded upon to present resource theories which are formed of a base SMC structure but where one is restricted to some agreed upon free processes. The common categorical theme will in later sections allow for the discussion of resource theories, namely the *resource theory of purity*, described within the framework of OPT, which we have shown, is itself a SMC seen through a certain lens. This fusion will prove useful in bringing together the resource perspective that allows for a description of isolated thermodynamic systems and the operational perspective useful for describing experiments in the abstract.

4
Rates

Once one has established the ability to transition between certain resources, or in particular, certain amounts of certain resources, one naturally starts to consider the question of rates of exchange, how much of one thing can be gotten by exchanging another. Consider transitions of the sort sending p lots of some resource ρ, to q lots of some resource σ, that is to say a transition $\rho^{\otimes p} \to \sigma^{\otimes q}$, this underlying concept behind all resource transitions. One can imagine two basic situations: trying to maximise the amount of σ one can *distil* per ρ, and trying to minimise the amount of ρ one needs to spend per σ, put alternatively, the *cost* per σ.

Before going into detail it is helpful consider the general philosophy of approaching rates of exchange this way. The work here as aimed at addressing the concerns of rates at the correct level of $'$granularity$'$ - what is meant by this is that the problem is analysed at the *clearest level of abstraction.* In Quantum literature resources are considered in variety of situations[18], and the problems tackled are often addressed in the context of a particular instantiation of a resource theory, as such computational and technical distractions abound. This attempt to reach an ideal level of conceptual clarity has been an implicit theme throughout the previous sections, from the categorical structure, and the abstraction of experimental procedures, to the arguments on rates presented here - by tackling a problem at the appropriate level of $'$granularity$'$ one sweeps aside complicating concerns while not abstracting to the point where the key observations are lost - in this way reasoning becomes easier and results are clarified through being presented in their correct context.

The thermodynamic perspective is also intimately related to rates. As one approaches the thermodynamic limit clearly one is then concerned with transitions in this limit. The issue is that in the limit it is no longer helpful to talk of individual transitions, the only directly meaningful relationships are the rates themselves as they are intensive quantities and so can remain sensibly finite while the potentially unbounded numbers of resources involved in the transitions wander off to infinity. As was mentioned in the introduction these rates also relate in a more phenomenological manner to the maximum extractable work from a Szillard-like machine as providing the amount of pure resources one can extract from other resources - which in turn can be converted to energy.

The first step to formalising an idea of rates is essentially a labelling exercise - one is required to take some appropriate mathematical structure capable of encapsulating an exchange and relate it to rates. Sequences of transition operations, called protocols, are used. As sequences provide a well trodden path to calculus and considering the limit there will be a natural route to asymptotic, or thermodynamic, rates. In line with earlier reasoning there are two types of protocol, which lead to two types of optimal rate.

Definition 35 (Distillation Protocol)**.** *Given two resources ρ and σ, a distillation protocol $\bar{\mathscr{F}}$ from the former to the latter is: a sequence of natural numbers $\{m_n\}$ which, for each $n \in \mathbb{N}$ the number of* initial resources*, gives the number of* output resources*, the elements of said sequence m_n; and a sequence $\mathscr{F} := \{\mathscr{F}_n\}$ of operations where $\mathscr{F}_n : \rho^{\otimes n} \to \sigma^{\otimes m_n}$ where n is also in $\mathbb{N}$, which provides these transitions.*

For each operation sequence $\mathscr{F}$ we define its n^{th} distillation rate $R_D(n, \mathscr{F})$ as the the ratio $\frac{m_n}{n}$, where n and m_n are those integers that correspond to the operation considered.

Definition 36 (Cost Protocol)**.** *Given two resources ρ and σ, a cost protocol $\bar{\mathscr{F}}$ from the former to the latter is: a sequence of natural numbers $\{m_n\}$ which, for each $n \in \mathbb{N}$ the number of* output resources*, gives the number of* input resources*, the elements of said sequence m_n; and a sequence $\mathscr{F} := \{\mathscr{F}_n\}$ of operations where $\mathscr{F}_n : \rho^{\otimes m_n} \to \sigma^{\otimes n}$ where n is also in $\mathbb{N}$, which provides these transitions.*

For each operation sequence $\mathscr{F}$ we define its n^{th} cost rate $R_C(n, \mathscr{F})$ as the the ratio $\frac{m_n}{n}$, where n and m_n are those integers that correspond to the operation considered.

4. Rates

With these protocols defined one is in a position to talk about optimal rates. For brevity the optimal forms will be discussed in unison, later it will be shown that they are actually inverses, and so separating them is a needless complication.

Definition 37. *The sequence of real numbers $R_{x-opt}(n)$, where x is either $'C'$ for cost or $'D'$ for distillation, is the optimum rate attainable from an operation acting on any of the possible x type protocols acting on n resources. For cost and distillation, optimum has a different meaning:*

$$R_{C-opt}(n) = \inf_{\mathscr{F}} R_C(n, \mathscr{F})$$
$$R_{D-opt}(n) = \sup_{\mathscr{F}} R_D(n, \mathscr{F})$$

The infimum minimises, while the supremum maximises, this is in line with the desire to minimises cost and maximise distillation.

Definition 38. *The* optimal rate $R_x(\rho \to \sigma)$*, where x is either $'C'$ for cost or $'D'$ for distillation and indicates the relevant protocol type, is the optimum rate possibly obtainable when one considers n in the limit for any transition*

$$R_C(\rho \to \sigma) = \liminf_{n\to\infty} R_{C-opt}(n) \tag{4.1}$$
$$R_D(\rho \to \sigma) = \limsup_{n\to\infty} R_{D-opt}(n). \tag{4.2}$$

It may not be obvious why the superior and inferior limit are invoked. The reason is the usual limit may not exist, there is no fundamental reason why the amount of a product obtained from an input would not oscillate continually. The superior and inferior limits allow one to take the limit set of the optimal rates, viewed as a sequence over n, and then choose the largest or smallest, depending on whether one seeks cost or distillation optimisation.

It will prove practical to note that the right hand side of equations 4.1 and 4.2 in the above definition can be rewritten as follows:

Lemma 4.0.1. *The limit superior(inferior) of the sequence formed of the optimum rate obtained for each n is equivalent to the supremum(infimum) over the protocols of the set of superior(inferior) limits of the rates for each protocol $\bar{\mathscr{F}}$, written explicitly*
For distillation one has

$$\limsup_{n\to\infty} R_{D-opt}(n) = \sup_{\mathscr{F}} \left(\limsup_{n\to\infty} R(n, \mathscr{F}) \right) \tag{4.3}$$

and for cost

$$\liminf_{n\to\infty} R_{C-opt}(n) = \inf_{\mathscr{F}} \left(\liminf_{n\to\infty} R(n, \mathscr{F}) \right) \tag{4.4}$$

Proof. For distillation

$$\limsup_{n\to\infty} R_{D-opt}(n) = \limsup_{n\to\infty} (\sup_{\mathscr{F}} R(n, \mathscr{F}))$$

Using the commutativity of the supremum and the limit superior (see appendix B.0.3)

$$\limsup_{n\to\infty} (\sup_{\mathscr{F}} R(n, \mathscr{F})) = \sup_{\mathscr{F}} (\limsup_{n\to\infty} R(n, \mathscr{F}))$$

For cost

$$\liminf_{n\to\infty} R_{C-opt}(n) = \liminf_{n\to\infty} (\inf_{\mathscr{F}} R(n, \mathscr{F}))$$

Using the commutativity of the infimum and the limit inferior (see appendix B.0.4)

$$\liminf_{n\to\infty} (\inf_{\mathscr{F}} R(n, \mathscr{F})) = \inf_{\mathscr{F}} (\liminf_{n\to\infty} R(n, \mathscr{F}))$$

□

In essence what has been done here is to move from considering the supremum(infimum) of the limit set of a sequence composed of the optimal rates attainable for each n over all the protocols, to a situation where one is collecting the superior(inferior) limit of the rate of every protocol and then taking the supremum(infimum) of these.

4. Rates

The ability to compose resources and the maps between them is a relatively simple principle which immediately implies a less obvious relationship between the supremum(infimum) of the optimal rate sequence and its superior(inferior) limits. If one has the optimal rate for some transition in the finite limit, from some $\rho^{\otimes a}$ to $\sigma^{\otimes b}$ by a transition $\mathscr{F}$, then by forming a tensor product of such transitions one can achieve this rate for every P multiple of the original a - this implies one can form a protocol where Pb is obtained from Pa which implies that the one can form a protocol where the optimal rate discussed is obtained a countably infinite number of times (once for every p multiple that exists). This means it is an accumulation point and so is in the limit set of the sequence of optimum rates - if this rate is known to be the largest(smallest) obtainable then one immediately knows it must be the largest(smallest) limit point, the superior(inferior) limit, but at the same time it was also the optimal rate - the supremum(infimum) of the rates over all n and protocols $\bar{\mathscr{F}}_x$, where x is either C for cost or D for distillation. Before this can be proved the following lemma is required:

Lemma 4.0.2. *If $R_x(\rho \to \sigma)$ is attained for some finite $n^* \in \mathbb{N}$ for some protocol $\bar{\mathscr{F}}^*$ then this rate is obtainable for every integer multiple of the input and output resources of this transition, explicitly*

$$R_{x-opt}(n^*) = R_x(\rho \to \sigma) \implies R_{x-opt}(Pn^*) = R_x(\rho \to \sigma) \text{ where } P \in \mathbb{N}.$$

where x is either D or C, for distillation and cost respectively.

Proof. Consider the n^{*th} operation of the protocol $\bar{\mathscr{F}}^*$, which achieves the rate $R_x(\rho \to \sigma)$ where x is either D or C, for distillation and cost respectively:
For distillation one has some transition $\rho^{\otimes n^*}$ to $\sigma^{\otimes m_{n^*}}$ and by composition one can form the operation

$$\mathscr{F}_{Pn^*} = \mathscr{F}_{n^*}^{\otimes P} : (\rho^{\otimes n^*})^P \to (\sigma^{\otimes m_{n^*}})^P \tag{4.5}$$

the rate for this operation is $\frac{Pm_n^*}{Pn^*} = \frac{m_n^*}{n^*} = R_D(\rho \to \sigma)$. As by definition $R_D(\rho \to \sigma)$ is the optimum obtainable rate, and one can construct some protocol with the above as the Pn^{*th} operation then $R_{D-opt}(Pn^*) = R_x(\rho \to \sigma)$.

For cost one has some transition $\rho^{\otimes m_{n^*}}$ to some $\sigma^{\otimes n^*}$ again by composition one can form the operation

$$\mathscr{F}_{Pn^*} = \mathscr{F}_{n^*}^{\otimes P} : (\rho^{\otimes m_{n^*}})^P \to (\sigma^{\otimes n^*})^P \tag{4.6}$$

the rate for this operation is $\frac{Pm_n^*}{Pn^*} = \frac{m_n^*}{n^*} = R_C(\rho \to \sigma)$. As by definition $R_C(\rho \to \sigma)$ is the optimum obtainable rate, and one can construct some protocol with the above as the Pn^{*th} operation then $R_{C-opt}(Pn^*) = R_x(\rho \to \sigma)$. □

With this lemma to hand, one can now prove the following:

Lemma 4.0.3. *If $R_x(\rho \to \sigma)$, where x has its usual meaning, is obtained at some finite n then it is also attained in the limit. Explicitly if the transition rate $R_x(\rho \to \sigma)$ is obtained at some finite n^* for some operation in protocol $\bar{\mathscr{F}}^*$ then this rate is also obtainable in the limit via the supremum or infimum, for distillation and cost respectively. Written formally*

$$\sup_{n,\mathscr{F}} R_D(n, \mathscr{F}) = \sup_{\mathscr{F}} (\limsup_{n\to\infty} R_D(n, \mathscr{F})) \tag{4.7}$$

$$\inf_{n,\mathscr{F}} R_C(n, \mathscr{F}) = \inf_{\mathscr{F}} (\liminf_{n\to\infty} R_D(n, \mathscr{F})) \tag{4.8}$$

Proof. Consider the sequence $R_{x-opt}(n)$, from lemma 4.0.2 one knows that $R_{x-opt}(Pn^*) = R_x(\rho \to \sigma)$ where $P \in \mathbb{N}$. This implies that $R_x(\rho \to \sigma) \in \mathbb{R}$ is such that there are an infinite number of natural numbers where $R_{x-opt}(n) = R_x(\rho \to \sigma)$. Therefore $R_x(\rho \to \sigma)$ is an accumulation point of the sequence.

By definition $R_D(\rho \to \sigma) \geq R_{D-opt}(n)\ \forall n \in \mathbb{N}$ and is therefore the largest possible accumulation point; $R_C(\rho \to \sigma) \leq R_{C-opt}(n)\ \forall n \in \mathbb{N}$ and is therefore the smallest possible accumulation point. One then has that

$$\sup_{n,\mathscr{F}} R_D(n, \mathscr{F}) = \limsup_{n\to\infty} R_{D-opt}(n) \tag{4.9}$$

$$\inf_{n,\mathscr{F}} R_C(n, \mathscr{F}) = \liminf_{n\to\infty} R_{C-opt}(n) \tag{4.10}$$

therefore using lemma 4.0.1 one has that

$$\sup_{n,\mathscr{F}} R_D(n, \mathscr{F}) = \sup_{\mathscr{F}} \limsup_{n\to\infty} R_D(n, \mathscr{F}) \tag{4.11}$$

$$\inf_{n,\mathscr{F}} R_C(n, \mathscr{F}) = \inf_{\mathscr{F}} \liminf_{n\to\infty} R_C(n, \mathscr{F}). \tag{4.12}$$

□

Remark 7. *The above shows that if an optimal rate exists for a finite number of resources it is also optimal in the limit. This may seem as if it leaves open the possibility that the optimal rate may only be obtained in the limit and so, depending on whether the resource theory is in a sense complete, the optimal rate may just be a bound not an actually obtainable rate. In fact, as was discussed in section 3.2,* any transition that can be approximated to arbitrary accuracy by free states is also free. *The implication is that for free transitions the optimal rate is actually obtainable by a free operation.*

With this in mind one arrives at the following corollary.

Corollary 4.0.3.1. *In bounded sets the supremum of the rate is equal to its superior limit. Also the infimum of the rate is equal to its inferior limit. The set of transitions is closed in a resource theory; any allowed transition is already in the theory. We can then say that the superior limit rate is equal to the supremum rate and the inferior limit rate is equal to the infimum rate.*

It seems reasonable that the optimum distillation rate based of the supremum of the ratio of product over input might be the inverse of the optimum cost rate, the infimum of the ratio of input over product. If true this would allow one to show that one need only consider one of these rates as they are inversely related. Indeed this is the case:

Proposition 4.0.4. *The distillation and cost rates are inversely related:*

$$R_D(\rho \to \sigma) = \frac{1}{R_C(\rho \to \sigma)} \tag{4.13}$$

Where $\rho, \sigma \in |D|$ for a resource theory D.

Note this applies only where this is defined, which implies finite rates.

Proof. Consider the family of ordered pairs $< a_i, b_i >$ with $a_i, b_i \in \mathbb{N}$ representing *every* transition that exists between resources ρ and σ, belonging to a resource theory D, such that $\rho^{a_i} \to \sigma^{b_i}$ where the index i runs over the set of all transitions between these resources.

By definition the distillation rates and cost rates are the supremum and infimum of the ratios $R_{D,i} = \frac{b_i}{a_i}$ and $R_{C,i} = \frac{a_i}{b_i}$ respectively. Explicitly we can write the following

$$R_D(\rho \to \sigma) = \sup_i \frac{b_i}{a_i} = \frac{b^*}{a^*}$$
$$R_C(\rho \to \sigma) = \inf_i \frac{a_i}{b_i} = \frac{a'}{b'}$$

where $a', a^* \in a_i$ and $b', b^* \in b_i$. By definition of the supremum one knows that

$$R_D(\rho \to \sigma) = \frac{b^*}{a^*} \geq \frac{b'}{a'} = \frac{1}{R_C(\rho \to \sigma)}$$

Now suppose that

$$\frac{b^*}{a^*} > \frac{b'}{a'}$$

this implies that

$$\frac{a'}{b'} > \frac{a^*}{b^*}$$

but $\frac{a'}{b'} = \inf_i \frac{a_i}{b_i}$, so the above is a contradiction. Therefore one can strengthen the relationship to equality. □

From here on the simplification $R_x(\rho \to \sigma)$ to $R(\rho \to \sigma)$ may be used where it seems practical, that is to say the assumption will be that the distillation rate is being discussed. The cost rate is taken to be the inverse fraction as discussed above; all work referring to distillation rates immediately also relates to cost rates. Similarly $R_{D-opt}(-)$ will be written $R_{opt}(-)$ unless otherwise stated. Finally protocols can be assumed to be distillation protocols.

4. Rates

With rates between resources defined, it is practical to consider the bounds on compound rates. One presumes that if their is a transition to some resource via a middle resource then the optimum rate is at the very least equal to the optimum rates of these composite transitions. This is the case, as is shown below.

Lemma 4.0.5.

$$R(\rho \to \sigma)R(\sigma \to \theta) \leq R(\rho \to \theta) \tag{4.14}$$

Proof. Consider the transitions

$$\rho^{\otimes n} \to \sigma^{\otimes m_n} \tag{4.15}$$

$$\sigma^{\tilde{n}} \to \theta^{\tilde{m}_{\tilde{n}}} \tag{4.16}$$

One can use these to construct a third transition

$$\rho^{\otimes n\tilde{n}} \to \sigma^{\otimes m_n\tilde{n}} \to \theta^{\otimes m_n\tilde{m}_{\tilde{n}}} \tag{4.17}$$

written more succinctly as

$$\rho^{\otimes n\tilde{n}} \to \theta^{\otimes m_n\tilde{m}_{\tilde{n}}}$$

By definition one knows if $\rho^{\hat{n}} \to \theta^{\hat{m}_{\hat{n}}}$ then as the largest possible such ratio $R(\rho \to \theta) \geq \frac{\hat{m}_{\hat{n}}}{\hat{n}}$. This implies that

$$R(\rho \to \theta) \geq \frac{m_n\tilde{m}_{\tilde{n}}}{n\tilde{n}} \tag{4.18}$$

Suppose now that the transitions given by equations 4.15 and 4.16 are the transitions giving the largest distillation rate possible for any protocol, which is to say

$$R(\rho \to \sigma) = \frac{m_n}{n} \text{ and } R(\sigma \to \theta) = \frac{\tilde{m}_{\tilde{n}}}{\tilde{n}}$$

Then by using equation 4.18 one can write

$$R(\rho \to \sigma)R(\sigma \to \theta) \leq R(\rho \to \theta) \tag{4.19}$$

□

There is a similar inequality for cost rates demonstrated in the following lemma.

Lemma 4.0.6.

$$R_C(\rho \to \theta) \leq R_C(\rho \to \sigma) R_C(\sigma \to \theta) \tag{4.20}$$

Proof. Taking equation 4.14 for the distillation equivalent of the result and using lemma 4.0.4 regarding the relation between cost and distillation rates one deduces that

$$\frac{1}{R_C(\rho \to \sigma) R_C(\sigma \to \theta)} \leq \frac{1}{R_C(\rho \to \theta)} \tag{4.21}$$

which trivially rearranges to the desired result. □

4.1 Monotones

In a more practical setting it seems natural that a rate of exchange relates somehow to *value* - in a sense the fact one resource can be converted into many others roughly defines value. In a resource theory one resorts to monotones in order to discuss value . In this section results relating the relative values of resources to the optimal transition between them will be demonstrated. In particular these relations will be considered in the limit; In this way one maintains a close relationship with the idea of thermodynamic exchanges where many particle systems transition between various states.

To begin the definition of a monotone is required, which takes a resource and assigns it a number value such that the resources it transitions to have a lower value, and the resources that transition to it have a higher one.

Definition 39 (Montone)**.** *Consider a resource theory D and a function $f : |D| \to \mathbb{R}$ sending resources to real numbers. This function is monotonic, or a monotone, if for all $\rho, \sigma \in |D|$ such that $\rho \to \sigma$ then $f(\rho) \geq f(\sigma)$. That is to say if it is preorder preserving.*

As mentioned the aim here is also to discuss asymptotic transitions. If value is to stay meaningful in the limit one must be able to move to a value density as the infinite composition of resources of non-zero value will themselves become asymptotic. For this reason regularised functions are introduced

Definition 40 (Regulisable function)**.** *A function $f : |D| \to \mathbb{R}$ acting on a resource theory D is called regularisable if the limit*

$$f^{\infty}(\rho) = \limsup_{n\to\infty} \frac{f(\rho^{\otimes n})}{n} \quad exists \quad \forall \rho \in |D| \text{ and is bounded.} \tag{4.22}$$

where f^{∞} is refereed to as the Regulised function.

These regularised functions clearly bear a density like form. This is ideal for attaining a value density - the question is whether a regularised monotone is itself a monotone. It turns out this is the case.

Proposition 4.1.1. *The Regulised monotones, of monotones that are regularisable, are themselves monotones.*

Proof. If one has a transition $\rho \to \sigma$ then by definition for any monotone f one has that $f(\rho) \geq f(\sigma)$. Taking the given transition labelled $\mathscr{F}_1$ one can form the composite

$$\mathscr{F}_1^{\otimes n} : \rho^{\otimes n} \to \sigma^{\otimes m_n}$$

which implies that $f(\rho)^{\otimes n} \geq f(\sigma)^{\otimes n}$. With this inequality in hand and using the monotonicity of the limit one arrives at

$$\lim_{n\to\infty} \frac{f(\rho^{\otimes n})}{n} \geq \lim_{n\to\infty} \frac{f(\sigma^{\otimes n})}{n}$$

which can be rewritten as

$$f^{\infty}(\rho) \geq f^{\infty}(\sigma)$$

□

These densities also have another property which is useful in later proofs: they are necessarily non-zero.

Lemma 4.1.2. *The Regulised monotones are non-negative.*

Proof. Free resources are defined as those which can be created at no cost. Therefore having one copy or many is should have the same value, which is to say that for a monotone function f and free resource $\theta \in |D|$ for a resource theory D

$$f(\theta) = f(\theta^{\otimes n}).$$

Now by considering the regularised monotone one can see

$$f^{\infty}(\theta) = \lim_{n\to\infty} \frac{f(\theta^{\otimes n})}{n} = \lim_{n\to\infty} \frac{f(\theta)}{n} = 0$$

As free resources are those of minimum value it must be the case that $f^{\infty} \geq 0$ for all $\rho \in |D|$. □

4. Rates

With regularised monotones defined one can show that the regularised monotonic value of an initial resource is larger than or equal to that of the final resource multiplied by the optimal rate. This cannot be presented directly, a few technical lemmas are required. Firstly one must observe that the sequence $\{m_n\}_{n\in\mathbb{N}}$ of some protocol $\bar{\mathscr{F}}$ is unbounded when the optimal rate is non-zero.

Lemma 4.1.3. *If $R(\rho \to \sigma) > 0$ for the transitions $\rho^n \to \sigma^{m_n}$, is given by a protocol $\bar{\mathscr{F}}$, then the sequence belonging to this protocol $\{m_n\}_{n\in\mathbb{N}}$ is unbounded.*

Proof. If m_n, which is positive by definition, is bounded then $\lim_{n\to\infty} m_n \leq L$ for some $L \in \mathbb{R}$. This then implies

$$\lim_{n\to\infty} \frac{m_n}{n} \leq \frac{L}{n} = 0$$

This implies that the limit must in fact be zero. As $R(\rho \to \sigma)$ is the limit superior of such ratios, which is here simply the limit, then it is zero - a contradiction. □

The second important technical observation is that from any protocol, provided one can discard resources, one can always construct from it something called its *Steady derived protocol*, the term *steady* indicates that it is non-oscillating. Crucially it can be shown that this new protocol has the same optimal rate as the original.

Lemma 4.1.4. *Suppose one has a protocol $\bar{\mathscr{F}}$, where its rate is given by $\limsup_{n\to\infty} \frac{m_n}{n}$ where $\{m_n\}_{n\in\mathbb{N}}$ is the sequence of output resources indexed by the input resources n. Then given that one can discard resources, one can construct a second protocol called the* steady derived protocol *$\bar{\mathscr{F}}'$ with a non-decreasing sequence giving the same rate.*

Proof. First take the protocol $\bar{\mathscr{F}}$ composed of a sequence $\{m_n\}_{n\in\mathbb{N}}$ and a operations $\{\mathscr{F}_n\}_{n\in\mathbb{N}}$ from this, one can then construct $\bar{\mathscr{F}}$ recursively as follows: for its sequence we let

$$m'_n = \begin{cases} m_1 \text{ if } n = 1 \\ m_n \text{ if } m_n \geq m'_{n-1} \\ m'_{n-1} \text{ if } m_n < m'_{n-1} \end{cases} \tag{4.23}$$

and for its operations $\mathscr{F}'_n$

$$\mathscr{F}'_n = \begin{cases} \mathscr{F}_1 \text{ if } n = 1 \\ \mathscr{F}_n \text{ if } m_n \geq m'_{n-1} \\ \mathscr{F}'_{n-1} \circ \left(u \otimes \mathscr{I}^{\otimes(n-1)}\right) \text{ if } m_n < m'_{n-1} \end{cases} . \tag{4.24}$$

where u represents the discarding operator. The new protocol takes the first operation in the original and then for every subsequent operation checks if its output is lower than the previous one. If it is, it discards one of the input resources and applies the previous operation, if not, it uses this operation. In this way one obtains a non-decreasing sequence.

Observe that this protocol will have the same optimal rate as the old protocol[1], as the ratio $\frac{m_{n*}}{n*} = \sup_n \frac{m_n}{n}$ can not be removed by this process. If it was removed this would imply $m'_{n*-1} > m_{n*}$ which would in turn imply that there exists m_p where is an integer where $p < n$ such that $\frac{m_p}{p} > \frac{m_{n*}}{n*}$ which is a contradiction. □

Corollary 4.1.4.1. *As $\bar{\mathscr{F}}'$ has a non-decreasing sequence its optimal rate obtained in the limit by lemma 4.0.3 is obtained as a limit proper* - not just a superior limit.

This *steady derived protocol* obeys the following, somewhat suggestive property, which will be key in reaching the final result.

Lemma 4.1.5. *For a steady derived protocol $\bar{\mathscr{F}}'$, composed of sequences $\{m'_n\}$ and $\{\mathscr{F}'_n\}$ derived from some protocol $\mathscr{F}$, composed of sequences composed of sequences $\{m_n\}$ and $\{\mathscr{F}_n\}$, the following relation holds*

$$\limsup_{n\to\infty} \frac{f(\rho^{\otimes n})}{n} \geq R(\rho \to \sigma) \limsup_{n\to\infty} \frac{f(\sigma^{\otimes m'_n})}{m'_n} \tag{4.25}$$

[1]Note this is not the same thing as the optimal rate of transition which is optimal over both the number of resources n and also over all the protocols.

Proof. Consider the protocol $\bar{\mathscr{F}}$ which contains the operation $\mathscr{F}_n$ which attains the optimal rate $R_D(\rho \to \sigma)$ for some transition $\rho^{\otimes n} \to \sigma^{\otimes m'_n}$. This transition then implies that for f, a regularisable monotone

$$f(\rho^{\otimes n}) \geq f(\sigma^{\otimes m'_n}) \tag{4.26}$$

from this one can see that

$$\frac{f(\rho^{\otimes n})}{n} \geq \frac{f(\sigma^{\otimes m'_n})}{n} = \frac{f(\sigma^{\otimes m'_n})}{m'_n}\frac{m'_n}{n} \tag{4.27}$$

one then take the limit superior

$$\limsup_{n\to\infty} \frac{f(\rho^{\otimes n})}{n} \geq \limsup_{n\to\infty} \frac{f(\sigma^{\otimes m'_n})}{m'_n}\frac{m'_n}{n} \tag{4.28}$$

Using corollary 4.1.4.1 one can pull the rate term out of the superior limit, as it converges, giving

$$\begin{aligned}
\limsup_{n\to\infty} \frac{f(\rho^{\otimes n})}{n} &\geq \limsup_{n\to\infty} \frac{f(\sigma^{\otimes m'_n})}{m'_n}\frac{m'_n}{n} \\
&= \lim_{n\to\infty} \frac{m'_n}{n} \limsup_{n\to\infty} \frac{f(\sigma^{\otimes m'_n})}{m'_n} \\
&= R(\rho \to \sigma) \limsup_{n\to\infty} \frac{f(\sigma^{\otimes m'_n})}{m'_n}
\end{aligned} \tag{4.29}$$

□

Lemma 4.1.6. *For a steady derived protocol $\bar{\mathscr{F}}'$ representing transitions $\rho^n \to \sigma^{m'_n}$, composed of sequences $\{m'_n\}$ and $\{\mathscr{F}'_n\}$ derived from some protocol $\mathscr{F}$ representing the transitions $\rho^n \to \sigma^{m_n}$, composed of sequences composed of sequences $\{m_n\}$ and $\{\mathscr{F}_n\}$ the following property holds for the regularised monotone of the output resource.*

$$\limsup_{n\to\infty} \frac{f(\sigma^{\otimes m'_n})}{m'_n} \geq f^{\infty}(\sigma) \tag{4.30}$$

Proof. Too see this one must make use of lemma 4.1.3 to see that the new protocol's sequence $\{m'_n\}_{n\in\mathbb{N}}$ is unbounded, which indicates we can use $\{m'_n\}_{n\in\mathbb{N}}$ construct a *strictly monotonic* subsequence $\{m'^s_n\}_{n\in\mathbb{N}}$

$$m'^s_n = \begin{cases} m'_1 \text{ if } n = 1 \\ m'_n \text{ if } m'_n > m'^s_{n-1} \\ m'_{n+p} \text{ if } m'_n \leq m'_{n-1} \text{ where } p = \inf\{p' \in \mathbb{N} | m'_{n+p'} > m'_n\} \end{cases} \tag{4.31}$$

one can view p as the number of terms one must skip in the original sequence to find one that is larger. Importantly as a strictly increasing monotonic sequence it is a subsequence of the natural numbers. The importance of this is that if we consider the limit set E_n of the sequence $\frac{f(\sigma^{\otimes n})}{n}$ then it contains the limit set $E_{m'^s_n}$ of the subsequence $\frac{f(\sigma^{\otimes m'^s_n})}{m'^s_n}$, that is to say

$$E_{m'^s_n} \subseteq E_n. \tag{4.32}$$

One also has that by construction $\{m'^s_n\}_{n\in\mathbb{N}}$ is a subsequence of $\{m'_n\}_{n\in\mathbb{N}}$ which implies that for the sequence $\frac{f(\sigma^{\otimes m'_n})}{n}$ one has that

$$E_{m'^s_n} \subseteq E_{m'_n}. \tag{4.33}$$

The key observation is that E_n contains one element, as the limit *must* exist by definition: due to the fact f is a regularisable monotone. One knows that any subsequence must converge to the same limit, or put another way one knows that $E_{m'_k}$ must have at least one element, however it is a subset of E_n which is a singleton - meaning it must also be a singleton with the same element - this directly implies that

$$E_{m'^s_n} = E_n \tag{4.34}$$

this allows one to rewrite

$$E_n \subseteq E_{m'_n}. \tag{4.35}$$

By the definition of the superior limit and the fact that the supremum of a set must be greater than or equal to those of its subsets one can see that

$$\limsup_{n\to\infty} \frac{f(\sigma^{\otimes m'_n})}{m'_n} = \sup E_{m'_n} \geq \sup E_n = \limsup_{n\to\infty} \frac{f(\sigma^{\otimes n})}{n} = f^\infty(\sigma) \tag{4.36}$$

□

4. Rates

Finally with the above lemma one is in a position to reach the main result:

Theorem 4.1.7. *For a resource theory D and a regularisable monotone f where there exists a resource $\theta \in |D|$ such that $f(\theta) > 0$ the following inequality holds*

$$f^{\infty}(\rho) \geq f^{\infty}(\sigma) R(\rho \to \sigma) \tag{4.37}$$

where $\rho, \sigma \in |D|$, provided there is such a transition between them.

Proof. Given some protocol achieving this rate, by lemma 4.1.4, one can construct a steady derived protocol which is non-decreasing and achieves this rate.

Then one knows, from lemma 4.1.6, that the following equation holds

$$\limsup_{n\to\infty} \frac{f(\sigma^{\otimes m'_n})}{m'_n} \geq f^{\infty}(\sigma) \tag{4.38}$$

Where $\{m'_n\}_{n\in\mathbb{N}}$ is the sequence of the steady derive protocol. Using this one can use the relation given by lemma 4.1.5 alongside the relation given above to show that

$$\begin{aligned} f^{\infty}(\rho) &\geq \limsup_{n\to\infty} \frac{f(\sigma^{\otimes m'_n})}{m'_n} \frac{m'_n}{n} \\ &= R(\rho \to \sigma) \limsup_{n\to\infty} \frac{f(\sigma^{\otimes m'_n})}{m'_n} \\ &\geq R(\rho \to \sigma) f^{\infty}(\sigma). \end{aligned}$$

where corollary 4.1.4.1 has been used, which showed the derived protocols rate is obtained as a limit and so can be removed from the superior limit as shown. □

This relationship allows one to propose useful bounds between certain cost and distillation rates.

Lemma 4.1.8. *For a regularisable monotone f we have the following relationship between the distillation and cost rate*

$$R_C(\sigma \to \rho) \geq \frac{f^\infty(\rho)}{f^\infty(\sigma)} \geq R_D(\rho \to \sigma) \tag{4.39}$$

when $f(\rho)$ and $f(\sigma)$ are non-zero.

Proof. consider the two arbitrary resources and some distillation and cost protocol between them

$$\mathscr{F}_D : \rho^{\otimes n} \to \sigma^{\otimes m_n}$$
$$\mathscr{F}_C : \sigma^{\otimes \tilde{m}_n} \to \rho^{\otimes n}$$

By equation 4.37 we can write

$$\frac{f^\infty(\rho)}{f^\infty(\sigma)} \geq R_D(\rho \to \sigma)$$

and similarly

$$\frac{f^\infty(\sigma)}{f^\infty(\rho)} \geq R_D(\sigma \to \rho)$$

taking this second equation and using the inverse relationship between $R_D(\sigma \to \rho)$ and $R_C(\sigma \to \rho)$ seen in lemma 4.0.4 gives us that

$$R_C(\sigma \to \rho) \geq \frac{f^\infty(\rho)}{f^\infty(\sigma)}$$

From which we can deduce the following relation

$$R_C(\sigma \to \rho) \geq \frac{f^\infty(\rho)}{f^\infty(\sigma)} \geq R_D(\rho \to \sigma)$$

□

In a practical sense this can be seen as the statement that the minimal cost of going from one resource to another is always less than or equal to the amount one can get back. If this was not the case it would be possible to continually extract more and more of a resource by moving in a cycle - the fact one cannot do this is in a sense an abstraction of Carnots theorem.

4. Rates

The final part of this section connects this work with that on rates found in the work of Coecke, Fritz and, Spekkens [11] where additive monotones were studied. These are defined below and then shown to be generalised by the concept of regularisable monotones. In this way all the above results apply in the additive setting

Definition 41. *A monotone $f : |D| \to \mathbb{R}$ is additive if $f(\rho^{\otimes n}) = nf(\rho)$.*

Lemma 4.1.9. *Additive monotones are regularisable*

Proof. If f is additive then by definition one has that for all resources ρ that

$$f(\rho^{\otimes n}) = nf(\rho)$$

and in this case

$$f^{\infty}(\rho) = \limsup_{n\to\infty} \frac{f(\rho^{\otimes n})}{n} = \limsup_{n\to\infty} \frac{nf(\rho)}{n} = f(\rho)$$

which is finite and exists. □

4.2 Currencies

It is a common feature of many systems that there may exist a resource which can transition to any other resource. In this way it can function as a link to all other resources and so acts as a medium of exchange.

Definition 42. *A currency is a resource $c \in |D|$ such that for all $\rho \in |D|$ there exists a transition*

$$c \to \rho \tag{4.40}$$

between these resources.

With a currency defined the rates associated with them take on a special role

Definition 43. *The distillation and cost rate in a resource theory D with a currency $c \in |D|$ is defined as follows*

$$R_D(\rho) := R_D(\rho \to c) \qquad R_C(\rho) := R_C(c \to \rho) \tag{4.41}$$

where $\rho \in |D|$ is some arbitrary resource.

These optimal rates are in one sense simply a particular case of those discussed above, however they now have a particular operational meaning. The currencies act as a mediator between all resources and therefore leads one to be able to define value as more than just something that exists between certain resources. It allows for a discussion of value as a global property that is induced by having this currency as a means to compare all resource. The advantage of this will become clear as this section progresses.

4. Rates

The nature of a currency has direct implications for the optimal rates too and from them. The rate of currency distillation from any other resource is at most one. Similarly one cannot spend currency at a rate any greater than one.

Lemma 4.2.1. *For an arbitrary resource $\rho \in |D|$ in a resource theory D with a currency, $R_D(\rho)$ and $R_C(\rho)$ are bounded above by one and below by zero*

$$0 \leq R_D(\rho) \leq 1 \text{ and } 0 \leq R_C(\rho) \leq 1$$

Proof. For $R_D(\rho)$

The lower bound is guaranteed by the fact rates cannot be negative by definition. To see the upper bound consider, once again, the family ordered pairs $< a_i, b_i >$ with $a_i, b_i \in \mathbb{N}$ representing *every* transition that exists between resources c and ρ, belonging to our resource theory, such that $c^{a_i} \to \rho^{b_i}$ where the index i runs over the set of all transitions between these resources

$$R_D(c \to \rho) = \sup_i \frac{b_i}{a_i}$$

To begin use the fact that $c \to \rho$ as c for all ρ in the resource theory as it is a currency. This implies that that $R_D(c \to \rho)$ must be at least 1, which is to say

$$R_D(c \to \rho) = \sup_i \frac{b_i}{a_i} \geq 1 \tag{4.42}$$

Now from lemma 4.1.8 one knows that

$$\frac{f^\infty(\rho)}{f^\infty(c)} \geq R_D(\rho) \tag{4.43}$$

$$\frac{f^\infty(c)}{f^\infty(\rho)} \geq R_D(c \to \rho) \tag{4.44}$$

From the second of these alongside equation 4.42 one can get the following inequality

$$\frac{f^\infty(c)}{f^\infty(\rho)} \geq R_D(c \to \rho) \geq 1$$

by rearranging this result and applying 4.44 one arrives at

$$1 \geq \frac{f^\infty(\rho)}{f^\infty(c)} \geq R_D(\rho)$$

which provides the desired inequality.

For $R_C(\rho)$

The lower bound is again guaranteed by the fact rates cannot be negative by definition. For the upper bound consider, once again, the family ordered pairs $< a_i, b_i >$, where one knows that

$$R_C(\rho) := R_C(c \to \rho) = \inf_i \frac{b_i}{a_i}$$

We again use the fact that $c \to \rho$ which here implies that $R_C(\rho)$ is at most 1, which is to say

$$R_C(\rho) \leq 1$$

There is a relationship between the cost and distillation rates showing the rate of cost is greater than or equal to that of distillation.

□

Proposition 4.2.2. *For an arbitrary resource $\rho \in |D|$ in a resource theory with a currency, $R_D(\rho)$ and $R_C(\rho)$ are related as follows*

$$1 \geq R_C(\rho) \geq R_D(\rho) \geq 0 \tag{4.45}$$

Proof. Firstly consider equation 4.39 in the context of currency resource cost and distillation which gives us

$$R_C(\rho) \geq \frac{f^\infty(\rho)}{f^\infty(c)} \geq R_D(\rho)$$

then using lemma 4.2.1 one can extend this equation to show that

$$1 \geq R_C(\rho) \geq \frac{f^\infty(\rho)}{f^\infty(c)} \geq R_D(\rho) \geq 0 \tag{4.46}$$

□

4. Rates

This proposition guarantees a certain coherence to the notion of a currency, it indicates a certain Carnot-esque, or *Carnotic property*: one cannot get more currency out than one puts in. Without it the system becomes trivial one one could cyclically generate more and more currency and accumulate value from nothing.

A final observation is that the act of taking the cost or distillation rate of some resource is the same of assigning a value to them - this process is is itself a monotonic function. All resources can be assigned a global value based on the rate at which one can transition from them or too them from a currency.

Proposition 4.2.3. *$R_D(-) : |D| \to \mathbb{R}$ and $R_C(-) : |D| \to \mathbb{R}$ are themselves monotonic functions.*

Proof. For $R_D(-)$

Suppose in our resource theory D one has that $\rho \to \sigma$ and $\sigma \to c$ and from these the associated distillation rates. We can then make use of the relation $R_D(\rho) \geq R_D(\rho \to \sigma)R_D(\sigma)$. As one has that $\rho \to \sigma$ one knows that $R_D(\rho \to \sigma) \geq 1$ which along with the fact $R_D(\rho \to \sigma) \geq 0$ allows us to deduce that

$$R_D(\rho) \geq R_D(\sigma)$$

Which shows $R_D(-)$ is a monotone function.

For $R_C(-)$

Suppose in our resource theory D one has that $\rho \to \sigma$ and $\sigma \to c$ and from these the associated cost rates. We can then make use of the relation $R_C(\rho) \leq R_C(\rho \to \sigma)R_C(\sigma)$. As one has that $\rho \to \sigma$ one knows that $R_C(\rho \to \sigma) \leq 1$ which along with the fact $R_C(\rho \to \sigma) \geq 0$ allows us to deduce that

$$R_C(\sigma) \leq R_C(\rho)$$

Which shows $R_C(-)$ is a monotone function.

□

This shows that the existence of a currency guarantees the ability to assign value; this is perhaps not so strange given that all resources are connected can be reached by the currency and they in relation to it only in the rates at which this happens.

4.3 Reversibility and the role of monotones

In a plain terms reversibility is the property of doing something and then undoing it to find oneself in the original position. In many physical situations this is not true, in mechanical settings friction will mean that in performing an action and then undoing it one has lost energy - reversibility marks something special. Indeed in macroscopic thermodynamics reversible processes guarantee entropy preservation and energy conservation. In resource terms one should be able to transition from a certain amount of one resource to another, and then be able to transition back and reclaim exactly the same amount of the original resource as one started with. To put it formally one has the following:

Definition 44. *A resource theory is described as reversible if* $R_D(\rho \to \sigma)R_D(\sigma \to \rho) = 1$ *for all* ρ *and* σ *in* $|D|$.

Given the existence of regularisable monotones one can usefully characterise a reversible theory as one where distillation is equal to cost.

Lemma 4.3.1. *A resource theory D is reversible iff*

$$R_D(\rho \to \sigma) = R_C(\sigma \to \rho) \quad \forall \sigma, \rho \in |D| \tag{4.47}$$

Proof. $\implies$

Presume $R_D(\rho \to \sigma)R_D(\sigma \to \rho) = 1$ then one can write

$$R_D(\rho \to \sigma)\frac{1}{R_C(\sigma \to \rho)} = 1 \implies R_D(\rho \to \sigma) = R_C(\sigma \to \rho) \tag{4.48}$$

where proposition 4.0.4 was use to introduce the cost rate. This proves the resourcement in this direction.

$\impliedby$

Going the other way one finds that

$$\begin{aligned} R_D(\rho \to \sigma)R_D(\sigma \to \rho) &= R_D(\rho \to \sigma)R_C(\rho \to \sigma) \\ &= R_D(\rho \to \sigma)\frac{1}{R_D(\rho \to \sigma)} \\ &= 1 \end{aligned}$$

This proves the resourcement in the other direction. □

Corollary 4.3.1.1. *In a resource theory with a regularisable monotone f, the pinching of lemma 4.1.8 by the above equality tells one $R_C(\sigma \to \rho) = \frac{f^\infty(\rho)}{f^\infty(\sigma)} = R_D(\rho \to \sigma)$*

In a resource theory with a currency reversibility imposes heavy restrictions on the possible regularised monotones. In fact it shows that they are unique up to a constant.

Theorem 4.3.2. *In a reversible resource theory with a currency all regularisable monotones are unique up to a constant. That is to say for two regularisable monotone functions $f_1, f_2 : |D| \to \mathbb{R}$*

$$f_1^\infty(\rho) = a f_2^\infty(\rho) \quad \textit{where} \quad a = \frac{f_1^\infty(c)}{f_2^\infty(c)} \tag{4.49}$$

where c is the currency.

Proof. Consider a reversible theory, by definition 44 one has that $R_D(\rho \to c)R_D(c \to \rho) = 1$ and from corollary 4.3.1.1 one knows that

$$R_D(\rho \to c) = R_C(c \to \rho) = \frac{f^\infty(\rho)}{f^\infty(c)}$$

This is true for all resources as the existence of a currency guarantees a connection to all resources, which is reciprocated due to reversibility. Note the equation is true for all such monotonic functions so one can take two such examples f_1^∞ and f_2^∞ and write that

$$\frac{f_1^\infty(\rho)}{f_1^\infty(c)} = R_D(\rho \to c) \qquad \frac{f_2^\infty(\rho)}{f_2^\infty(c)} = R_C(c \to \rho) \tag{4.50}$$

But one knows $R_D(\rho \to c) = R_C(c \to \rho)$ therefore

$$\frac{f_1^\infty(\rho)}{f_1^\infty(c)} = \frac{f_2^\infty(\rho)}{f_2^\infty(c)} \implies f_1^\infty(\rho) = \frac{f_1^\infty(c)}{f_2^\infty(c)} f_2^\infty(\rho) \tag{4.51}$$

□

Note that in the above proof the importance of a currency is in ensuring there exists a mapping to every other resource which allows for the construction of the constant, in general there is no requirement that resources are connected like this.

4.4 Asymptotic conversion

The rates considered hitherto consisted of exact transitions that exist for all n. The next step is to consider transitions that are only possible in the limit. These are transitions that are asymptotically exact. A complementary physical analogue of this is the thermodynamic system - where one is looking at the transition possibilities in the thermodynamic limit of many particles, here phase transitions are possible, something which doesn't exist on the microscopic scale[32] .

The broad approach to analysing this mathematically is to take some transition that exists between one desired initial resource and another that is distinct from the desired output resource, but that *converges to to it in the limit.* The first thing required is a notion of proximity, for this reason one uses a metric [2].

Definition 45. *A metric on a resource theory is D is a function $d : |D| \times |D| \to [0, \infty)$ such that for all $\rho, \sigma, \theta \in |D|$*

$$d(\rho, \sigma) \geq 0 \tag{4.52}$$

$$d(\rho, \sigma) = 0 \text{ iff } \rho = \sigma \tag{4.53}$$

$$d(\rho, \sigma) = d(\sigma, \rho) \tag{4.54}$$

$$d(\rho, \theta) \leq d(\rho, \sigma) + d(\sigma, \theta) \tag{4.55}$$

The notation $d(\rho, \sigma) = |\rho - \sigma|$ is also common and bears resemblance to the distance between resources in quantum theory - this notion will be used from here on.

Definition 46. *A metrised resource theory is a pair (D, d) where D is a resource theory and d is a metric acting on it.*

With a metric defined one needs to consider the functions which are compatible with this structure in the limit. The aim is to define a function that is continuous in the asymptotic limit, in this way the function will respect that when the resources converge so should their respective mappings under the function.

[2] In principle topological proximity could suffice, here however a metric is used to maintain a description close to quantum theory. In quantum mechanics one has a normed vector space which naturally induces a metric.

4. Rates

Definition 47. *A function $f : |D| \to \mathbb{R}$ is asymptomatically continuous if for a sequence of resources such that*

$$\lim_{n\to\infty} |\rho_n - \sigma_n| = 0 \tag{4.56}$$

then

$$\lim_{n\to\infty} |(f(\rho_n) - f(\sigma_n))| = 0 \tag{4.57}$$

These technical points may look to some degree contrived. They are actually very natural. In the asymptotic limit the valuation functions, that is the monotones, are often themselves unbounded; As was discussed earlier adding more resources with non-zero value implies greater value. Recall the conclusion was to move to intensive quantities in the form of value densities, the regularised monotones. If similar resources are to converge for these sequences in the limit, then the monotones themselves *must* also converge on these sequences.

The rates considered hitherto consisted of exact transitions between resources. The focus now moves to considering transitions that are only possible in the limit. Suppose there exists a sequence of resources $\{\sigma'_n\}_{n\in\mathbb{N}}$ such that there exists operations

$$\mathscr{F}_n : \rho^{\otimes n} \to \sigma'_n \tag{4.58}$$

where

$$\lim_{n\to\infty} |\sigma^{\otimes m_n} - \sigma'_n| = 0 \tag{4.59}$$

In this way one can define an asymptotic protocol that exists only in the limit. To have protocol by definition 35 one requires a sequence of operations $\{\mathscr{F}_n\}_{n\in\mathbb{N}}$ and integers $\{m_n\}_{n\in\mathbb{N}}$. Note however that this protocol is only meaningful in the asymptotic limit where the transitions $\rho^{\otimes n} \to {\sigma'}_n$ become $\rho^{\otimes n} \to \sigma^{m_n}$. In this way one has exact transitions only available in the limit. One can now define the transition rate in the asymptotic limit.

Definition 48. *In a metrised resource theory D the asymptotic transition rate between the resources $\rho, \sigma \in |D|$ is defined as follows: let $\mathscr{F}$ be a protocol such that*

$$\lim_{n\to\infty} |\mathscr{F}_n(\rho^{\otimes n}) - \sigma^{\otimes m_n}| = 0. \tag{4.60}$$

Then one can write

$$\begin{aligned} R(\rho \to \sigma) &= \sup_{\mathscr{F}} \limsup_{n\to\infty} \frac{m_n}{n} \\ &= \sup_{\mathscr{F},n} \frac{m_n}{n} \end{aligned}$$

where the simplification in the last line is due to lemma 4.0.3.

Notice that in most respects this transition rate is the same as that of the optimal distillation rate, there is a certain distinction due to the fact that this distillation rate only makes sense in the limit.

With asymptotic theories defined it is then possible to consider asymptotically reversible theories.

Definition 49. *An asymptotically reversible resource theory is a reversible resource theory where for any two $\rho, \sigma \in |D|$ which are* not free resources *$0 \leq R_D(\rho \to \sigma) \leq \infty$ and*

$$R(\rho \to \sigma)R(\sigma \to \rho) = 1. \tag{4.61}$$

Note that the bounds are inclusive as the infinite rate is possible when one wants to distil free resources out of non-free resources; conversely the rate is zero when one wants to distil non-free resources out of free resources. With these definitions outlined one can relate rates to the ratio of the input and outputs resources value densities, in other word their regularised monotones.

Theorem 4.4.1. *For a reversible resource theory D, if f is a regularisable monotone function and there exists a resource θ such that $f^\infty(\theta) > 0$ that is asymptomatically continuous where the asymptotic transition rate is such that $R_D(\rho \to \sigma) < \infty$ then one has the following equality:*

$$R(\rho \to \sigma) = \frac{f^\infty(\rho)}{f^\infty(\sigma)} \tag{4.62}$$

Proof. Take the valued resource θ which is defined such that $f(\theta) > 0$, one can then write that via lemma 4.1.7 that

$$f^\infty(\theta) R(\rho \to \theta) \leq f^\infty(\rho) \tag{4.63}$$

which is trivially equal to

$$R(\rho \to \theta) \leq \frac{f^\infty(\rho)}{f^\infty(\theta)}. \tag{4.64}$$

One also has that

$$f^\infty(\rho) R(\theta \to \rho) \leq f^\infty(\theta) \tag{4.65}$$

if one multiplies the latter by $R(\rho \to \theta)$ gives

$$f^\infty(\rho) R(\theta \to \rho) R(\rho \to \theta) \leq f^\infty(\theta) R(\rho \to \theta). \tag{4.66}$$

Reversibility implies the rates in the left cancel each other giving

$$\frac{f^\infty(\rho)}{f^\infty(\theta)} \leq R(\rho \to \theta). \tag{4.67}$$

this is the converse inequality to equation 4.64, which implies

$$\frac{f^\infty(\rho)}{f^\infty(\theta)} = R(\rho \to \theta). \tag{4.68}$$

This implies that $f^\infty(\rho) > 0$ for any resource of D. If one then applies the above logic to some arbitrary resource σ instead of θ multiplying the equalities, and again using reversibility, one obtains

$$R(\rho \to \sigma) = \frac{f^\infty(\rho)}{f^\infty(\sigma)} \tag{4.69}$$

□

Remark 8. *Careful analysis of this theorem reveals it would also be valid sans asymptotic continuity for standard non-asymptotic transitions rates. It is introduced in this form to draw parallels with previous work discussed below.*

Remark 9. *This last formula has particular importance when one is considering entropic functions. In quantum theory one has the notion of an entropy distance $S(\rho|\sigma)$[18] between two resources $S\rho, \sigma$*

$$S(\rho|\sigma) = Tr\rho\log\rho - Tr\rho\log\sigma \tag{4.70}$$

From this entropy distance one can talk of relative entropy distance

$$E_r(\rho) = \inf_{\rho\in States} S(\rho|\sigma) \tag{4.71}$$

which is gives the smallest entropic distance from the set of bound entangled resources (the free resources of entanglement theory).

This distance is regularisable[18] and so given that there exists resources θ such that $E_r(\theta) > 0$ then the optimal rate of transition from ρ to σ can be given by

$$R(\rho \to \sigma) = \frac{E_r^\infty(\rho)}{E_r^\infty(\sigma)} \tag{4.72}$$

In this way the above theorem can be considered a generalisation of the work of Horodecki and Oppenheim in [18].

The above theorem essentially proves that if one has a resource theory where one is considering transitions that only exist in the limit,and said theory is reversible one can get the optimum rate simply by dividing the asymptotically continuous regularised value of the input resource by that of the output resource. Their relative value densities fully characterise the rate of exchange: to know their values is to know their optimal transition rate.

5

The Resource Theory Of Purity

The focus of the resource theory of purity is entropy - the purer a state, the lower its entropy, the greater its value as a resource. Purity theory has a natural relation to thermodynamics. In fact one can consider thermodynamics as a theory of purity where one must preserve energy[1]. Conversely purity theory is thermodynamics at fixed energy[18], equivalently it is the thermodynamics of an isolated system. This has an intuitive feel to it when one considers that thermodynamics, be it quantum or classical, can be expressed as the transition rules where one has ever increasing entropy and a conserved Hamiltonian. There is also the relationship to information, in this respect purity relates to certainty of knowledge. To understand this in a thermodynamic manner one need only look to the third law of thermodynamics which can be stated as - *The entropy of a perfect crystal of any pure substance approaches zero as the temperature approaches absolute zero* - as temperature drops one becomes certain of where each atom in the crystal belongs in the lattice. When one is certain, where ones knowledge is 'pure', entropy is zero.

[1] As energy conservation is the result of a symmetry - symmetry in time. One can state thermodynamics as a theory of purity and asymmetry. For more on resource theories of asymmetry see []

One may feel that a notion of resource should be more practical - in this manner one can also find motivation. Consider again the Szillard engine [2]. The certainty of knowledge regarding, in the classical case, the position of a particle allows one to extract energy from the engine. Here the value of a resource, the low entropy, directly correlates to being able to extract energy - one is hard pushed to think of a more fundamentally property. In this way purity is a very natural quality to view as a resource.

Formally one can describe a resource theory of purity as a partitioned theory. The super-category is all the states and the possible transitions between them. The sub theory is generated by any one of a number of generating free processes, soon to be discussed. The functional aim will be to outline some set of free processes and analyse the result. Then by adding requirements regarding an invariant state with similar properties to the quantum mixed state one hopes to approach a realistic thermodynamic equivalent.

5.1 The Random Reversible Theory of Purity

In this theory the free transformations are *random reversible channels*[5]. These transformations are probabilistic selections of channels. There is a particular simplicity to this theory in that it requires only the axioms of an OPT.

Definition 50. *A random reversible (RaRe) channel is a probabilistic selection of reversible channels. For a random reversible channel $\mathscr{R}$ of the form $\mathscr{R} = \sum_i p_i \mathscr{U}$ where $\{p_i\}$ is a probability distribution, and $\mathscr{U}_i$ is a reversible channel for every i.*

As the identity transformation is a trivial RaRe channel, and the parallel and sequential composition of RaRe channels are Rare, these transformation can indeed form a set of free operations.

It does have the limitation of having no free states - such a thing requires a transformation from the trivial system to some other, however RaRe transformations go from a system to itself. It could equivalently be said that this theory posses only trivial free states. This issue motivates a broader search for suitable physically meaningful representations of thermodynamics.

5.2 Generalising Micro-Canonical Equilibrium [8]

To broaden the search for suitable resource theories one should return to the idea of minimum entropy being valuable. The contrapositive, that the maximally mixed state is the state with least value, suggests this state as an excellent candidate for the free state. This intuition finds justification in quantum thermodynamics proper; where one has the micro-canonical equilibrium state. A state which when reached is valueless - as one can no longer extract work - the functional purpose of thermodynamics.

In quantum thermodynamics the micro-canonical equilibrium state is the maximally-mixed state $\chi = \frac{I}{d}$, for a d dimensional system, where I is the identity operator. It is characterisable as the state that is invariant under unitary transformations[2]. Analogously one can define an invariant state:

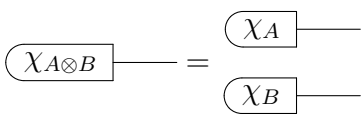
Definition 51 (Invariant State)**.** *A state $\chi \in St(A)$ is* invariant *if $\mathscr{U}\chi = \chi$, for every reversible channel $\mathscr{U}$. Where A is some arbitrary system.*

In microcanonical thermodynamics the system always has a definite energy, and such a system is presumed to have a well defined equilibrium which motivates the existence of an equilibrium like state in this framework. As the aim here is to define a framework for quantum theory and theories like it, the following requirements are proffered based on the work of [8]:

Requirement 1 – Every finite system has one unique deterministic invariant state. For a system A this state is written as χ_A. Mimicking quantum thermodynamics this state is considered free.

As the product of two free states must be a free state itself this demands a second requirement.

Requirement 2 – The invariant state of a composite system is the product of the invariant states of its components:



(5.1)

[2]For a geometric perspective consider that the equilibrium state is the only state unaffected by any rotation of the Bloch hypersphere

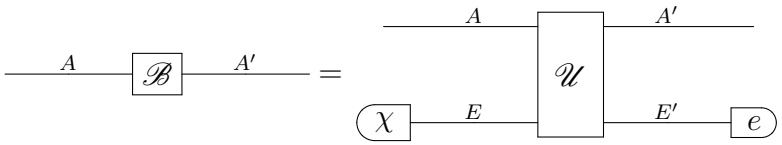
This second condition has a definite thermodynamic flavour in that the tensor product of a two equilibrium states is itself an equilibrium state. On these grounds equation 5.1 is referred to as the *equilibrium condition*. In later discussion of sharp theories with purification.

These requirements motivate the study of two further theories of purity.

5.2.1 The Noisy Resource Theory

Presented with the motivating requirements above it is natural to consider a theory where the free operations are those that allow one to generate invariant states and also dispose of outputs for free. In addition it seems natural that reversible channels should be costless. Formally one defines the following

Definition 52 (Basic Noisy Operation)**.** *A* basic noisy operation $\mathscr{B}$ *from a system* A *to a system* A' *is a channel that can be decomposed as*



(5.2)

Where E *and* E' *are suitable systems such that* $A \otimes E \equiv A' \otimes E'$, $\mathscr{U}$ *is a reversible channel, and* e *is a deterministic effect, representing one possible was to discard system* E*.*

Definition 53 (Noisy Operation)**.** *A channel* $\mathscr{N}$ *is a noisy operation if it is the limit of a sequence of basic noisy operations* $\{\mathscr{B}_n\}$*, that is, if for every reference system* R*, every state* $\rho \in St(A \otimes R)$*, and every effect* $E \in Eff(A' \otimes R)$ *the probabilities* $(E|\mathscr{B}_n \otimes \mathscr{I}_R|\rho)$ *converge to the probability* $(E|\mathscr{N} \otimes \mathscr{I}_R|\rho)$

This second definition is necessitated by the fact that the limit sequence of basic noisy operations may not be a basic noisy operation [8].

The equilibrium condition 5.1 enforces that the sequential and parallel composition of two noisy operations is a noisy operation. The identity channel is also trivial noisy operation. One can then deduce noisy operations are indeed a resource theory of purity.

5.2.2 The Unital Resource Theory

In this resource theory the free operations are defined only as those that transform invariant states to invariant states, In this way it is the broadest possible theory meeting the above requirements. Such transformations are denoted *unital channels*.

Definition 54. *A channel $\mathscr{D}$ from a system A to system B is called unital if $\mathscr{D}\chi_A = \chi_B$.*

The set of unital channels contains the identity and, once again, the equilibrium condition guarantees sequential and parallel composition of unital channels are themselves unital. This resource theory is called the unital resource theory of purity.

5.2.3 Containment Relations

There is a relation between these proposed resource theories, presented in [9], which will later allow us to define a necessary and sufficient condition for transitions between states. The first step is to outline the relative inclusions of the theories.

Proposition 5.2.1. *Every Rare channel is unital.*

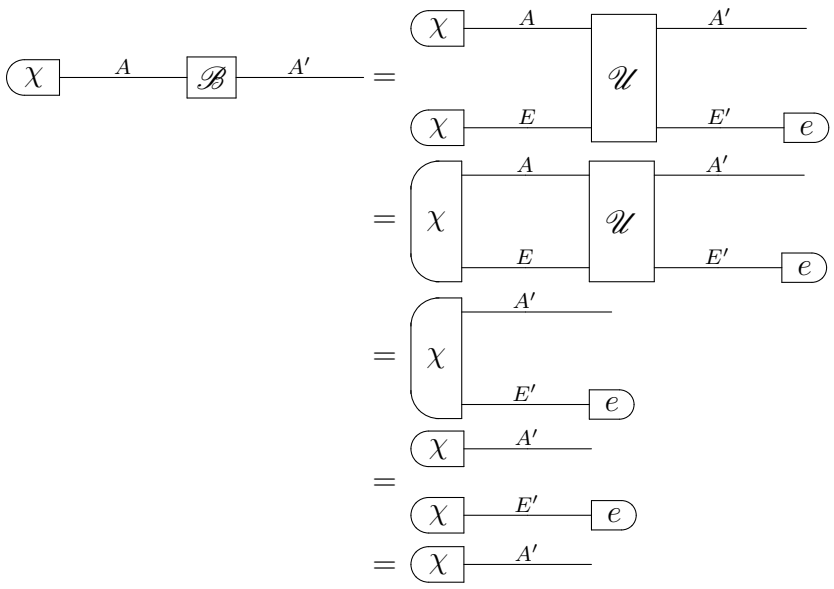
Proof. As all RaRe channels can be decomposed into a mixture of reversible channels which each preserve the invariant state then by definition

$$RaRe \subseteq Unital \tag{5.3}$$

□

Proposition 5.2.2. *Every noisy operation is unital, when requirement two is satisfied.*

Proof. [8] Suppose $\mathscr{B}$ is a basic noisy operation, decomposed as in equation 52, then one can see that



where requirement two allows the partition of the invariant state to be two invariant states. □

These relations hold for general probabilistic theories. In order to move closer to quantum theory further axioms must be introduced.

5.3 Informational Axioms [8]

It has been mentioned above how the fundamental axioms of quantum mechanics should be physical in nature. In [5], informational axioms are put forward - this is the natural approach to physical axioms of quantum mechanics. Direct human interaction with quantum systems comes almost exclusively through experiment; all one can know is the information that can be deduced about said systems via experiments. Information acquisition and manipulation are the fundamental phenomenological aspects of quantum mechanics.

The class of theories obeying these axioms are called *Sharp Theories with Purification*. These axioms are as follows:

1. Causality

   The probability that a transformation occurs is independent of the choice of tests performed on its output.

2. Purity Preservation

   Sequential and parallel compositions of pure transformations yield pure transformations.

3. Pure Sharpness

   Every system possesses at least one pure effect occurring with unit probability on some state.

4. Purification

   Every state has a purification and said purification is essentially unique - definition 23.

These can be given an informational flavour: causality implies that a signal cannot pass from the future to the past; Purity Preservation implies information is preserved when pure transformations are composed; Pure Sharpness implies certain knowledge is attainable; Purification implies incomplete knowledge is the result of 'forgetting' some information.

It will be profitable to consider causality under more formal conditions

Definition 55 (Causal theories)**.** *[4] A theory is causal if every preparation-test* $\{\rho_i\}_{i\in X}$ *and every observation-test* $\{a_j\}_{i\in Y}$ *on a system A the marginal probability* $p_i := \sum_{j\in Y}$*. Precisely, if* $\{a_j\}_{j\in Y}$ *and* $\{b_k\}_{k\in Z}$ *are two different observation-tests, then one has*

$$\sum_{j\in Y}(a_j|\rho_i)_A = \sum_{k\in Z}(b_k|\rho_i)_A \tag{5.4}$$

In the following section these informational axioms are utilised in the context of a general probabilistic theory to give a class of theories similar to and including quantum mechanics. These are called *sharp theories with purification.*

Definition 56. *A general probabilistic theory obeying the axioms of causality, purity preservation, pure sharpness, and purification, is a* sharp theory with purification*.*

5.4 Properties of Sharp Theories with Purification [4]

Restricting ones attention to Sharp theories with purification has a number of benefits, not least the replication of a number of philosophical and technical properties seen in quantum mechanics.

Causality alone introduces a number of restrictions on what is possible. One key technical observation is that the introduction of causality limits the deterministic effect to a single transformation.

Lemma 5.4.1. *In a causal theory the deterministic effect is unique for each system.*

Proof. If e and e' are two deterministic effects for system A. As $(e|\rho_i)_A = (e'|\rho_i)_A$ for every state ρ_i then $e = e'$.

Going the other way, if one has a unique deterministic effect and an observation test $\{a_j\}_{j\in Y}$ on system A then one can utilise coarse graining. This allows the creation of a single outcome test, with deterministic effect , which by uniqueness is equal to the original, that is to say,

$$(e|_A = (e'| = \sum_{j\in Y} (a_j|_A \,. \tag{5.5}$$

From this one can write that for every state ρ the probability $(e|\rho_i)_A = \sum_{j\in Y}(a_j|\rho_i)_A$, independently of the choice of observation tests $\{a_j\}_{j\in Y}$. The meaning of this is that the outcome probability is independent of what tests are performed in the future - the definition of causality. □

This lemma then restricts the marginal states of a system to being unique. This can be seen as guaranteeing that there is only one result of discarding information. This guarantees that it is physically meaningful to consider parts of a larger system as those components are uniquely defined. As considering sub-sets of larger systems is a procedure used repeatedly in thermodynamics, and physics in general, this serves as strong evidence in favour of sharp theories with purification being a suitable general framework for quantum thermodynamics.

Proposition 5.4.2. *The marginals of a bipartite state are unique.*

Proof. As the deterministic effect is unique all marginal states, definition 23, are unique. □

Mirroring quantum mechanics the notation $Tr[\rho] := (e|\rho)_A$ will be adopted for applying the deterministic state - which like the traditional trace in Hilbert spaces sends states to probabilities. Furthermore for some arbitrary pure state ρ_{AB} the marginal, where system B is discarded, will denoted by the partial trace $Tr_B[\rho_{AB}]$.

Remark 10. *If a theory obeys causality then a state ρ can be deterministically prepared iff $Tr[\rho] = 1$, which is to say it is normalised. The set of normalised states on some system A is denoted $St_1(A)$.*

A particularly striking and useful feature of sharp theories with purification is that one can also break and state apart into diagonalised components of perfectly distinguishable pure states - this serves both as a useful technical tool and a further evidence of a very close connection with quantum thoery.

Theorem 5.4.3 ([7])**.** *Every normalised state $\rho \in St_1(A)$ of every system A can be decomposed as*

$$\boxed{\rho}\!\!-\!\!- = \sum_{1=1}^{r} P_i \boxed{\alpha_i}\!\!-\!\!- \tag{5.6}$$

where r is an integer (called the rank of the state), $p_1 \geq p_2 \geq ... \geq p_r > 0$ are probabilities (called the eigenvalues), and $\{\alpha_i\}_{i=1}^{r}$ is a set of perfectly distinguishable pure states(called the eigenstates).

Proof. The proof of this is beyond the scope of this work but can be found in reference [6] and [7]. □

Remark 11. *The vector of eigenvalues $P = (p_1, ..., p_r)$ is called the spectrum. This spectrum is unique for sharp theories with purification[7], which is also shown for the above purity theories in the later corollary 5.5.1.2.*

The structure of sharp theories with purification also allows us to expand on the previous containment relations.

Theorem 5.4.4. *In every sharp theory with purification, RaRe channels are noisy operations.*

Proof. The proof of this is non-trivial and can be found as the proof of theorem 2 in [8]. □

This implies the following containment relation

$$RaRe \subseteq Noisy \subseteq Unital \tag{5.7}$$

.

5.4.1 State-effect Duality

A further relation with quantum mechanics is, in rough terms, is that one can know something with certainty by utilising the dual of a state - in quantum mechanics one knows for some state $|x\rangle$ that $\langle x|$ applied to $|x\rangle$ gives unity - one can be certain we have this pure state by applying its dual. This is also true in sharp theories with purification:

Proposition 5.4.5 (State-effect Duality)**.** *Let $PurSt_1(A)$ and $PurEff_1(A)$denote the set of pure states and effects on system A respectively. There is a bijective correspondence between the normalised pure states and normalised pure effects. Specifically, if $\alpha \in PurSt_1(A)$, where $PurSt_1(A)$,there exists a unique $\alpha^\dagger \in PurEff_1(A)$ such that $(\alpha^\dagger|\alpha) = 1$*

Proof. The proof of this is beyond the scope of this work but can be found in reference [6]. □

As discussed, the meaning of this *duality* is that for each pure state there is a means by which to identify it with certainty. Whether ones personal taste is for vector duals or 1-forms - the comparison with traditional quantum mechanics based on dual Hilbert spaces, particularly in Dirac notation, is apparent.

One can extend the notion of perfectly distinguishable states to the idea of distinguishable of maximal sets of said states, alongside the dual set of tests. To state this formally first note the following definition:

Definition 57. *An observation test $\{a_i\}_{i\in X}$ is called perfectly distinguishing if there exists a set of states $\{\rho_i\}_{i\in X}$, such that $(a_i|\rho_i) = \delta_{ij}$ for all i and j in X.*

One can now attain dual sets of pure states and effects:

Proposition 5.4.6. *The pure states $\{a_i\}_{i\in X}$ are a maximal set of perfectly distinguishable pure states if and only if the pure effects $\{a_i^\dagger\}_{i\in X}$ form an observation-test.*

Proof. The proof can be found in [7]. □

These maximal sets of perfectly distinguishable pure states have the same cardinality. One can see that

$$1 = \sum_{i \in X} \left(\alpha_i^\dagger | \chi_A \right) = \sum_{i \in X} \left(\alpha_0^\dagger | \mathscr{U}_i | \chi_A \right) = |X| \left(\alpha_0^\dagger | \chi_A \right) \tag{5.8}$$

where by proposition 5.4.8 α_0 is connected to all other α_i by a reversible transformation, which by state effect duality implies $\alpha_i^\dagger = \alpha_0^\dagger \mathscr{U}_i$ and definition 51 tells one that reversible channels do not alter the invariant state.

It is now meaningful to talk of a dimension of an entire system A denoted d_A which is tied to the dimension of the systems maximal set of distinguishing states. To understand this notice that any state on A can be diagonalised such that it is decomposed over the elements of a maximal set of distinguishable pure states - exactly like a linear combination of orthogonal basis vectors. As each can be selected by a maximal measurement set of the same size one can separate out the diagonal decomposition using these duals - the state really is a set of separate components - thus it is accurate to talk of dimension - the number of elements that must be specified to characterise the state.

This leads one into a position where one can talk about the dimension of composite systems relative to their components.

Proposition 5.4.7. *If $\{a_i\}_{i=1}^{d_A}$ is a maximal set of perfectly distinguishable pure states for system A and $\{\beta_j\}_{j=1}^{d_B}$ is a maximal set of perfectly distinguishable pure states for system B, then $\{\alpha_i \otimes \beta_j\}_{i \in \{1,\ldots,d_A\},\{1,\ldots,d_B\}}$ is a maximal set of perfectly distinguishable pure states for the composite system $A \otimes B$.*

Proof. By proposition we know $\{a_i\}_{i=1}^{d_A}$ is an observation test for A, as $\{\beta_j\}_{j=1}^{d_B}$ is for B. The product $\{\alpha_i \otimes \beta_j\}_{i \in \{1,\ldots,d_A\},\{1,\ldots,d_B\}}$ by definition is an observation test on the composite system $A \otimes B$. Moreover, each effect $\alpha_i \otimes \beta_j$ is pure, due to purity preservation. From proposition we know that $\{\alpha_i \otimes \beta_j\}_{i \in \{1,\ldots,d_A\},\{1,\ldots,d_B\}}$ is a maximal set. □

Corollary 5.4.7.1. *The dimension of a composite system is the product of the components. For a system A and B of dimension d_A and d_B, respectively, the dimension of $A \otimes B$ is $d_A d_B$.*

This fact is referred to as information locality by Hardy[14] for whom this proposition was itself proffered as a foundational axiom. It can be phrased as the fact that if one performs a maximal measurements on the components on the component od a composite one has performed one on the whole; the information was in this way locally obtainable and there is no necessity for

global operations. In a rough sense this comes back to the nature of what one can know of a system - and if one knows of all its components, one knows the state in its entirety. In this way the information within the parts are coherent with the whole.

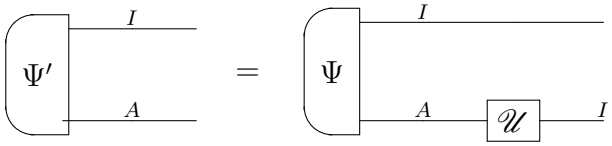
5.4.2 Uniqueness of the invariant state and the equilibrium condition

The fundamentals properties of sharp theories with purification outlined are sufficient to show that these theories satisfy the thwo requirements set out above. Recall that these were motivated from a desire to attain thermodynamic properties, in particular, those related, in a sense, to entropic value and the 'worthlessness' of the invariant state. Three propositions will suffice.

The first requirement was that every finite system has one unique deterministic invariant state, and indeed the following proposition is to hand:

Proposition 5.4.8. *For every system A and every pair of pure states α,$\alpha' \in PurSt(A)$ there exists a reversible transformation $\mathscr{U}$ such that $\alpha' = \mathscr{U}\,\alpha$*

Proof. [4] All systems are purifying systems for the trivial system, this is can be seen if one admits a trivial system wire, and one makes use of the purification axiom to show that



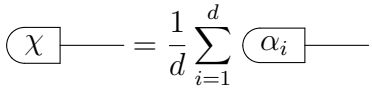
where Ψ and Ψ' are pure states. This is identical to

□

This tells us that if the invariant state exists, it must be unique and so the first requirement is satisfied.

By considering the diagonalisation of the invariant state for a single system and the composite one can show requirement two - that the invariant state of a composite system is the product of the invariant states of its components - is also satisfied.

Proposition 5.4.9. *For every maximal set of perfectly distinguishable pure states* $\{\alpha_i\}_{i=1}^{d_A}$ *one has the expression*



(5.9)

Proof. For a proof see reference [7] □

Through this decomposition of the invariant state one can consider the diagonalisation of the a composite state versus those of its components.

Proposition 5.4.10. *For every pair of systems* A *and* B*, one has* $\chi_{AB} = \chi_A \otimes \chi_B$

Proof. Suppose one has two maximal sets of perfectly distinguishable states for A and B $\{\alpha_i\}_{i=1}^{d_A}$ and $\{\beta_j\}_{j=1}^{d_B}$. The product set $\{\alpha_i \otimes \beta_j\}_{i\in\{1,\ldots,d_A\},j\in\{1,\ldots,d_B\}}$ is maximal for the composite system $A \otimes B$, by proposition 5.4.9, one obtains

$$\begin{aligned}\chi_{AB} &= \frac{1}{d_{AB}} \sum_{i=1}^{d_A}\sum_{j=1}^{d_B} \alpha_i \otimes \beta_j \\ &= \frac{1}{d_A d_B}\left(\sum_{i=1}^{d_A} \alpha_i\right) \otimes \left(\sum_{j=1}^{d_A} \beta_j\right) \\ &= \chi_A \otimes \chi_B\end{aligned}$$

where $d_{AB} = d_A d_B$ by information locality. □

In conclusion then, sharp theories with purification display properties consistent with the micro-canonical equilibrium in quantum thermodynamics.

5.5 State convertibility and Majorisation [8]

5.5.1 Introduction to Majorisation and its Necessity for state Conversion

The intuition behind majorisation is that it shows when one state is 'more mixed' than another. This is actually clearest when it is framed in terms of an equivalent statement[28] - suppose we say $\mathbf{p}$ majorises $\mathbf{q}$, setting aside this is not yet defined - this is equivalent to having D, a doubly stochastic matrix, where $D\mathbf{p} = \mathbf{q}$. If one considers the action of a such a matrix one notes that each component of $\mathbf{q}$ is an average over the elements of $\mathbf{p}$. It is in this way that one is more mixed than the other. The natural question at this point is if the equivalent statement is intuitively helpful for our purposes why consider the initial one at all? In practical terms Majorisation is simpler to check than attempting to find a suitable doubly stochastic matrix. The formal statement of majorisation is as follows:

Definition 58. *Let* $\mathbf{p}$ *and* $\mathbf{q}$ *be two vectors in* $\mathbb{R}^d$*, with the* components arranged in non-increasing order*. One says that* $\mathbf{p}$ *majorises* $\mathbf{q}$*, denoted* $\mathbf{p} \succeq \mathbf{q}$*, if*

$$\sum_{i=1}^{k} p_i \geq \sum_{i=1}^{k} q_i, \ \forall \ k < d \ \textit{and} \ \sum_{i=1}^{k} p_i \geq \sum_{i=1}^{k} q_i \tag{5.10}$$

The next step is to consider how majorisation relates to the theories of purity outlined above. For this to be possible some new notation is required.

Definition 59. *Let* ρ *and* σ *be two normalised states of the same system* A*, and let* F *be one of the sets RaRe, Noisy, and Unital. We say that* ρ *is purer than* σ *relative to the set* F*, denoted as* $\rho \succeq_F \sigma$*, if there exists a channel* $\mathscr{C} \in F$ *such that* $\mathscr{C}\rho = \sigma$*.*

If ρ *is purer than* σ*, and* σ *is purer than* ρ*, we say that they are equally pure relative to the set* F*, denoted* $\rho \simeq_F \sigma$

One should note that the inclusions outlined in subsection 5.2.3 implies relationships between the statements of purity as prescribed by the eponymous resource theories. In particular

$$\rho \succeq_{RaRe} \sigma \quad \Longrightarrow \quad \rho \succeq_{Unital} \sigma \quad \forall \sigma, \rho \in St(A)$$

and

$$\rho \succeq_{Noisy} \sigma \quad \Longrightarrow \quad \rho \succeq_{Unital} \sigma \quad \forall \sigma, \rho \in St(A)$$

for some any system A. In this way one can see unitality is the weakest relation. If one could show that majorisation is weaker still than unitality then majorisability would be a necessary condition for all the above purity relations; indeed showing this is the intention for the rest of this section.

Proposition 5.5.1. *Let ρ and σ be a normalised states, and let p and $\mathbf{q}$ be the vectors of their eigenvalues arranged in non-decreasing order. Then one has that ρ can be converted into σ by a unital channel iff $\mathbf{p}$ majorises $\mathbf{q}$. In formula*

$$\rho \succeq_{Unital} \sigma \quad \iff \quad \mathbf{p} \succeq \mathbf{q} \tag{5.11}$$

Proof. Let $\rho = \sum_{j=1}^{d} p_j \alpha_j$ and $\sigma = \sum_{j=1}^{d} q_j \alpha'_j$ be diagonalisations of ρ and σ, respectively. The first part of the proof is to show $\rho \succeq_{unital} \sigma$ implies $\mathbf{p} \succeq \mathbf{q}$. Suppose one has $\sigma = D\rho$ where D is a unital channel. This implies

$$\sum_{j=1}^{d} q_j \alpha'_j = \sum_{j=1}^{d} p_j D\alpha_j \tag{5.12}$$

Applying $\alpha_j'^{\dagger}$ to both sides, one obtains

$$q_i = \sum_{j=1}^{d} p_j \tag{5.13}$$

$$= \sum_{j=1}^{d} D_{ij} p_j \quad \text{where} \quad D_{ij} := \left(\alpha_i'^{\dagger} | D | \alpha_j\right). \tag{5.14}$$

The terms D_{ij} are the entries of a doubly stochastic by lemma A.0.1. This implies the above equation shows $\mathbf{p}$ majorises $\mathbf{q}$.

Going the other way suppose $\mathbf{p} \succeq \mathbf{q}$ and let D be a doubly stochastic matrix such that $\mathbf{q} = D\mathbf{p}$. Now let

$$D = \sum_{j=1}^{d} \rho_j \alpha_j^{\dagger} \tag{5.15}$$

$$\rho_j := \sum_{i=1}^{d} D_{ij} \alpha'_i. \tag{5.16}$$

By construction, one has

$$\begin{aligned} D\rho &= \sum_{j=1}^{d} \rho_j \left(\alpha_j^{\dagger}\right) \\ &= \sum_{i=1}^{d} \alpha'_i \sum_{j=1}^{d} D_{ij} p_j \\ &= \sum_{i=1}^{d} q_i \alpha'_i = \sigma \end{aligned}$$

As the channel D is unital by lemma A.0.2 ρ can be converted into σ by a unital channel. □

Due to the inclusion of the RaRe channels and noisy operations in the set of unital channels one can conclude the following:

Corollary 5.5.1.1. *Majorisation is a necessary condition for convertibility in the three theories of purity given here.*

There is a second important corollary of the proposition as majorisation provides one with an alternative way to see the uniqueness of the spectrum

Corollary 5.5.1.2. *Let $\rho = \sum_{i=1}^{d} p_i \alpha_i$ and $\rho = \sum_{i=1}^{d} p'_i \alpha'_i$ be two diagonalisations of the same state with the eigenvalues p and p' arranged in non-decreasing order. Then $p = p'$*

In this way majorisation establishes a preorder on the set of states - which as was the stated aim - is the weakest such ordering as it is implied by unitality and thereof the others. As a final corollary an important observation one can make at this point is that majorisation allows one to immediately see that pure states are a currency.

Corollary 5.5.1.3. *A pure state π acts as a currency in sharp theories with purification. Firstly for any other pure state π', its spectrum is trivially majorised by π. This implies by that all pure states majorise one another. Moreover any other state can be reached as a pure state necessarily majorises their eigenvalues.*

5.6 Purity Monotones [8]

Definition 60. *A Schur-convex function is a function* $f : \mathbb{R}^d \to \mathbb{R}$ *such that for all* $\mathbf{p}, \mathbf{q} \in \mathbb{R}$ *such that* $\mathbf{p}$ *majorises* $\mathbf{q}$*, one has* $f(\mathbf{p}) \leq f(\mathbf{q})$

Definition 61. *A monotone under the free operations* F *for a system* A *is a function* $f_F : St_1(A) \to \mathbb{R}$ *satisfying the condition*

$$f_F(\rho) \geq f_F(\sigma) \quad \forall \rho\sigma \in St_1(A), \rho \succeq_F \sigma \tag{5.17}$$

In sharp theories with purification one can talk of unital monotones. Written in terms of majorisation this gives that

$$f_{Unital}(\rho) \geq f_{Unital}(\sigma) \quad \forall \rho\sigma \in St_1(A), \mathbf{p} \succeq_{Unital} \mathbf{q} \tag{5.18}$$

where $\mathbf{p}$ and $\mathbf{q}$ are the respective eigenvalue spectrum's of ρ and σ.

Proposition 5.6.1. *[9] A function on the state* $f_{Unital} : St_1(A) \to \mathbb{R}$ *is a unital purity monotone iff there exists a schur-convex function* $f_s : \mathbb{R}^{d_A} \to \mathbb{R}$ *such that* $f_{Unital}(\rho) = f_s(\mathbf{p})$*, where* $\mathbf{p}$ *are the eigenvalues of* ρ

Proof. Let f_{Unital} be a unital purity monotone. Lemma 5.5.1 shows that majorisation, which depends only on eigenvalues, is a necessary and sufficient condition for unital transformations. This indicates unital conversion depends only on the eigenvalues, and in turn, the monotones too must only depend only on eigenvalues. Specifically $f_{Unital}(\rho)$ depends only on the eigenvalues of ρ, and not on its eigenstates.

This indicates there must exist a function $f_s : \mathbb{R}^{d_A} \to \mathbb{R}$ such that $f_{Unital}(\rho) = f_s(\mathbf{p})$, for every state ρ. Suppose now that $\mathbf{p}$ and $\mathbf{q}$ are two probability distributions satisfying $\mathbf{p} \succeq \mathbf{q}$. Then the sufficiency of majorisation criterion implies there exists a RaRe channel, which is itself a unitary channel, transforming $\rho = \sum_{i=1}^{d} p_i \alpha_i$ into the state $\sigma = \sum_{i=1}^{d} q_i \alpha_i$, for every pure maximal set $\{\alpha_i\}_{i=1}^{d}$. As a result, we obtain the relation

$$f_s(\mathbf{p}) = f_{Unital}(\rho) \geq f_{Unital}(\sigma) = f_s(\mathbf{q})$$

which proves that f_s is schur-convex.

Conversely, given a function schur-convex function f_s one can define a function $\tilde{f}_{Unital}$ on the resources, as $\tilde{f}_{Unital}(\rho) := f_s(\mathbf{p})$, where $\mathbf{p}$ is the spectrum ρ. As f_s is schur-convex this implies if $f_s(\mathbf{p}) \geq f_s(\mathbf{q})$ then $\mathbf{p}$ must majorise $\mathbf{q}$, by lemma 5.5.1 this then implies that there is a unital conversion from a state with spectrum $\mathbf{p}$ to a state with spectrum $\mathbf{q}$. □

The importance of such functions is that they encapsulate the notion of entropy in its various technical incarnations.

Consider the negative shannon entropy

$$H(\mathbf{p}) := -\sum_{i=1}^{d} p_i log p_i \tag{5.19}$$

The purity monotone that corresponds to this is the Von-Neumann entropy

$$S(\rho) := H(\mathbf{p}) \tag{5.20}$$

Further examples are the negative Rényi entropies[31]:

Definition 62. *The negative Rényi entropy of order alpha, where $\alpha \geq 0$ and $\alpha \neq 1$, is defined as*

$$H_\alpha(X) = -\frac{1}{1-\alpha} log_2 \left(\sum_{i=1}^{n} p_i^\alpha \right) \tag{5.21}$$

where X is a discrete random variable with outcomes in the set $\{1, 2, ..., n\}$ and associated probabilities $p_i = P(X = i)$ for $i = 1, ..., n$.

Corollary 5.6.1.1. *The existence of Rényi entropies for purity theories away from the asymptotic limit[13][7] imposes immediate restrictions on the theory of purity in the non asymptotic regime. Theorem 4.3.2 shows that for all reversible theories all entropies are unique up to a constant - Rényi entropies of different values of α are* not *equivalent up to a constant - which implies the theory of purity for exact conversions is* not *reversible.*

5.7 Asymptotic Reversibility In Sharp Theories with Purification

In this section a generalised version of the proof given in [17] will presented, applicable to all sharp theories with purification, of which quantum theory is one possibility.

The techniques applied in this section differs from those up until this point. It will be profitable to consider the state ρ partitioned via the diagonalisation seen in theorem 5.4.3, we have a set of eigenstates and their eigenvalues. If we now consider the tensor state $\rho^{\otimes n}$, and its diagonalisation in terms of its composite ρ, each eigenvector is a tensor product of one of the eigenvectors from each ρ. In this way the total set of eigenvectors of $\rho^{\otimes n}$ give all possible arrangements of the eigenvectors. The eigenvalues of these arrangements are the respective multiplied probabilities of the composite eigenvectors. For example consider $\rho = p_1\alpha_1 + p_2\alpha_2 + p_3\alpha_3$ where α_i are the eigenvectors and the $p_i \in [0, 1]$ are the eigenvalues whose sum equals unity. Then we can see that

$$\begin{aligned}\rho^{\otimes 2} = &(p_1p_1)\alpha_1 \otimes \alpha_1 + (p_2p_1)\alpha_2 \otimes \alpha_1 + (p_3p_1)\alpha_3 \otimes \alpha_1 \\ &+ (p_2p_1)\alpha_2 \otimes \alpha_1 + (p_2p_2)\alpha_2 \otimes \alpha_2 + (p_2p_3)\alpha_2 \otimes \alpha_3 \\ &+ (p_3p_1)\alpha_3 \otimes \alpha_1 + (p_3p_2)\alpha_3 \otimes \alpha_2 + (p_3p_3)\alpha_3 \otimes \alpha_3\end{aligned}$$

where we have avoided diagrammatic notation for compactness. One can view these eigenvectors as strings of symbols, where the eigenvectors of ρ are the symbols and consider the associated P_i as the probability associated with each of these symbols. In this way one can identify *typical* strings that - in the limit of large string length, or equivalently, many states - dominate in likelihood.

One should note that in the following proof the details of typicality and its formal underpinnings are outlined in appendix C and the associated proofs found in [27], though for clarity the basics will be discussed.

Lemma 5.7.1. *In sharp theories with purification there exists a protocol $\bar{\mathscr{F}}$ that allows one to transition from $\rho^{\otimes n}$ to $\pi^{n(1-S)}$, where π is a pure state, and there also exists an inverse protocol $\bar{\mathscr{F}}'$ from $\pi^{n(1-S)}$ to $\rho^{\otimes n}$. These are such that their rates multiplied gives unity.*

Proof. Consider $\rho^{\otimes n}$, where each ρ is of dimension d, with eigenvalues p_i^n, where the n labels exist to differentiate the term from the p_i^n eigenvalues of ρ that compose the state. Given $1 > \epsilon > 0$ and $\delta > 0$[3], there exists $n \in \mathbb{N}$ and the set of T eigenvalues such that

$$\sum_{P_i^n \in T} P_i^n \geq 1 - \epsilon, \tag{5.22}$$

indicating in the limit of large numbers these typical states are the support of the probability distribution. It can also be shown [30] that the probabilities of each typical string are bounded by

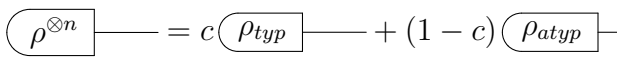
$$d^{-n(S+\delta)} \leq P_i^n \leq d^{-n(S-\delta)} \text{ for } p_i^n \in T. \tag{5.23}$$

where S is the Shannon entropy of the spectrum of ρ. This shows the typical states probability distribution is uniform in the limit.

The reason for this is that one can view the spectrum of the states $\rho^{\otimes n}$, as was discussed above, as a collection of all the possible strings formed of the perfectly distinguishable eigenstates - with an attending probability created by the multiplication of the composite eigenvalues. In line with the weak law of large numbers, one expects that in the limit of large n, the expected or *typical* strings making up set T would come to dominate in probabilistic likelihood. This motivates the idea of partitioning the state $\rho^{\otimes n}$ into a combination of typical and an atypical state, this can be seen diagrammatically by

(5.24)

where c is the normalisation given by $\sum_{p_i^n \in T} P_i$ and α_i^n are the perfectly distinguishable eigenstates from the diagonalisation of $\rho^{\otimes n}$ itself,*not those of the diagonalisation of ρ.* The original state then can be written in terms of these typical and atypical states



(5.25)

[3]In fact $\epsilon, \delta > 0$ is all that is required to produce typical states here. $1 > \epsilon$ is required for other aspects of our proof, however as ϵ is introduced in the argument with an aim to make it arbitrarily small this condition is merely a technicality.

Recall that c is bounded as in equation 5.22, this implies that by considering the operational metric 2.20 one can write

$$||\rho_{typ} - \rho^{\otimes n}|| \leq ||(1-c)(\rho_{typ} + \rho_{typ})|| \leq \epsilon||\rho_{typ} + \rho_{typ}|| \leq 2\epsilon, \tag{5.26}$$

as ϵ can be made arbitrarily close to zero this implies we can use ρ_{typ} to approximate $\rho^{\otimes n}$ to arbitrary accuracy in the limit.

The advantage of this comes from the fact one can construct a protocol that allows one to transition from the the typical state to approximately $n(N-S)$ pure qubits. To see this observe that the eigenvalues of ρ_{typ} denoted λ_i satisfy

$$\lambda_i \equiv \frac{P_i}{c} \geq \frac{1}{c} d^{-n(S+\delta)} > d^{n(S+\delta)}. \tag{5.27}$$

Consider a state ρ_{out} with eigenvalues

$$\{\underbrace{\frac{1}{D}, ..., \frac{1}{D}}_{D}, \underbrace{0, ..., 0}_{d^n - D}\} \text{ where } D = \left\lceil \frac{1}{d^{-n(S+\delta)}} \right\rceil. \tag{5.28}$$

such a state is reachable from ρ_{typ} via free processes through a noisy operation, which is unital by lemma 5.7.

Observe that the eigenvalues λ_i are all smaller than all the eigenvalues of ρ_{typ}. This combined with the fact that, by definition, both sets of eigenvalues sum to unity means that the spectrum of ρ_{typ} majorises the spectrum of this new state ρ_{out}; lemma 5.5.1 then proves that there is a unital transformation from the former to the latter.

Note that the fact the λ_i sum to unity also gives us the number of non zero eigenvalues in ρ_{out} which must be the inverse of the value of the eigenvalues, namely D.

Note now that we can use some other $\tilde{D} > D$ in the eigenvalues of ρ_{out} and the transition is still possible to such a state, as this merely diminishes the eigenvalues of ρ_{out} further against those of ρ_{typ}. In particular if one lets $\tilde{D}$ be larger than D and such that it is a power d then one finds that

$$\tilde{D} = \left\lceil d^{m(n(S+\delta))} \right\rceil = d^{\lceil m(n(S+\delta)) \rceil} \tag{5.29}$$

where $m \in \mathbb{N}$. Let us presume our $\tilde{D}$ is the smallest such term, that is to say $m = 1$. The result gives us the following formula

$$\log_d \tilde{D} = \lceil n(S+\delta) \rceil \leq n(S+\delta) + 1. \tag{5.30}$$

With this new $\tilde{D}$ the state ρ_{out} represents the tensor product of $\log_d \tilde{D}$ qubits in maximally mixed state and $n \log_d d - \log_d \tilde{D} = n - \log_d \tilde{D} \geq n(1 - S - \delta) - 1$ qubits in pure states.

Recall that noisy operations are unital and so one can use the free discarding operation to throw out the mixed qubits and retain the pure. This process will be referred to as protocol $\bar{\mathscr{F}}$, and the new state is denoted π_{out}. The rate of transition is the number of obtained pure states divided by n[4]

$$1 - S - \delta + \frac{1}{n} \tag{5.31}$$

In the limit of large n this becomes

$$1 - S - \delta \tag{5.32}$$

As δ can be made arbitrarily small, one obtains the final asymptotic rate

$$1 - S. \tag{5.33}$$

This however, is the transition from ρ_{typ}, one must apply this protocol $\bar{\mathscr{F}}$ with operations $\{\mathscr{F}_n\}_{n\in\mathbb{N}}$ to the state $\rho^{\otimes n}$. Using monotonicity of the operational norm[4] one can see that

$$||\mathscr{F}_n(\rho^{\otimes n}) - \pi_{out}|| = ||\mathscr{F}_n(\rho^{\otimes n}) - \mathscr{F}_n(\rho_{typ})|| \leq ||\rho^{\otimes n} - \rho_{typ}|| \leq \epsilon \tag{5.34}$$

In this way one shows that to arbitrary precision one can obtain π_{out} from $\rho^{\otimes n}$. In summary from an initial state $\rho^{\otimes n}$ one can obtain, via free unital processes, $n(1 - S)$ pure states at a rate of $1 - S$ per input state.

By showing the converse it becomes clear that this process is reversible. In order to achieve this the aim is to find a protocol from π_{out} to $\rho^{\otimes n}$. To begin one takes the state π_{out} which contains $1 - S$ pure qubits per copy of ρ.

Consider another state state $\tilde{\rho}_{out}$ is of the form

$$\{\underbrace{\frac{1}{D'}, ..., \frac{1}{D'}}_{D'}, \underbrace{0, ..., 0}_{d^n - D'}\} \text{ where } D' = \left\lfloor \frac{1}{d^{-n(S-\delta)}} \right\rfloor \tag{5.35}$$

notice that by equation 5.23 we have that

$$\lambda_i \equiv \frac{P_i}{c} \leq \frac{1}{c} d^{-n(S-\delta)} < d^{n(S-\delta)}. \tag{5.36}$$

[4]This division implicitly separates the obtained rate from any dependence on the base, here d, used in the typicality relationsC. The number of input copies is logarithmic in the size of the input vector so indeed this removes the particular dependence as all logarithms of different base are multiplicatively related.

recalling that λ_i are the eigenvalues of ρ_{typ}.

As the eigenvalues of $\tilde{\rho}_{out}$ majorises ρ_{typ} one can then transition from this state to ρ_{typ} by a unital transformation.

This will also be true for any $\tilde{D}'$ such that $\tilde{D}' < D$. In addition let $\tilde{D}'$ be a power of d. One then has that the state $\tilde{\rho}_{out}$ is a product of $\tilde{D}'$ qubits in the maximally mixed state and $n(1-S)$ pure state qubits.

Therefore one can transition from $n(1-S)$ pure states by adding $log_d\tilde{D}'$ mixed state qubits, which is a free operation, and then transitioning to the state ρ_{typ}, by a unital transition, which in the limit of large n can then be made arbitrarily close to $\rho^{\otimes n}$. □

Proposition 5.7.2. *The protocols given in proposition 5.7.1 achieve the optimal rate.*

Proof. Firstly remember that recall that by corollary 5.5.1.3 pure states are a currency, and that from proposition 4.2.2 currency cost and distillation rates are bounded as follows:

$$1 \geq R_C(\rho) \geq R_D(\rho) \geq 0 \tag{5.37}$$

This proof will demonstrate a contradiction with the above equation by supposing the rates given in proposition 5.7.1 are not optimal.

Take the initial protocol $\bar{\mathscr{F}}$ which represents the $\pi^{n(1-S)} \to \rho^{\otimes n}$, suppose one can achieve a cost transition more optimal than this, suppose one can achieve this rate with one less pure state $n(1-S)-1$, the smallest change possible. Label this new protocol $\bar{\mathscr{F}}'$. Then by considering these in series one can go from $\pi^{n(1-S)-1}$ to $\rho^{\otimes n}$ and then to $\pi^{n(1-S)-1}$ which indicates one can generate a states for free.

The implications of this are that the optimal distillation rate $R_D(\rho)$ can be made arbitrarily large. This is because any $\rho^{\otimes n}$ can generate any number of states by first moving to $n(1-S)$ pure states and then generating the state $\rho^{\otimes(n+1)}$ which in turn can generate $n(1-S)+1$ pure states from which the cycle can continue. Therefore the supremum of the ratio $\frac{m_n}{n}$ for transition $\rho^{\otimes n} \to \pi^{m_n}$ is unbounded.

Conversely given any state $\pi^{\otimes n(1-S)}$ of $n(1-S)$ tensored pure states the cost rate can be made arbitrarily close to zero. As $\pi^{\otimes n(1-S)}$ can be converted into $\rho^{\otimes(n+1)}$ which in turn can be converted into $\pi^{\otimes n(1-S)+1}$ which in turn can be converted into $\rho^{\otimes(n+2)}$ and so on. Thus Therefore

the infimum of the ratio $\frac{m_n}{n}$ for transition $\pi^{\otimes m_n} \to \rho^{\otimes n}$ can be made arbitrarily close to zero as $n \to \infty$ when one repeatedly extracts more qubits from which to generate larger quantities of ρ. The important thing to take from this is that the relationship between currency cost and distillation rates are

$$R_C(\rho) < R_D(\rho). \tag{5.38}$$

as pure states are a currency this is a direct contradiction of proposition 4.2.2. □

Theorem 5.7.3. *Sharp Theories with Purification are reversible in the asymptotic limit.*

Proof. Proposition 5.7.1 shows that in the asymptotic limit, for sharp theories with purification, there is a protocol such that one can transition to a product of pure states and a second protocol back to the original state. Proposition 5.7.2 shows the rates of said transitions are optimal.

Therefore the optimal distillation rate of going from some tensored state $\rho^{\otimes n}$ to $n(1-S)$ pure states is $1-S$, and the optimal cost rate going the other way, from $n(1-S)$ pure states to $\rho^{\otimes n}$, is the same. Lemma 4.3.1 then shows that this theory must be reversible. □

The importance of this result is that as sharp theories with purification are reversible in the thermodynamic limit theorem 4.3.2 shows all entropies in this theory are unique up to a constant. To summarise this means that as far as macroscopic thermodynamics is concerned there is only one entropy. This grants a certain coherence to the resource theory position, in a similar manner to Ehrenfest's theorem in quantum mechanics - where the expectation value of a quantum system follows Newtonian mechanics - in the asymptotic limit sharp theories with purification , based on resource theories, give the expected single macroscopic thermodynamic entropy.

6
Conclusion

What has been achieved here is focussed on the study of rates given in chapter 4 where new results regarding the transitions in a resource theory are presented in detail. Firstly it was shown how there are essentially two key types of optimal rates: the cost rate, where one tries to minimise how much one is spending to get the output; and the distillation rate, where one is trying to maximise the output. These rates were then shown to be inverses.

Montonic functions were then introduced as a means to assign value to different resources. In order to be meaningful in the limit however it was necessary to define a regularised montonic function. These functions act as a value densities which are appropriate where other monotonic functions may be unbounded, specifically in the thermodynamic limit of many resources. With the value density defined this author proved that there is a clear relationship between the monotonic value densities assigned to the resources through regularisable monotones and the optimal distillation rate between these states 4.1.7

$$f^{\infty}(\rho) \geq f^{\infty}(\sigma)R(\rho \to \sigma)$$

Where $R(\rho \to \sigma)$ is the distillation rate from a resource ρ to a resource σ and f^{∞} is the regularisable monotone sending these stares to the real numbers. What one is seeing here is the the value density of the original resource is greater than or equal to that of the output state times the rate at which this state can be extracted. This makes sense as the converse

would imply that per input one was extracting more value than one put in. This result was a generalisation, and the proof a correction, of Theorem 4 in [16].

Following this currencies were discussed as a resource which could transition to all other states. This was of particular importance for the later discussion of the study of quantum thermodynamics and purity theory, where pure states constitute a currency.

Reversibility in a resource theory was then discussed. This was defined as a resource theory where all the distillation rates from one state to another multiplied by the equivalent rate for the return trip gives unity. This is to say nothing is lost in transitioning there and back. The key result in this section was showing that for all such resource theories with a currency, all regularised monotones are unique up to a multiplicative constant. In a sense, there is only one method of evaluating value density in such theories. Formally this author showed that in a resource theory D for two regularisable monotonic functions $f_1, f_2 : |D| \to \mathbb{R}$

$$f_1^\infty(\rho) = a f_2^\infty(\rho) \quad where \quad a = \frac{f_1^\infty(c)}{f_2^\infty(c)}$$

where c is the currency resource and f_1^∞ and f_1^∞ are the regularised functions.

Following this, asymptotic conversion was introduced in an abstract manner, here the set of possible transitions was extended to include those that only exist in the limit. It was shown that for a regularisable monotonic function f that is asymptomatically continuous and where one has a transition between two states ρ and σ where $R_D(\rho \to \sigma) < \infty$, and there exists a state θ such that $f^\infty(\theta) > 0$ then

$$R(\rho \to \sigma) = \frac{f^\infty(\rho)}{f^\infty(\sigma)}$$

This authors result was a generalisation of the Work of Horodecki and Oppenheim in [18]. This result shows that in such theories the value densities of two states exactly dictate the optimum transition rate between them: the values in the limit completely characters the optimal rate of exchange.

The above result was then used in the context of the resource theory of purity. In particular this author showed that sharp theories with purification are asymptotically reversible in terms of the resource theory of purity. This result revealed that entropies in such theories are indeed compatible with thermodynamics, as in the limit their reversibility implies that up to constant

there is only one entropy function, just as in thermodynamics. This is due to the earlier result regarding the uniqueness of reversible theories regularised monotones. As entropy functions are an example of regularised monotonic functions one then knows they are unique up to a constant .

Following on from this work here, an immediate extension would be to consider transitions with error. This is a very natural physical situation, in practical experiments it is impossible to guarantee the fidelity of ones transition protocol. Suppose for instance that one only has an imprecise knowledge of what resource is the result of a transition. This could be expressed by saying that when transitioning from a resource $\rho^{\otimes n}$ and attempting to obtain a resource $\sigma^{\otimes m_n}$ what one actually gets is σ_n* such that $|\sigma^{\otimes m_n} - \sigma_n * | \leq \delta$ where δ would quantify the degree of error according to some metric. One would then seek to relate this transition to the values given by monotonic functions which would require the introduction of a kind of approximate continuity which would link the fixed error in the resources to a fixed difference in the error on their value. It is likely one would need this continuity to also preserve the relative size of errors - resources obtained with a smaller error margin than others should equivalently have a smaller error margin on their value.

A more exotic extension could be to take the rates work studies here into the causaloid framework which was devised to introduce the agent-centric approach to quantum gravity[15]. One should note that the rates material had no explicit need for causal structure. The idea of moving from one state to another need not be confined to a time-line. Conceptually one can view the causaloid structure as formed of devices where each use results in a card with three pieces of information on it; where the measurement is made in space-time, what is measured, and what the result of the measurement is. It has been shown in [25] that through causaloids one can conceive of a causally unbiased shannon-entropy though the mathematical properties of this entropy were unclear. It may prove fruitful to see how the relations between value and rates of exchange function in this causally peculiar situation.

Appendices

The law that entropy always increases holds, I think, the supreme position among the laws of Nature. If someone points out to you that your pet theory of the universe is in disagreement with Maxwell's equations — then so much the worse for Maxwell's equations.

If it is found to be contradicted by observation — well, these experimentalists do bungle things sometimes.

But if your theory is found to be against the second law of thermodynamics I can give you no hope; there is nothing for it but to collapse in deepest humiliation.

— Sir Arthur Stanley Eddington, The Nature of the Physical World (1927)

A

Unital channels and Doubly Stochastic Matrices

In a broad sense, unital channels are the generalisation of doubly stochastic matrices. This connection can be made more explicit

Lemma A.0.1. *Let $\mathscr{D}$ be a unital channel acting on system A and let $\{\alpha_i\}_{i=1}^d$ and $\{\alpha'_i\}_{i=1}^d$ be two pure maximal sets of system A. Then, the matrix D with entries*

$$D_{ij} := \left(\alpha_i^\dagger | \mathscr{D} | \alpha'_j\right) \tag{A.1}$$

is doubly stochastic.

Proof. [8] Every entry D_{ij} is a probability and therefore it is non-negative. Moreover, one has

$$\begin{aligned}\sum_{i=1}^d D_{ij} = \sum_{i=1}^d &= \sum_{i=1}^d \left(\alpha_i^\dagger | \mathscr{D} | \alpha'_j\right) \\ &= (u|\mathscr{D}|\alpha_j) \\ &= Tr[\alpha_j] \\ &= 1 \quad \forall j \in 1, ..., d\end{aligned}$$

having used the fact that the effects $\{\alpha_i'^{\dagger}\}$ form as observation-test and that $\mathscr{D}$ is a channel and therefore $u\mathscr{D} = u$ [4]. On the other hand, one has

$$\begin{aligned}\sum_{i=1}^{d} D_{ij} = \sum_{i=1}^{d} &= \sum_{i=1}^{d} \left(\alpha_i^{\dagger}|\mathscr{D}|\alpha_j'\right) \\ &= d\left(\alpha_i^{\dagger}|\mathscr{D}|\alpha_j'\right) \\ &= \left(\alpha_i^{\dagger}|\chi\right) \\ &= d \cdot \frac{1}{d} \\ &= 1 \qquad \forall j \in 1, ..., d\end{aligned}$$

Where one should remember unital channels by definition leave the invariant state unchanged, and we have used lemma A.0.2. The fact the summations over each variable shows that D is a doubly stochastic matrix. □

Lemma A.0.2. *Let D be a $d \times d$ doubly stochastic matrix and let $\{\alpha_i\}_{i=1}^{d}$ and $\{\alpha_i'\}_{i=1}^{d}$ be two pure maximal sets of system A. Then, the channel defined by*

$$D := \sum_{j=1}^{d} \rho_j \alpha_j^{\dagger} \quad \textit{with} \quad \rho_j := \sum_{i=1}^{d} D_{ij}\alpha_i'. \tag{A.2}$$

is unital.

Proof. [8] The transformation D is a channel of the measure and prepare form: it can be implemented by performing the observation test $\{\alpha_j^{\dagger}\}_{j=1}^{d}$ and by preparing the state ρ_j conditionally n outcome j. Moreover, one has

$$D\chi = \sum_{j=1}^{d} \rho_j \left(\alpha_j^{\dagger}|\chi\right) \tag{A.3}$$

$$= \frac{1}{d}\sum_{j=1}^{d}\sum_{i=1}^{d} D_{ij}\alpha_i' \tag{A.4}$$

$$= \frac{1}{d}\sum_{i=1}^{d} \alpha_i' \tag{A.5}$$

$$= \chi \tag{A.6}$$

the third equality follows from the definition of a doubly stochastic matrix, and the fourth from the diagonalisation of the invariant state χ in lemma A.0.2. □

B

Results from Real Analysis

Lemma B.0.1. *Supremum commute*

$$\sup_m \sup_n a_{n,m} = \sup_{n,m} a_{n,m} = \sup_n \sup_m a_{n,m}$$

where $a_{n,m}$ is considered a sequence over two indices.

Proof. Consider the left hand side, and suppose that

$$\sup_m \sup_n a_{n,m} < \sup_{n,m} a_{n,m} \tag{B.1}$$

This implies that there exists some $a_{\tilde{n},\tilde{m}}$ such that

$$\sup_m \sup_n a_{n,m} < a_{\tilde{n},\tilde{m}}$$

But by definition we have that

$$a_{\tilde{n},\tilde{m}} < \sup_n a_{n,\tilde{m}} < \sup_n \sup_m a_{n,m}$$

A direct contradiction. The inverse to equation B.1 would contradict the definition of supremum, therefore we have that

$$\sup_m \sup_n a_{n,m} = \sup_{n,m} a_{n,m}$$

The exact same argument applied to the inequality

$$\sup_n \sup_m a_{n,m} < \sup_{n,m} a_{n,m}$$

gives the right hand side equality. □

Lemma B.0.2. *Infimum commute*

$$\inf_m \inf_n a_{n,m} = \inf_{n,m} a_{n,m} = \inf_n \inf_m a_{n,m}$$

where $a_{n,m}$ is considered a sequence over two indices.

Proof. Consider the left hand side, and suppose that

$$\inf_m \inf_n a_{n,m} > \inf_{n,m} a_{n,m} \tag{B.2}$$

This implies that there exists some $a_{\tilde{n},\tilde{m}}$ such that

$$\inf_m \inf_n a_{n,m} > a_{\tilde{n},\tilde{m}}$$

But by definition we have that

$$a_{\tilde{n},\tilde{m}} > \inf_n a_{n,\tilde{m}} > \inf_n \inf_m a_{n,m}$$

A direct contradiction. The inverse to equation B.2 would contradict the definition of infimum, therefore we have that

$$\inf_m \inf_n a_{n,m} = \inf_{n,m} a_{n,m}$$

The exact same argument applied to the inequality

$$\inf_n \inf_m a_{n,m} > \inf_{n,m} a_{n,m}$$

gives the right hand side equality. □

B. Results from Real Analysis

Lemma B.0.3. *The supremum and the limit superior commute*

$$\sup_m \left(\limsup_{n\to\infty} a_{n,m} \right) = \limsup_{n\to\infty} \left(\sup_m a_{n,m} \right)$$

where $a_{n,m}$ is considered a sequence over two indices, where $n \in \mathbb{N}$ and m belongs to any, potentially uncountable, index set.

Proof. Let us write these the left hand and right hand side in terms of sub-sequential limits.

For the left hand side

$$\begin{aligned} \sup_m \left(\limsup_{n\to\infty} a_{n,m} \right) &= \sup_m \left(\sup_{a_m^\infty} a_m^\infty \right) \\ &= \sup_{m, a_m^\infty} a_m^\infty \end{aligned}$$

Where a_m^∞ is the limit set for each m. What one has done here is firstly taken the sequence over two indices $a_{n,m}$ and consider it as a set of, potentially uncountable, single index sequences, running over n, each labelled by a different index m, call them m-sequences. For each fixed m one takes the limit superior $\limsup_{n\to\infty} a_{n,m}$ which is to take the subsequence with the greatest limit, or equivalently take the largest accumulation point a_m^∞ of the sequence $a_{n,m}$, which is to say $\sup_{a_m^\infty} a_m^\infty$.As one does this *for each fixed* m-sequence one obtains a set of superior limits - one for each m.

Now consider the right hand side

$$\begin{aligned} \limsup_{n\to\infty} \left(\sup_m a_{n,m} \right) &= \limsup_{n\to\infty} a_{n,m^*} \\ &= \sup_{a_{m^*}^\infty} a_{n,m^*}^\infty \end{aligned}$$

Where a_{n,m^*} is the maximum $a_{n,m}$ for each n. We will call this the m-max series. In this case one takes each fixed n and for the series over m take the supremum. One now has for each n the largest associated m term; this is the m-max series a_{n,m^*}. One then takes the limit superior of this sequence.

The question is are these equations the same?

Suppose the superior limit of the m-max sequence was less than the largest point taken from all the limit points generated by each m-indexed series

$$\sup_{a^{\infty}_{m^*}} a^{\infty}_{m^*} \; < \; \sup_{a^{\infty}_{m},m} a^{\infty}_{m} \tag{B.3}$$

By definition then there must exist a point $\tilde{a}^{\infty}_{\tilde{m}}$ such that

$$\sup_{a^{\infty}_{m^*}} a^{\infty}_{m^*} \; < \; \tilde{a}^{\infty}_{\tilde{m}}$$

This is a limit point of one of the m-sequences which is larger than all the limit points of the m-max sequence.

However, by considerations regarding the supremum, it must also be the case that

$$\tilde{a}^{\infty}_{\tilde{m}} \; \leq \; \sup_{a^{\infty}_{\tilde{m}}} a^{\infty}_{\tilde{m}} \; \leq \; \sup_{a^{\infty}_{m},m} a^{\infty}_{m} \; = \; \sup_{a^{\infty}_{m}} a^{\infty}_{m^*}$$

What we outline here is that this point, that is limit point of some m-sequence, is by definition less than or equal to the supremum of the set of limit points of the m-sequence to which it belongs. The second relation then points out that this supremum, over the m-sequence limit points, is itself less than or equal to the supremum of the collection of all the largest limit points taken from each m-sequence. This final point is, by above arguments, equal to taking the largest m for each n and taking the superior limit of this sequence This implies a contradiction to equation B.3. The relationship must be a 'greater than or equal' to relation. However, the inverse to equation B.3 is a contradiction in terms - The supremum over both indices is by definition the largest point: The two must be equal. □

Lemma B.0.4. *The Infimum and the limit inferior commute*

$$\inf_{m} \left(\liminf_{n\to\infty} a_{n,m} \right) = \liminf_{n\to\infty} \left(\inf_{m} a_{n,m} \right)$$

where $a_{n,m}$ is considered a sequence over two indices, where $n \in \mathbb{N}$ and m belongs to any, potentially uncountable, index set.

Proof. This procedure is extremely similar to the above so the proof is presented without in depth discussion. Let us write these the left hand and right hand side in terms of sub-sequential limits.

For the left hand side

$$\begin{aligned}\inf_m \left(\liminf_{n\to\infty} a_{n,m} \right) &= \inf_m \left(\inf_{a_m^\infty} a_m^\infty \right) \\ &= \inf_{m, a_m^\infty} a_m^\infty\end{aligned}$$

Where a_m^∞ is the limit set for each m. Then consider the right hand side

$$\begin{aligned}\liminf_{n\to\infty} \left(\inf_m a_{n,m} \right) &= \liminf_{n\to\infty} a_{n,m^*} \\ &= \inf_{a_{m^*}^\infty} a_{n,m^*}\end{aligned}$$

Where a_{n,m^*} is the minimum $a_{n,m}$ for each n. Now suppose that

$$\inf_{a_{m^*}^\infty} a_{m^*}^\infty \; > \; \inf_{a_m^\infty, m} a_m^\infty \tag{B.4}$$

By definition then there exists a limit point $\tilde{a}_{\tilde{m}}^\infty$ such that

$$\inf_{a_{m^*}^\infty} a_{m^*}^\infty \; > \; \tilde{a}_{\tilde{m}}^\infty$$

However, by considerations regarding the infimum, it must also be the case that

$$\tilde{a}_{\tilde{m}}^\infty \; \geq \; \inf_{a_{\tilde{m}}^\infty} a_{\tilde{m}}^\infty \; \geq \; \inf_{a_m^\infty, m} a_m^\infty \; = \; \inf_{a_m^\infty} a_{m^*}^\infty$$

This is a contradiction. The inverse to equation B.4 is a contradiction in terms; the two must be equal.

□

C

Information Theory and Typicality

For further details and proofs see [27] and [30].

C.1 The weak law of large numbers

The practical use of statistics lies firmly on the fact that if enough trials occur statistical results should apply - nobody is shocked when twelve dice rolls to fail to give two sixes - one assumes something is wrong if twelve trillion rolls don't return around two trillion sixes. What one is implicitly assuming is that a sequence of independent events should converge in probability to an average.

Definition 63. *A sequence X_n of random variables* converges in probability *to X if, for any $\delta > 0$ one has*

$$\mathbb{P}\left(\|X_n - X| \leq \delta\right) \to 1 \tag{C.1}$$

as a sequence of real numbers.

. Essentially the probability X_n converges on X to within an error σ approaches 1.

The implication of this is that is one has some δ and ϵ such that $\delta, \epsilon > 0$, there exists some n_ϵ, such that for all $n \geq n_\epsilon$ one has

$$\mathbb{P}\left(\|X_n - X| \leq \delta\right) \to 1 - \epsilon \tag{C.2}$$

The statement of the weak law of large numbers is expressed in terms of convergence in probability.

Theorem C.1.1. *Let X_n be a sequence of independent and identically distributed random variables with finite expectation. Then*

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to \mathbb{E}[\mathbb{X}] \tag{C.3}$$

which states the average in converges to the expected value.

One can see that if X_n is a random variable counting the occurrences of a given outcome in the n^{th} round of some experiment then $\frac{1}{n}\sum_{i=1}^{n} X_i$ gives the relative frequency of that outcome after n rounds, while $\mathbb{E}[X]$ gives the probability of that same outcome.

In this context the weak law of large numbers is the statement that relative frequencies converge to probabilities "for n large enough". In general terms one can rephrase the weak law of large numbers a

$$\mathbb{P}\left(|statistical - probabilistic| \leq \delta\right) \to 1 \tag{C.4}$$

C.2 Typicality

Take the familiar case where Alice wishes to contact Bob. Suppose Alice has a source producing pure states $\delta_1, ..., \delta_d$ according to a probability distribution

$$\mathbf{P} = \begin{bmatrix} x_{p_1} \\ \vdots \\ x_{p_d} \end{bmatrix} \tag{C.5}$$

These states make up the message she wants to send to Bob. The aim is to send as few of these states as possible because Alice, much like anyone else, loves efficiency. In the classical case, knowing the pure state is equivalent to knowing the value od a random variable and the concept of weak typicality is useful.

Suppose the values of a random variable belong to a set $Var := x_1, ...x_d$. Consider a string formed of this set n letter long: $\mathbf{x} = \mathbf{x}_1 ... \mathbf{x}$

It is useful to consider the *sample entropy*

Definition 64. *Given an n-letter string* $\mathbf{x}$ *of symbols with probability* $\begin{bmatrix} x_{p_1} \\ \vdots \\ x_{p_d} \end{bmatrix}$*, the sample entropy is given by*

$$h(x) = -\frac{1}{n}\sum_{i=1}^{n} \log_a \mathbf{p}_i \tag{C.6}$$

where $a > 1$ *and* $\mathbf{p}_i$ *is the probability to produce the* i^{th} *letter in the string.*

The probabilistic equivalence of the quantity is the Shannon entropy

$$H(\mathbf{p}) = -\sum_{i=1}^{n} p_i \log_a p_i = \mathbb{E}[-log_a X] \tag{C.7}$$

then for any $\delta > 0$, one can define the weakly δ typical strings of n letters as follows.

Definition 65. *Given* $\delta > 0$ *a string* $\mathbf{x}$ *of n letters with probability distribution* $\mathbf{p}$ *is weakly* δ *typical, or* typical *if*

$$|h(\mathbf{x}) - H(\mathbf{p})| \leq \delta. \tag{C.8}$$

The set of weakly δ-typical sequences will be denoted by T_δ^n, while the set of all sequences by Var^n. The consequence of the weak law of large numbers is that "for n large enough" the probability of having a typical string is

$$\mathbb{P}(x \in T_\delta^n) \geq 1 - \epsilon \tag{C.9}$$

for any $\epsilon > 0$.

A related question is if one has a particular string $\mathbf{x} \in T_\delta^n$, what is the probability of this particular string?

Proposition C.2.1. *If* $\mathbf{x} \in T$ *one has*

$$a^{-n(H(p)+\delta)} \leq \mathbb{P}(\mathbf{x}) \leq a^{n(H(p)+\delta)} \tag{C.10}$$

for any $\delta > 0$*, and any* n*.*

As δ can be chosen to be very small, it becomes negligible if n is very large. This means that for large n, different typical strings have nearly the same probability.

How many typical strings are there? While being extremely likely, their cardinality is exponentially small compared to all possible strings.

Proposition C.2.2. *If "n is large enough" we have*

$$(1-\epsilon)d^{n(H(\mathbf{p}-\delta))} \leq |T_\delta^n| \leq d^{n(H(\mathbf{p}+\delta))} \tag{C.11}$$

for all $\epsilon, \delta > 0$.

The number of all possible strings is $|Var|^n = d^{nlog_a|Var|}$*,so*

$$\frac{typical}{all} = d^{n\delta} d^{n(H((\mathbf{p})|\mathbf{Var}|))}. \tag{C.12}$$

The argument of the round bracket is always negative because $H(\mathbf{p}) \leq log_a|Var|$ *with equality iff* $\mathbf{p}$ *is the uniform distribution on* Var*. If the distribution is non-uniform,* δ *can be chosen smaller than* $log_a|Var| - H(\mathbf{P})$*, so the ratio* $\frac{typical}{all}$ *vanishes. If the distribution is uniform, the cardinality of the typical set is comparable with the cardinality of the set of all strings.*

You can't win. You can't break even. You can't even get out of the game.

— Allen Ginsberg - The laws of Thermodynamics

Bibliography

[1] Samson Abramsky and Dominic Horsman. DEMONIC programming: a computational language for single-particle equilibrium thermodynamics, and its formal semantics. *Electronic Proceedings in Theoretical Computer Science*, 195(November):1–16, 2015.

[2] Charles H. Bennett. The thermodynamics of computation-a review. *International Journal of Theoretical Physics*, 21(12):905–940, 1982.

[3] Charles H. Bennett. Demons, Engines and the Second Law. *Scientific American*, 257(5):108–116, 1987.

[4] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti. Probabilistic theories with purification. *Physical Review A - Atomic, Molecular, and Optical Physics*, 81(6), 2010.

[5] Giulio Chiribella and Carlo Maria Scandolo. Entanglement and thermodynamics in general probabilistic theories. *New Journal of Physics*, 17(10), 2015.

[6] Giulio Chiribella and Carlo Maria Scandolo. Operational axioms for diagonalizing states. *Electronic Proceedings in Theoretical Computer Science*, 195(June):96–115, 2015.

[7] Giulio Chiribella and Carlo Maria Scandolo. Entanglement as an axiomatic foundation for statistical mechanics. *arXiv*, pages 1–29, 2016.

[8] Giulio Chiribella and Carlo Maria Scandolo. Microcanonical thermodynamics in general physical theories. pages 1–20, 2016.

[9] Giulio Chiribella and Carlo Maria Scandolo. Microcanonical thermodynamics in general physical theories. pages 1–20, 2016.

[10] Eric Chitambar, Debbie Leung, Laura Mančinska, Maris Ozols, and Andreas Winter. Everything You Always Wanted to Know About LOCC (But Were Afraid to Ask), 2014.

[11] Bob Coecke, Tobias Fritz, and Robert W. Spekkens. A mathematical theory of resources. *Information and Computation*, 250:59–86, 2016.

[12] Tobias Fritz. Resource convertibility and ordered commutative monoids. *Mathematical Structures in Computer Science*, 465(2108):1–89, 2015.

[13] Gilad Gour, Markus P. Müller, Varun Narasimhachar, Robert W. Spekkens, and Nicole Yunger Halpern. The resource theory of informational nonequilibrium in thermodynamics. *Physics Reports*, 583:1–58, 2015.

[14] Lucien Hardy. Quantum Theory From Five Reasonable Axioms. *arXiv*, page 34, 2001.

[15] Lucien Hardy. Probability Theories with Dynamic Causal Structure: A New Framework for Quantum Gravity. *arXiv*, page 68, 2005.

[16] Michal Horodecki. Entanglement Measures ¨. *QIC*, 1(1):3–26, 2001.

[17] Michał Horodecki, Paweł Horodecki, and Jonathan Oppenheim. Reversible transformations from pure to mixed states and the unique measure of information. *Physical Review A*, 67(6):10, 2003.

[18] Michał Horodecki and Jonathan Oppenheim. (Quantumness in the Context of) Resource Theories. *International Journal of Modern Physics B*, 27(01n03):1345019, 2012.

[19] Michał Horodecki and Jonathan Oppenheim. Fundamental limitations for quantum and nanoscale thermodynamics. *Nature Communications*, 4, 2013.

[20] Sang Wook Kim, Takahiro Sagawa, Simone De Liberato, and Masahito Ueda. Quantum szilard engine. *Physical Review Letters*, 106(7), 2011.

[21] Lea Kraemer and Lidia del Rio. Currencies in resource theories. page 13, 2016.

[22] James Ladyman and Katie Robertson. Going round in circles: Landauer vs. norton on the thermodynamics of computation. *Entropy*, 16(4):2278–2290, 2014.

[23] Ld Landau and Em Lifshitz. Statistical Physics, Part 1, 1980.

[24] Saunders Mac Lane. Categories for the Working Mathematician (BOOK), 1998.

[25] Sonia Markes and Lucien Hardy. Entropy for theories with indefinite causal structure. *Journal of Physics: Conference Series*, 306:012043, 2011.

[26] Ingo Muller. *A history of thermodynamics: The doctrine of energy and entropy*. 2007.

[27] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. 2010.

[28] Albert W. Marshall Olkin and Ingram. Inequalities: Theory of majorization and its applications. *Linear Algebra and Its Applications*, 199(SUPPL. 1):115–130, 1994.

[29] Asher Peres. Quantum Theory: Concepts and Methods. *New York*, page 464, 2002.

[30] John Preskill. Quantum Information and Computation. Lecture Notes. *Physics Today*, pages 24–30, 1998.

[31] Alfréd Rényi. On measures of entropy and information. *Fourth Berkeley Symposium on Mathematical Statistics and Probability*, 1(c):547–561, 1961.

[32] C. Yang and T. Lee. I Statistical Theory of Equations of State and Phase Transitions. I. Theory of Condensation, 1952.