Katholieke Universiteit
Leuven
Faculteit Wetenschappen
Instituut voor Theoretische
Fysica

Entropy, Information and Structure of Composite Quantum States

Milán Mosonyi

Promotor:
Prof. Dr. Mark Fannes
Prof. Dr. Dénes Petz

Proefschrift ingediend voor
het behalen van de graad van
Doctor in de Wetenschappen

Leuven 2005

Contents

Acknowledgements

I have started my PhD studies in 2000 in Budapest, Hungary, at the Mathematical Institute of the Budapest University of Technology and finished it in 2005 in Leuven, Belgium, at the Institute for Theoretical Physics at the Catholic University of Leuven. During these years I got so much help from so many people that it would be impossible to list them all in a few pages. Mentioning only a few, I would like to thank

my promotors, Prof. Dénes Petz and Prof. Mark Fannes for all their support and guidance;

my collaborators, Bart Haegeman, Caroline Vandenplas and Dimitri Vanpeteghem for the numerous hours of joyful discussion that deepened my understanding both in physics and mathematics;

Rob Heylen and Tim Jacobs for their patience and professionalism in solving my various problems with computers;

Joris Van den Bergh for his help in creating the Samenvatting of the thesis;

the members of the jury for carefully reading the manuscript and making useful comments on it;

the financial support of the grant G.0109.01N FWO-Vlaanderen that made my stay in Leuven possible;

all the members of I.T.F for creating a friendly, helpful and intellectually stimulating atmosphere that makes the Institute such an excellent place to be.

Chapter 1

Introduction

Statistical methods are used in physics because of incomplete knowledge of the observer about the system in study. Probability theory can then be used to describe the state or the evolution of the system. A slightly surprising fact is that in this framework a precise mathematical theory can be built around quantifying the lack of knowledge; this is information theory, originated by Shannon [101].

The key notion in information theory is the entropy of a finite probability distribution $\{p_i\}$, defined as

$$S(\{p_i\}) := -\sum_i p_i \log p_i \,. \tag{1.1}$$

The mathematical expression for the thermodynamic entropy of a quantum system in the state described by the density operator ρ was derived by John von Neumann in his fundamental work [78]. The form

$$S(\rho) = -\mathrm{Tr}\, \rho \log \rho \tag{1.2}$$

he obtained is equal to the Shannon entropy of the eigenvalue distribution of ρ, and gives a sharp lower bound on the Shannon entropies of possible measurement statistics, when projective measurements are performed on the system.

The equality of thermodynamical and information theoretic entropy is far from being a simple coincidence, and shows a deep relation between the two disciplines. Indeed, the theory of ensembles, for instance, can be built up by assuming maximal uncertainty about the state of the system, in the sense of maximizing informational entropy, conditioned on the available information on the system; the microcanonical (canonical, etc.) ensemble is obtained as the probability distribution that maximizes entropy under the constraint that the energy (expectation value of the energy, etc) is fixed. On the other hand, information doesn't exist independently of the medium carrying it ("Information is physical" as phrased by Rolf Landauer), hence the possible methods to manipulate, carry or to store information are governed by the laws of physics.

Information theory of quantum systems originates from the 70's (early works in the field are [41, 68, 45, 49, 50, 57, 52, 28]), and has gone through a rapid expansion in the last decade. Increasing interest is motivated by spectacular results such as quantum algorithms that work significantly faster than their classical equivalents [81], dense coding [13, 74], quantum teleportation [14], and quantum cryptography [11]. While theoretical results about quantum computation promise a breakthrough in computer science, the physical realization of a quantum computer seems to belong to the distant future. Quantum teleportation and quantum cryptography, on the other hand, are in the state of experimental realization [12, 20, 19, 39, 79], and seem to find applications in everyday commercial use. Apart from possible practical applications, the importance of quantum information theory lies in the renewed interest it induced in understanding the principles of quantum theory, and correlations in many-particle quantum systems.

The interpretation of the formalism of quantum mechanics is a matter of debate since the very beginning of its existence, and the question is still far from being settled [38]. It seems clear, however, that quantum physics is an intrinsically probabilistic theory, and the non-classical features of quantum systems are deeply related to the non-Kolmogorovian nature of the underlying probability theory. The main differences from the classical theory come from the significantly more difficult algebraic structure of quantum states (unlike in the classical case, it is not enough to give a set of weights to describe the state of a finite system, but one also has to specify a corresponding set of eigenvectors and this introduces a lot more degrees of freedom). This algebraic richness leads to the fact that a pure state of a composite quantum system is generally not a product of pure states of the subsystems, which is the simplest manifestation of the phenomenon of entanglement.

While entanglement often makes life difficult (or more interesting, as you wish), it is also shown to be the basic resource for the novel features of information processing that quantum information theory predicts. Moreover, experimental results show that entanglement influences macroscopically measurable quantities of a physical system, like magnetic susceptibility [40]. Hence a firm theoretical understanding of entanglement seems to be important in understanding the physical behavior of multi-component quantum systems. The first thing one can ask for is an appropriate quantification of entanglement. It turns out that rather simple criteria on a measure of entanglement lead to the fact that the only reasonable measure of entanglement for a pure bipartite state $|\psi\rangle$ is the entropy of its marginals:

$$\mathcal{E}(|\psi\rangle) = S\left(\mathrm{Tr}_2\ (|\psi\rangle\langle\psi|)\right) = S\left(\mathrm{Tr}_1\ (|\psi\rangle\langle\psi|)\right) . \tag{1.3}$$

The previous example suggests that entropy can not only be interpreted as a thermodynamic quantity, or as the information content of the state of the system, but it is also a useful tool in the study of the algebraic structure of multipartite states. For instance, a bipartite state being pure and entangled can be rephrased as $S(\rho_{\mathrm{AB}}) = 0$ and $S(\rho_{\mathrm{A}}) > 0$; the independence of the subsystems is equivalent to the saturation of the subadditivity inequality $S(\rho_{\mathrm{AB}}) \leq S(\rho_{\mathrm{A}}) + S(\rho_{\mathrm{B}})$, which in turn is equivalent to the density operator being a simple tensor product of its marginals; while the saturation of the strong subadditivity inequality $S(\rho_{\mathrm{ABC}}) + S(\rho_{\mathrm{B}}) \leq S(\rho_{\mathrm{AB}}) + S(\rho_{\mathrm{BC}})$ turns out to be equivalent to Markovianity, which property can again be characterized by a very special algebraic structure of the tripartite state ρ_{ABC} [44]. In this thesis we study the behavior of entropy in concrete physical models as well as its role in (quantum) information processing, together with implications that entropies of subsystems may give about the algebraic structure of multipartite (quantum) states.

Quantum channels

The mathematical model for a noisy quantum channel is a stochastic map Φ that maps states of an input Hilbert space $\mathcal{H}_{\mathrm{in}}$ to states of an output Hilbert space $\mathcal{H}_{\mathrm{out}}$. To transmit classical information through the channel, the sender encodes the set of possible messages $\mathcal{X}$ into quantum states $\{\rho_x : x \in \mathcal{X}\}$ of the input system. When a message x is sent, the receiver makes a measurement $\{M_x : x \in \mathcal{X}\}$ on the output system, and decodes the message as $y \in \mathcal{X}$ if the measurement outcome is y. An error occurs when $y \neq x$; the probability of this event is $\mathrm{Tr}\,\Phi(\rho_x)\,M_y$. The channel may be used several times; for n use of the channel the input and output Hilbert spaces are replaced to be $\mathcal{H}_{\mathrm{in}}^{\otimes n}$ and $\mathcal{H}_{\mathrm{out}}^{\otimes n}$; the set of messages to transmit is $\mathcal{X}_n$. Note that while the multiple use of the channel is modeled by the product $\Phi^{\otimes n}$, both the encoding quantum states and the measurement operators may be entangled.

The classical information carrying capacity of the channel is the maximal number of transmittable messages per channel use, when the probability of wrong decoding is required to vanish asymptotically. It was pointed out by Holevo, Schumacher and Westmoreland [53, 100] that this capacity is equal to

$$\overline{\chi}\,(\Phi) := \lim_{n\to\infty} \frac{1}{n}\,\chi\left(\Phi^{\otimes n}\right)\,, \tag{1.4}$$

where

$$\begin{aligned}\chi\left(\Phi^{\otimes n}\right) \;:=\; \sup\Big\{ & S\Big(\sum_{x\in\mathcal{X}_n} p_x\,\Phi^{\otimes n}\left(\rho_x\right)\Big) - \sum_{x\in\mathcal{X}_n} p_x\, S\left(\Phi^{\otimes n}\left(\rho_x\right)\right)\;:\;\{\rho_x\}\text{ states} \\ & \text{on } \mathcal{H}_{\mathrm{in}}^{\otimes n};\ \{p_x\}\text{ probability distribution on } \mathcal{X}_n\Big\}\,.\end{aligned}$$

The additivity question of channel capacity is whether $\overline{\chi}\,(\Phi)$ can be replaced by $\chi\,(\Phi)$, that is, if $\chi\,(\Phi^{\otimes n}) = n\chi\,(\Phi)$, or, more generally, if

$$\chi(\Phi_1 \otimes \Phi_2) = \chi(\Phi_1) + \chi(\Phi_2) \tag{1.5}$$

holds for two arbitrary channels Φ_1 and Φ_2. Obviously, additivity would mean a great computational simplification in evaluating the capacity of concrete channels. Another way to put the question is whether using entangled density operators for encoding can enhance the information carrying capacity of the channel (it is known that for certain channels the use of entangled measurement operators does enhance the capacity). Moreover, it was shown recently [8, 72, 103, 104] that additivity of channel capacity is equivalent to other important open additivity questions in quantum information theory, like the additivity of the entanglement of formation, which is a central question in entanglement theory. It was also shown to be equivalent to the additivity of the minimal entropy output

$$S_{\min}(\Phi) := \min\big\{\; S\big(\Phi(|\psi\rangle\langle\psi|)\big)\;:\;\psi\in\mathcal{H}_{\mathrm{in}},\;\|\psi\| = 1\big\}\,, \tag{1.6}$$

that measures the noise of the channel by the amount of mixedness it minimally creates on pure inputs.

In Chapter 3 we study the case of covariant channels, i.e. channels satisfying

$$\Phi\left(U^{*}\,\rho\,U\right)=\alpha(U)^{*}\,\Phi(\rho)\,\alpha(U) \tag{1.7}$$

for some representation α of $SU(\mathrm{d})$. After investigating the structure of such channels for general representations, we turn to the case when α is the conjugate representation $\alpha(U)=\overline{U}$, and show that the set of covariant channels is equal to the set of transpose depolarizing channels

$$\Lambda_t(A):=tA^{T}+(1-t)\,(\mathrm{Tr}\,A)\,\frac{1}{d}\,\mathbb{1}\,,\qquad -\frac{1}{\mathrm{d}-1}\leq t\leq\frac{1}{\mathrm{d}+1}, \tag{1.8}$$

where A^{T} is the transpose of A in the same base in which the U's are conjugated. We then prove the additivity

$$S_{\min}(\Lambda_t\otimes\Lambda_t)=2S_{\min}(\Lambda_t) \tag{1.9}$$

for the restricted range of parameters $-2/(\mathrm{d}^2-2)\leq t$.

Asymptotics of entropy on the spin chain

Let φ be a translation-invariant state on an infinite spin chain, such as the thermodynamical limit of thermal equilibrium states of a set of translation-invariant local Hamiltonians. For the local restrictions φ_N to N consecutive spins we can define an ε-typical subspace $\mathcal{H}_{N,\varepsilon}$ to be a subspace of minimal dimensionality such that $\varphi_N(P_{N,\varepsilon})>1-\varepsilon$, where $P_{N,\varepsilon}$ is the projection onto $\mathcal{H}_{N,\varepsilon}$. That is, for ε "small enough" the support of φ_N is not much different from $\mathcal{H}_{N,\varepsilon}$. Typical subspaces become important e.g. in numerical computations, supposedly their dimensions t_N grow significantly slower than these of the full supports. Heuristically one expects $S(\varphi_N)$ and t_N to be related to each other; a high t_N indicates a very scattered spectrum, just as a high entropy, while a well concentrated spectrum results in low t_N as well as low entropy. Indeed, it has been shown [46, 15] that for a wide class of states with a positive mean entropy $s(\varphi):=\lim_{N\to\infty}\frac{1}{N}\,S(\varphi_N)$ the size of the typical subspaces grow as

$$t_N\sim e^{S(\varphi_N)}\,, \tag{1.10}$$

where $S(\varphi_N)$ in turn grows linearly: $S(\varphi_N)\sim N\,s(\varphi)$.

For pure states, like the thermodynamical limit of the ground states of a set of translation-invariant local Hamiltonians, the mean entropy is in general zero, and no exact relation is known between the asymptotics of entropy and the size of the typical subspaces. The entropies $S(\varphi_N)$ can be viewed as a measure of entanglement between a finite part and the rest of the chain, and a diverging sequence of entropies indicates critical behavior in certain models [110].

In Chapter 7 we study the asymptotics of entropy for a class of pure quasi-free states. Such states can be characterized by a subset K of the interval $[0, 2\pi]$. We show that when K is a union of finitely many $(= M)$ disjoint intervals (the ground state of the XX-model, for instance, corresponds to the case of a single interval), then the entropies grow as

$$c_0 + c_1 M \log N \leq S_N \leq c_2 M \log^2 N , \tag{1.11}$$

where c_1 and c_2 are universal constants, while c_0 depends on the structure of K.

Using states with a more complicated characterizing set, a faster asymptotics can be obtained. If K is a Cantor type set, where in the n-th step of its construction 2^{n-1} intervals of length aq^n are removed, then the asymptotics

$$S_N \sim N^\alpha \tag{1.12}$$

can be shown, with $\alpha = -\dfrac{\log 2}{\log q}$.

Quantum data compression

The problems of quantum information theory fall into two related, but still different categories. The first kind of question one can ask is about the performance of information processing when a quantum mechanical medium is used to carry classical information; the problems related to the classical capacity of a quantum channel fall into this category. On the other hand, one can ask about what happens to the information content carried by a quantum state, when it is subjected to quantum operations. Schumacher's source coding theorem [99] shows that the entropy of a quantum state can be interpreted as a measure of its information content, the same way as it is in the classical case. That is, entropy gives the rate at which the quantum state can be compressed, when a large number of copies of the system is considered, and the recovery map is only required to be asymptotically perfect, when the number of copies tend to infinity.

In Chapter 6 we show that mean entropy plays the same role of asymptotical compressibility in a more general setting, namely for ergodic states of the spin chain, hence it can again be interpreted as the average information content of the state.

The setting of Schumacher's theorem can be modeled by storing classical information in a quantum mechanical system and compressing the quantum data afterwards. The letters of a long message are encoded one-by-one into a state of a quantum system with Hilbert space $\mathcal{H}$, therefore the code of the whole message is a product state on $\mathcal{H}^{\otimes n}$. Assuming a probability distribution on the set of possible messages (this we may call a language statistics), a statistical average of the codes can be taken. If there are no correlations between consecutive letters, then the language statistics is a product measure hence the statistical average of the codes is still a product state. This is the setting when Schumacher's theorem applies. However, in realistic situations the language statistics contains correlations and the average code may not be a product state. In section 6.1 we construct a model of an information source that results in a more general (ergodic) state, for which the data compression theorem can be proven.

Monotonicity of relative entropy

Positive mean entropy means an asymptotically linear growth of the sequence of entropies S_N for a shift-invariant state on the spin chain. The growth is exactly linear if and only if $S_{N+1} - S_N = S_2 - S_1$ for all N, i.e. the strong subadditivity inequality (SSA) $S_{N+1} + S_1 \leq S_N + S_2$ (valid for any shift-invariant state of the spin chain) is saturated. Strong subadditivity is a characteristic property of tripartite states, hence for a characterization of linear entropy growth one should study the case when (SSA) is saturated for a general tripartite state. This question was already studied in e.g. [91, 96], and in [44] a necessary and sufficient condition for the algebraic structure of saturating states was derived. Here we address a more general question, namely to give an algebraic characterization of states saturating the monotonicity inequality for relative entropy, from which the result of [44] follows as a special case.

The relative entropy of the state ρ with respect to the state σ is defined as

$$S(\rho, \sigma) := \mathrm{Tr}\, \rho\, (\log \rho - \log \sigma)\ , \tag{1.13}$$

and can be interpreted as a measure of distinguishability of the two states by measurements on a large number of copies of the system [46, 18]. This interpretation makes it heuristically clear that quantum operations should decrease relative entropy, which is called the monotonicity property of relative entropy:

$$S\left(\Phi\left(\rho\right), \Phi\left(\sigma\right)\right) \leq S(\rho, \sigma)\,. \tag{1.14}$$

All commonly used entropy inequalities (including (SSA)) can be derived from the monotonicity property, hence a characterization of its saturation, apart from being an interesting question on its own, yields characterization of saturation of other entropy inequalities as well.

It was also shown in [44] that saturation of (SSA) is equivalent to Markovianity of the tripartite state (where Markovianity of a quantum state is defined by an algebraic extension of the notion of Markovianity of classical states). This characterization can be extended to states of the spin chain with an exactly linear entropy growth. The structural characterization of tripartite Markovian states yields that Markovian states on the spin chain can be specified by a classical Markovian measure, and a finite set of quantum states, which in turn yields an explicit formula for the mean entropy of these states.

Structure of the thesis

To make the presentation of the material reasonably self-contained, two introductory **Chapters (2 and 5)** and three **Appendices (A, B and C)** are included. The material of this part can be found in the literature, quite scattered, however, in several books and research articles from different fields with different terminology and objectives, therefore it seemed to be essential to collect and present this material in a unified way, adopted to the purposes of the thesis. Throughout the thesis we use the C^*-algebraic formalism of quantum mechanics for two main reasons: first, it enables us to treat the classical and the quantum settings in the same formalism, and second, it provides a more convenient way to describe infinitely extended systems, like the quantum spin chin or the CAR algebra. For this we briefly summarize basic notions and results from the theory of C^*-algebras in **Appendix A**, focusing mainly on the finite dimensional case. The arguments in **Chapter 4** heavily rely on some more advanced results on operator algebras, that we also list, mainly in **Appendix A.5**. **Appendix A** also provides the necessary terminology to formulate the postulates of quantum mechanics in **Appendix B**, as we later use them. In **Appendix B** we also overview briefly how this terminology of quantum mechanics is related to the approach of standard textbooks. **Appendix C** contains minor computations that could be omitted from the main text, but make the presentation more complete.

The four remaining **Chapters, 3, 4, 6 and 7** contain the presentation of the results obtained during the work that finally led to the completion of this thesis, and consist basically of slightly extended and updated versions of the corresponding papers [36], [77], [93] and [35]. In **Chapter 4** a more transparent version of the structural characterization of saturating triples is given, including the structure of the stochastic map Φ. **Section 4.1** is an overview of applications of the structural characterization theorem, based mainly on the results of [44]. Recent results made it possible to state the source coding theorem in a stronger form in **Chapter 6** than in the original paper [93], only with a slight modification of the original argument. Also the model for a quantum information source in this chapter is an extra, compared to [93], and is an original idea, to the best knowledge of the author.

Chapter 2

Finite systems

In this chapter we summarize briefly some basic notions and technical tools from the theory of finite quantum systems and information theory, needed later in the presentation. The main emphasis is put on the notion of entanglement in bipartite quantum systems, and on properties of quantum relative entropy and derived informational quantities. The examples of quantum hypothesis testing and classical capacity of quantum channels support the informational interpretation of quantum relative entropy.

In order to treat classical and quantum systems in the same formalism, we use a C^*-algebraic language; for a brief introduction into the basic notions of C^*-algebras see Appendix A. States of a system are described by positive normalized functionals of the observable algebra. A measurements on a system is described by a finite set of positive semi-definite operators $\{M_i\}$, such that $\sum_i M_i = \mathbb{1}$. Such a measurement is called positive operator valued measurement (POVM). If the system is in state ρ then the outcome probabilities are $p_i = \rho(M_i)$. The change of the state of an open system (either as a result of a Hamiltonian evolution or an intervention into the system) cannot in general be described by a unitary evolution. A more suitable description can be given in terms of completely positive trace-preserving (CPTP) maps, also called quantum operations or stochastic maps, depending on the context. A CPTP map is of the form $\Phi : \rho \mapsto \sum_k X_k \rho X_k^*$, where the operators X_k are called the Kraus operators of Φ, and satisfy $\sum_k X_k^* X_k = \mathbb{1}$. For a brief overview on these notions, see Appendix A and B.

2.1 States and ensembles

A quantum system is called finite if it can be described using a finite-dimensional Hilbert space $\mathcal{H}$. The observable algebra is then a subalgebra of $\mathcal{B}(\mathcal{H})$ and is therefore finite dimensional, thus it is isomorphic to a direct sum of full matrix algebras. The two extremes are the classical and the purely quantum cases. In the first case all the direct summands are one-dimensional, hence the algebra is a commutative one and therefore isomorphic to the function space on some finite set $\mathcal{X}$; the states of the system can be given as probability measures

$$P = \sum_x p_x \delta_x \tag{2.1}$$

on $\mathcal{X}$ (δ_x denotes the Dirac measure concentrated on $\{x\}$). In the second case there is only one term in the direct sum, and the commutant of the observable algebra is trivial. States of a purely quantum system are in a one-to-one correspondence with density matrices on $\mathcal{H}$ (i.e. positive semi-definite matrices with trace 1). Unless a distinction is really needed, we use the same notation for the state and its density matrix, i.e. we use the slightly confusing but very convenient notation

$$\rho(A) = \operatorname{Tr} \rho A. \tag{2.2}$$

An important interpolating example is that of a coupled classical and quantum system. The observable algebra in this case is $\mathcal{A} = \mathcal{F}(\mathcal{X}) \otimes \mathcal{B}(\mathcal{H})$, and states of the system are given in the form

$$\rho = \sum_x p_x \, \delta_x \otimes \rho_x, \tag{2.3}$$

where $\{\rho_x : x \in \mathcal{X}\}$ is a set of density matrices on $\mathcal{H}$.

An **ensemble of states** is a measure μ on the state space $\mathcal{S}(\mathcal{A})$ of the observable algebra; it is called **discrete** if it is finitely supported and **pure** if the support is a finite set of pure states. An equivalent way of specifying a discrete ensemble is to give the set $\{(p_x, \rho_x)\}$ of weights and support points. The average of the ensemble is the state $\int_{\mathcal{S}(\mathcal{A})} \rho \, d\mu(\rho)$, or equivalently $\sum_x p_x \, \rho_x$ for a discrete ensemble. Expressing a given state ρ as the average of a discrete ensemble is also called a (convex) **decomposition** of ρ.

Note that in general there may be several ensembles yielding a given state ρ as the average; the only exceptions in the purely quantum case are the pure states. Let ρ be a given density operator with spectral decomposition $\rho = \sum_{i=1}^{r} \lambda_i \, |e_i\rangle\langle e_i|$. The possible pure decompositions of ρ are characterized by the following:

2.1.1. Theorem. The state ρ can be decomposed in the form

$$\rho = \sum_{j=1}^{n} p_j |v_j\rangle\langle v_j|$$

if and only if there exists an $r \times n$ matrix V with the property $VV^* = I$ such that

$$\sqrt{p_j}\, v_j = \sum_i \sqrt{\lambda_i}\, e_i\, V_{ij}$$

holds.

Note that the list of eigenvectors may be extended by $n - r$ zeros and V can be completed to be an $n \times n$ unitary such that the above formula still holds.

We call a vector $\boldsymbol{p}$ a probability vector if its components are non-negative and sum up to 1. The possible probability vectors $\boldsymbol{p}$ appearing in a pure decomposition of ρ can be described as those being more mixed than the eigenvalue vector $\boldsymbol{\lambda}$:

2.1.2. Definition. Let $\boldsymbol{p} = (p_1, \dots, p_m)$ and $\boldsymbol{q} = (q_1, \dots, q_n)$ be nonnegative vectors, and suppose their components are arranged in decreasing order. In case $n > m$ we can complete $\boldsymbol{p}$ by $n - m$ zeros, and therefore $n = m$ can be assumed without loss of generality. The vector $\boldsymbol{p}$ is called **more mixed** then $\boldsymbol{q}$, in notation $\boldsymbol{p} \succ \boldsymbol{q}$, if

$$\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} q_i \qquad \text{and} \qquad \sum_{i=1}^{k} p_i \leq \sum_{i=1}^{k} q_i \qquad k = 1, \dots, n-1 \tag{2.4}$$

holds. This relation is also referred to as **majorization** or **stochastic domination** in the literature.

2.1.3. Theorem. Given a state ρ and a probability vector $\boldsymbol{p} = (p_1, \ldots, p_n)$ there exist unit vectors $v_1, \ldots, v_n$ such that ρ can be decomposed in the form

$$\rho = \sum_{j=1}^{n} p_j |v_j\rangle\langle v_j| \qquad \text{if and only if} \qquad \boldsymbol{p} \succ \boldsymbol{\lambda}$$

holds, where $\boldsymbol{\lambda}$ is the eigenvalue vector of ρ.

A pure state ψ on an extended Hilbert space $\mathcal{H} \otimes \mathcal{K}$ is called a **purification** of the mixed state ρ, if ρ can be obtained as its first marginal, i.e.

$$\rho = \mathrm{Tr}_2 \, |\psi\rangle\langle\psi|. \tag{2.5}$$

To every pure decomposition $\rho = \sum_{j=1}^{n} p_j |v_j\rangle\langle v_j|$ of ρ there exists a purification, namely $\psi = \sum_{j=1}^{n} \sqrt{p_j}\, v_j \otimes f_j$, where $f_1, \ldots, f_n$ is an orthonormal set of vectors in some auxiliary Hilbert space $\mathcal{K}$. The purification $\psi = \sum_{i=1}^{r} \sqrt{\lambda_i}\, e_i \otimes e_i$ corresponding to the eigen-decomposition is in some sense canonical amongst all the purifications (even though it is not even unique, unless all the eigenvalues have multiplicity 1), and coincides with the GNS representing vector (Example A.3.8). As a consequence of Theorem 2.1.1, any two purifications ψ and $\hat{\psi}$ can be related by a unitary in the form $\hat{\psi} = (I_{\mathcal{H}} \otimes U)\, \psi$, where a smaller auxiliary Hilbert space may be embedded to a bigger one to have $\dim \mathcal{K} = \dim \hat{\mathcal{K}}$.

The **Schmidt decomposition** of a pure bipartite state ψ is dual to the canonical purification of a state, in the sense that it gives a 'canonical' (again not unique) way to express ψ in the eigenbases of its marginals. More precisely, given a pure state ψ in a composite Hilbert space $\mathcal{H} \otimes \mathcal{K}$ one can find orthonormal sets of vectors $\{e_1, \ldots, e_r\}$ and $\{f_1, \ldots, f_r\}$ in $\mathcal{H}$ and $\mathcal{K}$ such that

$$\psi = \sum_{i=1}^{r} \sqrt{\lambda_i}\, e_i \otimes f_i,$$

where $\boldsymbol{\lambda}$ is a probability vector, the components of which are called the **Schmidt coefficients** of ψ. The marginals of $|\psi\rangle\langle\psi|$ are then $\sum_i \lambda_i \, |e_i\rangle\langle e_i|$ and $\sum_i \lambda_i \, |f_i\rangle\langle f_i|$. As a consequence, the spectra of the marginals are the same, apart from multiplicities of 0.

2.2 Entanglement

A state ρ of a composite system (with observable algebra $\mathcal{A}_1 \otimes \mathcal{A}_2 \subset \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2)$) is called **separable** (or **non-entangled**), if it is decomposable into a convex combination of product states, i.e. if there exist states $\rho_i^1 \in \mathcal{S}(\mathcal{A}_1)$ and $\rho_i^2 \in \mathcal{S}(\mathcal{A}_2)$ together with a probability distribution $\{p_i\}$ such that

$$\rho = \sum_{i=1}^{n} p_i \, \rho_i^1 \otimes \rho_i^2 \,. \tag{2.6}$$

Note that a pure state is separable if and only if it is the product of two pure states and the terms appearing in decomposition (2.6) can obviously be chosen to be pure product states. Therefore the set of all separable states $\mathcal{S}_{\rm sep}(\mathcal{A}_1 \otimes \mathcal{A}_2)$ is a closed convex set, the extremal points of which are the pure product states. States that are not separable are called **entangled**.

Note that if one of the systems is classical, then any state ρ of the composite system is separable, as it can be expressed in the form of (2.3). On the other hand, if both subsystems are non-classical then the set of entangled states is always nonempty, thus entanglement is a characteristically quantum feature of composite states.

Since the set of separable states is convex, there exists a linear functional φ for every entangled state ρ, and a constant c such that $\varphi(\rho) < c$ and $\varphi(\sigma) > c$ holds for every $\sigma \in \mathcal{S}_{\rm sep}$. Using $\varphi(\rho) - c = \operatorname{Tr} F \rho - c \operatorname{Tr} \rho = \operatorname{Tr}(F - cI)\, \rho$ for some $F \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$, we obtain the existence of an operator $C \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ such that

$$\operatorname{Tr} C\rho < 0 \qquad \text{and} \qquad \operatorname{Tr} C\sigma > 0 \quad \forall \sigma \in \mathcal{S}_{sep} \,. \tag{2.7}$$

Obviously, C cannot be positive semidefinite, but the positivity criterion for separable states implies that $\langle x \otimes y,\, C\, x \otimes y \rangle > 0 \;\; \forall x \in \mathcal{H}_1,\, y \in \mathcal{H}_2$ has to hold, which is easily seen to be equivalent with C being the Choi matrix of a positive but not completely positive map $\Phi : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ (see Appendix A.4 for definition of Choi matrix and other notions). Note that $(\mathrm{id} \otimes \Phi)\,(\rho)$ is positive semi-definite for every positive map Φ and separable state ρ. On the other hand, $\operatorname{Tr} C\,(\Phi)\,\rho = \langle \psi,\, (\mathrm{id} \otimes \Phi^*)\,(\rho)\,\psi \rangle$ (see notations after Example A.4.10), thus (by replacing Φ by Φ^*) we obtain the following

2.2.1. Theorem. A bipartite state ρ is separable if and only if $(\mathrm{id} \otimes \Phi)\,(\rho)$ is positive semi-definite for all positivity-preserving maps Φ.

2.2.2. Example. The transposition map $\mathrm{T} : e_{ij} \mapsto e_{ji}$ corresponding to a base $\{e_1, \dots, e_d\}$ is positivity-preserving, but not completely positive (see Example A.4.1). Having positive partial transpose (PPT criterion) is a necessary condition for separability, i.e. for all separable bipartite states $(\mathrm{id} \otimes \mathrm{T})\,(\rho) \geq 0$ must hold [85]. This criterion is also sufficient when both of the Hilbert spaces are two dimensional [56].

2.3 Distance measures on states

There are several possible ways to quantify how close two quantum states are to each other. Here we list a few, and some relations among them; a more detailed presentation of the topic can be found in [81].

The dual space of $\mathcal{B}(\mathcal{H})$ (i.e. the space of linear functionals) can be identified with $\mathcal{B}(\mathcal{H})$ through the Hilbert-Schmidt inner product. The dual space norm on $\mathcal{B}(\mathcal{H})$ is easily seen to be the **trace norm**, given by $\|A\|_1 := \mathrm{Tr}\,|A|$. Since quantum states are functionals on the observable algebra, it is natural to equip the state space by the trace norm. It is common in probability theory to introduce an extra factor $\frac{1}{2}$, and define the **trace distance** by

$$D(\rho, \sigma) := \frac{1}{2}\,\mathrm{Tr}\,|\rho - \sigma|. \tag{2.8}$$

For classical probability distributions ρ and σ on a finite set $\mathcal{X}$, the trace distance can be expressed as

$$D(\rho, \sigma) = \max\{|\rho(H) - \sigma(H)| \,:\, H \subset \mathcal{X}\}\,, \tag{2.9}$$

i.e. it gives the maximal amount to which the two measures can differ on an arbitrary subset.

The trace distance is obviously symmetric and jointly convex in its variables; it is invariant under a common unitary rotation of the states, and takes values between 0 and 1; moreover $D(\rho, \sigma) = 0$ if and only if $\rho = \sigma$. It can be shown that the trace distance is monotonic under quantum operations, i.e.

$$D\left(\Phi(\rho), \Phi(\sigma)\right) \leq D(\rho, \sigma)\,, \qquad \forall \rho, \sigma. \tag{2.10}$$

It follows immediately that if a measurement $\{M_i\}$ yields the outcome probabilities $p_i := \rho(M_i)$ and $q_i := \sigma(M_i)$, then $D(\boldsymbol{p}, \boldsymbol{q}) \leq D(\rho, \sigma)$. The inequality can be made strict, i.e. there exists a measurement for which it is saturated; moreover, it can be chosen to be a projective measurement with two measurement operators, very much as in the classical characterization (2.9).

Instead of introducing a distance on the state space, i.e. measuring how far the states are from each other, we can can ask how close two states are. A standard measure of closeness for pure states ψ, φ in quantum mechanics is given by the transition probability $|\langle\psi, \varphi\rangle|^2$, or by the overlap $|\langle\psi, \varphi\rangle|$. This notion has several possible extension to mixed states, leading to different notions of **fidelity**. Uhlmann's fidelity [107] is defined as the maximal overlap between possible purifications of the mixed states ρ and σ, and leads to the formula

$$F_U(\rho, \sigma) = \mathrm{Tr}\,\sqrt{\rho^{1/2}\,\sigma\rho^{1/2}}\,. \tag{2.11}$$

By the very definition it is symmetric in its variables, invariant under common unitary rotations, takes values between 0 and 1, and $F_U(\rho, \sigma) = 1$ if and only if $\rho = \sigma$. It can be shown to be increasing under quantum operations, i.e.

$$F_U(\Phi(\rho), \Phi(\sigma)) \geq F_U(\rho, \sigma) \tag{2.12}$$

for any quantum operation Φ. It immediately implies joint concavity in its variables, and also that

$$F_U(\rho, \sigma) \leq F_U(\boldsymbol{p}, \boldsymbol{q})\ ; \qquad p_i := \rho(M_i),\ q_i := \sigma(M_i); \tag{2.13}$$

where $\{M_i\}$ is a measurement. The fidelity of classical probability distributions $\boldsymbol{p}$ and $\boldsymbol{q}$ can easily seen to be equal to the scalar product of the vectors $(\sqrt{p_1}, \ldots, \sqrt{p_m})$ and $(\sqrt{q_1}, \ldots, \sqrt{q_m})$; these are the images of the probability vectors $\boldsymbol{p}$ and $\boldsymbol{q}$ when the probability simplex is projected onto the unit sphere. Again the inequality can be saturated, hence

$$F_U(\rho, \sigma) = \min\{F_U(\{\rho(M_i)\}, \{\sigma(M_i)\}) \,:\, \{M_i\} \text{ is a measurement }\}\,. \tag{2.14}$$

Uhlmann's fidelity and the trace distance are related through the inequality

$$1 - F_U(\rho, \sigma) \leq D(\rho, \sigma) \leq \sqrt{1 - F_U^2(\rho, \sigma)}\,, \tag{2.15}$$

i.e. a high fidelity implies a small distance, and vice versa. The second inequality is easily seen to be saturated for pure states, i.e.

$$D(|\psi\rangle\langle\psi|, |\varphi\rangle\langle\varphi|) = \sqrt{1 - |\langle\psi, \varphi\rangle|^2}\,.$$

In concrete computations it may be easier to work with a different extension of the overlap for mixed states. Schumacher [99] used the notion of fidelity $F_S(\rho,\sigma) := \mathrm{Tr}\,\rho\,\sigma$ in the proof of his source coding theorem, while in [93] $F_P := \mathrm{Tr}\,\rho^{1/2}\,\sigma^{1/2}$ was introduced for the same purpose. These quantities are related by the inequality

$$F_S \le F_U \le F_P\,. \tag{2.16}$$

Indeed, $\mathrm{Tr}\,\rho\,\sigma = \mathrm{Tr}\,\rho^{1/2}\,\sigma\,\rho^{1/2} \le \mathrm{Tr}\,\sqrt{\rho^{1/2}\,\sigma\,\rho^{1/2}}$ yields the first inequality, while the Lieb-Thirring inequality $\mathrm{Tr}\,\left(\rho^{1/2}\,\sigma\,\rho^{1/2}\right)^p \le \mathrm{Tr}\,\rho^p\,\sigma^p$ with the choice $p = 1/2$ yields the second.

The amount to which a quantum operation Φ preserves a state ρ can be measured by the fidelity $F(\rho, \Phi(\rho))$, using any of the above notions of fidelity. If an ensemble $\{p_x, \rho_x\}$ is given, then a natural notion for the average amount to which the states are preserved is the ensemble fidelity

$$F(\{p_x,\rho_x\},\Phi) := \sum_x p_x\,F(\rho_x, \Phi(\rho_x))\,. \tag{2.17}$$

Following Uhlmann's previous idea, a mixed state can be viewed as the marginal of one of its purification, and the quantum operation Φ acts only on the first component of the system, hence the amount to which a quantum operation Φ preserves the state ρ is measured by

$$F(\rho,\Phi) := F_U^2\left(\,|\psi\rangle\langle\psi|,\,\Phi\otimes\mathrm{id}\,(\,|\psi\rangle\langle\psi|\,)\,\right)\,, \tag{2.18}$$

where ψ is an arbitrary purification of ρ (it doesn't matter which one). This quantity is called the **entanglement fidelity**. The square in the definition is there for mathematical convenience; with this choice the entanglement fidelity is convex in the variable ρ. Monotonicity of F_U under partial trace over the second subsystem yields $F(\rho,\Phi) \le F_U^2(\rho,\Phi(\rho))$, which, together with convexity implies

$$F\left(\sum_x p_x\rho_x, \Phi\right) \le \sum_x p_x\,F_U^2(\rho_x,\Phi(\rho_x)) \le F_U(\{p_x,\rho_x\},\Phi)\,. \tag{2.19}$$

A great advantage of entanglement fidelity is that it can easily be expressed as

$$F(\rho,\Phi) = \sum_j |\rho(X_j)|^2\,, \tag{2.20}$$

where $\{X_j\}$ is a set of Kraus operators for Φ (i.e. $\Phi(\rho) = \sum_j X_j\,\rho\,X_j^*$).

2.4 Quantum entropy

The notion of entropy plays a fundamental role in statistical physics, as well as in information theory and probability theory. The entropy of a quantum state ρ was given by John von Neumann in 1927 in the work [78] as

$$S(\rho) := -\mathrm{Tr}\,\rho \log \rho = -\sum_i \lambda_i \, \log \lambda_i, \tag{2.21}$$

where $\lambda_1, \lambda_2, \ldots$ are the eigenvalues of ρ, and the convention $0 \log 0 := 0$ is used. Almost twenty years later, in the work [101], C.E. Shannon defined the entropy of a finite probability distribution $P = \sum_{x\in\mathcal{X}} p_x \, \delta_x$ as

$$S(P) := -\sum_{x\in\mathcal{X}} p_x \log p_x, \tag{2.22}$$

as a measure of the information content of the distribution. Note that Shannon's formula is nothing else but the von Neumann entropy of the state of a classical system. For a heuristic interpretation of Shannon entropy, suppose we make a measurement on the system in state P. For simplicity, assume that the measurement has two possible outcomes x_1 and x_2 with probabilities $p_{x_1} < p_{x_2}$. A priori, we expect x_2 as the outcome of the measurement, therefore if we really get x_2 as outcome, it changes less our a priori knowledge than when we get x_1, hence it gives us less information. Thus it is a natural idea to measure the information we gain by getting outcome x_i as $f(p_{x_i})$, with some monotonic decreasing function f. The expectation value of the information provided by the measurement is then $\sum_i p_{x_i} f(p_{x_i})$. The choice $f := -\log$ is justified by the fact that quantity (2.22) obtains a natural interpretation as information content in coding theorems; for details, see Chapter 6.

An even more fundamental notion in (quantum) information theory is that of the relative entropy of two states:

$$S(\rho, \sigma) := \begin{cases} \mathrm{Tr}\,\rho\,(\log \rho - \log \sigma)\,, & \text{if supp } \rho \leq \text{supp } \sigma; \\ +\infty, & \text{otherwise}\,, \end{cases} \tag{2.23}$$

where supp ρ is the subspace spanned by the eigenvectors of ρ corresponding to strictly positive eigenvalues. For classical states supp $\rho \leq$ supp σ means that ρ is absolutely continuous with respect to σ, hence we will use the same notation $\rho \ll \sigma$ in the quantum case as well. Entropy can be obtained from relative entropy by choosing the second state to be the normalized trace state: $S(\rho) = \log d - S(\rho, \mathrm{tr})$, where d is the dimension of the Hilbert space of the system. In the classical case formula (2.23) reduces to

$$S(\{p_x\},\{q_x\}) = \sum_x p_x \log \frac{p_x}{q_x};$$

this form of relative entropy appears under several other names in the information theory literature, such as informational divergence or Kullback-Leibler distance.

In the following we summarize basic properties of entropy and relative entropy, and (also in later chapters) we show that von Neumann entropy and quantum relative entropy have the same information theoretic contents in the quantum setting as Shannon entropy and Kullback-Leibler distance have in the classical.

2.4.1 Basic properties

It is immediately clear from the definition of entropy that it is non-negative for any state and that it is equal to zero if and only if the state is pure. The relative entropy is also non-negative, as seen from the inequality ([83], Proposition 1.1):

$$\frac{1}{2}\,\mathrm{Tr}\,(\rho - \sigma)^2 \leq S(\rho, \sigma); \tag{2.24}$$

moreover, $S(\rho, \sigma) = 0$ if and only if the two states are equal. Choosing the second state to be the normalized trace, we obtain that

$$S(\rho) \leq \log d, \tag{2.25}$$

and equality holds if and only if $\rho = \mathrm{tr}$.

The strict positivity of relative entropy motivates to consider it as some kind of a distance measure on the set of states, even though it is not symmetric in its arguments and doesn't satisfy the triangle inequality. However, relative entropy can be interpreted as a measure of statistical distinguishability in a simple decision scheme (for details, see section 2.5.1).

The subsystems of a bipartite system are independent if and only if the state of the composite system is the product of its marginals, i.e. $\rho_{AB} = \rho_A \otimes \rho_B$. A possible way to quantify correlation between the subsystems is to measure it by the relative entropic distance of the state from the product of its marginals:

$$I(\mathrm{A}:\mathrm{B}) := S(\rho_{AB}, \rho_A \otimes \rho_B); \tag{2.26}$$

this quantity is called the **mutual information** of the two subsystems in information theory. Strict positivity of relative entropy implies that mutual information is strictly positive, unless the two marginals are independent. Since

$$S(\rho_{AB}, \rho_A \otimes \rho_B) = S(\rho_A) + S(\rho_B) - S(\rho_{AB}),$$

the nonnegativity of relative entropy implies

$$S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B), \tag{2.27}$$

which is called the **subadditivity of entropy**. The subadditivity inequality is satisfied with equality if and only if the two subsystems are independent. Applying the subadditivity inequality to the state $\sum_x p_x\, \delta_x \otimes \rho_x$ we obtain

$$\sum_x p_x\, S\left(\rho_x\right) \leq S\left(\sum_x p_x\, \rho_x\right), \tag{2.28}$$

i.e. entropy is a concave function on the state space. Consider now a purification of ρ_{AB} on the Hilbert space $\mathcal{H}_\mathrm{A} \otimes \mathcal{H}_\mathrm{B} \otimes \mathcal{H}'_\mathrm{A} \otimes \mathcal{H}'_\mathrm{B}$, where $\mathcal{H}'_\mathrm{A}$ and $\mathcal{H}'_\mathrm{B}$ are copies of $\mathcal{H}_\mathrm{A}$ and $\mathcal{H}_\mathrm{B}$. As the spectra of the marginals of a bipartite pure state are essentially the same (Section (2.1)), the subadditivity inequality, applied to the partition $\{\mathrm{A};\ \mathrm{BA'B'}\}$ yields

$$S(\rho_\mathrm{A}) = S(\rho_{\mathrm{BA'B'}}) \leq S(\rho_\mathrm{B}) + S(\rho_{\mathrm{A'B'}}) = S(\rho_\mathrm{B}) + S(\rho_{\mathrm{AB}}).$$

Repeating the same for the partition $\{\mathrm{B};\ \mathrm{AA'B'}\}$ we obtain the **triangle inequality** for entropy

$$|S(\rho_\mathrm{A}) - S(\rho_\mathrm{B})| \leq S(\rho_{\mathrm{AB}}). \tag{2.29}$$

Fannes's inequality [31] states that entropy is a continuous function on the state space:

2.4.1. Theorem. If $\|\rho - \sigma\|_1 \leq \frac{1}{3}$ holds for two states ρ and σ then

$$|S(\rho) - S(\sigma)| \leq \|\rho - \sigma\|_1 \log d - \|\rho - \sigma\|_1 \log \|\rho - \sigma\|_1 .$$

Here $\|A\|_1 := \operatorname{Tr}|A|$ is the **trace norm** of the operator A.

As $(\mathrm{id} \otimes \Phi)(\rho_{\mathrm{AB}})$ and $(\Phi \otimes \mathrm{id})(\rho_{\mathrm{AB}})$ are positive semi-definite for any *separable* bipartite state ρ_{AB}, the choice $\Phi(A) := (\operatorname{Tr} A)\, \mathbb{1}_{\mathcal{A}} - A$ yields

$$\rho_{\mathrm{AB}} \leq \rho_{\mathrm{A}} \otimes \mathbb{1}_{\mathcal{B}} \qquad \text{and} \qquad \rho_{\mathrm{AB}} \leq \mathbb{1}_{\mathcal{A}} \otimes \rho_{\mathrm{B}},$$

which implies

$$S(\rho_{\mathrm{A}}) \leq S(\rho_{\mathrm{AB}}) \qquad \text{and} \qquad S(\rho_{\mathrm{B}}) \leq S(\rho_{\mathrm{AB}}). \tag{2.30}$$

As classical states are separable, inequality (2.30) is valid for any classical bipartite state, in complete agreement with the heuristic view that a marginal of a distribution contains less information than the distribution itself. However, inequality (2.30) doesn't hold for a bipartite quantum state in general. A well-known example is the singlet state $\psi := \frac{1}{\sqrt{2}}\left(|01\rangle - |10\rangle\right)$, the first marginal of which is the trace state, thus $S(\rho_{\mathrm{A}}) = S(\mathrm{tr}) = \log 2 > 0 = S(|\psi\rangle\langle\psi|)$.

Applying inequality (2.30) to the separable state $\rho_{\mathrm{AB}} := \sum_x p_x\, \delta_x \otimes \rho_x$ yields

$$S\left(\sum_x p_x\, \rho_x\right) \leq \sum_x p_x\, S\left(\rho_x\right) + S\left(\{p_x\}\right), \tag{2.31}$$

complementing the concavity of entropy.

2.4.2 Entanglement measures

In section 2.2 we mentioned that the question whether a bipartite state is entangled or not, can be decided by the help of positivity-preserving maps. A more subtle question is to quantify the amount of entanglement contained in a bipartite state. A function $\mathcal{E}$ is called an **entanglement measure**, if it is defined on the state space of any bipartite system, and satisfies the requirements

(E1) $\mathcal{E}(\rho) = 0$ for any separable state ρ;

(E2) $\mathcal{E}$ is invariant under local unitary transformations, i.e.

$$\mathcal{E}\left((U_1 \otimes U_2)\ \rho\ (U_1^* \otimes U_2^*)\right) = \mathcal{E}\left(\rho\right)\ .$$

There are several different ways to introduce such a functional, all of which are natural from a given point of view. A geometric approach is to introduce a (weak) notion of distance on the state spaces of bipartite systems, i.e. a functional D such that $D(\rho,\rho)=0$ and $D(\rho,\sigma)>0$ if $\rho\neq\sigma$ holds. If D is unitarily invariant as well then we can define entanglement of a state as its distance from the set of separable states, i.e.

$$\mathcal{E}(\rho) := \inf\{D(\rho,\sigma)\ :\ \sigma \text{ is separable }\}. \tag{2.32}$$

Such a definition of entanglement satisfies a stronger version of (E1):

(E1') $\mathcal{E}(\rho)\geq 0$ and $\mathcal{E}(\rho)=0$ if and only if ρ is separable.

Another approach is to measure the entanglement of a state by its distinguishability by measurements from separable states. Based on the interpretation of relative entropy, given in section 2.5.1, this approach turns out to be a special case of the previous one, where the distance is the relative entropic distance; the entanglement measure $\mathcal{E}_S$ obtained this way is called **relative entropy of entanglement**. A valuable feature of this choice is that the infimum in (2.32) can be evaluated explicitly for pure states; if $\rho := |\psi\rangle\langle\psi|$ is a pure state with Schmidt decomposition $\psi=\sum_i\sqrt{\lambda_i}\,e_i\otimes f_i$, then $\mathcal{E}_S(\rho)$ turns out to be the entropy of its marginals [109], i.e.

$$\mathcal{E}_S(|\psi\rangle\langle\psi|) = S\left(\mathrm{Tr}_2\ (|\psi\rangle\langle\psi|)\right) = S\left(\mathrm{Tr}_1\ (|\psi\rangle\langle\psi|)\right) = -\sum_i \lambda_i\log\lambda_i. \tag{2.33}$$

The state that minimizes the distance from the set of separable states, is

$$\sigma_0 = \sum_i \lambda_i\,|e_i\rangle\langle e_i|\otimes|f_i\rangle\langle f_i|\,.$$

One can argue that entanglement measures the quantum correlations of the state, therefore classical correlations may be measured by the distance of the best separable approximant σ_0 from the product of the marginals $\rho_1\otimes\rho_2$. A straightforward computation shows that for pure states the total amount of correlations $C(\rho)$, as measured by the mutual information of the subsystems, is the sum of the quantum and the classical correlations, i.e.

$$C(\rho) = S(\rho,\rho_1\otimes\rho_2) = S(\rho,\sigma_0)+S(\sigma_0,\rho_1\otimes\rho_2) = C_{\text{quantum}}(\rho)+C_{\text{classical}}(\rho).$$

Another natural requirement an entanglement measure should satisfy, is additivity on independent states, i.e. if ρ is a state on $\mathcal{H}_{\mathrm{A}}\otimes\mathcal{H}_{\mathrm{B}}$, and σ is a state on $\mathcal{K}_{\mathrm{A}} \otimes \mathcal{K}_{\mathrm{B}}$, then the entanglement of $\rho \otimes \sigma$, when viewed as a state on $(\mathcal{H}_{\mathrm{A}} \otimes \mathcal{K}_{\mathrm{A}}) \otimes (\mathcal{H}_{\mathrm{B}} \otimes \mathcal{K}_{\mathrm{B}})$ satisfies

(E3) $\mathcal{E}(\rho \otimes \sigma) = \mathcal{E}(\rho) + \mathcal{E}(\sigma)$.

Unfortunately, the relative entropy of entanglement doesn't satisfy this criterion, even though its restriction to pure states does. However, the subadditivity of entropy (2.27) implies that it satisfies the weaker property of subadditivity

(E3') $\mathcal{E}(\rho \otimes \sigma) \leq \mathcal{E}(\rho) + \mathcal{E}(\sigma)$.

As a consequence of subadditivity, the limit

$$\mathcal{E}_S^{\infty}(\rho) := \lim_{n \to \infty} \frac{1}{n} \mathcal{E}_S\left(\rho^{\otimes n}\right) \tag{2.34}$$

exists for any bipartite state ρ, and $\mathcal{E}_S^{\infty}$ satisfies criteria (E1) and (E2), therefore it is an entanglement measure, which is called the **regularization** of $\mathcal{E}_S$. As an immediate consequence of its definition, it satisfies a weaker form of additivity:

(E3") $\mathcal{E}\left(\rho^{\otimes n}\right) = n\,\mathcal{E}(\rho)$.

It has been shown [29] that the only functional on pure bipartite states that satisfies (E1), (E2) and (E3"), together with a continuity criterion, and monotonicity under a class of stochastic maps, called LOCC (from local operations and classical communication), is a constant multiple of the relative entropy of entanglement. Therefore a new approach to define entanglement measures can be to extend the relative entropy of entanglement from pure states to mixed states in a well-behaved way. One candidate is the **entanglement of formation**, defined as

$$\mathcal{E}_f(\rho) := \inf\{\sum_i p_i\,\mathcal{E}_S(\rho_i)\ :\ \rho = \sum_i p_i\,\rho_i \text{ is a pure decomposition of } \rho\}. \tag{2.35}$$

It is easily seen to be an entanglement measure, that satisfies also the subadditivity property (E3'). It is also believed to be additive, but this conjecture remained unsolved up till now, despite several efforts to prove it. It has been shown recently [104] that the question of additivity for entanglement of formation is equivalent to the additivity question of the Holevo capacity (section 2.5) and of the minimal entropy output (section 3.1) of a quantum channel.

2.4.3 Entropy inequalities

As relative entropy is interpreted as statistical distinguishability of states, it is heuristically clear that if the states are subjected to a stochastic operation, their distinguishability, and therefore their relative entropy cannot increase. A proof is almost trivial for classical systems, while in the quantum case it requires rather involved techniques of matrix algebra and functional analysis.

2.4.2. Theorem. Let $\alpha : \mathcal{A} \to \mathcal{B}$ be a 2-positive unital map, $\Phi := \alpha^*$, and ρ and σ be states on $\mathcal{A}$ such that $\rho \ll \sigma$. Then

$$S\left(\Phi\left(\rho\right), \Phi\left(\sigma\right)\right) \leq S\left(\rho, \sigma\right). \tag{2.36}$$

Note that only 2-positivity of the map is required; that's because the proof relies on the 2-positivity inequality (A.12). Many important properties of information quantities can be derived from the monotonicity property of relative entropy.

2.4.3. Example. Let $\{M_i\}$ be a measurement with outcome probabilities $p_i := \rho(M_i)$ and $q_i := \sigma(M_i)$. Monotonicity implies

$$S\left(\{p_i\}, \{q_i\}\right) \leq S(\rho, \sigma),$$

i.e. quantum relative entropy gives an upper bound on the relative entropic distance of the measurement statistics.

2.4.4. Example. Let Φ be a bistochastic operation. Then

$$\log d - S\left(\Phi(\rho)\right) = S\left(\Phi(\rho), \mathrm{tr}\right) = S\left(\Phi(\rho), \Phi(\mathrm{tr})\right) \leq S\left(\rho, \mathrm{tr}\right) = \log d - S(\rho);$$

i.e. bistochastic maps increase entropy. A particular example is a projective measurement; monotonicity of entropy then states that the entropy of the outcome probabilities is bounded from below by the quantum entropy of the state. Equality holds when the measurement is chosen to consist of rank 1 eigenprojections of ρ.

2.4.5. Corollary. Joint convexity of relative entropy

Let ρ_1, ρ_2 and σ_1, σ_2 be states on $\mathcal{B}$ such that $\rho_k \ll \sigma_k$, and let $t \in [0,1]$. Choose $\mathcal{A} := \mathcal{B} \otimes \mathcal{M}_2$, $\rho := t\,\rho_1 \otimes e_{11} + (1-t)\,\rho_2 \otimes e_{22}$, $\sigma := t\,\sigma_1 \otimes e_{11} + (1-t)\,\sigma_2 \otimes e_{22}$, and $\Phi := \mathrm{Tr}_2$. Then

$$\Phi(\rho) = t\rho_1 + (1-t)\rho_2 \quad \text{and} \quad \Phi(\sigma) = t\sigma_1 + (1-t)\sigma_2,$$

implying

$$S\left(t\rho_1 + (1-t)\rho_2,\, t\sigma_1 + (1-t)\sigma_2\right) \le t\, S(\rho_1,\, \sigma_1) + (1-t)\, S(\rho_2,\, \sigma_2).$$

2.4.6. Corollary. Monotonicity of mutual information

Suppose we have a bipartite system with observable algebra $\mathcal{A} \otimes \mathcal{B}$ in the state ρ_{AB}, the second half of which is sent through the channel $\Lambda : \mathcal{S}(\mathcal{B}) \to \mathcal{S}(\mathcal{B}')$, i.e. $\rho_{\mathrm{AB}'} = (\mathrm{id} \otimes \Lambda)(\rho_{\mathrm{AB}})$. By choosing $\sigma := \rho_{\mathrm{A}} \otimes \rho_{\mathrm{B}}$ and $\Phi := \mathrm{id} \otimes \Lambda$ we have $\Phi(\sigma) = \rho_{\mathrm{A}} \otimes \rho_{\mathrm{B}'}$. Applying Theorem 2.4.2 we obtain

$$\begin{aligned}
\mathrm{I}(\mathrm{A} : \mathrm{B}') &= S\left(\rho_{\mathrm{AB}'},\, \rho_{\mathrm{A}} \otimes \rho_{\mathrm{B}'}\right) \\
&= S\left(\Phi(\rho_{\mathrm{AB}}),\, \Phi(\rho_{\mathrm{A}} \otimes \rho_{\mathrm{B}})\right) \\
&\le S\left(\rho_{\mathrm{AB}},\, \rho_{\mathrm{A}} \otimes \rho_{\mathrm{B}}\right) \\
&= \mathrm{I}(\mathrm{A} : \mathrm{B})
\end{aligned}$$

In classical information theory the monotonicity of mutual information is referred to as data processing inequality (DPE).

2.4.7. Example. The mutual information of the state $\rho_{\mathrm{AB}} := \sum_x p(x)\, \delta_x \otimes \rho_x$ of a coupled classical and quantum system is

$$I(\mathrm{A} : \mathrm{B}) = S\left(\sum_x p(x)\, \rho_x\right) - \sum_x p(x) S(\rho_x) =: \chi\left(\{p(x), \rho_x\}\right); \tag{2.37}$$

the quantity χ is called the Holevo-quantity of the ensemble $\{p(x), \rho_x\}$. Suppose now that the quantum part of the system is subjected to a quantum operation Φ, with resulting state $\sum_x p(x)\, \delta_x \otimes \Phi(\rho_x)$. Monotonicity of mutual information implies the monotonicity of the Holevo quantity, i.e.

$$\chi\left(\{p(x), \Phi(\rho_x)\}\right) \le \chi\left(\{p(x), \rho_x\}\right). \tag{2.38}$$

A case of special importance is when Φ is a measurement with measurement operators $\{M_y \,:\, y \in \mathcal{Y}\}$. Application of the measurement on the quantum system yields a probability distribution on the classical system described by $\mathcal{X} \times \mathcal{Y}$:

$$(\mathrm{id} \otimes \Phi) \left(\sum_x p(x)\, \delta_x \otimes \rho_x \right) = \sum_{xy} p(x)\, \delta_x \otimes p(y|\, x)\delta_y \,, \tag{2.39}$$

where $p(y|\, x) := \rho_x(M_y)$. Monotonicity of the Holevo quantity implies that the mutual information I between the marginals of this classical joint distribution is smaller than the Holevo quantity of the original ensemble, i.e.

$$I = \sum_{xy} p(x,y) \log \frac{p(x,y)}{p(x)\, p(y)} \leq \chi\left(\{p(x), \rho_x\}\right), \tag{2.40}$$

where $p(x,y) := p(x)\, p(y|\, x)$ and $p(y) := \sum_x p(x)\, p(y|\, x)$; this inequality is called the **Holevo bound**.

An important special case of monotonicity of mutual information is the so-called **strong subadditivity of entropy** (SSA):

2.4.8. Example. Strong subadditivity of entropy
Let $\rho := \rho_{\mathrm{ABC}}$ be a state on a tripartite system with observable algebra $\mathcal{A} \otimes \mathcal{B} \otimes \mathcal{C}$. Let Λ be the partial trace map from $\mathcal{B} \otimes \mathcal{C}$ to $\mathcal{C}$, and $\Phi := \mathrm{id}_{\mathcal{A}} \otimes \Lambda$. Choosing $\sigma := \rho_{\mathrm{A}} \otimes \rho_{\mathrm{BC}}$ we have $\Phi\,(\sigma) = \rho_{\mathrm{A}} \otimes \rho_{\mathrm{B}}$, and therefore we obtain

$$\mathrm{I}(\mathrm{A} : \mathrm{B}) \leq \mathrm{I}(\mathrm{A} : \mathrm{BC}). \tag{2.41}$$

The information theoretic interpretation is clear: Information about the composed system BC reveals more information about system A than the knowledge of only B.

Writing out both sides of (2.41) we obtain

$$S_{\mathrm{A}} + S_{\mathrm{B}} - S_{\mathrm{AB}} \leq S_{\mathrm{A}} + S_{\mathrm{BC}} - S_{\mathrm{ABC}},$$

or, after rearrangement

$$S_{\mathrm{ABC}} + S_{\mathrm{B}} \leq S_{\mathrm{AB}} + S_{\mathrm{BC}}. \tag{2.42}$$

Inequality (2.42) is called the strong subadditivity of entropy. Note that with the choice $\mathcal{B} := \mathbb{C}$, (SSA) reduces to the subadditivity of entropy.

2.5 Information theoretical questions

In the previous section entropy was introduced as a measure of the information content of a state; however, it is not yet clear why to favour entropy to other possible candidates for a measure of information. In this section two fundamental examples are presented to justify this choice; the distinguishability problem of quantum states, and the (classical) information carrying capacity of quantum channels. A third one is presented in Chapter 6; it gives an interpretation of entropy as the best possible compressibility rate for a sequence of signals.

2.5.1 Quantum hypothesis testing

A fundamental feature of probabilistic treatment of physical systems is that one can never learn the actual state of the system by making measurements (even if we accept that the state of the system represents some physical reality and is not only an entirely subjective notion, emerging from our limited knowledge). However, a preliminary knowledge on the possible set of states can help a lot of making a fairly good guess about the real state of the system; this is the subject of the theory of hypothesis testing.

The most basic setting is when we know that the state of the system is one of two given states, ρ and σ. Our initial guess is that the true state is ρ (this is called the null-hypothesis, and is denoted by H_0), and we want to check it against the counter-hypothesis H_1, that the true state is σ. To decide, we make a binary measurement with measurement operators M_0 and M_1, and reject the null-hypothesis if the outcome is 1. There are two kinds of wrong decisions we can make. The error of the first kind is to reject H_0 when it is actually true; the probability of this error is $\alpha := \rho(M_1)$. The error of the second kind is to keep H_0 when it is false; the probability now is $\beta := \sigma(M_0)$. The aim is to find a measurement that keeps these probabilities under some given value.

Let's first consider the case when $\rho = |\psi\rangle\langle\psi|$ and $\sigma = |\varphi\rangle\langle\varphi|$ are pure states of a quantum system. If the two states are orthogonal then we can distinguish them perfectly; the measurement $M_0 := \rho$ and $M_1 := \mathbb{1} - M_0$ yields $\alpha = \beta = 0$. It is easily seen that one of the error probabilities is always bounded away from 0 if the two states are not orthogonal. However, in this situation we can use the fact that ψ is 'asymptotically orthogonal' to φ. Suppose we are provided with several copies of the same system; the joint state of n copies is either $\rho^{\otimes n}$ or $\sigma^{\otimes n}$. The measurement $M_0 := \rho^{\otimes n}$ and $M_1 := \mathbb{1} - M_0$ gives $\alpha_n = 0$ and $\beta_n = |\langle\varphi,\psi\rangle|^{2n} = \exp\left(n \log|\langle\varphi,\psi\rangle|^2\right)$, i.e. the error of the second kind is not zero, but tends to zero exponentially, when the number of copies is increased.

When perfect distinguishability is not possible then it is standard to impose a bound ε on the error of the first kind, and look for an asymptotically optimal measurement, that minimizes the error of the second kind. The fundamental result in this setting is the following:

2.5.1. Theorem. Let $\rho \ll \sigma$ be states of a quantum system, and let

$$\beta_n(\varepsilon) := \inf\{\sigma^{\otimes n}(M) \,:\, \rho^{\otimes n}(\mathbb{1} - M) < \varepsilon, \quad 0 \leq M \leq \mathbb{1}\}. \tag{2.43}$$

Then

$$\lim_{n\to\infty} \frac{1}{n} \log \beta_n(\varepsilon) = -S(\rho,\sigma). \tag{2.44}$$

The infimum in (2.43) is easily seen to be a minimum, and the statement of the theorem is that for the minimizing measurement the error of the second kind goes exponentially to zero, with exponent $-S(\rho,\sigma)$, when an n-independent upper bound is imposed on the error of the first kind. Note that the exponent doesn't depend on ε. To keep both kinds of errors under given bounds ε and δ, we need approximately

$$-\frac{1}{S(\rho,\sigma)} \log \delta \tag{2.45}$$

copies of the system. That is, the smaller $S(\rho,\sigma)$ is, the more copies we need to reach a given precision in our decision process. This result motivates to view relative entropy as a measure of statistical distinguishability of states.

Note that relative entropy is an asymmetric 'distance' on the state space; its emergence here is related to the asymmetric way H_0 and H_1 is treated.

2.5.2 Capacity of quantum channels

Suppose two parties (Alice and Bob) want to communicate classical information through a quantum channel, described by a stochastic map Φ from $\mathcal{B}(\mathcal{H}_{\text{in}})$ to $\mathcal{B}(\mathcal{H}_{\text{out}})$. The set of possible messages sent through the channel forms a finite set $\mathcal{X}$, and the following protocol is used:

(i) Alice and Bob agree on a set of states $\mathcal{C} := \{\rho_x\}$, the elements of which are called the codewords for the message.

(ii) Bob fixes a set of positive operators $\mathcal{D} := \{M_x\}$ such that $\sum_x M_x \leq I_{\mathcal{H}_{\mathrm{out}}}$, and introduces $M_0 := I_{\mathcal{H}_{\mathrm{out}}} - \sum_x M_x$. The set $\mathcal{Y}$ is defined as $\mathcal{X} \cup \{0\}$.

(iii) When Alice wants to send the message x to Bob, she sends through the channel the state ρ_x. Bob then makes the measurement $\mathcal{M} := \mathcal{D} \cup \{M_0\}$ on $\Phi(\rho_x)$, and identifies the message to be y if the outcome of the measurement is $y \neq 0$, or neglects it, if the outcome is 0. Elements of $\mathcal{M}$ are called decoding operators.

The probability that Bob identifies the outcome as $y \in \mathcal{Y}$ when the message x was sent is $p(y|x) := \rho_x(M_y)$, therefore the probability of wrong identification is $e(x) = 1 - \rho_x(M_x)$. We can also assume that Alice chooses the message x according to a probability distribution $\{p(x)\}$; the average error probability is then $e = \sum_x p(x)e(x) = 1 - \sum_x p(x)\,\rho_x(M_x)$, which is obviously bounded from above by the maximal error probability $\hat{e} := \max_x e(x)$.

Monotonicity of the Holevo quantity (Example (2.4.7) implies that the classical mutual information Alice and Bob can share is always bounded from above by the Holevo quantity:

$$\begin{aligned} I = \sum_{xy} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} &\leq \chi\left(\{p(x), \Phi(\rho_x)\}\right) \\ &= S\left(\sum_x p(x)\,\Phi(\rho_x)\right) - \sum_x p(x) S\left(\Phi(\rho_x)\right) , \end{aligned}$$

where $p(x,y) = p(x)p(y|x)$ and $p(y) = \sum_x p(x,y)$. Note that the left-hand side depends both on the encoding $\mathcal{C}$ and the decoding $\mathcal{D}$, while the right-hand side depends only on the encoding. This inequality is in general strict, a saturating measurement exists only under very special circumstances. However, the upper bound is asymptotically achievable, if Alice and Bob can use the channel several times.

Suppose Alice and Bob can use the channel n times, i.e. Φ is replaced by $\Phi^{\otimes n}$, the codewords are states on $\mathcal{B}(\mathcal{H}_{\mathrm{in}})^{\otimes n}$, while the measurement operators are in $\mathcal{B}(\mathcal{H}_{\mathrm{out}})^{\otimes n}$. The set of messages to transmit is $\mathcal{X}_n$, with cardinality $|\mathcal{X}_n|$. We slightly change the notation, and replace the ensemble $\{p_x, \rho_x\}$ by a discrete distribution μ_n on the set of density matrices that are in the range of

$\Phi^{\otimes n}$, and denote the corresponding Holevo quantity by $\chi(\mu_n)$. The classical mutual information corresponding to an encoding μ and a measurement $\mathcal{M}_n$ is denoted by $I(\mu, \mathcal{M}_n)$. We define the following quantities:

$$\begin{aligned} C_n(\Phi) &:= \sup_{\{\mu_n, \mathcal{M}_n\}} \{I(\mu_n, \mathcal{M}_n)\} \qquad \text{and} \\ \chi_n(\Phi) &:= \sup_{\{\mu_n\}} \{\chi(\mu_n)\}. \end{aligned} \tag{2.46}$$

The Holevo bound implies $C_n(\Phi) \leq \chi_n(\Phi)$ for all n. Superadditivity of these quantities follow by taking product encoding and decoding with respect to the splitting $\mathcal{B}(\mathcal{H}_{\text{in}})^{\otimes(n+m)} = \mathcal{B}(\mathcal{H}_{\text{in}})^{\otimes n} \otimes \mathcal{B}(\mathcal{H}_{\text{in}})^{\otimes m}$, and the same for $\mathcal{H}_{\text{out}}$; that is

$$\begin{aligned} C_n(\Phi) + C_m(\Phi) &\leq C_{n+m}(\Phi) \qquad \text{and} \\ \chi_n(\Phi) + \chi_m(\Phi) &\leq \chi_{n+m}(\Phi). \end{aligned}$$

Hence the following limits exist:

$$\begin{aligned} \overline{C}(\Phi) &:= \lim_{n\to\infty} \frac{1}{n} C_n(\Phi) = \sup_n \frac{1}{n} C_n(\Phi) \qquad \text{and} \\ \overline{\chi}(\Phi) &:= \lim_{n\to\infty} \frac{1}{n} \chi_n(\Phi) = \sup_n \frac{1}{n} \chi_n(\Phi), \end{aligned} \tag{2.47}$$

and clearly $\overline{C}(\Phi) \leq \overline{\chi}(\Phi)$.

Now the aim of Alice and Bob is to maximize the amount of messages they can transmit per use of the channel with an asymptotically vanishing probability of error. The quantity

$$R := \limsup_{n\to\infty} \frac{1}{n} \log|\mathcal{X}_n| \tag{2.48}$$

is called the **rate** of the communication process. A rate R is called **achievable** if there exists a sequence $(\mathcal{C}_n, \mathcal{D}_n)$ of coding-decoding pairs such that its rate is R, and the maximal probability of error goes to zero. The supremum of all achievable rates is denoted by $\overline{R}$, and is called the **information carrying capacity** of the channel. The fundamental theorem of quantum data transmission (the Holevo-Schumacher-Westmoreland theorem ([53] and [100]) states

$$\overline{R} = \overline{C}(\Phi) = \overline{\chi}(\Phi). \tag{2.49}$$

Roughly speaking, the information carrying capacity of a quantum channel equals to the mutual information (quantum= $\overline{\chi}\,(\Phi)$, or classical= $\overline{C}\,(\Phi)$) attained by the best choice for coding and decoding.

For classical channels $C_n = n\,C_1 \;\; \forall n$, and therefore to compute the capacity of the channel it is enough to maximize the mutual information for one single use of the channel, which is clearly much easier from a computational point of view then to perform the task for all n, and take the limit (2.47). The question whether the same holds for the capacity of a quantum channel is still open, and would follow from the more general additivity property

$$\chi(\Phi_1 \otimes \Phi_2) = \chi(\Phi_1) + \chi(\Phi_2)\,, \tag{2.50}$$

where Φ_1 and Φ_2 are arbitrary quantum channels, and $\chi(\Phi) := \chi_1(\Phi)$. The proof of this conjecture seems to be one of the biggest challenges of quantum information theory, and was shown to be equivalent [72, 8, 103, 104] to other important additivity questions, like the additivity of the entanglement of formation (see section 2.4.2) and the minimal entropy output (see Chapter 3). If Alice and Bob are only allowed to use product operators for encoding and decoding, then the information carrying capacity of the channel is simply $\chi(\Phi)$, which may be strictly smaller than $\overline{\chi}(\Phi)$ if additivity doesn't hold. Therefore the additivity question can equivalently be formulated by asking whether entanglement in the encoding and decoding operators can increase the capacity of the channel or not.

References

The standard reference on quantum information theory is [81]; different distance measures are treated in Chapter 9. The results about decompositions of states can be found in [80]. An introductory survey on entanglement theory is [56]. Properties of quantum entropy are studied in the monograph [83]; information theoretic interpretations are given in [81]. Strong subadditivity of entropy was proven in [69], the proof of the monotonicity property of relative entropy appeared in [71] and [108]; an overview can be found in [91] and [96]. Various generalizations of Theorem 2.5.1 are treated in [46] and [18]. For an overview of additivity problems of quantum channel capacity, see http://www.imaph.nat.tu-bs.de/qi/problems/10.html.

Chapter 3

Covariant channels and an additivity problem

Symmetry properties can simplify computation of certain quantities of quantum channels. In this chapter we study the case when the channels are invariant under representations of the unitary group of the input Hilbert space. First a general overview on the structure of the set of such channels is given, after which we turn our attention to the case of the conjugate representation of the unitary group, and identify the set of invariant channels as a one-parameter family Λ_t, which we call transpose depolarizing channels. The range of the parameter t is a function of the dimension d of the Hilbert space: $-1/(d-1) \leq t \leq 1/(d+1)$. Using stochastic majorization among the output density matrices, we show the additivity of minimal entropy output for two copies of the same channel, for the restricted range of parameters $-2/(d^2-2) \leq t \leq 1/(d+1)$.

3.1 Minimal entropy output

A quantum channel Φ is a CPTP (completely positive, trace-preserving) map that maps states of the input Hilbert space $\mathcal{H}_{\text{in}}$ into states of the output Hilbert space $\mathcal{H}_{\text{out}}$. The minimal entropy output of the channel is

$$S_{\min}(\Phi) := \min\big\{\ S\big(\Phi(\rho)\big)\ :\ \rho \in \mathcal{S}(\mathcal{H}_{\text{in}})\big\},$$

where $\mathcal{S}(\mathcal{H}_{\text{in}})$ denotes the state space of the input system, i.e. the set of positive semi-definite operators with trace one, and $S(\sigma) := -\operatorname{Tr}\sigma\log\sigma$ is the von Neumann entropy of the density operator σ. By concavity of the entropy, $S_{\min}$ can be expressed as

$$S_{\min}(\Phi) := \min\big\{ S\big(\Phi(|\psi\rangle\langle\psi|)\big) : \psi \in \mathcal{H}_{\text{in}},\ \|\psi\| = 1\big\}.$$

Hence, $S_{\min}$ is a lower bound on the mixedness of the output state, measured by entropy, when the input is a pure state. It can therefore be interpreted as a measure of the noisiness of the channel.

A fundamental question of quantum information theory is whether the minimal entropy output is additive for product channels, i.e., given two channels $\Phi_1 : \mathcal{S}(\mathcal{H}^1_{\text{in}}) \to \mathcal{S}(\mathcal{H}^1_{\text{out}})$ and $\Phi_2 : \mathcal{S}(\mathcal{H}^2_{\text{in}}) \to \mathcal{S}(\mathcal{H}^2_{\text{out}})$, whether the equality

$$S_{\min}(\Phi_1 \otimes \Phi_2) = S_{\min}(\Phi_1) + S_{\min}(\Phi_2) \tag{3.1}$$

holds. This question was shown to be equivalent to the additivity of other quantities, like the Holevo capacity and the entanglement of formation [72, 8, 103, 104]. It has been solved in the special cases when one of the channels is the identical channel [5], the depolarizing channel [65], a unital qubit channel [64] or an entanglement breaking channel [102, 66].

The inequality

$$S_{\min}(\Phi_1 \otimes \Phi_2) \leq S_{\min}(\Phi_1) + S_{\min}(\Phi_2)$$

holds trivially, since $S\big(\Phi_1 \otimes \Phi_2(|\psi_1 \otimes \psi_2\rangle\langle\psi_1 \otimes \psi_2|)\big) = S_{\min}(\Phi_1) + S_{\min}(\Phi_2)$ for the states $|\psi_1\rangle\langle\psi_1|$ and $|\psi_2\rangle\langle\psi_2|$ at which the output entropies of Φ_1 and Φ_2 are minimized. Therefore, in order to show additivity, we have to show that $S\big(\Phi_1 \otimes \Phi_2(|\psi\rangle\langle\psi|)\big)$ attains its minimal value at $\psi = \psi_1 \otimes \psi_2$. Hence, the additivity question can be rephrased as whether entanglement has any influence on the minimal entropy output of a tensor product channel; if additivity holds, then the answer is no.

Using the Schmidt decomposition of ψ, we have to minimize

$$S\Bigg(\sum_{ij} \sqrt{\lambda_i \lambda_j}\, \Phi_1(|e_i\rangle\langle e_j|) \otimes \Phi_2(|f_i\rangle\langle f_j|)\Bigg), \tag{3.2}$$

where $\{e_i : i = 1, 2, \dots, \dim \mathcal{H}^1_{\text{in}}\}$ and $\{f_j : j = 1, 2, \dots, \dim \mathcal{H}^2_{\text{in}}\}$ are orthonormal bases of the input spaces $\mathcal{H}^1_{\text{in}}$ and $\mathcal{H}^2_{\text{in}}$ and where $\boldsymbol{\lambda} = \{\lambda_i : i = 1, \dots, \big(\dim \mathcal{H}^1_{\text{in}}\big) \wedge \big(\dim \mathcal{H}^2_{\text{in}}\big)\}$ is a probability distribution. Solving this problem in full generality is rather hopeless, however, it simplifies a great deal if the channels Φ_1 and Φ_2 possess the additional property of unitary covariance.

3.1.1. Definition. A channel $\Phi : \mathcal{S}(\mathcal{H}_{\rm in}) \to \mathcal{S}(\mathcal{H}_{\rm out})$ is unitarily covariant, if for every unitary U in $\mathcal{H}_{\rm in}$ there exists a unitary V in $\mathcal{H}_{\rm out}$ such that

$$\Phi \circ \mathrm{Ad}_U = \mathrm{Ad}_V \circ \Phi \tag{3.3}$$

holds, where Ad_X is the map $A \mapsto XAX^*$.

Note that the map Ad_V is unchanged when V is multiplied by a phase factor, and therefore the notion of unitary covariance remains unchanged if we require both U and V to have unit determinant.

Obviously, for unitarily covariant channels $S\big(\Phi(|\psi\rangle\langle\psi|)\big)$ is independent of the input state $|\psi\rangle\langle\psi|$ and is therefore equal to $S_{\min}(\Phi)$.

A significant simplification of the additivity problem for unitarily covariant maps is that the entropy in (3.2) depends only on the Schmidt coefficients of the input state ψ. Indeed, if ψ and φ have the same Schmidt coefficients, then there exist unitaries U_1 and U_2 in $\mathcal{H}_{\rm in}^1$ and $\mathcal{H}_{\rm in}^2$ such that $\varphi = U_1 \otimes U_2\, \psi$. Therefore

$$\Phi_1 \otimes \Phi_2(|\varphi\rangle\langle\varphi|) = (V_1 \otimes V_2)\, \Phi_1 \otimes \Phi_2(|\psi\rangle\langle\psi|)(V_1^* \otimes V_2^*)$$

for some unitaries V_1 and V_2 in $\mathcal{H}_{\rm out}^1$ and $\mathcal{H}_{\rm out}^2$, hence $\Phi_1 \otimes \Phi_2(|\varphi\rangle\langle\varphi|)$ and $\Phi_1 \otimes \Phi_2(|\psi\rangle\langle\psi|)$ have the same entropy. Now to prove additivity, it is sufficient to show that the function

$$\boldsymbol{\lambda} \mapsto S\Bigg(\sum_{ij} \sqrt{\lambda_i \lambda_j}\, \Phi_1(|e_i\rangle\langle e_j|) \otimes \Phi_2(|f_i\rangle\langle f_j|)\Bigg)$$

reaches its minimal value at the vertices of the simplex $\mathcal{S}$. Here, $\{e_i\}$ and $\{f_j\}$ are fixed orthonormal bases in $\mathcal{H}_{\rm in}^1$ and $\mathcal{H}_{\rm in}^2$ and $\mathcal{S}$ is the simplex of probability distributions on $\big(\dim \mathcal{H}_{\rm in}^1\big) \wedge \big(\dim \mathcal{H}_{\rm in}^2\big)$ points.

A trivial example of a unitarily covariant channel is the trace map

$$\tau(A) := (\mathrm{Tr} A)\, \frac{1}{\mathrm{d}_{\rm out}}\, \mathbb{1}_{\rm out},$$

where $\mathbb{1}_{\rm out}$ is the identity on $\mathcal{H}_{\rm out}$ and $\mathrm{d}_{\rm out} := \dim \mathcal{H}_{\rm out}$. In this case, the covariance relation (3.3) holds for any pair of unitaries (U, V). A more interesting situation is when V can be chosen such that $U \mapsto V$ gives a representation α of the group $\mathcal{G} = SU(\mathrm{d}_{\rm in})$ on $\mathcal{H}_{\rm out}$. More precisely, we are interested in channels that satisfy the following property:

3.1.2. Definition. A channel $\Phi : \mathcal{S}(\mathcal{H}_{\text{in}}) \to \mathcal{S}(\mathcal{H}_{\text{out}})$ is α-invariant, if

$$\Phi \circ \mathrm{Ad}_U = \mathrm{Ad}_{\alpha_{\text{out}}(U)} \circ \Phi$$

for all $U \in SU(\mathrm{d}_{\text{in}})$.

Such α-invariant channels were studied in [55], where it was also shown that for an irreducible representation α_{out}

$$\chi(\Phi) = \log \mathrm{d}_{\text{out}} - S_{\min}(\Phi).$$

Here, $\chi(\Phi)$ is the Holevo capacity of the channel (see section 2.5.2), given by the formula

$$\chi(\Phi) = \max\Big\{ S\Big(\sum_j p_j \Phi(\rho_j)\Big) - \sum_j p_j \, S\big(\Phi(\rho_j)\big)\Big\},$$

where the maximum is taken over all finitely supported measures $\mathbf{p}$ on the state space $\mathcal{S}(\mathcal{H}_{\text{in}})$ and $\mathbf{p}(\rho_j) =: p_j$. Therefore, for such channels, if additivity holds for the minimal entropy output it also holds for the Holevo capacity.

In section 3.2, we deal with the description of the set $\mathcal{B}_\alpha$ of α-invariant channels, focusing mainly on the case $\mathcal{H}_{\text{out}} = \mathcal{H}_{\text{in}}$ and α_{out} is either

(a) the identical representation, or

(b) the conjugate representation $U \mapsto \overline{U}$.

$\overline{U}$ is the entrywise conjugate of the matrix U in some fixed base. It turns out that

- in case (a), $\mathcal{B}_\alpha$ is the family of depolarizing channels, given by

  $$\Delta_t(A) := tA + (1-t)\tau(A), \qquad -\frac{1}{\mathrm{d}^2 - 1} \le t \le 1,$$

  where $\mathrm{d} = \mathrm{d}_{\text{in}} = \mathrm{d}_{\text{out}}$ and $[-1/(\mathrm{d}^2-1), 1]$ is the maximal range of t such that Δ_t is completely positive, while

- in case (b), we obtain the family

  $$\Lambda_t(A) := tA^T + (1-t)\tau(A), \qquad -\frac{1}{\mathrm{d}-1} \le t \le \frac{1}{\mathrm{d}+1},$$

  where A^T is the transpose of A in the same base in which the U's are conjugated.

The elements of this second class will be called transpose depolarizing channels.

Additivity of the minimal entropy output was shown for the depolarizing channels in [65], even when one of the channels is completely arbitrary, and for the extreme transpose depolarizing channel $\Lambda_{-1/(\mathrm{d}-1)}$ in [73], followed by proofs using different methods in [25] and [3]. In section 3.3, we shall show additivity for the subset of the transpose depolarizing channels determined by $-2/(\mathrm{d}^2-2) \leq t \leq 1/(\mathrm{d}+1)$.

3.2 Extreme invariant maps

We start out with the more general setting of a pair α of unitary representations α_{in} and α_{out} of an abstract group $\mathcal{G}$ on $\mathcal{H}_{\mathrm{in}}$ and $\mathcal{H}_{\mathrm{out}}$ respectively and define a map Φ to be α-invariant, if

$$\Phi \circ \mathrm{Ad}_{\alpha_{\mathrm{in}}(g)} = \mathrm{Ad}_{\alpha_{\mathrm{out}}(g)} \circ \Phi, \qquad g \in G. \tag{3.4}$$

Next, we fix a basis $\{e_i \,:\, i = 1, 2, \ldots, d_{\mathrm{in}}\}$ in $\mathcal{H}_{\mathrm{in}}$ with corresponding matrix units $e_{ij} := |e_i\rangle\langle e_j|$. The anti-unitary map J is defined by the property $J\, e_i = e_i$. The Choi matrix of Φ in this base is

$$C(\Phi) = \sum_{ij} e_{ij} \otimes \Phi(e_{ij}) \tag{3.5}$$

(see Appendix A.4 for details). A straightforward computation shows that Φ is trace-preserving if and only if $\mathrm{Tr}_2\, C(\Phi) = \mathbb{1}_{\mathrm{in}}$, and

$$\begin{aligned}
&\text{(i)} \quad C(\mathrm{Ad}_V \circ \Phi) = \mathrm{Ad}_{\mathbb{1}\otimes V}\big(C(\Phi)\big) = (\mathbb{1} \otimes V)\ C(\Phi)\ (\mathbb{1} \otimes V^*)\\
&\text{(ii)} \quad C(\Phi \circ \mathrm{Ad}_U) = \mathrm{Ad}_{JU^*J^*\otimes\mathbb{1}}\big(C(\Phi)\big)\\
&\qquad\qquad\qquad\quad = (JU^*J^* \otimes \mathbb{1})\ C(\Phi)\ (JUJ^* \otimes \mathbb{1}).
\end{aligned}$$

Properties (i) and (ii) imply that Φ is α-invariant if and only if

$$\big[C(\Phi)\,,\, J\alpha_{\mathrm{in}}(g)J^* \otimes \alpha_{\mathrm{out}}(g)\big] = 0, \qquad g \in \mathcal{G}.$$

The map $g \mapsto \kappa(g) := J\alpha_{\mathrm{in}}(g)J^* \otimes \alpha_{\mathrm{out}}(g)$ is again a unitary representation of $\mathcal{G}$, which decomposes into irreducible components

$$\kappa(g) \cong \bigoplus_k \kappa_k(g) \otimes \mathbb{1}_{\mathrm{m}(k)},$$

where k labels inequivalent irreducible representations of $\mathcal{G}$ on a $\mathrm{d}(k)$ dimensional Hilbert space and where $\mathrm{m}(k)$ is the multiplicity of the k-th representation. A general element of the commutant of $\kappa\big(SU(\mathrm{d_{in}})\big)$ is of the form

$$C = \bigoplus_k \mathbb{1}_{\mathrm{d}(k)} \otimes A_k,$$

and it is the Choi matrix of a CP map if and only if all $A_k \geq 0$. In the special case $\mathrm{m}(k) = 1\ \forall k$, we obtain that the map Φ is α-invariant if and only if its Choi matrix has the form

$$C(\Phi) = \sum_k c_k\, P_k, \tag{3.6}$$

where P_k is the projection onto the Hilbert space of the k-th irreducible component and where the c_k's are arbitrary complex numbers. Obviously, Φ is CP if and only if all the coefficients c_k are nonnegative. The maps $\mathcal{T}^* P_k$ span the extreme rays of the convex cone of α-invariant CP maps. However, $\mathcal{T}^* P_k$ need not be trace-preserving, therefore formula (3.6) is in general insufficient to describe the convex set $\mathcal{B}_\alpha$ of α-invariant CPTP maps.

Note that the map

$$\frac{1}{\mathrm{d_{in}}}\,\mathcal{T} : \Phi \mapsto \frac{1}{\mathrm{d_{in}}}\, C(\Phi)$$

gives an embedding of the convex set of CPTP maps into the state space of $\mathcal{H}_{\mathrm{in}} \otimes \mathcal{H}_{\mathrm{out}}$, its range, however, does not coincide with the whole of the state space because of the restriction $\mathrm{Tr}_2\, C(\Phi) = \mathbb{1}_{\mathrm{in}}$. Thus α-invariant CPTP maps can be identified with a subset of the set $\mathcal{S}_\alpha$ of α-invariant states, i.e. states that satisfy $\big[\rho\,,\ J\alpha_{\mathrm{in}}(g)J^* \otimes \alpha_{\mathrm{out}}(g)\big] = 0\ \forall g \in \mathcal{G}$. The characterization of $\mathcal{S}_\alpha$ for different symmetry groups was studied in [111].

Even if in general $\mathcal{S}_\alpha$ is strictly larger then the embedded image of $\mathcal{B}_\alpha$, both sets coincide when α_{in} is irreducible since the invariance

$$\begin{aligned}\mathrm{Tr}_2\, \rho &= \mathrm{Tr}_2\big(J\alpha_{\mathrm{in}}(g)J^* \otimes \alpha_{\mathrm{out}}(g)\big)\, \rho\, \big(J\alpha_{\mathrm{in}}(g)^* J^* \otimes \alpha_{\mathrm{out}}(g)^*\big)\\ &= \big(J\alpha_{\mathrm{in}}(g)J^*\big)\, \big(\mathrm{Tr}_2 \rho\big)\, \big(J\alpha_{\mathrm{in}}(g)^* J^*\big)\end{aligned}$$

implies $\mathrm{Tr}_2\, \rho = \frac{1}{\mathrm{d_{in}}} \mathbb{1}_{\mathrm{in}}$. Hence in this case the extremal points of $\mathcal{B}_\alpha$ can be identified as $\frac{\mathrm{d_{in}}}{\mathrm{d}(k)}\, \mathcal{T}^* P_k$, where $d(k) = \mathrm{rank}\ P_k$.

Now we turn to the case $\mathcal{H}_{\text{in}} = \mathcal{H}_{\text{out}} = \mathcal{H}$, α_{in} the identical representation of $\mathcal{G} = SU(\mathrm{d})$ and either $\alpha_{\text{out}}(U) = U$, case (a), or $\alpha_{\text{out}}(U) = \overline{U} = JUJ^*$, case (b).
In case (a), the invariant projections of $JUJ^* \otimes U$ are $P := |\psi\rangle\langle\psi|$ and $P^\perp$, where $\psi = \frac{1}{\sqrt{\mathrm{d}}} \sum_i e_i \otimes e_i$. The corresponding extreme maps are

$$\mathcal{T}^*(\mathrm{d}P) = \Delta_1 = \mathrm{id} \qquad \text{and} \qquad \mathcal{T}^*\Big(\frac{\mathrm{d}}{\mathrm{d}^2 - 1}\, P^\perp\Big) = \Delta_{-1/(\mathrm{d}^2-1)}.$$

In case (b), the invariant projections are P_{s} and P_{a}, the projections onto the symmetric and the antisymmetric subspaces of $\mathcal{H} \otimes \mathcal{H}$. The corresponding extreme maps are

$$\mathcal{T}^*\Big(\frac{2}{\mathrm{d}+1}\, P_{\mathrm{s}}\Big) = \Lambda_{1/(\mathrm{d}+1)} \qquad \text{and} \qquad \mathcal{T}^*\Big(\frac{2}{\mathrm{d}-1}\, P_{\mathrm{a}}\Big) = \Lambda_{-1/(\mathrm{d}-1)}.$$

For details on decomposing tensor products of representations of $SU(\mathrm{d})$, see e.g. [60].

3.3 The transpose depolarizing channel

In this section, we shall prove that

$$S_{\min}\big(\Lambda_t \otimes \Lambda_t\big) = 2\ S_{\min}\big(\Lambda_t\big), \qquad -\frac{2}{\mathrm{d}^2 - 2} \le t \le \frac{1}{\mathrm{d}+1}. \tag{3.7}$$

First note that the partial transpose of the state $\frac{1}{\mathrm{d}}C\big(\Lambda_t\big)$ is $\frac{1}{\mathrm{d}}C\big(\Delta_t\big)$, hence it is positive if and only if $t \in \big[-1/(\mathrm{d}^2-1)\,,\, 1/(\mathrm{d}+1)\big] =: R_1$. Having positive partial transpose is a necessary condition for a bipartite state to be separable [85] and was shown in [111] to be also sufficient for unitarily covariant states . Quantum channels with separable Choi matrix are called entanglement-breaking, and additivity of their minimal entropy output was proven in [102]. It holds even in the stronger form where one of the factors in (3.1) is entanglement-breaking and the other is completely arbitrary. Hence, additivity of minimal entropy output holds in the strong form

$$S_{\min}\big(\Lambda_t \otimes \Phi\big) = S_{\min}\big(\Lambda_t\big) + S_{\min}\big(\Phi\big), \tag{3.8}$$

when $t \in R_1$ and Φ is an arbitrary quantum channel. We can therefore assume t to be negative for the rest of the argument, even though the proof applies with a trivial modification to positive t's as well.

Let $\rho(\boldsymbol{\lambda}) := |\psi(\boldsymbol{\lambda})\rangle\langle\psi(\boldsymbol{\lambda})| = \sum_{ij} \sqrt{\lambda_i\lambda_j}\, |e_i\rangle\langle e_j| \otimes |e_i\rangle\langle e_j|$ denote the input state and $X(\boldsymbol{\lambda}) := \Lambda_t \otimes \Lambda_t\big(\rho(\boldsymbol{\lambda})\big)$ the output state for some fixed $t \leq 0$. An easy computation shows that

$$X(\boldsymbol{\lambda}) = t^2\,\rho(\boldsymbol{\lambda}) + \frac{t(1-t)}{\mathrm{d}}\big(\rho'(\boldsymbol{\lambda}) \otimes \mathbb{1} + \mathbb{1} \otimes \rho'(\boldsymbol{\lambda})\big) + \Big(\frac{1-t}{\mathrm{d}}\Big)^2\, \mathbb{1} \otimes \mathbb{1},$$

where $\rho'(\boldsymbol{\lambda}) := \mathrm{Tr}_2\, \rho(\boldsymbol{\lambda}) = \mathrm{Tr}_1\, \rho(\boldsymbol{\lambda})$. The basis $\{e_i\}$ is the one in which the transpose is taken.

The vectors $\{e_i \otimes e_j\ :\ i \neq j\}$ are eigenvectors of $X(\boldsymbol{\lambda})$ with corresponding eigenvalues $\eta_{ij} = t(1-t)(\lambda_i + \lambda_j)/\mathrm{d} + \big((1-t)/\mathrm{d}\big)^2$. The contribution of this part of the spectrum to the entropy is $S_1(\boldsymbol{\lambda}) = -\sum_{i\neq j} \eta_{ij} \ln \eta_{ij}$, which, being a concave function of $\boldsymbol{\lambda}$, takes its minimal value at the vertices of $\mathcal{S}$.

Now we consider the subspace spanned by $\{e_i \otimes e_i\ :\ i = 1, \ldots, \mathrm{d}\}$. The restricted output density has the form

$$\hat{X}(\boldsymbol{\lambda}) = t^2\, |\sqrt{\boldsymbol{\lambda}}\rangle\langle\sqrt{\boldsymbol{\lambda}}| + \frac{2t(1-t)}{\mathrm{d}}\mathrm{diag}(\boldsymbol{\lambda}) + \Big(\frac{1-t}{\mathrm{d}}\Big)^2 \mathbb{1} \otimes \mathbb{1},$$

where $|\sqrt{\boldsymbol{\lambda}}\rangle = \sum_j \sqrt{\lambda_j}\, e_j$ and where $\mathrm{diag}(\boldsymbol{\lambda})$ is the diagonal matrix with diagonal $\boldsymbol{\lambda}$. Note that, at the vertices of $\mathcal{S}$, $\lambda_j = \delta_{jk}$ for some k and that in this case $\hat{X}(\boldsymbol{\lambda})$ is diagonal in the basis $\{e_i \otimes e_i\}$. Let $S_2(\boldsymbol{\lambda})$ denote the entropy contribution of the eigenvalues of $\hat{X}(\boldsymbol{\lambda})$. In order to show that it takes its minimal value at the vertices of $\mathcal{S}$ it is sufficient to show that $\hat{X}(\boldsymbol{\lambda})$ is more mixed than $\hat{X}(\boldsymbol{\lambda}^*)$ for any $\boldsymbol{\lambda} \in \mathcal{S}$, where $\lambda^*_j = \delta_{j,1}$.

Recall that, for $n \times n$ hermitian matrices, A_1 is called more mixed than A_2, in notation $A_1 \succ A_2$, if

$$\begin{aligned}
\kappa_1(A_1) &\leq \kappa_1(A_2)\\
\kappa_1(A_1) + \kappa_2(A_1) &\leq \kappa_1(A_2) + \kappa_2(A_2)\\
&\ \vdots\\
\kappa_1(A_1) + \ldots + \kappa_{n-1}(A_1) &\leq \kappa_1(A_2) + \ldots + \kappa_{n-1}(A_2)\\
\mathrm{Tr}\, A_1 &= \mathrm{Tr}\, A_2.
\end{aligned}$$

Here, $\kappa_1(A_k) \geq \kappa_2(A_k) \geq \ldots$ are the eigenvalues of A_k in decreasing order. Note that $\sum_{i\neq j} \eta_{ij} = (1-t^2)(\mathrm{d}-1)/2\mathrm{d}$ for any $\boldsymbol{\lambda}$ and that therefore $\mathrm{Tr}\hat{X}(\boldsymbol{\lambda}) = 1 - \sum_{i\neq j} \eta_{ij}$ is independent of $\boldsymbol{\lambda}$. It is also clear that the $\boldsymbol{\lambda}$-independent diagonal term $\big((1-t)/\mathrm{d}\big)^2 \mathbb{1}\otimes\mathbb{1}$ doesn't influence the more mixedness relation. The eigenvalue vector of $\hat{X}(\boldsymbol{\lambda}^*) - \big((1-t)/\mathrm{d}\big)^2 \mathbb{1}\otimes\mathbb{1}$, arranged in decreasing order, is $\big(0, 0, \dots, 0, t(2 + (d-2)t)/\mathrm{d}\big)$. We have therefore to show that $\hat{X}(\boldsymbol{\lambda}) - \big((1-t)/\mathrm{d}\big)^2 \mathbb{1}\otimes\mathbb{1}$ is negative semidefinite. Writing out explicitly the corresponding quadratic form, this criterion becomes

$$\left|\sum_j \sqrt{\lambda_j}\, x_j\right|^2 \leq -\frac{2(1-t)}{t\mathrm{d}} \sum_j \lambda_j\, |x_j|^2, \qquad (x_1, \dots, x_{\mathrm{d}}) \in \mathbb{C}^{\mathrm{d}},$$

which, by the Schwarz inequality, holds if and only if $\mathrm{d} \leq -2(1-t)/t\mathrm{d}$, that is, $-2/(\mathrm{d}^2 - 2) \leq t$.

The case $\mathrm{d} = 2$ is special in the sense that the conjugate representation of $SU(\mathrm{d})$ is unitarily equivalent to the identical representation. As a consequence, the family of transpose depolarizing channels coincides with the family of depolarizing channels and additivity of the minimal entropy output holds in the stronger form (3.8) for all t allowed by complete positivity. Also, the irreducibility of α_{out} implies that an α-invariant channel is bistochastic when $\mathrm{d}_{\mathrm{in}} = \mathrm{d}_{\mathrm{out}}$ and therefore (3.8) follows also from [64] when $\mathrm{d} = 2$. Yet another proof of the weak version (3.7) of additivity can be given in this case by showing that

$$\boldsymbol{\lambda}' \succ \boldsymbol{\lambda} \qquad \text{implies} \qquad X(\boldsymbol{\lambda}') \succ X(\boldsymbol{\lambda}), \tag{3.9}$$

for any $\boldsymbol{\lambda}, \boldsymbol{\lambda}' \in \mathcal{S}$, which is a matter of straightforward computation and immediately yields (3.7).

Summary

In this chapter we studied quantum channels that are invariant under a pair of representations of the unitary group of the input Hilbert space, in the sense of (3.4). Transformation properties of the Choi matrix yield a decomposition of the set of invariant CP maps, related to the irreducible components of the product representation κ. In the case of the conjugate representation this decomposition enabled us to identify the set of invariant channels as an interval in the space of channels. Additivity of minimal entropy output was then proven for two copies of such a channel, on a restricted range of parameters. Recently, this additivity result was extended to the whole range of parameters [26], using a different approach, based on monotonicity of entropy as a function of symmetric polynomials [76].

Chapter 4

Entropy inequalities and Markovianity

Monotonicity of relative entropy (section 2.4.3) states that statistical distinguishability of states cannot increase when they are subjected to a quantum operation. It is natural to ask under what circumstances the distinguishability of the states remains the same, i.e. to describe triples $(\rho,\, \sigma,\, \Phi)$ such that

$$S\left(\Phi(\rho),\, \Phi(\sigma)\right) = S\left(\rho,\, \sigma\right) \tag{4.1}$$

holds. This is obviously the case when Φ is invertible with respect to the states ρ and σ, i.e. there exists a quantum operation $\hat{\Phi}$ such that

$$\hat{\Phi}\left(\Phi\left(\rho\right)\right) = \rho \qquad \text{and} \qquad \hat{\Phi}\left(\Phi\left(\sigma\right)\right) = \sigma\,, \tag{4.2}$$

since in this case

$$S\left(\rho,\, \sigma\right) \geq S\left(\Phi(\rho),\, \Phi(\sigma)\right) \geq S\left(\hat{\Phi}\left(\Phi(\rho)\right),\, \hat{\Phi}\left(\Phi(\sigma)\right)\right) = S\left(\rho,\, \sigma\right).$$

It has been shown [88] that this is actually the only case when the monotonicity inequality is saturated; moreover, $\hat{\Phi}$ can be chosen in a canonical way, as the map

$$\Phi^{\#} = \mathrm{Ad}_{\sigma^{1/2}} \circ \Phi^{*} \circ \mathrm{Ad}_{\Phi(\sigma)^{-1/2}}, \tag{4.3}$$

supposedly $\Phi(\sigma)$ is invertible. Therefore we obtain that a triple $(\rho,\, \sigma,\, \Phi)$ with invertible $\Phi(\sigma)$ saturates the monotonicity inequality if and only if

$$\rho = \sigma^{1/2}\, \Phi^{*}\left(\left(\Phi(\sigma)\right)^{-1/2}\, \rho\, \left(\Phi(\sigma)\right)^{-1/2}\right) \sigma^{1/2} \tag{4.4}$$

holds ($\hat{\Phi}\left(\Phi(\sigma)\right) = \sigma$ is automatically satisfied).

Even though criterion (4.4) can directly be checked for any triples (ρ, σ, Φ), it doesn't give a method to construct such triples. The aim of this chapter is to bridge this gap, and give a characterization of saturating triples via their algebraic structure, that makes explicit construction of such triples possible.

The monotonicity inequality for relative entropy has several important special cases (see section 2.4.3), and characterization of saturation in the general case yields characterization in these special cases. Of special importance is the strong subadditivity of entropy (2.42), the saturation of which is closely related to the notion of Markovianity in ordered tripartite systems.

4.1 Saturation of entropy inequalities

Probably the most important special case of monotonicity of relative entropy is the data processing inequality (Corollary 2.4.6), which in turn yields monotonicity of the Holevo quantity (Example 2.4.7) and the strong subadditivity of entropy (Example 2.4.8) as special cases. Note that both Φ and σ have a product form in (DPE); namely $\Phi = \mathrm{id} \otimes \Lambda$ and $\sigma = \rho_1 \otimes \rho_2$. An easy computation shows that in this case $\Phi^{\#}$ also factorizes, in the form

$$\Phi^{\#} = \mathrm{id} \otimes \Lambda^{\#}, \qquad \text{where} \qquad \Lambda^{\#} = \mathrm{Ad}_{\rho_2^{1/2}} \circ \Lambda^* \circ \mathrm{Ad}_{\Lambda(\rho_2)^{-1/2}}, \tag{4.5}$$

yielding immediately that equality holds in the data processing inequality if and only if there exists a map $\hat{\Lambda} : \mathcal{B}' \to \mathcal{B}$ such that

$$\left(\mathrm{id} \otimes \hat{\Lambda}\, \Lambda\right)(\rho) = \rho\,. \tag{4.6}$$

Applying it to the monotonicity of the Holevo quantity, we obtain that

$$\chi\left(\{p_x,\, \Lambda(\rho_x)\}\right) = \chi\left(\{p_x,\, \rho_x\}\right)$$

if and only if there exists a quantum operation $\hat{\Lambda}$ such that

$$\hat{\Lambda}\, \Lambda\left(\rho_x\right) = \rho_x \qquad \forall x. \tag{4.7}$$

An important consequence is that a general characterization of the invertibility of a quantum operation with respect to two states gives a characterization of invertibility with respect to any finite number of states.

To see the relation of the saturation in (SSA) to the structure of tripartite states, we start with the instructive example of classical systems. A classical tripartite state $\rho_{\mathrm{ABC}} = \sum_{x,y,z} p(x,y,z)\, \delta_x \otimes \delta_y \otimes \delta_z$ is called **Markovian**, if

$$p(x,z|y) = p(x|y)p(z|y) \tag{4.8}$$

holds for the conditional probabilities for all y such that $p(y) > 0$. The matrix Q with entries $Q_{yz} := p(z|y)$ is a stochastic one, and defines a stochastic map α from the state space of system B to the state space of system BC by

$$\delta_y \mapsto \sum_z Q_{yz}\, \delta_y \otimes \delta_z, \tag{4.9}$$

which is easily seen to extend the state ρ_{AB} to ρ_{ABC} in the way

$$\rho_{\mathrm{ABC}} = (\mathrm{id} \otimes \alpha)\,(\rho_{\mathrm{AB}})\,. \tag{4.10}$$

For each y with positive probability we can introduce bipartite states

$$\rho^y_{\mathrm{AC}} := \sum_{xz} p(x,z|y)\, \delta_x \otimes \delta_z \tag{4.11}$$

and obtain

$$\begin{aligned} S_{\mathrm{AB}} + S_{\mathrm{BC}} - S_{\mathrm{ABC}} - S_{\mathrm{B}} &= \sum_y p(y) \sum_{xz} p(x,z|y)\, [-\log p(x,y) - \log p(y,z) \\ &\qquad\qquad + \log p(x,y,z) + \log p(y)] \\ &= \sum_y p(y) \sum_{xz} p(x,z|y)\, [-\log p(x|y) - \log p(z|y) + \\ &\qquad\qquad + \log p(x,z|y)] \\ &= \sum_y p(y)\, [S\,(\rho^y_{\mathrm{A}}) + S\,(\rho^y_{\mathrm{C}}) - S\,(\rho^y_{\mathrm{AC}})]\,. \end{aligned}$$

Therefore (SSA) for classical systems follows from ordinary subadditivity of entropy, and equality holds in it if and only if the marginals of ρ^y_{AC} are independent, that is, ρ_{ABC} is Markovian.

Note that for multipartite quantum states we cannot define an analog of the conditional states (4.11), therefore the notion of Markovianity cannot be defined via conditional independence. Rather, an algebraic approach should be followed, defining Markovianity e.g. through formula (4.10):

4.1.1. Definition. A tripartite quantum state ρ_{ABC} is **Markovian** if there exists a stochastic map from $\mathcal{B}$ to $\mathcal{B}\otimes\mathcal{C}$ such that

$$\rho_{\mathrm{ABC}} = (\mathrm{id}_{\mathcal{A}}\otimes\alpha)\,(\rho_{\mathrm{AB}}) \tag{4.12}$$

holds.

As usual, there may be more than one natural-looking extension of a classical notion to the quantum world. An argument in favor of the above definition of Markovianity for quantum systems is that the set of Markovian states, defined this way, coincides with the set of states saturating (SSA), just like in the classical case. Indeed, (SSA) is a special case of (DPE) with $\Lambda = \mathrm{Tr}_{\mathcal{C}} : \mathcal{B}\otimes\mathcal{C}\to\mathcal{B}$, and therefore criterion (4.6) coincides with Markovianity, with $\hat{\Lambda}=\alpha$ and $\rho=\rho_{\mathrm{ABC}}$.

States of the form $\rho_{\mathrm{ABC}}=\rho_{\mathrm{A}}\otimes\rho_{\mathrm{BC}}$ and $\rho_{\mathrm{ABC}}=\rho_{\mathrm{AB}}\otimes\rho_{\mathrm{C}}$ are trivial examples of quantum Markovian states. Suppose now that the Hilbert space of system B splits in the form $\mathcal{H}_{\mathrm{B}}=\mathcal{H}_{\mathrm{B^L}}\otimes\mathcal{H}_{\mathrm{B^R}}$, and let $\rho_{\mathrm{AB^L}}$ and $\rho_{\mathrm{B^RC}}$ be states on $\mathcal{H}_{\mathrm{A}}\otimes\mathcal{H}_{\mathrm{B^L}}$ and $\mathcal{H}_{\mathrm{B^R}}\otimes\mathcal{H}_{\mathrm{C}}$, respectively. A straightforward computation shows that $\rho_{\mathrm{ABC}}=\rho_{\mathrm{AB^L}}\otimes\rho_{\mathrm{B^RC}}$ is Markovian. An immediate generalization is the following:

4.1.2. Proposition. Suppose the Hilbert space of system B can be decomposed in the form

$$\mathcal{H}_{\mathrm{B}} = \bigoplus_k \mathcal{H}_{\mathrm{B}^{\mathrm{L}}_k}\otimes\mathcal{H}_{\mathrm{B}^{\mathrm{R}}_k}. \tag{4.13}$$

Let $\rho_{\mathrm{AB}^{\mathrm{L}}_k}$ and $\rho_{\mathrm{B}^{\mathrm{R}}_k\mathrm{C}}$ be states on $\mathcal{H}_{\mathrm{A}}\otimes\mathcal{H}_{\mathrm{B}^{\mathrm{L}}_k}$ and $\mathcal{H}_{\mathrm{B}^{\mathrm{R}}_k}\otimes\mathcal{H}_{\mathrm{C}}$, respectively, and let $\{p_k\}$ be a probability distribution. Then the state

$$\rho_{\mathrm{ABC}} = \bigoplus_k p_k\,\rho_{\mathrm{AB}^{\mathrm{L}}_k}\otimes\rho_{\mathrm{B}^{\mathrm{R}}_k\mathrm{C}} \tag{4.14}$$

is Markovian.

The nontrivial part of the structural characterization of Markovian states is the converse of the above Proposition:

4.1.3. Theorem. A tripartite state ρ_{ABC} is Markovian if and only if it admits a decomposition of the form (4.14).

The form (4.14) was already conjectured in [96], and Theorem 4.1.3 was proven in [44], using a result of [67]. Here we will derive this result as a special case of the structural characterization of triples that saturate the monotonicity of relative entropy.

4.2 Structure of saturating triples

In this section we give a structural characterization of triples (ρ, σ, Φ) that saturate the monotonicity inequality for relative entropy, under the condition that all states ρ, σ, $\Phi(\rho)$ and $\Phi(\sigma)$ are faithful, and Φ is a trace-preserving 2-positive map from $\mathcal{A}_{\mathcal{H}}$ to $\mathcal{A}_{\mathcal{K}}$, where $\mathcal{A}_{\mathcal{H}}$ and $\mathcal{A}_{\mathcal{K}}$ are finite dimensional C^*-algebras of $\mathcal{B}(\mathcal{H})$ and $\mathcal{B}(\mathcal{K})$. For compactness of presentation we slightly change the notation, and introduce $D_1 := \rho$ and $D_2 := \sigma$.

4.2.1. Theorem. The triple (D_1, D_2, Φ) is saturating if and only if the followings hold:

(i) The Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ decompose in the form

$$\mathcal{H} \cong \bigoplus_{k=1}^{K} \mathcal{H}_k^{\mathrm{L}} \otimes \mathcal{H}_k^{\mathrm{R}} \quad \text{and} \quad \mathcal{K} \cong \bigoplus_{k=1}^{K} \mathcal{K}_k^{\mathrm{L}} \otimes \mathcal{K}_k^{\mathrm{R}}, \tag{4.15}$$

where $\dim \mathcal{H}_k^{\mathrm{L}} = \dim \mathcal{K}_k^{\mathrm{L}}$.

(ii) Φ maps $\mathcal{B}\left(\mathcal{H}_k^{\mathrm{L}} \otimes \mathcal{H}_k^{\mathrm{R}}\right)$ to $\mathcal{B}\left(\mathcal{K}_k^{\mathrm{L}} \otimes \mathcal{K}_k^{\mathrm{R}}\right)$, and its restriction to $\mathcal{B}\left(\mathcal{H}_k^{\mathrm{L}} \otimes \mathcal{H}_k^{\mathrm{R}}\right)$ factorizes in the form

$$\Phi_k \cong \mathrm{Ad}_{U_k} \otimes \Phi_k^{\mathrm{R}}, \tag{4.16}$$

where $U_k : \mathcal{H}_k^{\mathrm{L}} \to \mathcal{K}_k^{\mathrm{L}}$ are unitaries, and Φ_k^{R} are stochastic maps from $\mathcal{B}\left(\mathcal{H}_k^{\mathrm{R}}\right)$ to $\mathcal{B}\left(\mathcal{K}_k^{\mathrm{R}}\right)$.

(iii) The states decompose as

$$D_j \cong \bigoplus_{k=1}^{K} p(j,k)\, D(j,k,L) \otimes D(k,R), \qquad j = 1, 2; \tag{4.17}$$

where all the operators are density operators, and $\{p(j,k)\}$ is a probability distribution for a fixed j. Note that the second terms in the decompositions don't depend on j.

Proof: Decomposition of D_j and Φ implies

$$\log \Phi(D_j) = \bigoplus_{k=1}^{K} \log p(j,k) I_{\mathcal{K}_k} + \left[U\ \left(\log D(j,k,L)\right)\ U^*\right] \otimes I_{\mathcal{K}_k^{\mathrm{R}}} + I_{\mathcal{K}_k^{\mathrm{L}}} \otimes \Phi_k^{\mathrm{R}}\left(D(k,R)\right),$$

with $\mathcal{K}_k := \mathcal{K}_k^{\mathrm{L}} \otimes \mathcal{K}_k^{\mathrm{R}}$, therefore

$$\begin{aligned} S\big(\Phi(D_1),\, \Phi(D_2)\big) &= \operatorname{Tr}\Phi(D_1)\big[\log\Phi(D_1) - \log\Phi(D_2)\big] \\ &= \sum_{k=1}^{K} p(1,k)\left[\log p(1,k) - \log p(2,k)\right] \\ &\quad + p(1,k)\operatorname{Tr} D(1,k,L)\big[\log D(1,k,L) - \log D(2,k,L)\big] \\ &\quad + p(1,k)\operatorname{Tr} D(k,R)\big[\log D(k,R) - \log D(k,R)\big] \\ &= S\left(p(1,\,.\,),\, p(2,\,.\,)\right) + \sum_{k=1}^{K} p(1,k)\, S\left(D(1,k,L),\, D(2,k,L)\right) \\ &= S\left(D_1,\, D_2\right)\,, \end{aligned}$$

and sufficiency of conditions (i), (ii) and (iii) is proven.

To prove the converse, we introduce the maps

$$\beta := \mathrm{Ad}_{\Phi(D_2)^{-1/2}} \circ \Phi \circ \mathrm{Ad}_{D_2^{1/2}}\,,$$

and its adjoint with respect to the Hilbert-Schmidt inner product

$$\Phi^{\#} := \mathrm{Ad}_{D_2^{1/2}} \circ \Phi^* \circ \mathrm{Ad}_{\Phi(D_2)^{-1/2}}\,.$$

β is easily seen to be unital and 2-positive, hence its adjoint in turn is trace-preserving and 2-positive.

Next we define the 2-positive unital maps $\gamma_{\mathcal{H}} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ and $\gamma_{\mathcal{K}} : \mathcal{B}(\mathcal{K}) \to \mathcal{B}(\mathcal{K})$ as

$$\gamma_{\mathcal{H}} := \left(\Phi^{\#} \circ \Phi\right)^* = \Phi^* \circ \beta \quad \text{and} \quad \gamma_{\mathcal{K}} := \left(\Phi \circ \Phi^{\#}\right)^* = \beta \circ \Phi^*.$$

$\gamma_{\mathcal{H}}$ leaves the faithful state corresponding to D_2 invariant; indeed,

$$\gamma_{\mathcal{H}}^*\left(D_2\right) = \Phi^{\#}\left(\Phi(D_2)\right) = D_2\,,$$

by the unitality of Φ^*. Thus, by the arguments in Appendix A.5, the fixed point set $\mathcal{D}_{\mathcal{H}}$ of $\gamma_{\mathcal{H}}$ is a C^*-algebra. The same applies o the fixed point set $D_{\mathcal{K}}$ of $\gamma_{\mathcal{K}}$, the faithful invariant state in this case is $\Phi(D_2)$.

Note that

$$\beta\left(\mathcal{D}_{\mathcal{H}}\right) \subset \mathcal{D}_{\mathcal{K}} \qquad \text{and} \qquad \Phi^*\left(\mathcal{D}_{\mathcal{K}}\right) \subset \mathcal{D}_{\mathcal{H}}. \tag{4.18}$$

Indeed, for $x \in \mathcal{D}_{\mathcal{H}}$ we get $\gamma_{\mathcal{K}}(\beta(x)) = \beta \circ (\Phi^* \circ \beta)(x) = \beta(x)$; the other inclusion follows the same way.

Let $\mathcal{A}_{\Phi^*}$ and $\mathcal{A}_\beta$ be as defined in (A.13); then we have

$$\mathcal{D}_{\mathcal{H}} \subset \mathcal{A}_\beta \qquad \text{and} \qquad \mathcal{D}_{\mathcal{K}} \subset \mathcal{A}_{\Phi^*}. \tag{4.19}$$

Indeed, for any $x \in \mathcal{D}_{\mathcal{H}}$ also $x^*x \in \mathcal{D}_{\mathcal{H}}$, and Schwarz inequality $\beta(x^*x) \leq \beta(x)^*\,\beta(x)$ implies

$$x^*x = \Phi^*(\beta(x^*x)) \leq \Phi^*(\beta(x)^*\beta(x)) \leq \Phi^*(\beta(x))^*\,\Phi^*(\beta(x)) = x^*x,$$

therefore all inequalities are satisfied with equality. Hence

$$\beta(x^*x) = \beta \circ \Phi^*(\beta(x)^*\beta(x)) = \gamma_{\mathcal{K}}(\beta(x)^*\beta(x)) = \beta(x)^*\beta(x)\,,$$

where in the last step we used (4.18) and the fact that $\mathcal{D}_{\mathcal{K}}$ is an algebra. $\beta(xx^*) = \beta(x)\,\beta(x^*)$ follows by the same argument, as well as the other inclusion $\mathcal{D}_{\mathcal{K}} \subset \mathcal{A}_{\Phi^*}$.

By the definitions of $\mathcal{A}_\beta$ and $\mathcal{A}_{\Phi^*}$, (4.19) implies that β restricted to $\mathcal{D}_{\mathcal{H}}$ and Φ^* restricted to $\mathcal{D}_{\mathcal{K}}$ are C^*-algebra morphisms, moreover, they are inverses of each other, yielding that $\mathcal{D}_{\mathcal{H}}$ and $\mathcal{D}_{\mathcal{K}}$ are isomorphic to each other. Now we can apply Corollary (A.2.3), and obtain (i) of the Theorem.

Minimal central projections[1] of $\mathcal{D}_{\mathcal{H}}$ and $\mathcal{D}_{\mathcal{K}}$ are of the form

$$P_k = \bigoplus_r \delta_{r,k}\, I_{\mathcal{H}_r^{\mathrm{L}}} \otimes I_{\mathcal{H}_r^{\mathrm{R}}} \qquad \text{and} \qquad Q_l = \bigoplus_r \delta_{r,l}\, I_{\mathcal{K}_r^{\mathrm{L}}} \otimes I_{\mathcal{K}_r^{\mathrm{R}}}.$$

Since Φ^* is a C^*-isomorphism between $\mathcal{D}_{\mathcal{K}}$ and $\mathcal{D}_{\mathcal{H}}$, it maps minimal central projections to minimal central projections, i.e.

$$\Phi^*(Q_k) = P_k\,, \qquad\qquad k = 1, \ldots, K.$$

Now if A is a self-adjoint element in $Q_k\,\mathcal{B}(\mathcal{K})\,Q_k$ then $-\|A\|\; Q_k \leq A \leq \|A\|\; Q_k$ implies $-\|A\|\; P_k \leq \Phi^*(A) \leq \|A\|\; P_k$, hence $A \in P_k\,\mathcal{B}(\mathcal{H})\,P_k$. For a general $A \in Q_k\,\mathcal{B}(\mathcal{K})\,Q_k$ also $A^* \in Q_k\,\mathcal{B}(\mathcal{K})\,Q_k$, hence the decomposition $A = \frac{1}{2}(A + A^*) + \frac{1}{2}(A - A^*)$ together with the previous argument yields

$$\Phi^*\left(\, Q_k\,\mathcal{B}(\mathcal{K})\,Q_k\,\right) \subset P_k\,\mathcal{B}(\mathcal{H})\,P_k\,. \tag{4.20}$$

[1]A projection p is *central* in the C^*-algebra $\mathcal{A}$ if it commutes with all $a \in \mathcal{A}$. It is *minimal central* if there is no proper subprojection of p that is also central.

Now let Φ_k^* be the restriction of Φ^* to $Q_k\,\mathcal{B}(\mathcal{K})\,Q_k$ and let $\mathcal{D}'_{\mathcal{K}}$ denote the commutant of $\mathcal{D}_{\mathcal{K}}$. Note that the commutant of $Q_k\,\mathcal{D}_{\mathcal{K}}\,Q_k$ in $Q_k\,\mathcal{B}(\mathcal{K})\,Q_k$ is equal to $Q_k\,\mathcal{D}'_{\mathcal{K}}\,Q_k$. Elements of $Q_k\,\mathcal{D}_{\mathcal{K}}\,Q_k$ have the form $A\otimes I_{\mathcal{K}_k^{\mathrm{R}}}$ while elements in $Q_k\,\mathcal{D}'_{\mathcal{K}}\,Q_k$ are of the form $I_{\mathcal{K}_k^{\mathrm{L}}}\otimes B$ with $A\in\mathcal{K}_k^{\mathrm{L}}$ and $B\in\mathcal{K}_k^{\mathrm{R}}$. Inclusion (4.20) implies the existence of a unital 2-positive map $\eta_1:\ \mathcal{B}(\mathcal{K}_k^{\mathrm{L}})\to\mathcal{B}(\mathcal{H}_k^{\mathrm{L}})$ such that $\Phi^*(A\otimes I_{\mathcal{K}_k^{\mathrm{R}}})=\eta_1(A)\otimes I_{\mathcal{H}_k^{\mathrm{R}}}$. On the other hand, by (4.19)

$$\begin{aligned}\Phi_k^*\left(\left(A\otimes I_{\mathcal{K}_k^{\mathrm{R}}}\right)X\right) &= \Phi_k^*\left(A\otimes I_{\mathcal{K}_k^{\mathrm{R}}}\right)\Phi_k^*(X) \qquad \text{and}\\ \Phi_k^*\left(X\left(A\otimes I_{\mathcal{K}_k^{\mathrm{R}}}\right)\right) &= \Phi_k^*(X)\,\Phi_k^*\left(A\otimes I_{\mathcal{K}_k^{\mathrm{R}}}\right)\end{aligned}$$

for any $X\in Q_k\,\mathcal{B}(\mathcal{K})\,Q_k$. Applying it to $I_{\mathcal{K}_k^{\mathrm{L}}}\otimes B$ we obtain $\Phi_K^*\left(I_{\mathcal{K}_k^{\mathrm{L}}}\otimes B\right)\in P_k\,\mathcal{D}'_{\mathcal{H}}\,P_k$ hence there exists a unital 2-positive map $\eta_2:\ \mathcal{B}(\mathcal{K}_k^{\mathrm{R}})\to\mathcal{B}(\mathcal{H}_k^{\mathrm{R}})$ such that $\Phi^*(I_{\mathcal{K}_k^{\mathrm{L}}}\otimes B)=I_{\mathcal{H}_k^{\mathrm{L}}}\otimes\eta_2(B)$. Putting it together, we get the form $\Phi_k^*=\eta_1\otimes\eta_2$. Moreover, as Φ^* is an algebra isomorphism from $\mathcal{D}_{\mathcal{K}}$ to $\mathcal{D}_{\mathcal{H}}$, also η_1 has to be an algebra isomorphism from $\mathcal{B}\left(\mathcal{K}_k^{\mathrm{L}}\right)$ to $\mathcal{B}\left(\mathcal{H}_k^{\mathrm{L}}\right)$, and Theorem (A.2.1) implies the existence of a unitary $U_k:\ \mathcal{H}_k^{\mathrm{L}}\to\mathcal{K}_k^{\mathrm{L}}$ such that $\eta_1=\mathrm{Ad}_{U_k^*}$. Hence (ii) of the Theorem is proven.

Since D_2 is the density of a faithful $\gamma_{\mathcal{H}}$-invariant state, the corresponding modular group leaves $\mathcal{D}_{\mathcal{H}}$ invariant (see Appendix A.5 for details). Now we use the result of [88], namely that the saturation property implies $\Phi^{\#}\left(\Phi(D_1)\right) = D_1$. This in turn yields that D_1 is the density of a faithful $\gamma_{\mathcal{H}}$-invariant state as well, hence its modular group also leaves $\mathcal{D}_{\mathcal{H}}$ invariant; that is we have

$$D_j^{it}\,\mathcal{D}_{\mathcal{H}}\,D_j^{-it}\subset\mathcal{D}_{\mathcal{H}}\,;\qquad t\in\mathbb{R},\quad j=1,2\,.$$

Now we can use the following structure theorem, proved in section 4.2.1:

4.2.2. Theorem. Let $\mathcal{D}$ be a subalgebra of $\mathcal{B}(\mathcal{H})$ for a finite dimensional Hilbert space $\mathcal{H}$, with corresponding decomposition of the Hilbert space

$$\mathcal{H}\cong\bigoplus_{(m,d)}\mathcal{H}_d\otimes\mathcal{H}_m\otimes\mathcal{H}_{(m,d)}, \tag{4.21}$$

as in Corollary (A.2.3), and let U be a unitary such that

$$U\,\mathcal{D}\,U^*\subset\mathcal{D} \tag{4.22}$$

holds. Then U is of the form

$$U \cong \bigoplus_{(m,d)} \sum_{i=1}^{K(m,d)} v(m,d,i) \otimes w(m,d,i) \otimes E^{(m,d)}_{\pi(m,d,i),\, i} \tag{4.23}$$

where $v(m,d,i)$ and $w(m,d,i)$ are unitaries and $\pi(m,d,\,.\,)$ are permutations on the sets $\{1,\dots,K(m,d)\}$. Moreover,

$$U \mapsto \pi(m,d,\,.\,)$$

is a group homomorphism on the group of unitaries satisfying (4.22), for all (m,d).

In the case of the modular groups, the permutations also depend on t, and the map $t \mapsto \pi(m,d,t,\,.\,)$ is a continuous group homomorphism from $\mathbb{R}$ to the permutation group on $\{1,\dots,K(m,d)\}$, and therefore it is constant, i.e. all the permutations $\pi(m,d,t,\,.\,)$ equal to the identity. It enables us to form one index k from the triple (m,d,i), and use decompositions (A.4) and (A.5) instead of (A.6) and (A.7). In these decompositions elements of $\mathcal{D}_{\mathcal{H}}$ and $\mathcal{D}_{\mathcal{K}}$ are of the form

$$A \cong \bigoplus_{k=1}^{K} A_k \otimes I_{\mathcal{H}_k^{\mathrm{R}}}\,, \qquad \text{and} \qquad X \cong \bigoplus_{k=1}^{K} X_k \otimes I_{\mathcal{K}_k^{\mathrm{R}}}\,, \tag{4.24}$$

while D_j^{it} decomposes into

$$D_j^{it} \cong \bigoplus_{k=1}^{K} v_j(k,t) \otimes w_j(k,t), \tag{4.25}$$

where $t \mapsto v_j(k,t)$ and $t \mapsto w_j(k,t)$ are groups of unitaries.

Since $D_1^{it} D_2^{-it} \in \mathcal{D}_{\mathcal{H}}$ (see Appendix A.5), we obtain that $w_j(k,t)$ is actually independent of j. The generator of $t \mapsto v_j(k,t) \otimes w(k,t)$ is of the form $G_j(k,L) \otimes I_{\mathcal{H}_k^{\mathrm{R}}} + I_{\mathcal{H}_k^{\mathrm{L}}} \otimes G(k,R)$, while the generator of $t \mapsto D_j^{it}$ is $\ln D_j$, that proves (iii) of the Theorem. $\square$

4.2.3. Example. Structure of Markovian states

In the case of (SSA), we have

$$\mathcal{H} = \mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}} \otimes \mathcal{H}_{\mathrm{C}} \qquad \text{and} \qquad \mathcal{K} = \mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}};$$

while

$$D_1 = \rho_{\mathrm{ABC}} \qquad \text{and} \qquad D_2 = \rho_{\mathrm{A}} \otimes \rho_{\mathrm{BC}}.$$

Φ is of the form $\mathrm{id}_{\mathcal{A}} \otimes \tau$, where τ is the partial trace operation from $\mathcal{B}(\mathcal{H}_{\mathrm{B}} \otimes \mathcal{H}_{\mathrm{C}})$ to $\mathcal{B}(\mathcal{H}_{\mathrm{B}})$. We obtain

$$\beta = \mathrm{id}_{\mathcal{A}} \otimes \left(\mathrm{Ad}_{\rho_{\mathrm{B}}^{-1/2}} \circ \tau \circ \mathrm{Ad}_{\rho_{\mathrm{BC}}^{1/2}} \right) \qquad \text{and} \qquad \gamma_{\mathcal{K}} = \mathrm{id}_{\mathcal{A}} \otimes \kappa,$$

where

$$\kappa(B) = \rho_{\mathrm{B}}^{-1/2} \, \tau \left(\rho_{\mathrm{BC}}^{1/2} \, (B \otimes I_{\mathcal{H}_{\mathrm{C}}}) \, \rho_{\mathrm{BC}}^{1/2} \right) \, \rho_{\mathrm{B}}^{-1/2}.$$

ρ_{B} is easily seen to be a faithful invariant state of the CP map κ, therefore its fixpoint set $\mathcal{D}_\kappa$ is a C^*-subalgebra of $\mathcal{B}(\mathcal{H}_{\mathrm{B}})$, with corresponding decomposition

$$\mathcal{H}_{\mathrm{B}} = \sum_{k=1}^{K} \mathcal{H}_{\mathrm{B}}^{\mathrm{L}} \otimes \mathcal{H}_{\mathrm{B}}^{\mathrm{R}}.$$

Since $\mathcal{D}_{\mathcal{K}} = \mathcal{B}(\mathcal{H}_{\mathrm{A}}) \otimes \mathcal{D}_\kappa$ and $\mathcal{D}_{\mathcal{H}} = \mathcal{D}_{\mathcal{K}} \otimes I_{\mathcal{H}_{\mathrm{C}}}$, the elements of $\mathcal{D}_{\mathcal{H}}$ are of the form

$$\bigoplus_{k=1}^{K} X_k \otimes I_{\mathcal{H}_{\mathrm{B}}^{\mathrm{R}} \otimes \mathcal{H}_{\mathrm{C}}},$$

which immediately implies the decomposition in (4.14).

We can also identify the components of T_k from (4.16): $U_k = I_{\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}^{\mathrm{L}}}$, while Φ_K^{R} is the partial trace from $\mathcal{B}(\mathcal{H}_{\mathrm{B}}^{\mathrm{R}} \otimes \mathcal{H}_{\mathrm{C}})$ to $\mathcal{B}(\mathcal{H}_{\mathrm{B}}^{\mathrm{R}})$.

4.2.4. Example. Classical case

Let D_1, D_2 be two probability measures on a finite space $\mathcal{X}$, and Φ be a stochastic map, that maps probability measures on $\mathcal{X}$ to probability measures on $\mathcal{Y}$. The Hilbert space in this case is $\mathcal{H} = \mathbb{C}^{|\mathcal{X}|}$, and $\mathcal{A}_{\mathcal{H}}$ is the commutative algebra of diagonal matrices. Since $\mathcal{D}_\mathcal{H} \subset \mathcal{A}_\mathcal{H}$ is also commutative, we have $\dim \mathcal{H}_k^{\mathrm{L}} = 1$ for all k, therefore $D(j,k,L)$ is a one-dimensional density operator, i.e. the number 1. Thus

$$\mathcal{H} \cong \bigoplus_{k=1}^{K} \mathcal{H}_k, \qquad \text{and} \qquad D_j \cong \bigoplus_{k=1}^{K} p(j,k)\, D(k)\,.$$

We can equivalently say that there exists a partition $\mathcal{X} = \bigcup_{k=1}^K \mathcal{X}_k$, and probability distributions $D(k)$ on $\mathcal{X}_k$ such that the restrictions of D_1 and D_2 to $\mathcal{X}_k$ are both constant multiples of $D(k)$.

Repeating the same argument for $T(D_1)$ and $T(D_2)$, we obtain a decomposition $\mathcal{Y} = \bigcup_{k=1}^K \mathcal{Y}_k$. Criterion (ii) of Theorem 4.2.1 implies that T^* maps functions supported in $\mathcal{Y}_k$ to functions supported in $\mathcal{X}_k$, or equivalently, that the matrix of Φ is blockdiagonal with respect to the partitions $\{\mathcal{X}_k\}$ and $\{\mathcal{Y}_k\}$.

4.2.1 Structure of certain unitaries

This section is entirely devoted to the proof of Theorem 4.2.2. We start from decompositions (A.6) and (A.7), corresponding to the algebra $\mathcal{D}$, and define $P_{m,d}$ to be the projection corresponding to multiplicity m and dimension d, that is

$$P_{m,d}(m',d',i) := \delta_{m',m}\delta_{d',d}\, I_{\mathcal{H}_d}\,, \qquad 1 \le i \le K(m,d); \tag{4.26}$$

and

$$P_m := \oplus_d P_{m,d} \tag{4.27}$$

to be the projection corresponding to the multiplicity m. We denote by $\mathcal{H}_m$ the range of P_m. Note that $P_{m,d}$ and hence P_m commutes with elements of $\mathcal{D}$, thus

$$\mathcal{D}_{m,d} := P_{m,d}\mathcal{D} \qquad \text{and} \qquad \mathcal{D}_m := P_m\mathcal{D} \tag{4.28}$$

are unital algebras with unit $P_{m,d}$ and P_m, respectively.

We fix an orthonormal basis

$$\{\, e(m,d,k) \;:\; 1 \le k \le mdK(m,d)\,\} \tag{4.29}$$

in the range of $P_{m,d}$ for every possible m and d, such that $|e(m,d,k)\rangle\langle e(m,d,k)| \in \mathcal{D}$.

4.2.5. Proposition. Let U be a unitary in $\mathcal{B}(\mathcal{H})$ such that the map $A \mapsto U^*AU$ leaves $\mathcal{D}$ invariant. Then U commutes with P_m for every multiplicity m.

Proof: The statement is trivial when only one multiplicity exists, and we apply mathematical induction in the number of multiplicities.

Note that the rank of any minimal projection in $P_m\mathcal{D}$ is m. Let m_1 denote the smallest multiplicity, and let $q \in P_{m_1}\mathcal{D}$ be a minimal projection, then q is of rank m_1. P_mU^*qU is a projection again and its rank is at most m_1. Since all non-zero projections in $P_m\mathcal{D}$ has rank at least m, we can conclude that $P_mU^*\,q\,U = 0$ if $m > m_1$. Every element of $P_{m_1}\mathcal{D}$ is a linear combination of the above q's, hence we have $P_mU^*aU = 0$ if $m > m_1$ and $a \in P_{m_1}\mathcal{D}$.

Since $|e(m_1, d, k)\rangle\langle e(m_1, d, k)| \leq P_{m_1}$ for every possible d, k, it follows that

$$P_m\,U^*\,|e(m_1, d, k)\rangle\langle e(m_1, d, k)|\,UP_m \leq P_m\,(U^*P_{m_1}U) = 0$$

if $m > m_1$. So we can conclude that $U^*\,e(m_1, d, k) \in \mathcal{H}_{m_1}$ which gives that both $\mathcal{H}_{m_1}$ and its orthogonal complement are invariant subspaces for U, that is, $P_{m_1}U = UP_{m_1}$. Now we can restrict the whole problem to the orthogonal complement of $\mathcal{H}_{m_1}$ and use induction hypothesis in the number of multiplicities. $\square$

We have obtained that U has a blockdiagonal structure $U = \bigoplus_m P_mUP_m$, and to explore the finer structure of U, we can restrict our attention to the case when all the multiplicities are the same number m, i.e. the elements of $\mathcal{D}$ have the form

$$A = \bigoplus_d \sum_{i=1}^{K(d)} A\,(d, i) \otimes I_m \otimes E_{ii}^{(d)}\,. \tag{4.30}$$

As before, we can define projections $P_{d,i}$ $(1 \leq i \leq K(d))$ by the formula

$$P_{d,i}(d', i') := \delta_{d,\,d'}\,\delta_{i,\,i'}\,I_{\mathcal{H}_d},$$

and the projection corresponding to dimension d is

$$P_d := \bigoplus_{i=1}^{K(d)} P_{d,i}.$$

Again, all these projections commute with all elements of $\mathcal{D}$.

4.2.6. Proposition. In the above setting U commutes with P_d for every dimension d, and so U has the blockdiagonal structure

$$U = \bigoplus_d P_d U P_d.$$

Proof: Since $\mathcal{D}$ is isomorphic to $\mathcal{M} := \bigoplus_d \bigoplus_{i=1}^{K(d)} \mathcal{M}_d$, then Ad_U induces an automorphism γ of $\mathcal{M}$ (here $\mathcal{M}_d$ denotes the algebra of $d \times d$ complex matrices). The inclusion matrix of γ (see Appendix A.2 for definition) is a permutation matrix corresponding to a permutation τ of the set of all possible pairs (d, i) such that

$$\tau(d, i) = (d', i') \quad \text{gives} \quad d = d'.$$

This implies that

$$P_{d'}\, U^*\, P_d\, U = 0 \text{ for } d \neq d',$$

which, by the same argument as in Proposition 4.2.5 implies the desired statement. $\square$

In the view of the above Propositions we can suppose that all the blocks in $\mathcal{D}$ have the same multiplicity m and the same dimension d, consequently $\mathcal{D}$ is isomorphic to $\bigoplus_{i=1}^{K} \mathcal{M}_d$. In this case $\dim \mathcal{H} = mdK$, and $\mathcal{B}(\mathcal{H})$ is isomorphic to $\mathcal{M}_d \otimes \mathcal{M}_m \otimes \mathcal{M}_K$. Elements of $\mathcal{D}$ have the form

$$\sum_i A_i \otimes I_m \otimes E_{ii},$$

where A_i is an element of $\mathcal{M}_d$ and $\{E_{ij} \;:\; 1 \leq i, j \leq K\}$ are the standard matrix units of $\mathcal{M}_K$. It is easily seen that in this representation U has the form

$$U = \sum_{i=1}^{K} U_i \otimes E_{\pi(i)i},$$

where π is a permutation of the set $\{1, 2, \ldots, K\}$, and the U_i 's are unitary elements of $\mathcal{M}_d \otimes \mathcal{M}_m$ that leave the subalgebra $\mathcal{M}_d \otimes I_m$ invariant.

The final step is to describe the structure of a possible unitary $U \in \mathcal{M}_d \otimes \mathcal{M}_m$ such that Ad_U leaves the subalgebra $\mathcal{M}_d \otimes I_m$ invariant. Since Ad_U induces an automorphism of $\mathcal{M}_d \otimes I_m$, by Theorem (A.2.1) we have a unitary v such that

$$\mathrm{Ad}_U(A \otimes I_m) = (v \otimes I_m)^*(A \otimes I_m)(v \otimes I_m).$$

The automorphism $\mathrm{Ad}_U \circ \mathrm{Ad}_{v \otimes I_m}^{-1}$ leaves the subalgebra fixed and is induced by a unitary $W \in \mathcal{M}_d \otimes \mathcal{M}_m$. Hence W must be in the commutant of the subalgebra, that is, $W = I_d \otimes w$. From this we conclude that $U = v \otimes w$.

We arrived at the following:

4.2.7. Proposition. In the case when all the multiplicities and all the dimensions are the same, U must be of the form

$$U = \sum_{i=1}^{K} v_i \otimes w_i \otimes E_{\pi(i)i},$$

where $v_i \in \mathcal{M}_d$ and $w_i \in \mathcal{M}_m$ are unitaries and π is a permutation of the set $\{1, 2, \ldots, K\}$.

Propposition (4.2.5), (4.2.6) and (4.2.7) together give the structure claimed in Theorem (4.2.2).

Note that any unitary U that satisfies (4.22) also satisfies the seemingly stronger criterion $U\mathcal{D}U^* = \mathcal{D}$. Indeed, Ad_U is a unitary, when $\mathcal{B}(\mathcal{H})$ is equipped with the Hilbert-Schmidt inner product, and therefore all of its invariant subspaces are invariant for Ad_{U^*} as well, from which the statement follows. It also implies that the set of unitaries satisfying (4.22) forms a group.

Now let U and V be two unitaries for which (4.22) holds, with decompositions

$$\begin{aligned} U &\cong \bigoplus_{(m,d)} \sum_{i=1}^{K(m,d)} v(U,m,d,i) \otimes w(U,m,d,i) \otimes E^{(m,d)}_{\pi(U,m,d,i),\, i} \qquad \text{and} \\ V &\cong \bigoplus_{(m,d)} \sum_{i=1}^{K(m,d)} v(V,m,d,i) \otimes w(V,m,d,i) \otimes E^{(m,d)}_{\pi(V,m,d,i),\, i}\,. \end{aligned}$$

Then

$$\begin{aligned} UV &\cong \bigoplus_{(m,d)} \sum_{i,j=1}^{K(m,d)} v(U,m,d,i)\, v(V,m,d,j) \otimes w(U,m,d,i)\, w(V,m,d,j) \\ &\qquad \otimes E^{(m,d)}_{\pi(U,m,d,i),\,i} E^{(m,d)}_{\pi(V,m,d,j),\,j} \\ &= \bigoplus_{(m,d)} \sum_{i,j=1}^{K(m,d)} v(U,m,d,i)\, v(V,m,d,j) \otimes w(U,m,d,i)\, w(V,m,d,i) \\ &\qquad \otimes E^{(m,d)}_{\pi(U,m,d,i),\,j} \delta_{\pi(V,m,d,j),i} \\ &= \bigoplus_{(m,d)} \sum_{j=1}^{K(m,d)} v(U,m,d,\pi(V,m,d,j))\, v(V,m,d,j) \\ &\qquad \otimes w(U,m,d,\pi(V,m,d,j))\, w(V,m,d,j) \otimes E^{(m,d)}_{\pi(U,m,d,\pi(V,m,d,j)),\,j}, \end{aligned}$$

hence the map $U \mapsto \pi(U,m,d,\,.\,)$ is a group homomorphism, while $U \mapsto v(U,m,d,i) \otimes w(U,m,d,i)$ is in general not. However, if all the permutations are trivial, then also $U \mapsto v(U,m,d,i)$ and $U \mapsto w(U,m,d,i)$ are group homomorphisms.

Summary

In this chapter we have shown that the monotonicity inequality of relative entropy is saturated for a triple (D_1, D_2, Φ) only if both the states and the CPTP map have a particularly simple structure, given in Theorem 4.2.1. The main ingredient of the proof is the observation that the modular groups of D_1 and D_2 leave a certain subalgebra of the original algebra invariant, which in turn is a consequence of the characterization of saturation given in [88]. In section 4.2.1 we gave a necessary and sufficient condition on the structure of a unitary U that leaves a given subalgebra invariant. The combination of the two results yields the structural characterization of saturating triples. As an application we have shown that the result of [44] about the structure of tripartite states that saturate the strong subadditivity inequality follows from Theorem 4.2.1. The use of the modular groups requires the invertibility of the densities, which constraint doesn't seem to be necessary heuristically. However, it is not clear at the moment how this restriction could be removed.

Chapter 5

Infinite systems

One of the simplest models for magnetization of matter in quantum statistical physics is the one-dimensional spin chain, which is also the basic example for a quantum stochastic process, or a quantum information source. In realistic situations the chain consists of a number of spins in the order of 10^{23}, hence its description as an infinitely extended system seems to be a good mathematical approximation. Also, the study of asymptotic questions in quantum probability and information theory requires a mathematically rigorous model of a chain with an infinite number of sites.

First we briefly describe the mathematical construction of the observable algebra of the infinite spin chain, and general properties of its states. We mention the XY model as a concrete physical model for a one dimensional spin system. To approximate the true equilibrium state of a physical system, one needs classes of states with an explicit prescription for their mathematical construction, together with computable local and asymptotic quantities. Finitely correlated and Markovian states are examples for such classes.

Quasi-free states appear naturally as equilibrium states of interaction-free fermionic systems. A construction of Araki, described in section 5.3, makes it possible to redefine these states as states on the spin chain algebra. That it is not merely an abstract mathematical construction, but quasi-free states play an important role in the physics of the spin chain, is well illustrated by the fact that the thermodynamical limit of the ground states of the XX model belongs to this class. A great mathematical convenience in the study of this class is that many of the properties of a quasi-free state can be analyzed by the study of a single operator, called the symbol of the state.

The entropies $S(\varphi_n)$ of finite-size restrictions yield valuable information on properties of a shift-invariant state φ on the spin chain. The related notions of mean entropy and typical subspaces are introduced in section 5.4. The rest of the thesis is devoted to the study of these notions in concrete models. In Chapter 6 we show that the mean relative entropy can be interpreted as the asymptotical compressibility of the information content of an ergodic state. In Chapter 7 the asymptotic behavior of the sequence $S(\varphi_n)$ is studied for a class of pure quasi-free states.

5.1 The spin chain

The observable algebra

In this section we give a description of the observable algebra of an infinite spin system. To avoid technical difficulties of working with infinite tensor product of Hilbert spaces, we choose a C^*-algebraic description instead of working on the Hilbert space level. An additional advantage of this choice is that quantum and classical chains can be treated within the same formalism.

Let $\mathcal{C}_0$ denote the observable algebra of a single spin, located at one site of the infinite lattice $\mathbb{Z}$. For a classical spin-$\frac{1}{2}$ system $\mathcal{C}_0$ is the commutative algebra $\mathcal{F}(\{-1,1\})$, while for a quantum spin-s system $\mathcal{C}_0 = \mathcal{B}(\mathcal{H})$ where $\mathcal{H}$ is a $2s+1$-dimensional Hilbert space. The observable algebra of the spins located at the sites of a finite subset $\Lambda \subset \mathbb{Z}$ is $\mathcal{C}_\Lambda := \otimes_{k\in\Lambda}\mathcal{C}_0$. For any pair $\Lambda \subset \Lambda'$ the observable algebra $\mathcal{C}_\Lambda$ can naturally be viewed as a subalgebra of $\mathcal{C}_{\Lambda'}$. More precisely, there exists a unity-preserving embedding $\iota_{\Lambda',\Lambda}$ of $\mathcal{C}_\Lambda$ into $\mathcal{C}_{\Lambda'}$, given by the formula

$$\iota_{\Lambda',\Lambda}(A) := A \otimes \left(\underset{l\in\Lambda'\setminus\Lambda}{\otimes} \mathbb{1} \right) \qquad \text{for} \quad A \in \mathcal{C}_\Lambda.$$

The union of the C_Λ's can be equipped with a natural equivalence relation; for $A \in \mathcal{C}_{\Lambda_1},\ \ B \in \mathcal{C}_{\Lambda_2}$ we say that A and B are equivalent, if there exists a bigger subset Λ such that the embedded images of A and B in $\mathcal{C}_\Lambda$ are the same, i.e.

$$\iota_{\Lambda,\Lambda_1}(A) = \iota_{\Lambda,\Lambda_2}(B).$$

Factorization with this equivalence relation yields a normed $*$-algebra, in which also the C^*-property is satisfied. This algebra is called the **algebra of local observables**, and its norm completion is by definition the spin chain algebra $\mathcal{C}$. By identifying elements in different $\mathcal{C}_\Lambda$'s with their equivalence classes, we get an embedding of the finite algebras into the infinite one, which is compatible with the embeddings of the finite algebras into each other, i.e. we have embeddings $\iota_\Lambda : \mathcal{C}_\Lambda \to \mathcal{C}$ such that $\iota_\Lambda = \iota_{\Lambda'} \circ \iota_{\Lambda',\Lambda}$ holds. For the rest we identify $\mathcal{C}_\Lambda$ with its embedded image and in the following we use the same notation $\mathcal{C}_\Lambda$ for it. With this identification the local algebra can be written as the union of all the $\mathcal{C}_\Lambda$'s. Elements of the local algebra can be pictured as linear combinations of elements of the form

$$\dots \mathbb{1} \otimes \mathbb{1} \otimes A_{i_1} \otimes \dots \otimes A_{i_n} \otimes \mathbb{1} \otimes \mathbb{1} \otimes \dots ,$$

where $i_1, \dots , i_n \in \mathbb{Z}$ and $A_{i_k} \in \mathcal{C}_0$.

Note that the local algebra can also be given as the union of all subalgebras of the form $\mathcal{C}_{\{-n,\dots,n\}}$. The half-infinite chain $\mathcal{C}_{\mathbb{N}}$ is defined as the C^*-subalgebra of $\mathcal{C}$, generated by the union of all subalgebras of the form $\mathcal{C}_{\{0\dots,n\}};\quad n \geq 0$.

5.1.1. Example. The classical chain

In the classical case, a spin-$\frac{1}{2}$ particle is modeled by a two-level system with configuration space $\{-1, +1\}$. The confuguration space of a system of spins located at the sites corresponding to a finite subset $\Lambda \subset \mathbb{Z}$ is $\Omega_\Lambda := \{-1, +1\}^\Lambda$; the observable algebra $\mathcal{F}(\Omega_\Lambda)$ is isomorphic to $\mathcal{C}_\Lambda = \otimes_{k\in\Lambda} \mathcal{F}(\{-1, 1\})$. If we endow the finite set $\{-1, +1\}$ with the discrete topology, then the configuration space of the infinite system $\Omega := \{-1, +1\}^{\mathbb{Z}}$ becomes a topological space with the product topology, which is homeomorphic to $C \times C$, where C is the Cantor set

$$C := \left\{ \sum_{n\in\mathbb{N}} \frac{x_n}{3^n} \ :\ x_n \in \{0, 2\},\ n \in \mathbb{N} \right\} \subset [0, 1];$$

an explicit homeomorphism is given by

$$h(\underline{\omega}) := \left(\sum_{n=-1}^{-\infty} \frac{\omega_n + 1}{3^{-n}} ; \sum_{n=0}^{+\infty} \frac{\omega_n + 1}{3^{n+1}} \right); \qquad\qquad \underline{\omega} \in \Omega.$$

The Cantor set can be written in the form

$$C = \bigcap_{n\in\mathbb{N}} \bigcup_{k=0}^{2^n-1} I_{k,n},$$

where $I_{k,n}$ is an interval of the form $\left[\sum_{l=1}^{n}\frac{x_l}{3^l},\sum_{l=1}^{n}\frac{x_l}{3^l}+\frac{1}{3^n}\right]$ with $\sum_{l=1}^{n}x_l 2^{l-1}=k$.

The algebra $\mathcal{C}_\Lambda$ can be identified with functions on Ω that depend only on the coordinates corresponding to Λ; these functions are obviously continuous on Ω. In the case $\Lambda=[-n,n-1]$ the elements of $\mathcal{C}_\Lambda$ can also be pictured as functions on $C\times C$ that are constants on the sets $(I_{k,n}\times I_{l,n})\cap(C\times C)$. Since the collection of these functions for all n forms a dense set in $\mathcal{F}_c(C\times C)$ then the observable algebra $\mathcal{C}$ can be identified by the commutative C^*-algebra $\mathcal{F}_c(C\times C)$. The algebra of the left (as well as the right) half-chain in this picture is naturally isomorphic to $\mathcal{F}_c(C)$. Note that since $C\times C$ is homeomorphic to C then also $\mathcal{C}$ is isomorphic to $\mathcal{F}_c(C)$.

The quantum spin chain algebra can also be given as the universal C^*-algebra[1] corresponding to a set of generators and relations. Two canonical choices are the **spin operators** $\{S_j^n : n\in\mathbb{Z};\ j=1,2,3\}$, satisfying relations

$$\text{(R1)}\quad (S_j^n)^*=S_j^n \qquad \forall j\ \forall n \tag{5.1}$$

$$\text{(R2)}\quad [S_j^n,S_k^m]=\begin{cases}0\,, & \text{if } n\neq m;\\ i\epsilon_{jkl}S_l^n\,, & \text{if } n=m,\end{cases} \tag{5.2}$$

and the **commuting matrix units** $\{e_{ij}^n : n\in\mathbb{Z};\ i,j=1,\dots,d\}$, satisfying relations

$$\text{(R1')}\quad (e_{ij}^n)^*=e_{ji}^n \tag{5.3}$$

$$\text{(R2')}\quad e_{ij}^n\,e_{kl}^m=\begin{cases}e_{kl}^m\,e_{ij}^n\,, & \text{if } n\neq m;\\ \delta_{jk}\,e_{il}^n\,, & \text{if } n=m\end{cases} \tag{5.4}$$

$$\text{(R3')}\quad \sum_{i=1}^{2s+1}e_{ii}=\mathbb{1}. \tag{5.5}$$

Here $[.,.]$ denotes the commutator, δ_{ij} is the Kronecker delta, and ϵ_{jkl} is the Levi-Civita symbol. A set of matrix units can be given in $\mathcal{C}_\Lambda$ by the definition

$$e_{\boldsymbol{i}\boldsymbol{j}}:=\prod_{n\in\Lambda}e_{i_n j_n}; \qquad \boldsymbol{i},\boldsymbol{j}\in\{1,\dots,d\}^\Lambda. \tag{5.6}$$

[1]For the definition of a universal C^*-algebra see Appendix A.1.

For the quantum spin-$\frac{1}{2}$ case it is convenient to introduce the **Pauli operators**

$$\sigma_x^n := \sigma_1^n := 2\,S_1^n; \qquad \sigma_y^n := \sigma_2^n := 2\,S_2^n; \qquad \sigma_z^n := \sigma_3^n := 2\,S_3^n\,; \tag{5.7}$$

a commuting set of matrix units can then be given by

$$\begin{aligned} e_{11}^n &:= \frac{1}{2}\left(\mathbb{1} + \sigma_z^n\right); \qquad & e_{12}^n &:= \tfrac{1}{2}\left(\sigma_x^n + i\sigma_y^n\right); \\ e_{21}^n &:= \frac{1}{2}\left(\sigma_x^n - i\sigma_y^n\right); \qquad & e_{22}^n &:= \tfrac{1}{2}\left(\mathbb{1} - \sigma_z^n\right). \end{aligned} \tag{5.8}$$

The **right shift** is defined on the generators as $\gamma\left(S_k^n\right) := S_k^{n+1}$, and is easily seen to extend to an automorphism of the spin chain. For the half-infinite chain the right shift is defined the same way, and its effect on local elements can be pictured as

$$\gamma:\ A_1 \otimes \ldots \otimes A_n \otimes I \otimes \ldots \mapsto I \otimes A_1 \otimes \ldots \otimes A_n \otimes I \otimes \ldots\ .$$

States on the spin chain

For a state φ on the spin chain the restrictions $\varphi_\Lambda := \varphi \restriction_{\mathcal{C}_\Lambda}$ form a **compatible family**, i.e. $\varphi_{\Lambda'}(A) = \varphi_\Lambda(A)$ holds when $\Lambda \subset \Lambda'$ and $A \in \mathcal{C}_\Lambda$. On the other hand, given a compatible family $\{\varphi_\Lambda\}$ of states, there exists a unique state φ on $\mathcal{C}$ such that $\varphi_\Lambda = \varphi \restriction_{\mathcal{C}_\Lambda} \quad \forall\Lambda$ holds. To determine a state of the spin chain it is actually sufficient to know all the restrictions to subalgebras of the form $\mathcal{C}_{\{-n,\ldots,n\}}$ (or $\mathcal{C}_{\{0,\ldots,n\}}$ in the half-infinite case). The density matrices of the states φ_Λ are given by

$$[\varphi_\Lambda]_{\boldsymbol{i},\boldsymbol{j}} = \varphi_\Lambda\left(e_{\boldsymbol{j}\boldsymbol{i}}\right); \qquad \boldsymbol{i},\boldsymbol{j} \in \{1,2\}^\Lambda.$$

A state φ on the (half-)infinite spin chain is called **translation-invariant** (or **shift-invariant**), if

$$\varphi \circ \gamma = \varphi$$

holds for the right shift γ. To specify a shift-invariant state it suffices to know all the restrictions

$$\varphi_n := \varphi \restriction_{\mathcal{C}_{\{0,\ldots,n-1\}}}$$

even for the two-sided infinite chain $\mathcal{C}$. Note that the restriction to $\mathcal{C}_{\mathbb{N}}$ of a shift-invariant state on $\mathcal{C}$ gives a shift-invariant state, while a shift-invariant state on $\mathcal{C}_{\mathbb{N}}$ can be uniquely extended to a shift-invariant state of $\mathcal{C}$, giving a one-to one correspondence between shift-invariant states of the full chain and the right half chain.

The set of translation-invariant states is convex; its extremal points are called **ergodic states**. Note that a pure translation-invariant state is extremal in the convex set of all states, and so it is also ergodic. An equivalent characterization of ergodicity is the following [3]:

5.1.2. Theorem. A state φ is ergodic if and only if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi\left(a\gamma^k(b)\right) = \varphi(a)\varphi(b)\,, \qquad a, b \in \mathcal{C}. \tag{5.9}$$

Note that by continuity it is enough to check (5.9) for local elements a, b.

5.1.3. Example. Given a state ρ on $\mathcal{C}_0$, we can define a compatible family of states by

$$\varphi_n(A_0 \otimes \ldots \otimes A_{n-1}) = \rho(A_0) \cdot \ldots \cdot \rho(A_{n-1}); \qquad A_0, \ldots, A_{n-1} \in \mathcal{C}_0.$$

The resulting state on $\mathcal{C}$ is denoted by $\rho^{\otimes\infty}$. States of this form are called (shift-invariant) **product states**.

A product state is always ergodic, as easily seen from the characterization (5.9), and is pure if and only if ρ is a pure state on $\mathcal{C}_0$.

By Example (5.1.1) and Example (A.3.3) a pure state on the classical spin-$\frac{1}{2}$ chain is a Dirac measure concentrated on an infinite configuration $\underline{\omega}$, and therefore it is shift-invariant if and only if $\omega_k = 1 \;\; \forall k$ or $\omega_k = -1 \;\; \forall k$. On the quantum spin chain, however, there is an abundance of shift-invariant pure states, showing a far richer structure of the state space of the quantum chain.

It is in general rather difficult to check properties (e.g. purity or ergodicity) of a state, resulting from a compatible family $\{\rho_n\}$. Two important classes with well handable criteria are the quasi-free states, presented in section 5.3, and the finitely correlated states, that contain quantum Markov states as a subclass, which in turn contains classical Markov states.

5.1.4. Example. Finitely correlated states

Let $\mathcal{B}$ be a finite dimensional C^*-algebra with a state ρ on it, and $\mathbb{E} : \mathcal{A} \otimes \mathcal{B} \to \mathcal{B}$ be a unital CP map, where $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ is the one-site algebra of the spin chain; hence $\mathbb{E}^* : \mathcal{B} \to \mathcal{A} \otimes \mathcal{B}$ is a stochastic map. $\mathbb{E}$ is related to ρ such that

$$\mathrm{Tr}_{\mathcal{A}}\, \mathbb{E}^*(\rho) = \rho \tag{5.10}$$

holds. Let

$$\begin{aligned} \omega_1 &:= \mathbb{E}^*(\rho) \\ \omega_2 &:= (\mathrm{id}_{\mathcal{A}} \otimes \mathbb{E}^*) \circ \mathbb{E}^*\ (\rho) \\ &\ \vdots \\ \omega_n &:= \left(\mathrm{id}_{\mathcal{A}}^{\otimes(n-1)} \otimes \mathbb{E}^*\right) \circ \ldots \circ (\mathrm{id}_{\mathcal{A}} \otimes \mathbb{E}^*) \circ \mathbb{E}^*\ (\rho)\ ; \end{aligned}$$

here ω_n is defined on $\mathcal{A}^{\otimes n} \otimes \mathcal{B}$. To obtain a family on the spin chain, we take

$$\varphi_n := \mathrm{Tr}_{\mathcal{B}}\, \omega_n;$$

on simple product operators it takes the value

$$\varphi_n(A_1 \otimes \ldots \otimes A_n) = \rho\left(\mathbb{E}\left(A_1 \otimes \mathbb{E}\left(A_2 \otimes \ldots \mathbb{E}\left(A_n \otimes \mathbb{1}_{\mathcal{B}}\right) \ldots\right)\right)\right) .$$

Compatibility of this family is guaranteed by the unitality of $\mathbb{E}$, while shift-invariance follows from (5.10). The induced state φ on the spin chain is called a finitely correlated state. By lemma 2.5, [33], $\mathcal{B}$ can always be taken to be $\mathcal{B}(\mathcal{K})$ for some finite dimensional Hilbert space $\mathcal{K}$, and ρ can be assumed to be faithful. Note that the construction of the sequence ω_n resembles very much to that of Markov states (Example (5.1.5)). However, the construction of the sequence φ_n doesn't follow the scheme for Markovian states, as φ_{n+1} is derived from ω_{n+1} instead of φ_n. As a consequence, the resulting state φ is in general not Markovian.

A finitely correlated state is purely generated, if $\mathbb{E} = \mathrm{Ad}_{V^*}$ for an isometry $V : \mathcal{K} \to \mathcal{H} \otimes \mathcal{K}$. A purely generated state is pure if and only if the spectrum of $\hat{\mathbb{E}} : b \mapsto \mathbb{E}(\mathbb{1}_{\mathcal{A}} \otimes b)$ intersects the complex unit circle only at $\{1\}$, and 1 is a simple eigenvalue.

A finitely correlated state is ergodic, if and only if 1 is a simple eigenvalue of $\hat{\mathbb{E}}$ (Proposition 3.1, [33]).

A special subclass is when the algebra $\mathcal{B}$ is a commutative one, isomorphic to $\mathcal{F}(\mathcal{X})$ for some finite set $\mathcal{X}$. $\mathbb{E}^*$ is specified by its values on the Dirac measures δ_x, and since the result is a state on $\mathcal{A} \otimes \mathcal{F}(\mathcal{X})$, it can uniquely be written as $\sum_y p_{xy}\, \rho_{xy} \otimes \delta_y$, where $\{p_{xy}\}$ is a probability distribution for a fixed x, and ρ_{xy} are states on $\mathcal{A}$. The state ρ is of the form $\rho = \sum_x p_x\, \delta_x$, and (5.10) is equivalent to $\{p_x\}$ being an invariant measure of the stochastic map with matrix $T_{x,y} := p_{xy}$. The resulting states are of the form

$$\varphi_n = \sum_{\{x_1,\dots,x_{n+1}\}} \mu(x_1, \dots , x_{n+1})\, \rho_{x_1 x_2} \otimes \dots \otimes \rho_{x_n x_{n+1}}\,,$$

where μ is the classical Markov measure, generated by $\{p_x\}$ and T. In this case $\hat{\mathbb{E}}$ is the linear map on $\mathcal{F}(\mathcal{X})$ with matrix T, thus φ is ergodic if and only if 1 is a simple eigenvalue of T, which is equivalent to the ergodicity of the classical measure μ. We obtain a special form when ρ_{xy} is independent of x; in this case

$$\varphi_n = \sum_{\{x_1,\dots,x_n\}} \mu(x_1, \dots , x_n)\, \rho_{x_1} \otimes \dots \otimes \rho_{x_n}\,. \tag{5.11}$$

5.1.5. Example. Markov states

A shift-invariant state ρ on the infinite spin chain is called a Markov state if there exists a stochastic map $\alpha : \mathcal{A} \to \mathcal{A} \otimes \mathcal{A}$ such that

$$\rho_{n+1} = \left(\mathrm{id}_{\mathcal{A}}^{\otimes(n-1)} \otimes \alpha\right)\, (\rho_n)\,. \tag{5.12}$$

Note that the criterion for Markovianity doesn't provide a method to construct Markovian states, as for a given map α and state φ_1 the sequence
$\varphi_1 := \alpha\, (\varphi_1)\,;$
$\varphi_2 := (\mathrm{id}_{\mathcal{A}} \otimes \alpha) \circ \alpha\, (\varphi_1)\, ; \dots$ is in general not a compatible one.

Markov states are finitely correlated. Indeed, the choice

$$\mathcal{B} := \mathcal{A}; \qquad \mathbb{E} := \alpha^* \qquad \text{and} \qquad \rho := \varphi_1$$

yields $\omega_n = \varphi_{n+1}$, and $\mathrm{Tr}_{\mathcal{B}}\, \omega_n = \mathrm{Tr}_{n+1}\, \varphi_{n+1} = \varphi_n$. Compatibility and shift-invariance implies

$$\mathrm{Tr}_1\, \alpha\, (\varphi_1) = \mathrm{Tr}_2\, \alpha\, (\varphi_1) = \varphi_1\,.$$

Note that (5.12) is equivalent to ρ_{n+2} saturating (SSA) with respect to the partition $\mathcal{A}^{\otimes n} \otimes \mathcal{A} \otimes \mathcal{A}$ (see Chapter 4 for details), therefore the algebraic characterization (Theorem 4.1.3) can be used to construct Markov states. We can specify a Markov state by

- a splitting $\mathcal{H} = \bigoplus_{k=1}^{K} \mathcal{H}_k^{\mathrm{L}} \otimes \mathcal{H}_k^{\mathrm{R}}$;
- a set of states $\{\rho_{kl} \,:\, k, l = 1, \ldots, K\}$, where ρ_{kl} is a state on $\mathcal{H}_k^{\mathrm{R}} \otimes \mathcal{H}_l^{\mathrm{L}}$;
- a classical shift-invariant Markov measure μ on $\{1, \ldots, K\}^\infty$.

Let

$$\rho_l^{\mathrm{L}} := \sum_k \frac{\mu(k,l)}{\mu(l)} \operatorname{Tr}_{\mathcal{H}_k^{\mathrm{R}}} \rho_{kl} \qquad \text{and} \qquad \rho_k^{\mathrm{R}} := \sum_l \frac{\mu(k,l)}{\mu(k)} \operatorname{Tr}_{\mathcal{H}_l^{\mathrm{L}}} \rho_{kl}; \quad (5.13)$$

Then

$$\begin{aligned} \rho_1 &= \bigoplus_k \mu(k)\, \rho_k^{\mathrm{L}} \otimes \rho_k^{\mathrm{R}}, \qquad \text{and} \\ \rho_n &= \bigoplus_{\{k_1, \ldots, k_n\}} \mu(k_1 \,\ldots, k_n)\, \rho_{k_1}^{\mathrm{L}} \otimes \rho_{k_1 k_2} \otimes \ldots \otimes \rho_{k_{n-1} k_n} \otimes \rho_{k_n}^{\mathrm{R}}. \end{aligned} \quad (5.14)$$

Markovianity of μ implies $\mu(k_1, \ldots, k_n) = \mu(k_{n-1}) \frac{\mu(k_1, \ldots, k_{n-1})}{\mu(k_{n-1})} \frac{\mu(k_{n-1}, k_n)}{\mu(k_{n-1})}$, hence

$$\rho_n = \bigoplus_{k_{n-1}} \mu(k_{n-1}) \left[\bigoplus_{\{k_1, \ldots, k_{n-2}\}} \frac{\mu(k_1, \ldots, k_{n-1})}{\mu(k_{n-1})} \rho_{k_1}^{\mathrm{L}} \otimes \rho_{k_1 k_2} \otimes \ldots \otimes \rho_{k_{n-2} k_{n-1}} \right]$$

$$\bigotimes \left[\bigoplus_{k_n} \frac{\mu(k_{n-1}, k_n)}{\mu(k_{n-1})} \rho_{k_{n-1} k_n} \otimes \rho_{k_n}^{\mathrm{R}} \right],$$

therefore it satisfies the structural criterion (4.14) for all $n \geq 2$.

Symmetries

As in the general theory of dynamical systems, symmetries are described by the same mathematical tool as the dynamical evolution of the system, i.e. by (a group of) automorphisms. Symmetries can be local, that is, automorphisms of a local algebra $\mathcal{C}_\Lambda$, or global, i.e. automorphisms of $\mathcal{C}$. If a global automorphism preserves the local structure of the spin chain then it can be restricted to give a local automorphism on each local algebra. As for states, the thermodynamical limit of a family of local automorphisms can be defined. If a given family of local automorphisms leaves the local Hamiltonians invariant, then in an ideal case the same holds for the thermodynamical limit automorphism and the thermodynamical limit of the equilibrium states. Here we describe in some detail the automorphism group of rotations, as it is going to have some importance later on.

The map

$$\kappa_w : \; \sigma_x^k \mapsto \cos\theta \; \sigma_x^k + \sin\theta \; \sigma_y^k; \quad \sigma_y^k \mapsto -\sin\theta \; \sigma_x^k + \cos\theta \; \sigma_y^k; \quad \sigma_z^k \mapsto \sigma_z^k$$

for $w = e^{i\theta}$; $\theta \in [0, 2\Pi)$ describes a rotation of angle θ in the xy plane, and extends to an automorphism of the infinite chain for every w. Note that the definition of κ_w depends on the choice of spin operators σ_x, σ_y, σ_z, i.e. on the choice of the basis. The map $w \mapsto \kappa_w$ gives a representation of the complex unit circle $\mathcal{T}$ in the automorphism group of the spin chain, which we will call the **rotation group** for the sake of simplicity. Since $\kappa_w\left(\mathcal{C}_\Lambda\right) = \mathcal{C}_\Lambda$, it can be restricted to an automorphism $\left(\kappa_w\right)_\Lambda$ of the algebra $\mathcal{C}_\Lambda$.

The set $\{-1, 1\}$ forms an order two subgroup of the unit circle; the order two automorphism κ_{-1} we call the **parity automorphism**. It describes a rotation of angle $180°$ in the xy plane. The elements of its fixed point algebra are called **even**, while elements for which $\kappa_{-1}(a) = -a$ are called **odd**. Every element can uniquely be decomposed into a sum of an even and an odd element in the form

$$a = a_+ + a_-; \qquad a_+ = \frac{1}{2}\left(a + \kappa_{-1}(a)\right), \qquad a_- = \frac{1}{2}\left(a - \kappa_{-1}(a)\right).$$

Product of even elements is even again, therefore the set of even elements forms a subalgebra $\mathcal{C}_+$, while the set of odd elements is not closed under multiplication, and therefore not an algebra. A matrix unit e_{ij}^n is even, if $i = j$, and odd, if $i \neq j$; thus matrix units in $\mathcal{C}_\Lambda$ of the form (5.6) are even (odd) if and only if $\sum_{n\in\Lambda}\left(i_n - j_n\right)$ is even (odd).

We call a state φ **rotation-invariant**, if $\varphi \circ \kappa_w = \varphi \;\; \forall w \in \mathcal{T}$, and **even**, if $\varphi \circ \kappa_{-1} = \varphi$. Note that a state is even if and only if it vanishes on odd elements, therefore for the entries of the density matrices of an even state we have

$$[\varphi_\Lambda]_{\boldsymbol{i},\boldsymbol{j}} = 0, \qquad \text{if} \qquad \sum_{k\in\Lambda}\left(i_k - j_k\right) \text{ is odd.}$$

Physical models

Mathematical models describing effectively one-dimensional spin systems are in general specified by a sequence of local Hamiltonians $H_N \in \mathcal{C}_{[-N,N]}$. The thermal equilibrium state of a finite subsystem at inverse temperature β is then the Gibbs state $\rho_{N,\beta}$ with density $e^{-\beta H_N} \big/ \operatorname{Tr} e^{-\beta H_N}$. A state ρ is a ground state, if it minimizes the expectation value of H_N. It is easily seen that ρ is a ground state if and only if its support is contained in the lowest-energy eigensubspace of H_N. The zero-temperature limit $\rho_{N,\infty} := \lim_{\beta\to\infty} \rho_{N,\beta}$ of the Gibbs states gives a particular ground state. The thermodynamical limit of the states $\rho_{N,\beta}$ is defined on local elements as

$$\rho_\beta(A) := \lim_{N\to\infty} \rho_{N,\beta}(A); \qquad\qquad A \in \mathcal{C}_{\text{loc}} \,. \tag{5.15}$$

For a detailed analysis of conditions that guarantee the existence of the above limit, see [22].

5.1.6. Example. The XY model describes nearest-neighbor interaction of spins in the presence of a homogeneous z-directional magnetic field h. It is specified by a sequence of local Hamiltonians of the form

$$H_N := \sum_{k=-N}^{N-1} \left(\frac{1+\delta}{2}\, \sigma_x^k \sigma_x^{k+1} + \frac{1-\delta}{2}\, \sigma_y^k \sigma_y^{k+1} \right) + h \sum_{k=-N}^{N} \sigma_z^k \tag{5.16}$$

where δ is a real parameter. Special cases are the Ising model, where $|\delta| = 1$, and the XX-model, where $\delta = 0$.

The parity automorphism leaves the local Hamiltonians of the XY-model invariant, i.e. κ_{-1} is a symmetry of the local systems, and the local Gibbs states (as well as the ground states) are even, and these two properties carry through to the thermodynamical limit. The local Hamiltonians of the XX-model are invariant under the whole rotation group, and so the local Gibbs and the ground states are rotation-invariant, the and the same holds after taking the thermodynamical limit.

The first solution of the ground state problem of the XY model was given in [70] for the case $h = 0$. The case $h \neq 0$ was studied for the Ising model in [94], and for the general XY model in [10]. In Appendix C.1 we present an explicit computation, following the method of [6], that shows that the ground state of the XX model can be identified as a pure translation-invariant quasi-free state (see section 5.3 for quasi-free states on the spin chain).

5.2 The CAR algebra

The observable algebra of an indeterminate number of Fermions with one-particle Hilbert space $\mathcal{H}$ is called the CAR (**c**anonical **a**nticommutation **r**elations) algebra on $\mathcal{H}$; in notation $CAR(\mathcal{H})$. Mathematically it can be described as the universal C^*-algebra given by the set of generators $\{c(x) : x \in \mathcal{H}\}$ satisfying relations

$$\text{(R1”)} \quad \text{The map } x \mapsto c(x) \text{ is complex anti-linear;} \tag{5.17}$$

$$\text{(R2”)} \quad \{c(x), c(y)\} = 0; \tag{5.18}$$

$$\text{(R3”)} \quad \{c^*(x), c(y)\} = \langle y, x \rangle \, \mathbb{1}; \tag{5.19}$$

where $\{\, . \, , \, . \,\}$ denotes the anti-commutator.

The map $x \mapsto c(x)$ is easily seen to be norm-preserving, which, by (anti-)linearity implies that $CAR(\mathcal{H})$ can equally be specified as the universal C^*-algebra with generators $\{c_n : n = 1, \dots, \dim \mathcal{H}\}$ satisfying

$$\{c_n, c_m\} = 0 \qquad \text{and} \qquad \{c_n^*, c_m\} = \delta_{nm} \mathbb{1} \, . \tag{5.20}$$

The relation between the two pictures is that c_n equals to $c(e_n)$, where $\{e_n : n = 1, \dots, \dim \mathcal{H}\}$ is an orthonormal base in $\mathcal{H}$.

A possible realization of relations (5.20) can be obtained by defining $\hat{c}_n := \left(\prod_{j=0}^{n-1} \sigma_z^j \right) \sigma_+^n$ in the spin-$\frac{1}{2}$ quantum chain, with $\sigma_+^n := \frac{1}{2}\left(\sigma_x^n + i \sigma_y^n \right)$. The set $\{\hat{c}_n : n \in \mathbb{N}\}$ is easily seen to generate the half-infinite spin chain, thus $\mathcal{C}_{\mathbb{N}}$ is a homomorphic image of $CAR(\mathcal{H})$. On the other hand, the elements

$$\begin{array}{lllllll} E_{11}^n & := & c_n c_n^* & \qquad & E_{22}^n & := & c_n^* c_n \\ E_{12}^n & := & \left(\prod\limits_{j=0}^{n-1} \left(1 - 2c_j^* c_j\right) \right) c_k & \qquad & E_{21}^n & := & \left(\prod\limits_{j=0}^{n-1} \left(1 - 2c_j^* c_j\right) \right) c_k^* \end{array}$$

in $CAR(\mathcal{H})$ satisfy relations (R1'), (R2') and (R3') of commuting matrix units, moreover, they generate $CAR(\mathcal{H})$, therefore $CAR(\mathcal{H})$ is isomorphic to $\mathcal{C}_{\mathbb{N}}$. An explicit isomorphism is obtained by extending the map

$$\mathbb{J} : E_{ij}^n \mapsto e_{ij}^n,$$

where $\{e_{ij}^n\}$ is an arbitrary set of commuting matrix units in $\mathcal{C}_{\mathbb{N}}$. This isomorphism is called the **Jordan-Wigner isomorphism**.

A particular representation of the CAR algebra is the **Fock representation**. The Hilbert space carrying it is the antisymmetric Fock space

$$\mathcal{F}(\mathcal{H}) := \bigoplus_{n\in\mathbb{N}} \wedge^n \mathcal{H}\,, \tag{5.21}$$

where $\wedge^n\mathcal{H}$ is the n-th antisymmetric tensor power of $\mathcal{H}$, spanned by vectors of the form

$$x_1 \wedge \ldots \wedge x_n := \frac{1}{\sqrt{n!}} \sum_{\sigma\in S_n} \varepsilon(\sigma) x_{\sigma(1)} \otimes \ldots \otimes x_{\sigma(n)}\,.$$

Here $\varepsilon(\sigma)$ is the sign of the permutation σ. Representants of $c(y)^*$ are the creation operators, defined by

$$a^*(y)\,(x_1 \wedge \ldots \wedge x_n) := y \wedge x_1 \wedge \ldots \wedge x_n\,.$$

The adjoints $a(y)$ are called the annihilation operators. The number operator N is the unbounded operator that equals to $n\, I_{\wedge^n\mathcal{H}}$ when restricted to $\wedge^n\mathcal{H}$. It can also be expressed in the form $N = \sum_k a_k^* a_k$, where $a_k = a(e_k)$ for some orthonormal base $\{e_k\}$ in $\mathcal{H}$.

5.2.1 Quasi-free morphisms

An isometry $V : \mathcal{H} \to \mathcal{K}$ between the Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ defines a map

$$\gamma_V\,(c(x)) := c\,(Vx)$$

from $CAR(\mathcal{H})$ to $CAR(\mathcal{K})$. It preserves the CAR relations, i.e.

$$\{\gamma_V\,(c^*(x))\,, \gamma_V\,(c(y))\} = \langle y, x\rangle\,\mathbb{1} \tag{5.22}$$

$$\{\gamma_V\,(c(x))\,, \gamma_V\,(c(y))\} = 0 \tag{5.23}$$

holds. As a consequence, γ_V extends to a homomorphism from $CAR(\mathcal{H})$ to $CAR(\mathcal{K})$, with range $CAR\,(V\,(\mathcal{H}))$. Such a homomorphism is called a **quasi-free homomorphism**. In particular, a unitary U defines a quasi-free automorphism of $CAR(\mathcal{H})$.

5.2.1. Example. The **parity automorphism** α is the quasi-free automorphism generated by the unitary $-I$, i.e.

$$\alpha(c(x)) = c(-x).$$

The fixed point algebra of α is

$$CAR(\mathcal{H})_+ := \{b \in CAR(\mathcal{H}) \ : \ \alpha(b) = b\};$$

its elements are called **even**, while elements with the property $\alpha(b) = -b$ are called **odd**. $CAR(\mathcal{H})_+$ is generated by elements of the form

$$c^*_{i_1} \dots c^*_{i_n} c_{j_m} \dots c_{j_1}; \qquad n+m \text{ is even },$$

while the subset of the odd elements is the closure of the linear span of elements of the form

$$c^*_{i_1} \dots c^*_{i_n} c_{j_m} \dots c_{j_1}; \qquad n+m \text{ is odd }.$$

Any element $b \in CAR(\mathcal{H})$ can uniquely be decomposed into a sum of an even and an odd element; the decomposition is of the form

$$b = b_+ + b_- \,; \qquad b_+ = \frac{1}{2}\left(b + \alpha(b)\right), \quad b_- = \frac{1}{2}\left(b - \alpha(b)\right).$$

A functional φ on $CAR(\mathcal{H})$ is called **even**, if $\varphi \circ \alpha = \varphi$, or equivalently, if φ vanishes on odd elements.

5.2.2. Example. The **gauge group** of $CAR(\mathcal{H})$ is the group of automorphisms $(\kappa_w)_{w\in\mathcal{T}}$, where κ_w is the quasi-free automorphism given by the unitary wI on the one-particle space, and $\mathcal{T}$ is the complex unit circle. The subalgebra

$$\{b \ : \ \kappa_w(b) = b \ \ \forall w \in \mathcal{T}\}$$

is called the gauge-invariant part of $CAR(\mathcal{H})$, and is generated by the monomials

$$c^*_{i_1} \dots\dots c^*_{i_n} c_{j_m} \dots c_{j_1}; \qquad n = m.$$

A state φ is called gauge-invariant, if $\varphi \circ \kappa_w = \varphi$ holds for all $w \in \mathcal{T}$.

Note that $\kappa_{-1} = \alpha$ is the parity automorphism.

5.2.3. Example. The shift operator

$$S\,\delta_k := \delta_{k+1}$$

on $\mathcal{H} := l^2(\mathbb{Z})$, where $\{\delta_k \,:\, k \in \mathbb{Z}\}$ is the standard base in $\mathcal{H}$, is a unitary operator; the quasi-free automorphism generated by it is called the **shift automorphism**, and is denoted by γ. The shift automorphism is characterized by the property

$$\gamma(c_k) = c_{k+1}$$

with $c_k = c(\delta_k)$.

A state ω on $CAR\,(l^2\,(\mathbb{Z}))$ is called **shift-invariant**, if

$$\omega = \omega \circ \gamma$$

holds.

5.2.4. Example. In the Fock representation every quasi-free automorphism β with one-particle unitary U is of the form $\beta = \mathrm{Ad}_{U_{\mathrm{F}}}$, where $U_{\mathrm{F}} := \bigoplus_n \wedge^n U$ is a unitary on the Fock space. The unitary $\wedge^n U$ on the n-particle space is defined by $(\wedge^n U)\;(x_1 \wedge \ldots \wedge x_n) := Ux_1 \wedge \ldots \wedge Ux_n$.

5.2.5. Example. The unitary evolution of an interaction-free Fermionic system is the typical physical example of a quasi-free evolution. The second-quantized Hamiltonian $\hat{H}$ is the direct sum of the restrictions of $H_n := H \otimes I \otimes \ldots \otimes I + I \otimes H \otimes I \otimes \ldots \otimes I + I \otimes \ldots \otimes I \otimes H$ onto the antisymmetrized n-particle spaces, where H is the one-particle Hamiltonian. The corresponding unitary group is $U(t) := e^{-it\hat{H}}$, and the evolution of the annihilation operators in the Fock representation is

$$c(x)(t) = e^{it\hat{H}}\,c(x)\,e^{-it\hat{H}} = c\left(e^{itH}x\right) \tag{5.24}$$

5.2.2 Quasi-free states

Suppose an interaction-free fermionic system is described by a Hamilton operator H with discrete spectrum, such that $e^{-\beta\hat{H}}$ is a trace-class operator for all positive β, where $\hat{H}$ is the second-quantized Hamiltonian. Then the Gibbs state ρ_β of the system in the Fock representation has density operator $\frac{e^{-\beta\hat{H}}}{\mathrm{Tr}\,e^{-\beta\hat{H}}}$, and its value on monomials can be expressed as

$$\rho_\beta\;(c(x_1)^* \ldots c(x_n)^* c(y_m) \ldots c(y_1)) = \delta_{m,n} \det\{\langle y_i, Q\;x_j\rangle\,, \tag{5.25}$$

where Q has the same eigenvectors as H with corresponding eigenvalues $\frac{e^{-\beta h_i}}{1+e^{-\beta h_i}}$, where the h_i's are the eigenvalues of H.

In general, given an operator Q on the one-particle Hilbert space, a functional defined on monomials by the right-hand side of (5.25) extends to a state of the CAR algebra if and only if $0 \leq Q \leq I$; the resulting state is called a **quasi-free state** with **symbol** Q. A quasi-free state is pure if and only if its symbol Q is a projection. The two extremes are the **Fock state** and the **anti-Fock state**, with symbol 0 and I; these two states are completely determined by

$$\omega_{\mathrm{F}}(c(x)^* c(x)) = 0 \qquad \text{and} \qquad \omega_{\mathrm{aF}}(c(x) c(x)^*) = 0\,. \tag{5.26}$$

The Fock representation is the GNS representation of the Fock state; the representing vector is the vacuum vector $\Omega \in \wedge^0 \mathcal{H}$.

For a subspace $\mathcal{K} \subset \mathcal{H}$ let $\mathcal{A}_{\mathcal{K}}$ be the subalgebra of $CAR(\mathcal{H})$, generated by $\{c(x) \,:\, x \in \mathcal{K}\}$. The restriction of a quasi-free state to $\mathcal{A}_{\mathcal{K}}$ is easily seen to be a quasi-free state itself, with symbol PQP, where P is the projection onto $\mathcal{K}$.

For a finite dimensional Hilbert space $\mathcal{H}$ the symbol Q yields a base of the space, consisting of its eigenvectors. The corresponding Jordan-Wigner isomorphism sends the quasi-free state into a product state on the spin chain, with density

$$D = \bigotimes_{k=1}^{\dim \mathcal{H}} \begin{bmatrix} q_k & 0 \\ 0 & 1 - q_k \end{bmatrix} \tag{5.27}$$

where the q_k are the eigenvalues of Q.

A quasi-free state ω_Q on $CAR\,(l^2\,(\mathbb{Z}))$ is translation-invariant if and only if its symbol is, i.e. it commutes with the shift operator of $l^2\,(\mathbb{Z})$. Shift-invariant operators are also called Toeplitz operators. A shift-invariant symbol is uniquely determined by a sequence $q \,:\, \mathbb{Z} \to \mathbb{C}$, such that the matrix of Q in the standard base of $l^2\,(\mathbb{Z})$ is of the form $Q_{k,l} = q(k-l)$. Shift-invariance implies that the Fourier transformation maps the symbol of a shift-invariant quasi-free state to a multiplication operator $M_{\hat{q}}$ on $\mathcal{L}^2\,(\mathcal{T})$, where $\mathcal{T}$ is the one dimensional torus and $\hat{q}$ is a real-valued measurable function on $\mathcal{T}$. By identifying $\mathcal{T}$ with the interval $[0, 2\pi)$, the function $\hat{q}$ can be viewed as a function on $[0, 2\pi)$, satisfying $0 \leq \hat{q}(\theta) \leq 1$ for almost all $\theta \in [0, 2\pi)$. The defining sequence q is then the sequence of Fourier coefficients of $\hat{q}$:

$$q(k) = \frac{1}{2\pi} \int_0^{2\pi} e^{-ik\theta}\, \hat{q}(\theta)\; d\theta.$$

The state is pure if and only if $\hat{q}$ is the characteristic function χ_K of a subset $K \subset [0, 2\pi]$.

5.3 Quasi-free states on the spin chain

Let $\mathcal{A} := CAR\,(l^2(\mathbb{Z}))$ and $\mathcal{C}$ be the two-sided quantum spin-$\frac{1}{2}$ chain with corresponding shift automorphisms $\gamma_{\mathcal{A}}$ and $\gamma_{\mathcal{C}}$. Our aim in this section is to give translation-invariant states on the two-sided spin chain, by carrying shift-invariant states on $\mathcal{A}$ to $\mathcal{C}$. Since these algebras are isomorphic to each other, we can carry any state φ on $\mathcal{A}$ to a state $\varphi \circ \tau^{-1}$ on $\mathcal{C}$, where τ is any isomorphism from $\mathcal{A}$ onto $\mathcal{C}$. However, the image of a shift-invariant state need not be shift-invariant on $\mathcal{C}$, unless τ is compatible with the shifts of the two algebras, i.e.

$$\gamma_{\mathcal{C}} \circ \tau = \tau \circ \gamma_{\mathcal{A}}$$

holds. Note that the Jordan-Wigner isomorphism gives an isomorphism between $\mathcal{A}$ and the one-sided chain $\mathcal{C}_{\mathbb{N}}$, but the procedure described there cannot be modified to obtain an isomorphism onto $\mathcal{C}$. Moreover, the Jordan-Wigner isomorphism is not compatible with the shift of $\mathcal{A}$ and the shift of the one-sided chain.

Not having an isomorphism at hand that is compatible with the shifts, we use a more involved method, developed by Araki in [6], to obtain shift-invariant states on $\mathcal{C}$ from shift-invariant states on $\mathcal{A}$. The basic idea is to embed the algebra $\mathcal{A}$ into a bigger algebra $\tilde{A}$, extend the shift $\gamma_{\mathcal{A}}$ to an automorphism $\tilde{\gamma}$ of $\tilde{A}$ and the shift-invariant state φ to a $\tilde{\gamma}$-invariant state $\tilde{\varphi}$ on $\tilde{A}$, and then to find a $\tilde{\gamma}$-invariant subalgebra $\hat{\mathcal{A}}$ of $\tilde{\mathcal{A}}$, restrict $\tilde{\gamma}$ and $\tilde{\varphi}$ to $\hat{\mathcal{A}}$, and, as a last step, to find an explicit isomorphism between $\hat{\mathcal{A}}$ and $\mathcal{C}$ that is compatible with the restriction $\hat{\gamma}$ and $\gamma_{\mathcal{C}}$.

The algebra

We work in the Fock representation of $CAR\,(l^2\,(\mathbb{Z}))$, and identify $\mathcal{A}$ with the algebra generated by the Fock space annihilation operators $\{a(x) \,:\, x \in l^2\,(\mathbb{Z})\}$. We introduce the unbounded self-adjoint operator

$$N_- := \sum_{k<0} a_k^* a_k,$$

and, by means of functional calculus, the bounded operator

$$T := (-1)^{N_-} ,$$

which is a self-adjoint unitary, i.e.

$$T^* = T; \qquad T^2 = I. \tag{5.28}$$

Evaluating the product Ta_kT on vectors of the form $x_1 \wedge \ldots \wedge x_n$, and taking into account that the linear span of vectors of this form is dense in the Fock space, we get

$$Ta_kT = \begin{cases} a_k & \text{if } k \geq 0; \\ -a_k & \text{if } k < 0. \end{cases} \tag{5.29}$$

In other words, the map $b \mapsto TbT$ is the quasi-free automorphism generated by the unitary

$$U = \sum_{k \geq 0} |\delta_k\rangle\langle\delta_k| - \sum_{k<0} |\delta_k\rangle\langle\delta_k| . \tag{5.30}$$

Now we define $\tilde{\mathcal{A}}$ to be the C^*-algebra generated by $\mathcal{A}$ and T. It is shown in Appendix C.2 that T is not an element of $\mathcal{A}$, therefore $\tilde{\mathcal{A}}$ is strictly bigger than $\mathcal{A}$. Clearly, every element of $\tilde{\mathcal{A}}$ can uniquely be written in the form $a + Tb$ with $a, b \in \mathcal{A}$, that is,

$$\tilde{\mathcal{A}} = \mathcal{A} + T\mathcal{A}.$$

A straightforward computation verifies that the following maps are automorphisms of $\tilde{\mathcal{A}}$:

$$\begin{aligned} \beta\,(a + Tb) &:= a - Tb; \\ \tilde{\gamma}\,(a + Tb) &:= \gamma(a) + T\hat{\sigma}_z^0\gamma(b), \\ \tilde{\kappa}_w\,(a + Tb) &:= \kappa_w(a) + T\kappa_w(b), \end{aligned}$$

with $\hat{\sigma}_z^0 := \mathbb{1} - 2a_0^*a_0$, and that $\tilde{\gamma}$ is an extension of the shift automorphism $\gamma_{\mathcal{A}}$, while $\tilde{\kappa}_w$ extends the gauge group to $\tilde{\mathcal{A}}$. We use the notation $\tilde{\alpha}$ for the extension of the parity automorphism $\alpha = \kappa_{-1}$.

The fixed point set of the automorphism $\beta^{-1}\tilde{\alpha}$ forms a C^*-algebra, which we denote by $\hat{\mathcal{A}}$, i.e.

$$\begin{aligned}\hat{\mathcal{A}} &= \{a + Tb \in \tilde{\mathcal{A}} \,:\, \tilde{\alpha}(a+Tb) = \beta(a+Tb)\} \\ &= \{a + Tb \in \tilde{\mathcal{A}} \,:\, \alpha(a) = a,\, \alpha(b) = -b\} \\ &= \mathcal{A}_+ + T\mathcal{A}_-,\end{aligned}$$

where $\mathcal{A}_+$ and $\mathcal{A}_-$ are the even and the odd parts of $\mathcal{A}$.

For an odd element $b \in \mathcal{A}$ the element $\sigma_z^0 b$ is odd again, and since $\gamma(\mathcal{A}_+) = \mathcal{A}_+$ and $\gamma(\mathcal{A}_-) = \mathcal{A}_-$, we have

$$\tilde{\gamma}\left(\hat{\mathcal{A}}\right) = \hat{\mathcal{A}},$$

and so $\tilde{\gamma}$ can be restricted to an automorphism of $\hat{\mathcal{A}}$, which we denote by $\hat{\gamma}$, and call the shift automorphism of the algebra $\hat{\mathcal{A}}$. A similar argument shows that $\tilde{\kappa}_w$ can also be restricted to $\hat{\mathcal{A}}$; the restriction is again denoted by $\hat{\kappa}_w$.

Using the notation $\hat{\sigma}_z^k := \mathbb{1} - 2a_k^* a_k$ we define the elements

$$A_k := \begin{cases} \prod_{\ell=0}^{k-1} \hat{\sigma}_z^l\,, & \text{if } k > 0; \\ \mathbb{1}\,, & \text{if } k = 0; \\ \prod_{\ell=k}^{-1} \hat{\sigma}_z^l\,, & \text{if } k < 0. \end{cases}$$

and

$$\hat{E}_{11}^k := a_k a_k^*, \quad \hat{E}_{22}^k := a_k^* a_k, \quad \hat{E}_{12}^k := T A_k a_k, \quad \hat{E}_{21}^k := T A_k a_k^*. \tag{5.31}$$

The elements $\hat{E}_{ij}^k$ are easily seen to be in the subalgebra $\hat{\mathcal{A}}$, and a direct computation shows that they satisfy the relations (R1'), (R2') and (R3') of commuting matrix units, moreover,

$$\hat{\gamma}\left(\hat{E}_{ij}^k\right) = \hat{E}_{ij}^{k+1}\,.$$

This implies that the map

$$\Pi : \hat{E}_{ij}^k \mapsto e_{ij}^k$$

extends to an isomorphism between the algebras $\hat{\mathcal{A}}$ and $\mathcal{C}$, and that

$$\gamma_{\mathcal{C}} \circ \Pi = \Pi \circ \hat{\gamma}, \tag{5.32}$$

i.e. Π intertwines the shifts of the two algebras.

With these notations we have

$$\Pi^{-1}\left(\sigma_z^k\right) = \hat{\sigma}_z^k,$$

and

$$\hat{\sigma}_x^k := \Pi^{-1}\left(\sigma_x^k\right) = \hat{E}_{12}^k + \hat{E}_{21}^k; \qquad \hat{\sigma}_y^k := \Pi^{-1}\left(\sigma_y^k\right) = -i\hat{E}_{12}^k + i\hat{E}_{21}^k\,. \tag{5.33}$$

Since

$$\hat{\kappa}_w\left(\hat{E}_{11}^k\right) = \hat{E}_{11}^k; \quad \hat{\kappa}_w\left(\hat{E}_{22}^k\right) = \hat{E}_{22}^k; \quad \hat{\kappa}_w\left(\hat{E}_{12}^k\right) = \overline{w}\hat{E}_{12}^k; \quad \hat{\kappa}_w\left(\hat{E}_{21}^k\right) = w\hat{E}_{21}^k;$$

then we have

$$\hat{\kappa}_w\left(\hat{\sigma}_x^k\right) = \cos\theta\;\hat{\sigma}_x^k + \sin\theta\;\hat{\sigma}_y^k; \;\; \hat{\kappa}_w\left(\hat{\sigma}_y^k\right) = -\sin\theta\;\hat{\sigma}_x^k + \cos\theta\;\hat{\sigma}_y^k; \;\; \hat{\kappa}_w\left(\hat{\sigma}_z^k\right) = \hat{\sigma}_z^k;$$

with $w = e^{i\theta}$, that is,

$$\Pi \circ \hat{\kappa}_w = \kappa_w^{\mathcal{C}} \circ \Pi, \tag{5.34}$$

where now $\kappa_w^{\mathcal{C}}$ denotes the rotation group of the spin chain. In particular, the even part of the spin chain is identified with the subalgebra

$$\hat{\mathcal{A}}_+ := \{a + Tb \in \hat{\mathcal{A}} \;:\; \hat{\alpha}\,(a + Tb) = a + Tb\} = \mathcal{A}_+,$$

while the odd part is mapped by Π^{-1} to

$$\hat{\mathcal{A}}_- := \{a + Tb \in \hat{\mathcal{A}} \;:\; \hat{\alpha}\,(a + Tb) = -\,(a + Tb)\} = T\mathcal{A}_-.$$

Since the decomposition of an element into the sum of an even and an odd element is unique both in the CAR algebra and the spin chain, the map $W : \mathcal{A} \to \hat{\mathcal{A}}$;

$$Wb := b_+ + Tb_-$$

is well-defined, and gives a linear isomorphism between the two algebras, with the locality property

$$W\left(\mathcal{A}_\Lambda\right) = \hat{\mathcal{A}}_\Lambda \qquad \text{for an interval } \;\Lambda = [a,b]; \quad a \leq 0 \leq b,$$

where $\hat{\mathcal{A}}_\Lambda$ is the algebra generated by $\{\hat{E}^k_{ij} \;:\; i,j \in \{1,2\};\; k \in \Lambda\}$. Note, however, that $W(\mathcal{A}_\Lambda) = \hat{\mathcal{A}}_\Lambda$ doesn't hold for a general Λ. W is also compatible with the translations, and the gauge (rotation) groups of the algebras:

$$\hat{\gamma} \circ W = W \circ \gamma; \qquad \hat{\kappa}_w \circ W = W \circ \kappa_w.$$

The restriction onto the intersection $\mathcal{A} \cap \hat{\mathcal{A}} = \mathcal{A}_+$ is simply the identity of $\mathcal{A}_+$.

The algebra $\mathcal{A}_{[0,N-1]}$ is also generated by a set of matrix units $\left\{E^{(N)}_{\boldsymbol{ij}} := \prod_{k=0}^{N-1} E^k_{ij} \;:\; \boldsymbol{i},\boldsymbol{j} \in \{1,2\}^N\right\}$, where

$$E^k_{11} := a_k a_k^*, \quad E^k_{22} := a_k^* a_k, \quad E^k_{12} := A_k a_k, \quad E^k_{21} := A_k a_k^*.$$

Such a matrix unit is even (odd) if and only if $\sum_{k=0}^{n-1}(i_k - j_k)$ is even (odd). We can define matrix units the same way in $\hat{\mathcal{A}}_{[0,N-1]}$, i.e. $\hat{E}^{(N)}_{\boldsymbol{ij}} := \prod_{k=0}^{N-1} \hat{E}^k_{ij}$. Relations (5.28) and (5.29) imply that

$$\hat{E}^{(N)}_{\boldsymbol{ij}} = \begin{cases} E^{(N)}_{\boldsymbol{ij}}, & \text{when } \sum_{k=0}^{n-1}(i_k - j_k) \text{ is even;} \\ TE^{(N)}_{\boldsymbol{ij}}, & \text{when } \sum_{k=0}^{n-1}(i_k - j_k) \text{ is odd;} \end{cases}$$

therefore we have

$$W\left(E^{(N)}_{\boldsymbol{ij}}\right) = \hat{E}^{(N)}_{\boldsymbol{ij}}.$$

States

If φ_0 is a state on $\mathcal{A}_+$, then

$$\varphi(b) := \varphi_0(b_+)$$

defines a linear functional on $\mathcal{A}$, which is the trivial extension of φ_0. Since

$$|\varphi(b)| = \left|\varphi_0\left(\frac{1}{2}\left(b + \alpha(b)\right)\right)\right| \leq \left\|\frac{1}{2}\left(b + \alpha(b)\right)\right\| \leq \frac{1}{2}\left(\|b\| + \|\alpha(b)\|\right) = \|b\|$$

then $\|\varphi\| = 1 = \varphi(\mathbb{1})$, and so, by lemma (A.3.2), the extension is a state, which is obviously even. On the other hand, any even state on $\mathcal{A}$ can be recovered by this method from its restriction to $\mathcal{A}_+$, so we have a one-to-one correspondence between states on $\mathcal{A}_+$ and even states of $\mathcal{A}$. Repeating the same argument for $\hat{\mathcal{A}}$ we get a one-to one correspondence between states of $\mathcal{A}_+$ and even states of $\hat{\mathcal{A}}$. We denote the even extension of φ_0 to $\hat{\mathcal{A}}$ by $\hat{\varphi}$. Note that $\hat{\varphi}$ is the restriction to $\hat{\mathcal{A}}$ of the state $\tilde{\varphi}$ on $\tilde{\mathcal{A}}$, defined by

$$\tilde{\varphi}\,(a + Tb) := \varphi(a).$$

The positivity of $\tilde{\varphi}$ follows from

$$\tilde{\varphi}\left((a + Tb)^* \,(a + Tb)\right) = \varphi(a^*a + b^*b) \geq 0.$$

We can also express the one-to-one correspondence between even states φ of $\mathcal{A}$ and even states of $\hat{\mathcal{A}}$ by the help of the operator W from the previous section:

$$\varphi = \hat{\varphi} \circ W.$$

The states $\varphi, \hat{\varphi}$ and φ_0 not only determine each other uniquely, but they are essentially the same in the sense that both φ and $\hat{\varphi}$ are trivial extensions of φ_0. Note that the same statements hold when we consider the local algebras $\mathcal{A}_N$, $\hat{\mathcal{A}}_N$ and $(\mathcal{A}_N)_+$.

Since W intertwines the shifts of $\mathcal{A}$ and $\hat{\mathcal{A}}$, and $\gamma(\mathcal{A}_+) = \mathcal{A}_+$, then the shift-invariance of any of the above three states implies the shift-invariance of the other two.

The density matrices of the restrictions $\varphi_N := \varphi \restriction_{\mathcal{A}_{[0,N-1]}}$ and $\hat{\varphi}_N := \hat{\varphi} \restriction_{\hat{\mathcal{A}}_{[0,N-1]}}$ are the same:

$$[\varphi_N]_{\boldsymbol{i},\boldsymbol{j}} = \varphi\left(E^{(N)}_{\boldsymbol{ji}}\right) = \hat{\varphi} \circ W\left(E^{(N)}_{\boldsymbol{ji}}\right) = \hat{\varphi}\left(\hat{E}^{(N)}_{\boldsymbol{ji}}\right) = [\hat{\varphi}_N]_{\boldsymbol{i},\boldsymbol{j}}\,.$$

In the shift-invariant case these restrictions determine the state completely. As a consequence, the purity of φ implies the purity of $\hat{\varphi}$ for even shift-invariant states.

5.3.1. Example. Let φ be the Fock state on $\mathcal{A}$, with corresponding symbol $Q = 0$. The restriction φ_N to $\mathcal{A}_{[0,N-1]}$ is again a quasi-free state with symbol 0_N, where 0_N is the $N \times N$ zero operator on the subspace spanned by $\{\delta_0, \ldots, \delta_{N-1}\}$. Since the symbol of φ_N is a projection then φ_N is pure, implying that $\hat{\varphi}_N$ is also pure for every N, which immediately yields that $\hat{\varphi}$

is a pure product state on $\mathcal{C}$. To identify it, it is enough to know the density of $\hat{\varphi}_1$. Since the state is even, $[\hat{\varphi}_1]_{12} = \hat{\varphi}_1(\hat{E}^1_{21}) = 0 = \hat{\varphi}_1(\hat{E}^1_{12}) = [\hat{\varphi}_1]_{21}$. For the diagonal matrix elements we have $[\hat{\varphi}_1]_{11} = \hat{\varphi}_1(\hat{E}^1_{11}) = \varphi(a_1 a_1^*) = 1$ and $[\hat{\varphi}_1]_{22} = \hat{\varphi}_1(\hat{E}^1_{22}) = \varphi(a_1^* a_1) = 0$. The density is then

$$[\hat{\varphi}_1] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = |\uparrow\rangle\langle\uparrow|,$$

with $|\uparrow\rangle := \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, and so $\hat{\varphi}$ is the pure product state $|\uparrow\rangle\langle\uparrow|^{\otimes\infty}$.

A completely similar argument shows that if φ is the anti-Fock state, then $\hat{\varphi} = |\downarrow\rangle\langle\downarrow|^{\otimes\infty}$.

5.4 Asymptotical behavior of states

5.4.1 Mean entropy

Let φ be a translation-invariant state on the infinite spin chain $\mathcal{C}$, with restrictions to the finite subchains $\varphi_\Lambda := \varphi \restriction_{\mathcal{C}_\Lambda}$. The translation-invariance of the state implies

$$S(\varphi_{\{k,\dots,k+n-1\}}) = S(\varphi_{\{0,\dots,n-1\}}) \quad \text{for every } \ k \text{ and } n,$$

while from the subadditivity of entropy we get

$$S(\varphi_{\{0,\dots,n+m-1\}}) \le S(\varphi_{\{0,\dots,n-1\}}) + S(\varphi_{\{n,\dots,n+m-1\}}).$$

With the notation $S_n := S(\varphi_{\{0,\dots,n-1\}})$, the two formulas above give

$$S_{n+m} \le S_n + S_m\,,$$

i.e. the sequence of entropies is a subadditive one, therefore the sequence $\frac{1}{n}S_n$ is convergent, and

$$\lim_{n\to+\infty} \frac{1}{n}S_n = \inf_{n\in\mathbb{N}} \frac{1}{n}S_n\,.$$

This quantity is called the **mean entropy** or **entropy rate** of the state φ, and is denoted by $s(\varphi)$. Note that

$$s(\varphi) \le S_1 \le \log\dim\mathcal{H}\,,$$

and property (2.31) of entropy implies that $s : \varphi \mapsto s(\varphi)$ is an affine function on the state space of the spin chain.

It can also be shown (see [3], Theorem 9.14.) that the sequence $S_{n+1} - S_n$ is monotonic decreasing, and

$$\lim_{n\to+\infty} (S_{n+1} - S_n) = s(\varphi)\,.$$

As a consequence, we have

$$s(\varphi) + S_n \leq S_{n+1},$$

thus the sequence S_n is monotonically increasing. The case $s(\varphi) > 0$ corresponds to an asymptotically linear growth, with slope $s(\varphi)$, while $s(\varphi) = 0$ means a sublinear growth of the entropies.

5.4.1. Example. For a product state $\varphi = \rho^{\otimes\infty}$ we have $S_n = nS_1$ by additivity, and so $s = S_1 = S(\rho)$.

For a pure product state $s(\varphi) = 0$, and it is a commonly believed, yet unproven conjecture that the mean entropy is zero for any pure translation-invariant states. Note that the converse is not true; the convex combination of states with zero mean entropy is a mixed state, and affinity of the mean entropy yields that its mean entropy is zero.

For a pure state φ the restrictions φ_n are in general mixed, due to entanglement between the subchain $\mathcal{C}_{\{0,\dots,n-1\}}$ and the rest of the chain, hence $S_n \neq 0$ is an indication of entanglement. Moreover, the results for finite systems (section 2.4.2) suggest that S_n is an appropriate measure of this entanglement.

The thermodynamical limit of the ground states of a translation-invariant family of local Hamiltonians is a translation-invariant state on the spin chain, which is pure if the ground state is unique. For all known examples the mean entropy is zero in this case. However, there are two essentially different situations; the sequence of entropies can saturate, or diverge with a sublinear speed, which is interpreted as a sign of criticality. Numerical computations in [110] suggest that the ground state of the XY-model behaves the first way for parameter values $(\delta, h) \in (-1, 1) \times (-1, 1)$, while the entropies diverge logarithmically for $\delta = 0$ (XX model) or $|h| = 1$. The logarithmic divergence was proven for the XX model in [59]; in section 7.2 we prove a logarithmic lower bound for a more general set of states.

Another importance of the mean entropy is in defining the thermal equilibrium state of a physical model, given by a sequence of local Hamiltonians $H_N \in \mathcal{C}_{\{-N,\dots,N\}}$. One way to do it is to define it as the thermodynamical limit of the local Gibbs states $\rho_{N,\beta}$; the local Gibbs state is the unique state minimizing the local free energy

$$F_{N,\beta}(\varphi) := \varphi(H_N) - \frac{1}{\beta}\, S(\varphi)\,; \qquad \varphi \in \mathcal{S}(\mathcal{C}_{\{-N,\dots,N\}})\,. \tag{5.35}$$

An alternative approach could be to define an asymptotic version of free energy, as

$$f_\beta(\varphi) := h(\varphi) - \frac{1}{\beta}\, s(\varphi) := \lim_{N\to\infty} \frac{1}{N}\left(\varphi(H_N) - \frac{1}{\beta}\, S(\varphi_N)\right)\,; \qquad \varphi \in \mathcal{S}(\mathcal{C})\,, \tag{5.36}$$

and define the equilibrium state of the system as the one minimizing this mean free energy. Even though the minimization may be hopeless to carry out in general, one can try to make an approximation of the true equilibrium state by restricting the minimization to a certain class of states. Hence classes of states for which the mean entropy can be explicitly computed become important. We have already seen that for the independent spin approximation this can be done; below we show a few other classes of states that may give a more realistic approximation (or even an exact solution) for certain models.

5.4.2. Example. Quasi-free states

Let φ be a quasi-free state on $CAR\,(l^2(\mathbb{Z}))$ with symbol Q, and $\hat\varphi$ the corresponding quasi-free state on the spin chain. The restrictions $\hat\varphi_n$ are also quasi-free states, with symbol $P_n\, Q\, P_n$, where P_n is the projection onto the subspace spanned by $\{\delta_0, \dots, \delta_{n-1}\}$, where $\{\delta_k\}$ is the standard base of $CAR\,(l^2(\mathbb{Z}))$. Then φ_n can be represented as in (5.27), hence its entropy is

$$S_n = \mathrm{Tr}\, \tilde\eta\,(Q_n)\,,$$

where $\tilde\eta(x) := -x\log x - (1-x)\log(1-x)$ is the binary entropy function.

Now the mean entropy can be explicitly evaluated, using a classical result of Szegö about the asymptotics of the eigenvalue distribution of truncated Toeplitz operators.

Let a be a real-valued bounded measurable function on $[0, 2\pi]$, and A be the self-adjoint operator on $l^2(\mathbb{N})$ with matrix elements

$$A_{k,l} := \frac{1}{2\Pi} \int_0^{2\pi} e^{-i(k-l)x} a(x) dx.$$

Let $A^{(n)}$ denote the $n \times n$ matrix with entries $A_{k,l}$, $k, l = 0, \ldots, n-1$, and $\lambda_1^{(n)}, \ldots, \lambda_n^{(n)}$ its eigenvalues. The eigenvalue distribution of $A^{(n)}$ is the probability measure $\frac{1}{n} \sum_{k=1}^{n} \delta_{\lambda_k^{(n)}}$.

Szegö's theorem [43] states that the eigenvalue distribution of $A^{(n)}$ converges weakly to the measure $\lambda \circ a^{-1}$, where λ is the normalized Lebesgue measure on the torus, and $a^{-1}(H) = \{x \in [0, 2\pi] \ : \ a(x) \in H\}$ for a subset H of $\mathbb{R}$. That is, we have

$$\frac{1}{n} \sum_{k=1}^{n} f\left(\lambda_k^{(n)}\right) = \int_{\mathbb{R}} f \ \mathrm{d} \left(\frac{1}{n} \sum_{k=1}^{n} \delta_{\lambda_k^{(n)}} \right) \xrightarrow[n \to +\infty]{} \int_0^{2\pi} (f \circ a) \ \mathrm{d}\lambda$$

for every continuous bounded function on $\mathbb{R}$.

Applying the theorem in the situation described above, with the choice

$$f(x) := \begin{cases} \tilde{\eta}(x) & \text{if } x \in [0,1]; \\ 0 & \text{otherwise} \end{cases}$$

we get

$$\lim_{n \to +\infty} \frac{1}{n} S_n = \int_{\mathcal{T}} (\tilde{\eta} \circ \hat{q}) \ \mathrm{d}\lambda.$$

In the case of a pure state the function $\hat{q}$ is a characteristic function, and so $\tilde{\eta} \circ \hat{q}$ is constant zero, giving

$$s = \lim_{n \to +\infty} \frac{1}{n} S_n = 0.$$

5.4.3. Example. Markov states

As we have seen, a positive mean entropy means an asymptotically linear growth of the sequence of entropies. An exactly linear behavior $S_n = a + bn$ is equivalent to

$$S_{n+1} - S_n = S_2 - S_1\,, \qquad\qquad n \geq 2. \tag{5.37}$$

This, on the other hand, is equivalent to φ_{n+1} saturating (SSA) with respect to the partition $\mathcal{C}_0^{\otimes(n-1)} \otimes \mathcal{C}_0 \otimes \mathcal{C}_0$ for all $n \geq 1$, which in turn is equivalent to the Markovianity of φ (see Example 5.1.5). The mean entropy in this case is

$$s(\varphi) = \lim_{n\to\infty} (S_{n+1} - S_n) = S_2 - S_1 \,, \tag{5.38}$$

hence

$$S_n = S_1 + (n-1)\, s(\varphi) \,. \tag{5.39}$$

These quantities can explicitly be computed using the algebraic structure (5.14) of Markov states, and we obtain

$$\begin{aligned} S_1 &= S(\mu_1) + \sum_k \mu_1(k) \ \left[S(\rho_k^{\mathrm{L}}) + S(\rho_k^{\mathrm{R}})\right] \qquad \text{and} \\ s(\varphi) &= S(\mu_2) - S(\mu_1) + \sum_{\{k_1,k_2\}} \mu_2(k_1,k_2)\, S(\rho_{k_1,k_2}) \,. \end{aligned}$$

Note that $S(\mu_2) - S(\mu_1)$ is the mean entropy of the classical Markov state μ.

5.4.2 Typical subspace

Let φ be a translation-invariant state on the spin chain, with finite-size restrictions φ_n. If all the φ_n are faithful then the dimension of their support equals to the total dimension of the space, i.e. $(\dim \mathcal{H})^n$. However, it may be possible to restrict the state to a considerably smaller subspace such that any measurement on the restricted state yields approximately the same outcome probabilities as with the original state. Let ε be a positive number, and

$$\varphi_n = \sum_{k=1}^{d^n} \lambda_{k,n}\, |e_{k,n}\rangle\langle e_{k,n}| \tag{5.40}$$

be an eigendecomposition of φ_n, where $d = \dim \mathcal{H}$. Let the eigenvalues be ordered decreasingly and let $t(n)$ be the smallest integer such that

$$\sum_{k=1}^{t(n)} \lambda_{k,n} > 1 - \varepsilon \,. \tag{5.41}$$

The subspace spanned by the corresponding eigenvectors is called an ε-typical subspace; note that there may be more than one ε-typical subspaces if some of the eigenvalues have multipilicity greater than one. However, they all have the same dimension $t(n)$. Note also that for any projections P such that $\operatorname{rank} P \leq t(n)$ we have $\varphi_n(P) \leq \sum_{k=1}^{t(n)} \lambda_{k,n}$, therefore the above choice of the ε-typical subspace is optimal in the sense that its dimension is the smallest possible such that the total mass of the restriction of φ_n to its orthocomplement is less than ε. Slightly generalizing this concept, we call a sequence P_n of projections a sequence of typical projections if there exists a sequence ε_n, monotonically decreasing to 0 such that

(i) $\varphi_n(P_n) > 1 - \varepsilon_n$ and

(ii) $\operatorname{rank} P_n = \min\{\operatorname{rank} P \,:\, P \in \mathcal{B}(\mathcal{H}^{\otimes n}),\quad \varphi_n(P) > 1 - \varepsilon_n\}$.

It is heuristically clear that the more concentrated the eigenvalue distributions of the φ_n's are, the slower the dimension of the typical subspaces grow. Since entropy is a measure of disorder, one can expect that there is a strong relation between the asymptotics of $S(\varphi_n)$ and $\operatorname{rank} P_n$ for a sequence of typical projections. The most concentrated eigenvalue distribution we obtain from a pure product state $|\psi\rangle\langle\psi|^{\otimes\infty}$; this case φ_n is supported on the one-dimensional subspace spanned by $\psi^{\otimes n}$, and therefore $\dim P_n = 1$ for an optimal sequence of typical projections, and $S(\varphi_n) \equiv 0$. The other extreme is the tracial state of the spin chain, for which $\operatorname{rank} P_n$ grows asymptotically as d^n, and $S(\varphi_n) = n \log d$.

5.4.4. Lemma. Let P_n be a sequence of typical projections for φ. Then

$$s(\varphi) \leq \liminf \frac{1}{n} \operatorname{rank} P_n \,.$$

Proof: Let $p_n := \sum_{k=1}^{\operatorname{rank} P_n} \lambda_{k,n}$ and $q_n := 1 - p_n$. Convexity of the function $-x \log x$ implies

$$-\sum_{k=\operatorname{rank} P_n+1}^{d^n} \lambda_{k,n} \log \lambda_{k,n} \leq -q_n \log \frac{q_n}{\operatorname{rank} P_n} \leq -\varepsilon_n \log \varepsilon_n + n\,\varepsilon_n \,,$$

while

$$-\sum_{k=1}^{\operatorname{rank} P_n} \lambda_{k,n} \log \lambda_{k,n} \leq p_n \log \frac{p_n}{\operatorname{rank} P_n} \leq \frac{1}{e} \log e + \log \operatorname{rank} P_n \,.$$

Thus

$$S(\varphi_n) \leq -\varepsilon_n \log \varepsilon_n + n\varepsilon_n + \frac{1}{e} \log e + \log \operatorname{rank} P_n \,,$$

and the statement follows. $\square$

Hence a linear growth of the entropies implies an exponential growth of the dimension of any sequence of typical subspaces, while a slower-than-exponential growth of the dimension of the typical subspaces implies a slower-than-linear growth of the entropy sequence. For ergodic states a stronger relation can be established, which is called the **typical subspace theorem**:

5.4.5. Theorem. Let φ be an ergodic state on the spin chain. Then for any sequence of typical projections

$$s(\varphi) = \lim_{n\to\infty} \frac{1}{n} \log \operatorname{rank} P_n \,. \tag{5.42}$$

A linear growth of the entropy sequence is therefore equivalent to an exponential growth of the size of the typical subspaces, and so the typical subspace theorem describes completely the relation between the asymptotics of the entropy sequence and the size of the typical subspaces. However, such an explicit relation is not known when $s(\varphi) = 0$.

References

An introduction to the algebraic formalism of the spin chain and fermionic systems can be found in [3] and [32]. More detailed treatments of the topic are presented in [21, 22] and [30]. The construction of section 5.3 can also be found in [30]; it originally appeared in [6]. Properties of finitely correlated states are extensively studied in the papers [33, 34, 47]. For a proof of the typical subspace theorem, see [83] or [46].

Chapter 6

Quantum data compression

Suppose two parties (Alice and Bob, as usual) play the following game: Alice thinks of an element probability finite set $\mathcal{X}$ with probability p_x, and Bob has to find out x by asking as few yes/no questions as possible. The sequence of questions leading to x can be viewed as a binary code of x, and therefore this problem is equivalent to finding a binary encoding of $\mathcal{X}$ with the smallest average (with respect to $\boldsymbol{p} = \{p_x\}$) codelength. A classical result of information theory [101] is that the average codelength is bounded from below by the entropy $H(\boldsymbol{p})$ of the distribution $\{p_x\}$, and, apart from very special cases, the inequality is strict. However, if Alice and Bob play several games simultaneously (i.e. Bob has to find out sequences $(x_1, \ldots, x_n)$ instead of individual elements of $\mathcal{X}$) then the lower bound can be reached asymptotically, that is, if L_n denotes the optimal average codelength for sequences of length n, then $\lim_{n\to\infty} \frac{1}{n} L_n = H(\boldsymbol{p})$. Note that this coding is of variable length, i.e. the codelengths for different sequences may be different; the heuristically good strategy is to attach codes with a shorter length to sequences with higher probability. Note that the number of sequences grows as $\exp(n \log |\mathcal{X}|)$, while the average length of the codewords grows as $\exp(nH(\boldsymbol{p}))$, and $H(\boldsymbol{p}) < \log |\mathcal{X}|$ unless $\boldsymbol{p}$ is the uniform distribution on $\mathcal{X}$, therefore we can say that the information carried by the original sequences is compressed by the coding procedure; the rate of compression is $H(\boldsymbol{p})/\log |\mathcal{X}|$.

In practical applications long messages are split up into blocks of length n for a sufficiently large n, these blocks are encoded as described above, and the codeword for the whole message is simply the concatenation of the codewords of the consecutive blocks. A serious disadvantage of this method is that when errors can occur then an error in one single bit of the code may make the code of the whole message undecodable, since it may not be possible anymore to determine where the codeword of the next block is starting. A solution of this problem can be to encode the blocks into codewords of the same length. However, it can easily be seen that a perfectly decodable code with fixed codelengths exists only if sequences of all blocks are encoded by binary sequences of length $n \log |\mathcal{X}|$, therefore the compression feature of the previous protocol is lost.

It was Shannon's idea to combine the advantages of the two coding scheme by relaxing the requirement of perfect decodability, and allowing a small probability of error in the decoding. His classical source coding theorem states that sufficiently long messages can be encoded using asymptotically $nH(\boldsymbol{p})$ bits (fixed codelength), while the probability of a wrong decoding goes to zero. On the other hand, the converse of the source coding theorem states that the messages cannot be compressed further, i.e. there is no coding-decoding scheme with asymptotically vanishing error probability such that messages of length n are encoded by codewords asymptotically shorter than $nH(\boldsymbol{p})$.

The source coding theorem gives a plausible interpretation of entropy as the best attainable compressibility rate of long messages when the letters are drawn independently according to some distribution $\boldsymbol{p}$. The smaller the entropy is, the more the message can be compressed, which can be interpreted as a small information content of the message.

The coding theorem for quantum sources is analogous to the classical case in that the entropy (or rather the mean entropy) appears as the best attainable compressibility rate, and therefore gives a justification of interpreting entropy as the information content of a quantum state. In the following sections we formalize precisely the notion of a quantum information source, and prove the source coding theorem for a class of ergodic sources.

6.1 Quantum information sources

The notion of a quantum information source can easily be presented through the following simple example. Suppose Alice is reading a book in a language with a finite alphabet $\mathcal{X}$. Each letter $x \in \mathcal{X}$ has a certain frequency in the given language, that gives a probability distribution μ_1 on $\mathcal{X}$. However, this individual frequency of the letters doesn't describe the statistics of the language completely. For instance, the combination 'dne' appears with a significantly smaller frequency in English language than 'end'. To build a statistical model of texts in the given language, we should specify a probability measure on $\mathcal{X}^n$ for all $n \in \mathbb{N}$, given by the frequency of appearance of different sequences of length n. Even though real texts are always of finite length, and it doesn't make too much sense to measure the frequency of the complete text of Hamlet in English texts, for mathematical convenience we describe the language statistics by a probability measure μ on the set of infinite sequences $\mathcal{X}^\infty$.

Suppose Alice wants to store the content of the book in the state of a quantum system. To this end, she is provided with several copies of a quantum system with Hilbert space $\mathcal{H}$, and a set of states $\{\rho_x : x \in \mathcal{X}\}$ of the given system. After reading a new letter, she takes a new copy of the system, and prepares it in the state ρ_x. When she has finished the book that can be described by a sequence $\underline{x} := (x_1, \dots, x_N) \in \mathcal{X}^N$, she has a big system consisting of N copies of the original system, in the state $\rho_{\underline{x}} := \rho_{x_1} \otimes \dots \otimes \rho_{x_N}$.

After a book has been finished, Bob gets the resulting physical system, and the task to represent its quantum state in a more efficient way. He knows the statistics of the language, and the set of states $\{\rho_x\}$ used to encode the book. However, he doesn't know which book was encoded, and therefore from his point of view the state of the system is

$$\rho_N = \sum_{\underline{x} \in \mathcal{X}^N} \mu(\underline{x}) \rho_{x_1} \otimes \dots \otimes \rho_{x_N}. \tag{6.1}$$

The strategy he follows goes like this: He splits up the big system into blocks of n systems, and compresses the state of each such block independently. The encoding process is described by a completely positive map $\mathcal{E}_n : \mathcal{B}(\mathcal{H})^{\otimes n} \to \mathcal{B}(\mathcal{K}_n)$, and is basically the composition of a projection onto a subspace and a unitary operation. The decoding is a completely positive map $\mathcal{D}_n : \mathcal{B}(\mathcal{K}_n) \to \mathcal{B}(\mathcal{H})^{\otimes n}$, which in turn is the composition of a unitary and the embedding of a subspace into a bigger Hilbert space. The efficiency of the restoration of the original state is measured by the average fidelity

$$F_n := \sum_{\underline{x} \in \mathcal{X}^N} \mu(\underline{x})\, F_U(\rho_{\underline{x}}, \mathcal{D}_n \circ \mathcal{E}_n\,(\rho_{\underline{x}})), \tag{6.2}$$

where F_U is the fidelity, introduced in section (2.3).

As a mathematical idealization we assume that the books can be considered to be infinitely long. The sequence of states defined in (6.1) are easily seen to form a compatible family, and therefore they extend to a state of the half-infinite spin chain, with one-site algebra $\mathcal{B}(\mathcal{H})$. It is natural to assume that the language statistics μ is translation-invariant, which immediately implies the translation-invariance of ρ. Moreover, the following statement holds:

6.1.1. Lemma. If μ is ergodic, then so is ρ.

Proof: We use criterion (5.9) for ergodicity, i.e. we have to show

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \rho\left(A\gamma^k(B)\right) = \rho(A)\rho(B) \tag{6.3}$$

for all elements A, B of the spin chain. However, it is enough to check the criterion for local elements, moreover, bilinearity implies that both A and B can be assumed to be product operators, i.e. $A = A_1 \otimes \ldots \otimes A_m$ and $B = B_1 \otimes \ldots \otimes B_m$. Because of the $\frac{1}{n}$ factor, we can neglect the first $m+1$ terms of the sum, since in the limit they give 0, and so we can assume $A\gamma^k(B) = A_1 \otimes \ldots \otimes A_m \otimes I \otimes \ldots \otimes I \otimes B_1 \otimes \ldots \otimes B_m$. Substituting these into the ergodicity criterion, we obtain that the left-hand side of (6.3), after neglecting the first m terms, takes the form

$$\sum_{x_1,\dots,x_m} \rho_{x_1}(A_1)\cdot\ldots\cdot\rho_{x_m}(A_m) \sum_{y_1,\dots,y_m} \rho_{y_1}(B_1)\cdot\ldots\cdot\rho_{y_m}(B_m)\cdot$$
$$\cdot\, c_n(x_1,\ldots x_m;\, y_1,\ldots,y_m)\,, \qquad \text{where} \quad c_n \quad \text{is}$$

$$\frac{1}{n} \sum_{k=m+1}^{n-1} \sum_{\{x_{m+1},\dots,x_{k-m}\}} \mu(x_1,\ldots,x_m;\, x_{m+1},\ldots,x_k;\, y_1,\ldots y_m)\,. \tag{6.4}$$

Ergodicity of μ implies that (6.4) converges to $\mu(x_1,\ldots,x_m)\mu(y_1,\ldots,y_m)$, which finishes the proof. $\square$

In case the language statistics is assumed to be Markovian, the state ρ is a finitely correlated state, given in (5.11).

The **compression rate** is defined as

$$R := \limsup_{n\to\infty} \frac{1}{n} \log\dim\mathcal{K}_n\,; \tag{6.5}$$

the smaller the rate, the better the compression is. A compression scheme is **reliable**, if the sequence of fidelities goes to 0 as n goes to infinity; a compression rate R is **attainable** if there exists a reliable compression scheme with rate R. The data compression theorem states that the best attainable compression rate equals to the mean entropy of the source:

6.1.2. Theorem.

$$\inf\{R\,:\,R \text{ is attainable }\} = s(\rho). \tag{6.6}$$

This statement is also called the source coding theorem, and a proof of a generalization of it is given in the next section.

6.2 The source coding theorem

The aim of this section is to prove the following

6.2.1. Theorem. Let ρ be an ergodic state on the infinite quantum spin chain with one-site algebra $\mathcal{B}(\mathcal{H})$, and let s be its entropy rate. Then

(i) There exists a sequence of encodings $\mathcal{E}_n\,:\,\mathcal{B}(\mathcal{H})^{\otimes n} \to \mathcal{B}(\mathcal{K}_n)$ and decodings $\mathcal{D}_n\,:\,\mathcal{B}(\mathcal{K}_n) \to \mathcal{B}(\mathcal{H})^{\otimes n}$, such that

$$\limsup_{n\to\infty} \frac{1}{n} \log\dim\, \mathcal{K}_n = s\,, \tag{6.7}$$

and

$$\lim_{n\to\infty} \sum_i \lambda_i\, F_U\left(\rho_i^{(n)}, \mathcal{D}_n \circ \mathcal{E}_n\,(\rho_i^{(n)})\right) = 1\,, \tag{6.8}$$

where $\rho_n = \sum_i \lambda_i^{(n)}\,\rho_i^{(n)}$ is an arbitrary convex decomposition of ρ_n.

(ii) For any sequence of encodings $\mathcal{E}_n\,:\,\mathcal{B}(\mathcal{H})^{\otimes n} \to \mathcal{B}(\mathcal{K}_n)$ and decodings $\mathcal{D}_n\,:\,\mathcal{B}(\mathcal{K}_n) \to \mathcal{B}(\mathcal{H})^{\otimes n}$, such that

$$\limsup_{n\to\infty} \frac{1}{n} \log\dim\, \mathcal{K}_n < s \tag{6.9}$$

and any convex decomposition $\rho_n = \sum_i \lambda_i^{(n)}\,\rho_i^{(n)}$ we have

$$\lim_{n\to\infty} \sum_i \lambda_i\, F_U\left(\rho_i^{(n)}, \mathcal{D}_n \circ \mathcal{E}_n\,(\rho_i^{(n)})\right) = 0\,. \tag{6.10}$$

The data compression theorem of the previous section obviously follows from the theorem above. Note, however, that this theorem is more general in at least two senses. First, it states that the existence of a good compression scheme doesn't depend on the ensemble to which the state ρ corresponds to as the ensemble average; it only depends on ρ. Moreover, the decompositions of ρ_n and ρ_m for $n \neq m$ can be completely independent. As we will see from the proof, the construction of the coding scheme depends also only on ρ, therefore it is the same for all ensembles with the same average state.

Second, the above theorem holds for all ergodic states, while the model for a quantum information source, as described in the previous section, yields a sequence ρ_n of separable states. However, the restrictions of an ergodic state need not in general be separable.

Proof: To prove part (i), we fix positive numbers ε and δ. By the typical subspace theorem (Theorem 5.4.5) there exists $N(\varepsilon, \delta)$ such that for all $n > N(\varepsilon, \delta)$ we can find a projection P_n satisfying

$$\operatorname{Tr} P_n < 2^{n(s+\delta)} \qquad \text{and} \qquad \operatorname{Tr} \rho_n \, P_n > 1 - \varepsilon \,. \tag{6.11}$$

Now let $\mathcal{K}_n :=$ ran P_n, and ψ be a fixed unit vector in $\mathcal{K}_n$. We define the encoding as

$$\mathcal{E}_n(A) := P_n \, A \, P_n + |\psi\rangle\langle\psi| \operatorname{Tr} A \, (\mathbb{1} - P_n) \ , \tag{6.12}$$

where the result is meant as an operator on $\mathcal{K}_n$. The decoding is simply the identical embedding of $\mathcal{E}_n(A)$ into $\mathcal{B}(\mathcal{H})^{\otimes n}$. The encoding map is completely positive and trace-preserving; a set of Kraus operators can be given as

$$V_0 := P_n; \quad V_i := |\psi\rangle\langle e_i| \,, \quad i = 1, \dots, \text{rank} \ (\mathbb{1} - P_n) \,; \tag{6.13}$$

where $\{e_i\}$ is an orthonormal base in the range of $\mathbb{1} - P_n$. Now we can estimate the average fidelity from below by the entanglement fidelity (2.18), and using (2.19) we obtain

$$\begin{aligned}
\sum_i \lambda_i \, F_U \left(\rho_i^{(n)}, \mathcal{D}_n \circ \mathcal{E}_n \,(\rho_i^{(n)})\right) \;&\geq\; F\,(\rho_n, \mathcal{D}_n \circ \mathcal{E}_n) \\
&=\; \sum_i \,|\mathrm{Tr}\,\rho_n\, V_i|^2 \\
&=\; (\mathrm{Tr}\,\rho_n\, P_n)^2 + \sum_{i\geq 1} \,|\langle \psi, \rho_n\, e_i\rangle|^2 \\
&\geq\; (\mathrm{Tr}\,\rho_n\, P_n)^2 \\
&>\; (1-\varepsilon)^2 .
\end{aligned}$$

Now letting $\varepsilon, \delta \to 0$ we get the desired result.

Suppose now that there exists a coding scheme with rate strictly smaller than s; then there exists a positive number η such that for sufficiently large n

$$\dim \mathcal{K}_n < 2^{n(s-\eta)} \,. \tag{6.14}$$

Since the dimension of the support of the densities cannot increase during the decoding process, we obtain the existence of a projection Q_n in $\mathcal{H}^{\otimes n}$, such that $\mathrm{rank}\, Q_n < 2^{n(s-\eta)}$, and all the densities $\hat{\rho}_i^{(n)} := \mathcal{D}_n \circ \mathcal{E}_n(\rho_i^{(n)})$ are supported in its range, therefore $Q_n\, \hat{\rho}_i^{(n)} = \hat{\rho}_i^{(n)}$ for all i.

Now we can use the fact $F_U \leq F_P$ (see section 2.3) and the Schwarz inequality in the form

$$\Big|\sum_i p_i \,\mathrm{Tr}\, X_i Y_i\Big| \leq \Big[\sum_i p_i \,\mathrm{Tr}\, X_i^* X_i\Big]^{1/2} \Big[\sum_i p_i \,\mathrm{Tr}\, Y_i^{\,*} Y_i\Big]^{1/2} \tag{6.15}$$

to give an upper bound on the average fidelity:

$$\begin{aligned}
\sum_i \lambda_i \, F_U \left(\rho_i^{(n)}, \hat{\rho}_i^{(n)}\right) \;&\leq\; \sum_i \lambda_i \,\mathrm{Tr}\,\rho_i^{(n)}\, \hat{\rho}_i^{(n)} \\
&=\; \sum_i \lambda_i \,\mathrm{Tr}\,\rho_i^{(n)}\, Q_n\, \hat{\rho}_i^{(n)} \\
&\leq\; \Big[\sum_i \lambda_i \,\mathrm{Tr}\,\rho_i^{(n)}\, Q_n\Big]^{1/2} \Big[\sum_i \lambda_i \,\mathrm{Tr}\,\hat{\rho}_i^{(n)}\Big]^{1/2} \\
&=\; (\mathrm{Tr}\,\rho_n\, Q_n)^{1/2} \;.
\end{aligned}$$

The typical subspace theorem states that $\mathrm{Tr}\,\rho_n\, Q_n$ tends to zero, hence the statement is proven. □

Summary

In this chapter we have proven that for ergodic states the mean entropy of the state is equal to the amount to which the state can be compressed in an asymptotically lossless way. The proof is based on the typical subspace theorem (Theorem 5.4.5) and the idea of block-coding, that originates from Shannon. In section 6.1 we constructed a plausible information-theoretic model that results in an ergodic state of the spin chain, hence the general compression theorem is applicable to it.

The first proof of the quantum source coding theorem appeared in [99], where it was proven for the case when ρ is a product state, and the decomposition of ρ_n is a pure one. The first proof for non-product sources appeared in [93]; the presentation given here follows the lines of that work. However, a few things have been changed:

- In [93], a weaker version of the typical subspace theorem was used, which required the source to be completely ergodic instead of the weaker criterion of ergodicity, and it enabled only to prove that the fidelity is bounded away from 1 in part (ii), instead of the stronger statement that it converges to 0.

- Part (i) of the source coding theorem was proven only for pure decompositions, mainly because of the difficulty to give a lower bound for the fidelity F_P.

The stronger version of the typical subspace theorem was proven in [15], while the idea of lower bounding the average fidelity by the entanglement fidelity comes from [81]. The general version of the source coding theorem was also considered in [17].

Chapter 7

Entropy growth of pure shift-invariant quasi-free states

In section 5.4.2 we have seen that the mean entropy of a shift-invariant pure quasi-free state is zero, yielding a sublinear growth of the entropies of the restricted states. In this chapter we examine the entropy growth of two subclasses of quasi-free states, namely when the defining subset of the torus is (a) a disjoint union of finitely many (=M) intervals, or (b) a Cantor type set. First we develop some general methods to analyze the growth for a general subset $K \subset \mathcal{T}$, and then we apply these methods to show that the entropy asymptotics is (a) $c_0 + c_1\, M\, \log N \le S_N \le c_2\, M\, \log^2 N$ with universal constants c_1 and c_2 and (b) $S_N \sim N^\alpha$ for some $\alpha \in (0, 1)$, determined by the structure of the Cantor type set.

7.1 General results

7.1.1 Bounds on the entropy function

The symbol of the state φ with defining function $\hat{q} = \chi_K$ is the operator Q with entries

$$Q_{k,l} = \frac{1}{2\pi} \int_0^{2\pi} e^{-i(k-l)x}\, \chi_K(x)\, \mathrm{d}x = \frac{1}{2\pi} \int_K e^{-i(k-l)x}\, \mathrm{d}x, \qquad k, l \in \mathbb{Z},$$

while the symbols of the restricted states φ_N are the $N \times N$ matrices Q_N with entries

$$(Q_N)_{k,l} = Q_{k,l}, \qquad k, l = 0, \ldots, N-1.$$

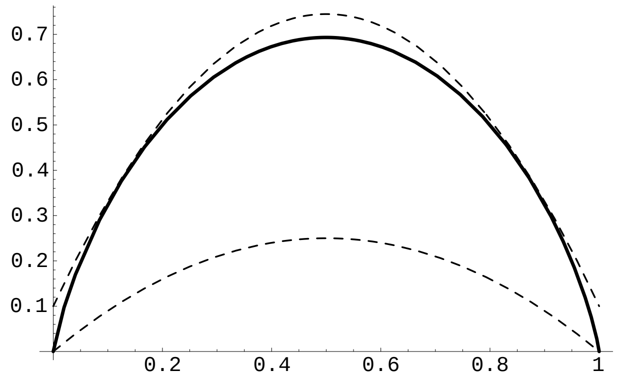

Figure 7.1: A quadratic upper and lower bound for the entropy function $\tilde{\eta}$.

For the entropies we have

$$S_N = \mathrm{Tr}\tilde{\eta}\,(Q_N)\ , \tag{7.1}$$

where $\tilde{\eta}(x) = -x\log x - (1-x)\log(1-x)$ is the binary entropy function. Note that the symbol Q of the state φ is a projection, and so its spectrum is the set $\{0,1\}$ (except for the trivial cases when $Q=0$ or $Q=1$). However, the restricted symbols are not projections any more, and so the analytic calculation of expression (7.1) is a nontrivial task. Instead of it, we give bounds on S_N by bounding the function $\tilde{\eta}$ with simpler functions. Such a bound is

$$x(1-x) \le \tilde{\eta}(x) \le \epsilon - c\ (\log\epsilon)\, x(1-x), \qquad 0 \le x \le 1; \tag{7.2}$$

see Figure 7.1. The upper bound for $\tilde{\eta}$ holds for c a constant independent of $0 < \epsilon$. Therefore,

$$\mathrm{Tr}\, Q_N(\mathbb{1} - Q_N) \le S_N \le \epsilon N - c\ (\log\epsilon)\, \mathrm{Tr}\, Q_N(\mathbb{1} - Q_N).$$

By choosing a function $\epsilon(N)$ for which $\epsilon \to 0$ as $N \to \infty$, we obtain bounds for the entropy S_N in terms of $\mathrm{Tr}\, Q_N(\mathbb{1} - Q_N)$. E.g., putting $\epsilon(N) := \frac{1}{N}$,

$$\mathrm{Tr}\, Q_N(\mathbb{1} - Q_N) \le S_N \le 1 + c\ (\log N)\, \mathrm{Tr}\, Q_N(\mathbb{1} - Q_N). \tag{7.3}$$

The main value of approximation (7.3) is that $\mathrm{Tr}\, Q_N(\mathbb{1} - Q_N)$ is easily expressible by the Fourier coefficients of $\hat{q}$:

$$
\begin{aligned}
\operatorname{Tr} Q_N(\mathbb{1}-Q_N) &= Nq(0) - \sum_{n,m=0}^{N-1} |q(n-m)|^2 \\
&= Nq(0) - N \sum_{n=-N+1}^{N-1} \Big(1-\frac{|n|}{N}\Big) |q(n)|^2 \\
&= Nq(0) - \frac{N}{4\pi^2} \int d\theta_1 \int d\theta_2\, \chi_K(\theta_1)\chi_K(\theta_2) \sum_{n=-N+1}^{N-1} \Big(1-\frac{|n|}{N}\Big) e^{in(\theta_1-\theta_2)}.
\end{aligned}
$$

By introducing the Dirichlet kernel

$$
k_N(\varphi) := \sum_{n=-N+1}^{N-1} \Big(1-\frac{|n|}{N}\Big) e^{in\varphi} = 1 + 2\sum_{n=1}^{N-1} \frac{N-n}{N} \cos n\varphi \;=\; \frac{1}{N}\frac{\sin^2(N\varphi/2)}{\sin^2(\varphi/2)},
$$

we get

$$
\begin{aligned}
&\operatorname{Tr} Q_N(\mathbb{1}-Q_N) = \\
&\quad = N\Big[\frac{1}{2\pi}\int d\theta\, \chi_K(\theta) - \frac{1}{4\pi^2}\int d\theta \int d\varphi\, \chi_K(\theta)\chi_K(\theta-\varphi) k_N(\varphi)\Big] \\
&\quad = \frac{N}{4\pi^2}\int d\theta \int d\varphi\, \chi_K(\theta)\,[1-\chi_K(\theta-\varphi)]\, k_N(\varphi) \\
&\quad = \frac{N}{4\pi^2}\int d\varphi\, k_N(\varphi) \int d\theta\, \chi_K(\theta)\,[1-\chi_K(\theta-\varphi)] \\
&\quad = \frac{N}{4\pi^2}\int d\varphi\, k_N(\varphi)\, |K \setminus (K+\varphi)|, \qquad (7.4)
\end{aligned}
$$

where $|H|$ denotes the Lebesgue measure of the set $H \subset \mathcal{T}$. Note that both $S(Q_N)$ and $\operatorname{Tr} Q_N(\mathbb{1}-Q_N)$ are invariant for the replacement of Q_N by $\mathbb{1}-Q_N$. As a consequence, Equation (7.4) can be written in the form

$$
\operatorname{Tr} Q_N(\mathbb{1}-Q_N) = \frac{N}{4\pi^2}\int d\varphi\, k_N(\varphi)\, |K^c \setminus (K^c+\varphi)|, \qquad (7.5)
$$

where $K^c = \mathcal{T} \setminus K$.

7.1.2 Subadditivity

Suppose our set K is the union of two disjoint sets K_1 and K_2, with corresponding symbols Q, Q_1 and Q_2. Note that $Q = Q_1 + Q_2$. We prove the subadditivity property

$$\operatorname{Tr}\tilde{\eta}(Q_N) \leq \operatorname{Tr}\tilde{\eta}((Q_1)_N) + \operatorname{Tr}\tilde{\eta}((Q_2)_N)\,. \tag{7.6}$$

This property has the consequence that if we can decompose our set K into a disjoint union of finitely many sets $K_1, \ldots, K_m$, and manage to derive upper bounds for the corresponding entropies $S_N^{(1)}, \ldots, S_N^{(m)}$, then ($m$ times) the highest among the bounds will also give an upper bound for the entropy S_N.

To simplify notation, we introduce $R := Q_N$, $R_1 := (Q_1)_N$ and $R_2 := (Q_2)_N$; then $Q = Q_1 + Q_2$ implies $R = R_1 + R_2$.

We assume $R_1 > 0$, $R_2 > 0$ and $R_1 + R_2 < \mathbb{1}$; otherwise, we can introduce operators $\tilde{R}_1 := (1-\epsilon)R_1 + \frac{\epsilon}{2}\mathbb{1}$ and $\tilde{R}_2 := (1-\epsilon)R_2 + \frac{\epsilon}{2}\mathbb{1}$, prove the subadditivity for these two operators and then take the limit $\epsilon \to 0$.

Using the operator identity

$$\frac{\mathrm{d}}{\mathrm{d}t}\operatorname{Tr} f(A+tB) = \operatorname{Tr} B f'(A+tB),$$

we get

$$\begin{aligned}
\operatorname{Tr}\tilde{\eta}(R_1+R_2) - \operatorname{Tr}\tilde{\eta}(R_1) &= \operatorname{Tr}\tilde{\eta}(R_1+tR_2)\Big|_{t=0}^{1} \\
&= \int_0^1 \mathrm{d}t\, \frac{\mathrm{d}}{\mathrm{d}t}\operatorname{Tr}\tilde{\eta}(R_1+tR_2) \\
&= \int_0^1 \mathrm{d}t\, \operatorname{Tr} R_2\big[\log(\mathbb{1}-R_1-tR_2) - \log(R_1+tR_2)\big]\,.
\end{aligned} \tag{7.7}$$

Since the logarithm is operator increasing, we have

$$\log(\mathbb{1}-R_1-tR_2) - \log(R_1+tR_2) \leq \log(\mathbb{1}-tR_2) - \log tR_2\,,$$

hence (7.7) can be estimated from above by

$$\int_0^1 \mathrm{d}t\, \operatorname{Tr} R_2\left[\log(\mathbb{1}-tR_2) - \log tR_2\right] = \int_0^1 \mathrm{d}t\, \frac{\mathrm{d}}{\mathrm{d}t}\operatorname{Tr}\tilde{\eta}\,(tR_2) = \operatorname{Tr}\tilde{\eta}(R_2)\,,$$

and this finishes the proof.

7.2 Finitely many intervals

Now we examine the case when the subset K is the union of disjoint intervals $I_1, \dots, I_M$. In section 7.2 we derive a logarithmic lower bound for the entropies S_N, while in section 7.2 an upper bound of the magnitude $(\log N)^2$ is given. However, numerical results suggest a logarithmic growth:

$$S_N = a \log N + b$$

for some constants a and b.

Lower bound

By Equation (7.3) we have to bound $\operatorname{Tr} Q_N(\mathbb{1} - Q_N)$ from below. Let $\delta > 0$ be a fixed number which is smaller than any of the intervals and the holes between two such intervals. Therefore, $|K \setminus (K + \varphi)| \geq M\varphi$ for $0 \leq \varphi \leq \delta$. Equation (7.4) becomes,

$$\begin{aligned}
S_N &\geq \operatorname{Tr} Q_N(\mathbb{1} - Q_N) \\
&= \frac{N}{4\pi^2} \int \mathrm{d}\varphi\, k_N(\varphi)\, |K \setminus (K + \varphi)| \\
&\geq \frac{NM}{2\pi^2} \int_0^\delta \mathrm{d}\varphi\, k_N(\varphi)\, \varphi \\
&= \frac{NM}{2\pi^2} \int_0^\delta \mathrm{d}\varphi\, \varphi \left[1 + 2 \sum_{n=1}^{N-1} \frac{N-n}{N} \cos n\varphi \right] \\
&= \frac{NM}{\pi^2} \left[\frac{\delta^2}{4} + \delta \sum_{n=1}^{N-1} \frac{\sin n\delta}{n} - 2 \sum_{n=1}^{N-1} \frac{\sin^2(n\delta/2)}{n^2} \right] \\
&\quad + \frac{M}{\pi^2} \left[-\delta \sum_{n=1}^{N-1} \sin n\delta - \sum_{n=1}^{N-1} \frac{\cos n\delta}{n} + \sum_{n=1}^{N-1} \frac{1}{n} \right].
\end{aligned}$$

Using the identities[1]

$$\sum_{n=1}^{\infty} \frac{\sin n\delta}{n} = \frac{\pi - \delta}{2} \qquad \text{and} \qquad \sum_{n=1}^{\infty} \frac{\sin^2(n\delta/2)}{n^2} = \frac{\delta(2\pi - \delta)}{8},$$

[1] The values of these and the following trigonometric sums can be found e.g. in [62].

we obtain

$$S_N \geq \frac{NM}{\pi^2}\left[-\delta\sum_{n=N}^{\infty}\frac{\sin n\delta}{n} + 2\sum_{n=N}^{\infty}\frac{\sin^2(n\delta/2)}{n^2}\right]$$
$$+\frac{M}{\pi^2}\left[-\delta\sum_{n=1}^{N-1}\sin n\delta - \sum_{n=1}^{N-1}\frac{\cos n\delta}{n} + \sum_{n=1}^{N-1}\frac{1}{n}\right]. \tag{7.8}$$

Next, we estimate the different terms in (7.8). The first term on the first line,

$$\left|\sum_{n=N}^{\infty}\frac{\sin n\delta}{n}\right| = \left|\sum_{n=N}^{\infty}\frac{\cos(n+\frac{1}{2})\delta - \cos(n-\frac{1}{2})\delta}{2n\sin(\delta/2)}\right| \leq \frac{1}{N}\frac{1}{|\sin(\delta/2)|}\,.$$

The second term on the first line,

$$\left|\sum_{n=N}^{\infty}\frac{\sin^2(n\delta/2)}{n^2}\right| \leq \sum_{n=N}^{\infty}\frac{1}{n^2} \leq \int_{N-1}^{\infty}\,\mathrm{d}x\,\frac{1}{x^2} = \frac{1}{N-1}.$$

The first term on the second line,

$$\left|\sum_{n=1}^{N-1}\sin n\delta\right| \leq \frac{1}{|\sin(\delta/2)|}.$$

The second term on the second line, for any $\epsilon > 0$ and N sufficiently large,

$$\left|\sum_{n=1}^{N-1}\frac{\cos n\delta}{n}\right| \leq -\log[2\sin\delta] + \epsilon.$$

Finally, the last term on the last line,

$$\sum_{n=1}^{N-1}\frac{1}{n} \geq \int_{1}^{N}\,\mathrm{d}x\,\frac{1}{x} = \log N.$$

Putting everything together in (7.8), we find that there exists a constant $c_1(\delta) > 0$ that only depends on δ such that

$$S_N \geq \frac{M}{\pi^2}\log N + M\,c_1(\delta)\,. \tag{7.9}$$

Note that the coefficient of the $\log N$ term only depends on the number of intervals.

Upper bound

Due to subadditivity, it is enough to prove an upper bound for a set K consisting of a single interval. We assume that the length of this interval is not bigger than π; otherwise, we can work with K^c. By Equation (7.3) we have to bound $\operatorname{Tr} Q_N(\mathbb{1} - Q_N)$. In this case,

$$|K \setminus (K+\varphi)| = \begin{cases} \varphi & \text{for } |\varphi| \leq |K|, \\ |K| & \text{for } |K| \leq |\varphi| \leq \pi. \end{cases}$$

By Equation (7.4),

$$\begin{aligned} \operatorname{Tr} Q_N(\mathbb{1} - Q_N) &= \frac{N}{2\pi^2} \int_0^\pi \mathrm{d}\varphi \left[1 + 2\sum_{n=1}^{N-1} \frac{N-n}{N} \cos n\varphi\right] |K \setminus (K+\varphi)| \\ &= \frac{N}{4\pi^2}\left[|K|(2\pi - |K|) - 8\sum_{n=1}^{N-1} \frac{\sin^2(n|K|/2)}{n^2}\right] \\ &\quad + \frac{2}{\pi^2}\sum_{n=1}^{N-1} \frac{\sin^2(n|K|/2)}{n}. \end{aligned}$$

Using the identity

$$\sum_{n=1}^{\infty} \frac{\sin^2(n|K|/2)}{n^2} = \frac{|K|(2\pi - |K|)}{8},$$

we obtain

$$\operatorname{Tr} Q_N(\mathbb{1} - Q_N) = \frac{2N}{\pi^2}\sum_{n=N}^{\infty} \frac{\sin^2(n|K|/2)}{n^2} + \frac{2}{\pi^2}\sum_{n=1}^{N-1} \frac{\sin^2(n|K|/2)}{n}. \tag{7.10}$$

Bounds for the first term and the second term are

$$\sum_{n=N}^{\infty} \frac{\sin^2(n|K|/2)}{n^2} \leq \sum_{n=N}^{\infty} \frac{1}{n^2} \leq \int_N^\infty \mathrm{d}x \, \frac{1}{x^2} = \frac{1}{N}.$$

and

$$\sum_{n=1}^{N-1} \frac{\sin^2(n|K|/2)}{n} \leq \sum_{n=1}^{N-1} \frac{1}{n} \leq 1 + \int_1^{N-1} \mathrm{d}x \, \frac{1}{x} = 1 + \log(N-1).$$

Putting everything together in (7.10), we get

$$\operatorname{Tr} Q_N(\mathbb{1} - Q_N) \leq \frac{4}{\pi^2} + \frac{2}{\pi^2} \log N\,,$$

hence, by Equation (7.3),

$$S_N \leq 1 + \frac{4c}{\pi^2} \log N + \frac{2c}{\pi^2} \log^2 N\,.$$

Finally, for a set consisting of M intervals we obtain

$$S_N \leq M\, c_2 \log^2 N\,, \tag{7.11}$$

where c_2 is a universal constant, independent of the set K.

Numerical example

Analytically, we determined the asymptotics of the entropy S_N between the $\log N$ lower bound (7.9) and the $\log^2 N$ upper bound (7.11). In Figure 7.2 we present the results of a numerical calculation. The set K consists of one interval of length $|K| = \pi$. The figure shows clearly the $\log N$ dependence. By the subadditivity property (7.6), we expect the same behavior for all sets K consisting of a finite number of intervals.

7.3 Infinitely many intervals

A Cantor-like construction

The standard Cantor set is constructed recursively by removing in step m a fixed fraction of the set obtained in step $m-1$. This would leave us with a set of zero Lebesgue measure. To avoid this, we remove a fraction in step m which decreases with m, such that the limit set has strictly positive Lebesgue measure.

We start with the unit circle. Remove in the first step in a symmetrical way two open arcs of the unit circle with lengths $(1-\gamma(1))\pi$. The resulting set K_1 consists of two closed arcs each of length $\pi\gamma(1)$. In the second step, for each of these two arcs, a fraction $1-\gamma(2)$ is removed in their middle. This leaves us with a set K_2 consisting of four closed arcs of length $\pi\gamma(1)\gamma(2)/2$.

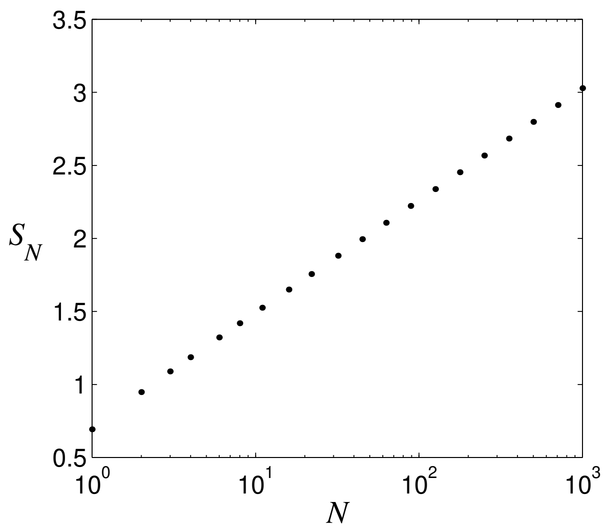

Figure 7.2: The entropy S_N as a function of the length N of the restriction, for an interval of length π as set K. Notice the logarithmic scale.

Continuing this procedure, in step m we obtain a set K_m of 2^m closed arcs of length

$$\ell_{\mathrm{a}}(m) := \frac{\pi}{2^{m-1}} \prod_{n=1}^{m} \gamma(n). \tag{7.12}$$

For $m \geq 2$ there are 2^m holes in between such intervals, 2^{m-1} of which are created in step m. The latter have a length

$$\ell_{\mathrm{h}}(m) := \ell_{\mathrm{a}}(m-1)(1-\gamma(m)) = \frac{\pi}{2^{m-2}}(1-\gamma(m)) \prod_{n=1}^{m-1} \gamma(n). \tag{7.13}$$

The Lebesgue measure of the limit set K is then $2\pi \prod_{n=1}^{\infty} \gamma(n)$.

To construct an explicit example, we have to fix a function $m \mapsto \gamma(m)$. We can as well specify the function $m \mapsto \ell_{\mathrm{h}}(m) := aq^m$, where $0 < q < \frac{1}{2}$ and a is chosen such that

$$2\pi > 2aq + \sum_{m=2}^{\infty} 2^{m-1} a q^m = 2a\,\frac{q(1-q)}{1-2q}.$$

The resulting set has Lebesgue measure $2\pi - 2a\,\dfrac{q(1-q)}{1-2q}$.

Lower bound

To bound $\operatorname{Tr} Q_N(\mathbb{1} - Q_N)$ from below, we start from Equation (7.4). As before, K_m denotes the set obtained after m steps in the construction of the Cantor-like set K. Then K_m is the union of 2^m arcs, each of length $\ell_{\mathrm{a}}(m)$. Because $K \subset K_m$, (7.4) can be estimated by

$$\begin{aligned}
\operatorname{Tr} Q_N(\mathbb{1} - Q_N) &= \frac{N}{4\pi^2} \int_{-\pi}^{\pi} d\varphi\, k_N(\varphi) |K \setminus (K + \varphi)| \\
&\geq \frac{N}{4\pi^2} \int_{-\pi}^{\pi} d\varphi\, k_N(\varphi) |K \setminus (K_m + \varphi)|,
\end{aligned}$$

and since $k_N(\varphi) \geq 4N/\pi^2$ when $|\varphi| \leq \pi/N$,

$$\begin{aligned}
\operatorname{Tr} Q_N(\mathbb{1} - Q_N) &\geq \frac{N^2}{\pi^4} \int_{-\frac{\pi}{N}}^{\frac{\pi}{N}} d\varphi\, |K \setminus (K_m + \varphi)| \\
&= \frac{2N^2}{\pi^4} \int_{0}^{\frac{\pi}{N}} d\varphi\, |K \setminus (K_m + \varphi)|.
\end{aligned}$$

For given N, take m such that

$$\ell_{\mathrm{h}}(m) \geq \frac{\pi}{N} > \ell_{\mathrm{h}}(m+1). \tag{7.14}$$

As K_m consists of 2^m translations of the arc $[0, \ell_{\mathrm{a}}(m)]$, the Cantor-like set K consists of 2^m translations of another Cantor-like set $\tilde{K} \subset [0, \ell_{\mathrm{a}}(m)]$. Let us denote these translations by $x_\ell + [0, \ell_{\mathrm{a}}(m)]$ and $x_\ell + \tilde{K}$ for $\ell = 1, \ldots, 2^m$.

For $\varphi \in [0, \pi/N]$, $\varphi \leq \ell_{\mathrm{h}}(k)$ for all $k = 1, \ldots, m$. This means that a translation by φ of an arc of length $\ell_{\mathrm{a}}(m)$ in K_m will never bridge the hole of length $\ell_{\mathrm{h}}(k), k = 1, \ldots, m$, between this arc and the next. Therefore, every $x_\ell + \tilde{K}$ will overlap with one and only one $x_{\tilde{\ell}} + [0, \ell_i(m)] + \varphi$, namely the one with $\tilde{\ell} = \ell$. As a consequence,

$$\left| K \setminus (K_m + \varphi) \right| = 2^m \left| \tilde{K} \setminus ([0, \ell_{\mathrm{a}}(m)] + \varphi) \right| = 2^m \left| \tilde{K} \setminus [\varphi, \ell_{\mathrm{a}}(m)] \right| = 2^m \left| \tilde{K} \cap [0, \varphi] \right|.$$

This quantity has to be estimated from below. For $\varphi \in (0, \pi/N]$, take n such that $\ell_{\mathrm{a}}(n) \geq \varphi > \ell_{\mathrm{a}}(n+1)$. Then,

$$\left| \tilde{K} \cap [0, \varphi] \right| \geq \ell_{\mathrm{a}}(n+1) \prod_{k=n+2}^{\infty} \gamma(k) = \frac{\pi}{2^n} \prod_{k=1}^{\infty} \gamma(k),$$

and

$$\varphi \leq \ell_{\mathrm{a}}(n) = \frac{\pi}{2^{n-1}} \prod_{k=1}^{n} \gamma(k),$$

which gives

$$\frac{\left|\tilde{K} \cap [0, \varphi]\right|}{\varphi} \geq \frac{1}{2} \prod_{k=n+1}^{\infty} \gamma(k) \geq \frac{|K|}{4\pi},$$

and so

$$\left|\tilde{K} \cap [0, \varphi]\right| \geq \frac{|K| \varphi}{4\pi},$$

which does not depend any longer on n. It follows that

$$\operatorname{Tr} Q_N(\mathbb{1} - Q_N) \geq 2^m \frac{2N^2}{\pi^4} \frac{|K|}{4\pi} \int_0^{\frac{\pi}{N}} d\varphi \, \varphi = \frac{2^m |K|}{4\pi^3}.$$

This is an estimate from below of $\operatorname{Tr} Q_N(\mathbb{1} - Q_N)$ where N and m are coupled by (7.14). >From the latter we also have that $N < \pi/\ell_{\mathrm{h}}(m+1)$. Therefore,

$$\frac{\log \operatorname{Tr} Q_N(\mathbb{1} - Q_N)}{\log N} > \frac{\log\left(\frac{1}{4\pi^3} 2^m |K|\right)}{-\log\left(\frac{1}{\pi} \ell_{\mathrm{h}}(m+1)\right)}.$$

The limit $N \to \infty$ corresponds to the limit $m \to \infty$. Using the explicit form $\ell_{\mathrm{h}}(m) = aq^m$, we finally get

$$\liminf_{N \to \infty} \frac{\log \operatorname{Tr} Q_N(\mathbb{1} - Q_N)}{\log N} \geq \frac{\log 2}{-\log q}. \tag{7.15}$$

Upper bound

To get an upper bound for $\operatorname{Tr} Q_N(\mathbb{1} - Q_N)$, we start from Equation (7.5). With $C(\varphi) := |K^c \setminus (K^c + \varphi)|$, it reads

$$\operatorname{Tr} Q_N(\mathbb{1} - Q_N) = \frac{N}{4\pi^2} \int d\varphi \, k_N(\varphi) C(\varphi),$$

For $\theta > 0$, take m such that

$$\ell_{\mathrm{h}}(m) \geq \theta > \ell_{\mathrm{h}}(m+1). \tag{7.16}$$

We bound $C(\theta)$ from above,

$$\begin{aligned}
C(\theta) &\leq \sum_{k=1}^{\infty} 2^{k-1} \min\{\theta, \ell_{\mathrm{h}}(k)\} \\
&= \sum_{k=1}^{m} 2^{k-1}\theta + \sum_{k=m+1}^{\infty} 2^{k-1}\ell_{\mathrm{h}}(k) \\
&\leq (2^m - 1)\ell_{\mathrm{h}}(m) + \sum_{k=m+1}^{\infty} 2^{k-1}\ell_{\mathrm{h}}(k) \\
&\leq 4\pi(1-\gamma(m)) \prod_{n=1}^{m-1} \gamma(n) + 2\pi \sum_{k=m+1}^{\infty} (1-\gamma(k)) \prod_{n=1}^{k-1} \gamma(n).
\end{aligned} \tag{7.17}$$

Obviously, this bound increases with θ. The kernel $k_N(\varphi)$ satisfies

$$k_N(\varphi) \leq \begin{cases} N & \text{for } |\varphi| \leq \theta, \\ \frac{\pi^2}{N}\frac{1}{\varphi^2} & \text{for } |\varphi| \geq \theta, \end{cases}$$

and so we find

$$\begin{aligned}
\operatorname{Tr} Q_N(\mathbb{1} - Q_N) &\leq \frac{N^2}{4\pi^2} \int_{|\varphi| \leq \theta} d\varphi\, C(\varphi) + \frac{1}{4} \int_{|\varphi| \geq \theta} d\varphi\, C(\varphi)\, \frac{1}{\varphi^2} \\
&\leq \frac{N^2}{2\pi^2} \theta C(\theta) + \frac{1}{4} \sum_{k=0}^{m} (\ell_{\mathrm{h}}(k) - \ell_{\mathrm{h}}(k+1)) \frac{C(\ell_{\mathrm{h}}(k))}{\ell_{\mathrm{h}}(k+1)^2} \\
&\leq \frac{N^2}{2\pi^2} \ell_{\mathrm{h}}(m) C(\theta) + \frac{1}{4} \sum_{k=0}^{m} \frac{\ell_{\mathrm{h}}(k) C(\ell_{\mathrm{h}}(k))}{\ell_{\mathrm{h}}(k+1)^2}.
\end{aligned} \tag{7.18}$$

Take again the explicit form $\ell_{\mathrm{h}}(m) = aq^m$. Then

$$(1-\gamma(m)) \prod_{n=1}^{m-1} \gamma(n) = \frac{2^{m-2}}{\pi} \ell_{\mathrm{h}}(m) = \frac{a}{4\pi}(2q)^m,$$

and (7.17) becomes

$$
\begin{aligned}
C(\theta) &\leq 4\pi(1-\gamma(m))\prod_{n=1}^{m-1}\gamma(n) + 2\pi\sum_{k=m+1}^{\infty}(1-\gamma(k))\prod_{n=1}^{k-1}\gamma(n) \\
&= a(2q)^m + \frac{a}{2}\sum_{k=m+1}^{\infty}(2q)^k = \frac{a(1-q)}{1-2q}(2q)^m. \qquad (7.19)
\end{aligned}
$$

If $\theta = \ell_{\mathrm{h}}(k)$, then by (7.16), we have to put $m = k$, and so

$$
C(\ell_{\mathrm{h}}(k)) \leq \frac{a(1-q)}{1-2q}(2q)^k. \qquad (7.20)
$$

Substituting inequalities (7.19) and (7.20) into (7.18), we find

$$
\begin{aligned}
\operatorname{Tr} Q_N(\mathbb{1}-Q_N) &\leq \frac{N^2}{2\pi^2}aq^m\frac{a(1-q)}{1-2q}(2q)^m + \frac{1}{4}\sum_{k=0}^{m}\frac{aq^k}{(aq^{k+1})^2}\frac{a(1-q)}{1-2q}(2q)^k \\
&= \frac{N^2}{2\pi^2}aq^m\frac{a(1-q)}{1-2q}(2q)^m + \frac{1}{4}\frac{1}{q^2}\frac{2(1-q)}{1-2q}(2^{m+1}-1) \\
&\leq \frac{1-q}{2(1-2q)}\left[\frac{N^2a^2}{\pi^2}(2q^2)^m + \frac{1}{q^2}2^m\right] \\
&=: c_1N^2(2q^2)^m + c_2 2^m \qquad (7.21)
\end{aligned}
$$

where c_1 and c_2 are independent of N.

To get an upper bound as a function of N, we have to fix a function $m(N)$ and plug it into (7.21). Let

$$
\gamma := \frac{\log 2}{-\log q} = \frac{1}{-\log_2 q},
$$

then choose m to be

$$
m = \lfloor \log_{2q^2} N^{\gamma-2} \rfloor \leq \log_{2q^2} N^{\gamma-2} = \frac{\log_2 N^{\gamma-2}}{\log_2 2q^2},
$$

where $\lfloor a \rfloor$ denotes the integer part of the number a. Then

$$
N^2(2q^2)^m \leq N^\gamma,
$$

and

$$2^m \leq \left(2^{\log_2 N^{\gamma-2}}\right)^{\frac{1}{\log_2 2q^2}} = N^{\frac{\gamma-2}{1+\log_2 q^2}} = N^{\gamma},$$

and so

$$\operatorname{Tr} Q_N(\mathbb{1} - Q_N) \leq (c_1 + c_2)N^{\gamma},$$

from which we get the upper bound

$$\limsup_{N\to\infty} \frac{\log \operatorname{Tr} Q_N(\mathbb{1} - Q_N)}{\log N} \leq \gamma = \frac{\log 2}{-\log q}. \tag{7.22}$$

Combining the results (7.15) and (7.22) we see that $\lim_{N\to\infty} \log \operatorname{Tr} Q_N(\mathbb{1} - Q_N)/\log N$ exists, which implies that also $\lim_{N\to\infty} \log S_N/\log N$ exists, and

$$\alpha = \lim_{N\to\infty} \frac{\log S_N}{\log N} = \lim_{N\to\infty} \frac{\log \operatorname{Tr} Q_N(\mathbb{1} - Q_N)}{\log N} = \frac{\log 2}{-\log q}.$$

Since q can be any number in the interval $(0, 1/2)$, the growth exponent α can take any value in $(0, 1)$.

Summary

In this chapter we studied the behavior of the sequence $S(\varphi_N)$ where φ_N is the restriction of a pure quasi-free state on N consecutive spins. For motivations of studying this type of question we refer to section 5.4. The base of our investigations were the estimation S_N in terms of $\operatorname{Tr} Q_N(\mathbb{1} - Q_N)$ (7.3) and the explicit expression (7.4) for the latter.

For a special construction with a Cantor type defining set K we managed to determine the asymptotics explicitly and obtained $S_N \sim N^{\alpha}$ where α is determined by the structure of the Cantor type set, and can take values in the interval $(0, 1)$.

In the case of a single interval we could only locate the asymptotics between a $\log N$ lower bound and a $\log^2 N$ upper bound. Our numerical computations, however, as well as the numerical results of [110] suggest a $\log N$ asymptotics. This was indeed verified in [59] using a different approach from ours, based on the asymptotics of truncated Toeplitz determinants.

The subadditivity property of entropy, proven in section 7.1.2 made it possible to extend the results for one interval to sets consisting of finitely many intervals. The asymptotics in this case was shown to be $c_0 + c_1 M \log N \leq S_N \leq c_2 M \log^2 N$, where M is the number of intervals, c_1 and c_2 are universal constants while c_0 depends on the structure of K.

Chapter 8

Conclusion

We studied the behavior of entropy in different models, with a special emphasis on the insight the entropy of subsystems can give about the structure of a multipartite quantum state. Obviously, several questions remained open, that could be the subject of further study.

In Chapter 3 we proved the additivity of the minimal entropy output

$$S_{\min}\big(\Lambda_t \otimes \Lambda_t\big) = 2\, S_{\min}\big(\Lambda_t\big)$$

for the family of transpose depolarizing channels, in the parameter range $-2/(d^2-2) \leq t \leq 1/(d+1)$, which is strictly smaller than the maximal possible range of parameters on which this family of channels can be defined. The most obvious question is whether this result can be extended to the remaining part $-1/(d-1) \leq t \leq -2/(d^2-2)$. This question was positively answered recently [26], using the monotonicity of entropy as a function of symmetric polynomials [76]. A more challenging task is to show the more general additivity

$$S_{\min}\big(\Lambda_t \otimes \Phi\big) = S_{\min}\big(\Lambda_t\big) + S_{\min}\big(\Phi\big)$$

for an arbitrary channel Φ and paramater t; by taking $\Phi = \Lambda_t^{\otimes n}$ this would imply the reduction of the expression $\lim \frac{1}{n}\, \chi\left(\Lambda_t^{\otimes n}\right)$ for the channel capacity to $\chi(\Lambda_t)$. The main difficulty in this direction is the lack of a suitable generalization of the Schmidt decomposition for systems with more than two components, that would allow to utilize the unitary invariance to simplify the problem as in the case of two components. An other direction of further research could be to investigate the case of channels that are invariant under a higher dimensional irreducible representation of $SU(\mathrm{d_{in}})$.

In Chapter 4 a characterization of triples (ρ, σ, Φ) saturating the monotonicity inequality

$$S\left(\Phi(\rho), \Phi(\sigma)\right) \leq S\left(\rho, \sigma\right)$$

was given in terms of their algebraic structure. A drawback of the algebraic techniques used in the proof is that faithfulness of all the densities ρ, σ, $\Phi(\rho)$ and $\Phi(\sigma)$ has to be assumed. It is not clear how this constraint could be removed; saturation is most likely to fail under the slightest perturbation, hence the standard tricks like adding a constant multiple of identity to make the states faithful, and then let the constant go to zero don't work here.

The source coding theorem of Chapter 6 gives an interpretation of positive mean entropy of ergodic states of the one dimensional quantum spin chain as the best asymptotic compressibility rate under the criterion that a certain measure of fidelity of the reconstruction procedure has to tend to one. A straightforward way of generalization could be to investigate the case when ergodicity of the state is not assumed. A (probably much more interesting) case is when the mean entropy of the state is zero (all pure states are believed to belong to this class). The main idea of the source coding protocol is the projection of the state onto its typical subspace; the efficiency of the protocol is provided by the typical subspace theorem, that tells that the dimension t_n of the typical subspace is related to the entropy as

$$t_n \;\sim\; e^{S(\varphi_n)} \;\sim\; e^{n\, s(\varphi)}$$

when $s(\varphi) > 0$. Such a result is not known in the case of zero mean entropy, and would not only complete the source coding theorem, but would also be very welcome in the analysis of the efficiency of numerical techniques like DMRG computations [110].

A wide class of examples for shift-invariant pure states of the spin chain with vanishing mean entropy is provided by a method of Araki, that enables one to transfer quasi-free states of the CAR algebra to the spin chain [6]. These states are characterized by subsets of the interval $[0, 2\pi]$; a single interval for instance describes the thermodynamical limit ground state of the XX model, where the length of the interval is determined by the ratio of the coupling constant and the outer magnetic field. In Chapter 7 we derived the asymptotics

$$c_0 + c_1\, M\, \log N \leq S_N \leq c_2\, M\, \log^2 N$$

when the subset is the union of M intervals, and

$$S_N \sim N^{\alpha}$$

for a Cantor type subset, where α can take values in the interval $(0,1)$. These results raise a lot of further questions. First of all, numerical computations [110] as well as theoretical results [59] show that the growth is actually logarithmic in the finite interval case; to show this, a more refined upper bound should be used for the entropy function than the one used in (7.2).

As we already mentioned the asymptotics of the size of the typical subspaces is a highly interesting question even in concrete examples, as there is no general result connecting it with the entropy asymptotics. However, to address this question analytically in the specific examples of quasi-free states we have studied, a much better control on the spectrum of the local density matrices is needed.

Zero mean entropy means a sublinear entropy asymptotics; it may be an interesting question from the theoretical point of view whether there exists a highest sublinear growth attainable by pure states, or any sublinear speed can be attained, possibly by quasi-free states with a more complex characterizing set.

The entropy S_N can be interpreted as the entanglement between a finite part with increasing size and the rest of the chain, and its behavior seems to be related to quantum critical phenomena [110], hence its study in other physical models may be relevant as well.

Appendix A

C^*-algebra basics

A.1 Basic examples

A.1.1. Definition. Let $\mathcal{A}$ be an algebra with a norm $\|.\|$ on it. If $(\mathcal{A}, \|.\|)$ as a vector space is a Banach space, and the norm satisfies the submultiplicative property

$$\|ab\| \leq \|a\| \, \|b\| \qquad \forall a, b \in \mathcal{A}$$

then it is called a **Banach algebra**.

An **involution** on the algebra is a map $^* : \mathcal{A} \to \mathcal{A}$ with the following properties:

(i) $(a^*)^* = a$

(ii) $(a+b)^* = a^* + b^*$

(iii) $(\lambda a)^* = \overline{\lambda} a^*$

(iv) $(ab)^* = b^* a^*$

for all $\lambda \in \mathbb{C}$ and $a, b \in \mathcal{A}$.

A Banach algebra with an involution is called a Banach-* algebra, if the involution is an isometry, i.e.

$$\|a^*\| = \|a\| \, ; \qquad \forall a \in \mathcal{A}.$$

If the involution and the norm are also related by the so-called C^*-property

$$\|a^* a\| = \|a\|^2 \, ; \qquad \forall a \in \mathcal{A}, \tag{A.1}$$

then the algebra is called a **C*-algebra**.

A.1.2. Definition. A C^*-algebra $\mathcal{A}$ is **unital**, if there exists an element $\mathbb{1} \in \mathcal{A}$ such that $\mathbb{1}a = a\mathbb{1} = a \quad \forall a \in \mathcal{A}$.

In the following, under a C^*-algebra we always mean a unital one. The following examples give standard forms of certain classes of C^*-algebras.

A.1.3. Example. Let X be a compact Hausdorff topological space, and $\mathcal{F}_c(X)$ the set of all continuous complex-valued functions on X. $\mathcal{F}_c(X)$ forms an algebra with respect to the pointwise operations. If we define the norm to be the supremum norm

$$\|f\| := \sup\{|f(x)| \ : \ x \in X\}$$

and the involution to be the pointwise conjugation, then $\mathcal{F}_c(X)$ becomes a commutative C^*-algebra.

A.1.4. Theorem. (Gelfand and Naimark)
Every commutative C^*-algebra is isomorphic to $\mathcal{F}_c(X)$ for some compact Hausdorff space X.

A.1.5. Example. Let $\mathcal{B}(\mathcal{H})$ denote the set of all bounded operators on a Hilbert space $\mathcal{H}$. $\mathcal{B}(\mathcal{H})$ is a noncommutative C^*-algebra with the standard (pointwise) operations and operator norm and the image of $A \in \mathcal{B}(\mathcal{H})$ under involution being the adjoint of A.

Every subalgebra of $\mathcal{B}(\mathcal{H})$, which is closed under the involution and closed in norm, is again a C^*-algebra. Such an algebra is called a C^*-subalgebra of $\mathcal{B}(\mathcal{H})$.

A.1.6. Theorem. (Gelfand, Naimark and Segal)
Every C^*-algebra is isomorphic to a C^*-subalgebra of $\mathcal{B}(\mathcal{H})$ for some Hilbert space $\mathcal{H}$.

A C^*-algebra is called finite dimensional if it is finite dimensional as a vector space.

A.1.7. Theorem. Every finite dimensional C^*-algebra is isomorphic to $\bigoplus_{k=1}^{K} \mathcal{B}(\mathcal{H}_k)$ for some finite dimensional Hilbert spaces $\mathcal{H}_1, \dots, \mathcal{H}_K$.
The above form we will call the **canonical representation** of the finite dimensional C^*-algebra. A finite dimensional C^*-algebra $\mathcal{A}$ is commutative if and only if all the Hilbert spaces $\mathcal{H}_1, \dots, \mathcal{H}_K$ are one-dimensional in its canonical representation; in this case $\mathcal{A}$ is isomorphic to $\mathcal{F}(\{1, \dots, K\})$.

A C^*-algebra can be specified as the minimal C^*-algebra, the element of which satisfy certain relations. This idea can be formalized precisely by the following

A.1.8. Definition. Let $\mathcal{G} = \{a_i : i \in I\}$ be a set of abstract elements and $\mathcal{R}$ be a set of relations among these elements. A C^*-algebra $\mathcal{A}$ is called a **realization** of these relations if there exists an injective map $\iota : \mathcal{G} \to \mathcal{A}$ such that

(i) if $r(a_{i_1}, \dots, a_{i_n}) = 0$ is a relation in $\mathcal{R}$ then $r(\iota(a_{i_1}), \dots, \iota(a_{i_n})) = 0$ holds as well, and

(ii) the set $\{\iota(a_i) : i \in I\}$ generates $\mathcal{A}$.

A realization $\mathcal{A}$ of the relations $\mathcal{R}$ is called a **universal C^*-algebra** generated by $\mathcal{R}$ if for any other realization $\hat{\mathcal{A}}$ with representants $\hat{\iota}(a_i)$, the map $\kappa : \iota(a_i) \mapsto \hat{\iota}(a_i)$ extends uniquely to a morphism from $\mathcal{A}$ to $\hat{\mathcal{A}}$. Obviously, the universal C^*-algebra is unique up to isomorphism.

A.2 Morphisms

A morphism between unital C^*-algebras is a map that preserves all C^*-algebraic operations and is unital. Automorphisms of full operator algebras have a particularly simple structure:

A.2.1. Theorem. Let γ be an automorphism of $\mathcal{B}(\mathcal{H})$. Then there exists a unitary U such that

$$\gamma(A) = U^* A U \qquad A \in \mathcal{B}(\mathcal{H}). \tag{A.2}$$

Consider now the case when $\mathcal{A} = \bigoplus_{k=1}^{K} \mathcal{B}(\mathcal{H}_k)$ is a finite dimensional C^*-algebra, $\mathcal{B} = \mathcal{B}(\mathcal{H})$ for some finite dimensional Hilbert space $\mathcal{H}$ and γ is a morphism from $\mathcal{A}$ to $\mathcal{B}$. The following is a reformulation of Theorem I.10.7, [27]:

A.2.2. Theorem. Let $\mathcal{A}$, $\mathcal{B}$ and γ be as above. Then there exist

(i) natural numbers $m_1, \dots, m_K$;

(ii) unitaries $U_k : \mathcal{H}_k \to \mathcal{H}_k$, $k = 1, \dots, K$;

(iii) a unitary $U:\ H \to \bigoplus_{k=1}^{K} \mathcal{H}_k \otimes \mathbb{C}^{m_k}$

such that

$$\gamma\left(\bigoplus_{k=1}^{K} A_k\right) = U^*\left(\bigoplus_{k=1}^{K} (U_k A_k U_k^*) \otimes I_{m_k}\right) U. \tag{A.3}$$

A.2.3. Corollary. Any finite dimensional sub-C^*-algebra $\mathcal{A}$ of $\mathcal{B}(\mathcal{H})$ induces a decomposition

$$\mathcal{H} \cong \bigoplus_{k=1}^{K} \mathcal{H}_k^{\mathrm{L}} \otimes \mathcal{H}_k^{\mathrm{R}}, \tag{A.4}$$

in which elements of $\mathcal{A}$ have the form

$$A \cong \bigoplus_{k=1}^{K} A_k \otimes I_{\mathcal{H}_k^{\mathrm{R}}}, \qquad\qquad A_k \in \mathcal{B}\left(\mathcal{H}_k^{\mathrm{L}}\right). \tag{A.5}$$

We could group the terms with the same dimension $d = \dim \mathcal{H}_k^{\mathrm{L}}$ and multiplicity $m = \dim \mathcal{H}_k^{\mathrm{R}}$ together, and obtain an equivalent decomposition

$$\mathcal{H} \cong \bigoplus_{(m,d)} \mathcal{H}_d \otimes \mathcal{H}_m \otimes \mathcal{H}_{(m,d)}, \tag{A.6}$$

with

$$\dim \mathcal{H}_d = d; \qquad \dim \mathcal{H}_m = m; \quad \text{and} \quad \dim \mathcal{H}_{(m,d)} = K(m,d),$$

where $K(m,d)$ is the number of k's with multiplicity m and dimension d. Elements of $\mathcal{A}$ in this representation have the form

$$A \cong \bigoplus_{(m,d)} \sum_{i=1}^{K(m,d)} A(m,d,i) \otimes I_m \otimes E_{i,i}^{(m,d)}, \qquad\qquad A(m,d,i) \in \mathcal{H}_d; \tag{A.7}$$

where $\{E_{i,j}^{(m,d)}\}$ is a set of matrix units in $\mathcal{H}_{(m,d)}$.

Now let $\mathcal{A} = \bigoplus_{k=1}^{K} \mathcal{B}(\mathcal{H}_k)$ and $\mathcal{B} = \bigoplus_{l=1}^{L} \mathcal{B}(\mathcal{K}_l)$ be finite dimensional C^*-algebras in there canonical representations, and γ be a morphism between them. Then

$$\pi_l : \bigoplus_{l'=1}^{L} B_{l'} \mapsto B_l$$

is a morphism from $\mathcal{B}$ to $\mathcal{B}(\mathcal{K}_l)$, and therefore $\pi_l \circ \gamma$ is a morphism from $\mathcal{A}$ to $\mathcal{B}(\mathcal{K}_l)$. By Theorem (A.2.2) there exists a sequence of multiplicities $(m(1,l), \ldots, m(K,l))$ corresponding to the morphism $\pi_l \circ \gamma$. The $K \times L$ matrix μ with entries $\mu_{kl} := m(k,l)$ is called the **inclusion matrix** of γ. Note that the inclusion matrix determines γ, up to unitary rotations in the Hilbert spaces.

A.3 States

A.3.1. Definition. A **state** on the unital C^*-algebra $\mathcal{A}$ is a linear functional $\varphi : \mathcal{A} \to \mathbb{C}$, which is positive, i.e. $\varphi(a^*a) \geq 0 \ \ \forall a \in \mathcal{A}$, and takes the value 1 on the unit: $\varphi(\mathbb{1}) = 1$.

We will denote the set of all states on $\mathcal{A}$ by $\mathcal{S}(\mathcal{A})$.

A positive linear functional φ always satisfies the Schwarz inequality

$$\varphi(b^*a) \leq [\varphi(b^*b)]^{1/2} \, [\varphi(a^*a)]^{1/2} \ . \tag{A.8}$$

Substituting $b = \mathbb{1}$ implies $\|\varphi\| \leq |\varphi(\mathbb{1})|$. A considerably stronger statement is also true:

A.3.2. Lemma. A functional φ is positive if and only if $\|\varphi\| = \varphi(\mathbb{1})$.

Thus, a functional is a state if and only if $\|\varphi\| = \varphi(\mathbb{1}) = 1$.

It is easy to check that if φ_1 and φ_2 are states on the algebra $\mathcal{A}$, then $t\varphi_1 + (1-t)\varphi_2$ is again a state on $\mathcal{A}$ for all $t \in [0,1]$, so $\mathcal{S}(\mathcal{A})$ is a convex subset of the dual space $\mathcal{A}^*$ of $\mathcal{A}$. The extremal points of this convex set are called **pure states**; that is, a state is pure, if and only if it cannot be decomposed into a nontrivial convex combination of two states.

A.3.3. Example. If μ is a probability measure on the Borel sets of a compact Hausdorff space X, then

$$\varphi_\mu(f) := \int f \, d\mu$$

defines a state on $\mathcal{F}_c(X)$. The Riesz representation theorem ([95], Theorem IV. 18.) gives that all the sates on $\mathcal{F}_c(X)$ are given in the above form.

It can easily be seen that such a state is pure if and only if μ is concentrated at one point; in this case $\varphi(f) = f(x)$, where $\{x\}$ is the support of μ. Such a μ is called a **Dirac measure**, and we denote it by δ_x.

A.3.4. Example. Let $\mathcal{A}$ be represented as a subalgebra of $\mathcal{B}(\mathcal{H})$, and let ρ be a positive trace-class operator on $\mathcal{H}$ with unit trace. Then

$$a \mapsto \operatorname{Tr} \rho a$$

defines a state on $\mathcal{A}$.

States that can be given by a density operator in a given representation are called **normal** with respect to that representation. The following example shows that all states are normal with respect to the canonical representation of a finite dimensional C^*-algebra.

A.3.5. Example. Let $\mathcal{A}$ be a finite dimensional C^*-algebra with canonical representation $\bigoplus_{k=1}^n \mathcal{B}(\mathcal{H}_k)$, and let $\mathcal{H} := \bigoplus_{k=1}^n \mathcal{H}_k$. The Hilbert-Schmidt scalar product on $\mathcal{B}(\mathcal{H})$, given by

$$\langle B, A \rangle := \operatorname{Tr} B^* A,$$

can be restricted to $\mathcal{A}$, and makes $\mathcal{A}$ a finite dimensional Hilbert space. Since any state φ on $\mathcal{A}$ is continuous with respect to the norm induced by the Hilbert-Schmidt scalar product, the Riesz representation theorem gives the existence of an element $D_\varphi \in \mathcal{A}$ such that

$$\varphi(A) = \langle D_\varphi, A \rangle = \operatorname{Tr} D_\varphi A\,, \qquad\qquad A \in \mathcal{A}$$

holds. The positivity of φ implies the positivity of D_φ, and $\varphi(\mathbb{1}) = 1$ implies $\operatorname{Tr} D_\varphi = 1$.

The easiest example of a density operator is a rank one projection; the state defined by it is called a vector state. The universal representation of a C^*-algebra $\mathcal{A}$ is a faithful representation, with the property that all states of $\mathcal{A}$ can be given as vector states in this representation (Theorem I.9.12, [27]). However, for each state of $\mathcal{A}$ there exists a canonical representation of $\mathcal{A}$, associated to the state φ, the properties of which are in intimate relation with the properties of the state.

A.3.6. Definition. A **representation** of a C^*-algebra $\mathcal{A}$ is a C^*-algebra morphism $\Pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ with some Hilbert space $\mathcal{H}$. The representation is called **faithful**, if $\ker \Pi = \{0\}$.

A representation $\Pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ is called **irreducible**, if there is no nontrivial subspace of $\mathcal{H}$ invariant under $\Pi(\mathcal{A})$, which statement is equivalent with $\Pi(\mathcal{A})' = \mathbb{C}I$, where $\Pi(\mathcal{A})'$ denotes the commutant of $\Pi(\mathcal{A})$.

Two representations $\Pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ and $\hat{\Pi} : \mathcal{A} \to \mathcal{B}(\hat{\mathcal{H}})$ are called **equivalent**, if there exists a C^*-algebra isomorphism between $\Pi(\mathcal{A})$ and $\hat{\Pi}(\mathcal{A})$, and **unitarily equivalent**, if there exists a unitary $U : \mathcal{H} \to \hat{\mathcal{H}}$ such that $\hat{\Pi}(a) = U\Pi(a)U^*$ holds for all $a \in \mathcal{A}$.

A vector $x \in \mathcal{H}$ is called **cyclic** for the representation, if $\{\Pi(a)x \ : \ a \in \mathcal{A}\}$ is dense in $\mathcal{H}$, and **separable** if $\Pi(a)x = 0$ implies $a = 0$.

A.3.7. Theorem. For a given state φ on a C^*-algebra $\mathcal{A}$ there exists a representation Π_φ of $\mathcal{A}$ on a Hilbert space $\mathcal{H}_\varphi$, and a unit vector $\Omega_\varphi \in \mathcal{H}_\varphi$ which is cyclic for the representation, such that

$$\varphi(a) = \langle \Omega_\varphi, \Pi_\varphi(a)\Omega_\varphi \rangle$$

holds for all $a \in \mathcal{A}$.

The representation Π_φ is irreducible if and only if φ is pure, and Ω_φ is separable if and only if φ is **faithful**, i.e. $\varphi(a^*a) = 0$ implies $a = 0$.

This representation is unique up to unitary equivalence, and any such representation is called a **GNS representation** of the state φ.

A.3.8. Example. Let $\mathcal{A} = \mathcal{B}(\mathcal{H})$ for some finite dimensional $\mathcal{H}$, and let D be the density operator of the state φ with non-zero eigenvalues $\lambda_1, \dots, \lambda_r$ and corresponding orthonormal eigenvectors $e_1, \dots, e_r$. Let $\mathcal{K}$ denote the subspace spanned by $e_1, \dots, e_r$ (i.e. the support of φ). Then the GNS representation of φ can be given in the following form:

$$\mathcal{H}_\varphi := \mathcal{H} \otimes \mathcal{K}, \qquad \Pi_\varphi(A) := A \otimes I_{\mathcal{K}}, \qquad \Omega_\varphi := \sum_{i=1}^{r} \sqrt{\lambda_i}\, e_i \otimes e_i.$$

The GNS representation of a state φ is in general 'much smaller' than the universal representation; in particular, not all states ω can be given by a density operator in this representation. If, in the contrary, ω is normal with respect to the GNS representation of φ then we say that ω is normal with respect to φ. The following example shows that it need not be the case for two arbitrarily chosen states even if the algebra is finite dimensional.

A.3.9. Example. Let

$$\mathcal{A} := \mathcal{B}(\mathcal{H}) \oplus \mathcal{B}(\mathcal{K}); \quad D_\omega := |x\rangle\langle x| \oplus 0; \quad D_\varphi := 0 \oplus |y\rangle\langle y|;$$

where x and y are unit vectors in $\mathcal{H}$ and $\mathcal{K}$, respectively. The corresponding GNS representations are

$$\mathcal{H}_\omega := \mathcal{H} \qquad \Pi_\omega(A \oplus B) := A; \qquad \Omega_\omega := x;$$

and

$$\mathcal{H}_\varphi := \mathcal{K}; \qquad \Pi_\varphi(A \oplus B) := B; \qquad \Omega_\varphi := y.$$

We can see that no density matrix D can represent the state ω in the GNS representation of φ, simply because the first summand of a general algebra element is not present in the expression $\mathrm{Tr} D \Pi_\varphi (A \oplus B) = \mathrm{Tr} DB$.

Simple considerations show that the above situation cannot happen when $\mathcal{A} = \mathcal{B}(\mathcal{H})$, i.e. any state is normal with respect to any other state in this case.

A.4 Positivity and complete positivity

An element a in the C^*-algebra $\mathcal{A}$ is called positive, if it can be written in the form $a = b^* b$ for some $b \in \mathcal{A}$. The set of positive elements $\mathcal{A}_+$ forms a convex cone with the property $\mathcal{A}_+ \bigcap (-\mathcal{A}_+) = \{0\}$. The notion of positivity induces a partial ordering on the set of self-adjoint elements; $a \geq b$ holds if $a - b$ is positive. The positive cone is also generating, in the sense that every element can be decomposed into a linear combination of positive elements.

A map $\alpha : \mathcal{A} \to \mathcal{B}$ is positive, if it preserves positivity of elements, i.e. $\alpha(\mathcal{A}_+) \subset \mathcal{B}_+$. Convex combinations and compositions of positive maps are positive again. However, positivity is not preserved under forming tensor product of maps.

A.4.1. Example. Let $\{E_i : i = 1, \ldots, d\}$ and $\{e_i : i = 1, 2\}$ be the standard set of base vectors in $\mathbb{C}^d$ and $\mathbb{C}^2$, with corresponding set of matrix units $\{E_{ij} : i, j = 1, \ldots, d\}$ and $\{e_{ij} : i, j = 1, 2\}$. The matrix $M := \sum_{i,j=1}^2 E_{ij} \otimes e_{ij}$ is positive in $\mathcal{M}_d \otimes \mathcal{M}_2$, as it equals to $|\psi\rangle\langle\psi|$ with $\psi = E_1 \otimes e_1 + E_2 \otimes e_2$ (here $\mathcal{M}_d$ denotes the algebra of $d \times d$ complex matrices).

The transposition map $\alpha(E_{ij}) := E_{ji}$ is positive. However, $\alpha \otimes \mathrm{id}_2$ exchanges the off-diagonal blocks of a 4×4 block matrix, hence $\alpha \otimes \mathrm{id}_2(M) = \sum_{i,j=1}^{2} E_{ji} \otimes e_{ij}$ is easily seen to be non-positive. Therefore the map $\alpha \otimes \mathrm{id}_2$ is non-positive even though both factors are positive.

Motivated by this example, the following subclasses of positive maps can be introduced:

A.4.2. Definition. A map $\alpha : \mathcal{A} \to \mathcal{B}$ is called d-positive, if

$$\mathrm{id}_d \otimes \alpha : \mathcal{M}_d \otimes \mathcal{A} \to \mathcal{M}_d \otimes \mathcal{B} \tag{A.9}$$

is positive.

If α is d-positive for all $d \geq 1$ then it is called completely positive (CP).

It is easily seen that d-positivity implies d'-positivity for all $d' \leq d$, and that 1-positivity coincides with positivity, therefore all d-positive maps are positive as well. Example (A.4.1) shows that the transposition map is not 2-positive, indicating that Definition (A.4.2) gives a non-trivial classification of positive maps. However, for commutative C^*-algebras all the above defined classes coincide:

A.4.3. Theorem. If $\mathcal{A}$ or $\mathcal{B}$ is a commutative C^*-algebra then any positive map $\Phi : \mathcal{A} \to \mathcal{B}$ is completely positive as well.

Positive, d-positive and completely positive maps were extensively studied for general C^*-algebras. Here we briefly summarize some results about finite dimensional C^*-algebras; first we start with a few basic examples, and then we turn to standard representations of completely positive maps.

A.4.4. Example. C^*-algebra morphisms are completely positive. In particular, if $\mathcal{A}$ is a sub-C^*-algebra of $\mathcal{B}(\mathcal{H})$ and U is a unitary on $\mathcal{H}$ then

$$\mathrm{Ad}_U(A) := UAU^*$$

is completely positive.

A.4.5. Example. Sums, positive multiples and compositions of d-positive maps are d-positive again.

A.4.6. Example. Let $\mathcal{A}$ and $\mathcal{B}$ be finite dimensional C^*-algebras, and $\alpha^* : \mathcal{B} \to \mathcal{A}$ be the adjoint of $\alpha : \mathcal{A} \to \mathcal{B}$ with respect to the Hilbert-Schmidt inner products. The property

$$A \geq 0 \iff \langle B, A \rangle \geq 0 \;\; \forall B \geq 0$$

of the Hilbert-Schmidt inner product, together with the fact $(\mathrm{id}_d \otimes \alpha)^* = \mathrm{id}_d \otimes \alpha^*$ easily implies

$$\alpha \;\text{ is } d\text{-positive} \quad \iff \quad \alpha^* \;\text{ is } d\text{-positive}\,.$$

A.4.7. Example. Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be finite dimensional C^*-algebras, ι_1 and ι_2 be the canonical embeddings

$$\iota_1(A) := A \otimes \mathbb{1}_{\mathcal{B}}; \qquad \iota_2(B) := \mathbb{1}_{\mathcal{A}} \otimes B.$$

The adjoints of these embeddings are the partial trace operations, given by

$$\mathrm{Tr}_2 \left(\sum_k A_k \otimes B_k \right) = \sum_k A_k \,\mathrm{Tr}\, B_k\,; \;\; \mathrm{Tr}_1 \left(\sum_k A_k \otimes B_k \right) = \sum_k B_k \,\mathrm{Tr}\, A_k.$$

The embeddings are C^*-algebra morphisms and therefore completely positive, hence by Example (A.4.6) the partial trace operations are completely positive as well.

A.4.8. Definition. A map $\Phi : \mathcal{A} \to \mathcal{B}$ between finite dimensional C^*-algebras is called

(i) **unital**, if $\Phi\left(\mathbb{1}_{\mathcal{A}}\right) = \mathbb{1}_{\mathcal{B}}$;

(ii) **trace-preserving**, if $\mathrm{Tr}\,\Phi(A) = \mathrm{Tr}\,A \quad \forall A \in \mathcal{A}$;

(iii) **stochastic**, if it is completely positive and trace-preserving (CPTP);

(iv) **bistochastic**, if it is stochastic and unital.

It is easily seen that a map is unital if and only if its adjoint is trace-preserving, and vice versa.

A.4.9. Example. Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be subalgebras of the finite dimensional Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, and $U : \mathcal{H}_1 \otimes \mathcal{H}_2 \to \mathcal{H}_1 \otimes \mathcal{H}_2$ be a unitary. For a fixed density operator $\rho \in \mathcal{B}$ the map

$$A \mapsto A \otimes \rho$$

is easily seen to be stochastic, and therefore, by the previous examples, the map

$$A \mapsto \mathrm{Tr}_2\, U\,(A \otimes \rho)\, U^*$$

is stochastic as well.

A.4.10. Example. Let $\mathcal{A}$ and $\mathcal{B}$ be subalgebras of the finite dimensional Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$, and let $\{V_k : k = 1, \dots n\}$ be linear operators from $\mathcal{H}$ to $\mathcal{K}$. The map

$$A \mapsto \sum_k V_k A V_k^*$$

is easily seen to be completely positive.

To investigate properties of maps $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ we fix a base $\{e_i : i = 1, \dots, \dim \mathcal{H}\}$ in $\mathcal{H}$ with corresponding matrix units $e_{ij} := |e_i\rangle\langle e_j|$, and introduce J as the unique anti-unitary extension of the map $Je_i := e_i$, and $\psi := \sum_i e_i \otimes e_i$. The map

$$T : \; |y\rangle\langle x| \mapsto Jx \otimes y$$

is well-defined on $\mathcal{B}(\mathcal{H}, \mathcal{K})$ and gives a linear isomorphism between $\mathcal{B}(\mathcal{H}, \mathcal{K})$ and $\mathcal{H} \otimes \mathcal{K}$. For a general element $A \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ we obtain

$$T(A) = \sum_i e_i \otimes Ae_i = (I \otimes A)\psi \quad \text{and} \quad T^*\left(\sum_i e_i \otimes y_i\right) = \sum_i |y_i\rangle\langle e_i|.$$

Similarly, the map

$$\mathcal{T} : \; |B\rangle\langle A| \mapsto JAJ^* \otimes B \tag{A.10}$$

is well-defined on $\mathcal{B}(\mathcal{B}(\mathcal{H}), \mathcal{B}(\mathcal{K}))$, and gives a linear isomorphism between $\mathcal{B}(\mathcal{B}(\mathcal{H}), \mathcal{B}(\mathcal{K}))$ and $\mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{K})$. For a general map Φ we obtain

$$\mathcal{T}(\Phi) = \sum_{ij} e_{ij} \otimes \Phi(e_{ij}) = \mathrm{id} \otimes \Phi(|\psi\rangle\langle\psi|).$$

The operator $\mathcal{T}(\Phi)$ is called the **Choi matrix** of the map Φ, and we denote it by $C(\Phi)$. An alternative way to identify the above spaces could be by using the map $\tilde{\mathcal{T}}: \ |B\rangle\langle A| \mapsto A^* \otimes B$; this form is called Jamiolkowski duality [58]. The resulting operators $\mathcal{T}(\Phi)$ and $\tilde{\mathcal{T}}(\Phi)$ have slightly different properties; for our present purposes it is more convenient to use Choi's approach.

The concrete form of the Choi matrix depends on the base chosen. However, if we fix another base $\{\hat{e}_i \ : \ i = 1, \ldots, \dim \mathcal{H}\}$, and denote the corresponding Choi matrix of Φ by $\hat{C}(\Phi)$ then a straightforward computation shows that

$$\hat{C}(\Phi) = [(UJU^*J^*) \otimes I] \ C(\Phi) \ [(JUJ^*U^*) \otimes I]$$

where U is the unitary defined by $Ue_i = \hat{e}_i$, i.e. $C(\Phi)$ and $\hat{C}(\Phi)$ are unitarily equivalent to each other.

Let $X \mapsto X^{\mathrm{T}_1}$ be the first partial transpose operation on $\mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{K})$, defined by $(e_{ij} \otimes B)^{\mathrm{T}_1} := e_{ji} \otimes B, \quad B \in \mathcal{B}(\mathcal{K})$. The following properties of the Choi matrix can be checked by a straightforward computation:

(i) $A = \mathrm{Tr}_2 \ C(\Phi)^{\mathrm{T}_1} (A \otimes I)$

(ii) $\Phi(A) = \mathrm{Tr}_1 \ C(\Phi)^{\mathrm{T}_1} (A \otimes I)$

(iii) $C(\Phi)^* = C(\Phi) \Longleftrightarrow \Phi(A)^* = \Phi(A^*) \quad \forall A \in \mathcal{A}$

(iv) $\mathrm{Tr}_1 C(\Phi) = \Phi(\mathbb{1})$

(v) $\mathrm{Tr}_2 C(\Phi) = \sum_{ij} e_{ij} \, \mathrm{Tr} \, \Phi(e_{ij})$

(vi) Φ is unital $\Longleftrightarrow \mathrm{Tr}_1 C(\Phi) = \mathbb{1}$
Φ is trace-preserving $\Longleftrightarrow \mathrm{Tr}_2 C(\Phi) = \mathbb{1}$

(vii) If Φ is unital or trace-preserving then $\mathrm{Tr} \, C(\Phi) = d$.

(viii) $C(\Phi) = |\psi\rangle\langle\psi|$ for some $\psi \in \mathcal{H} \otimes \mathcal{K}$ if and only if $\Phi(A) = VAV^*$ where $V = T^*|\psi\rangle$.

Property (viii) together with Example (A.4.10) implies

A.4.11. Theorem. Let $\mathcal{A}$ and $\mathcal{B}$ be subalgebras of the finite dimensional Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$, and $\Phi: \ \mathcal{A} \to \mathcal{B}$. The followings are equivalent:

(i) Φ is completely positive

(ii) $C(\Phi)$ is positive semi-definite

(iii) There exist operators $V_k : \mathcal{H} \to \mathcal{K}; \quad k = 1, \dots, n$ such that

$$\Phi(A) = \sum_k V_k A V_k^*; \qquad A \in \mathcal{A} \quad \text{holds.} \tag{A.11}$$

The form (A.11) is called a **Kraus representation** of Φ, and the operators V_k the corresponding **Kraus operators**.

The representing operators are not at all unique; any decomposition $\mathcal{C}(\Phi) = \sum_k |\psi_k\rangle\langle\psi_k|$ yields a corresponding decomposition, with Kraus operators $V_k := T^*|\psi_k\rangle$. The minimal number of Kraus operators is easily seen to be equal to the rank of the Choi matrix; a minimal set of Kraus operators is $\{T^*|\psi_k\rangle\}$ for a set of orthonormal eigenvectors $\{|\psi_k\rangle\}$.

A completely positive map is stochastic if and only if $\sum_{k=1}^n V_k^* V_k = I_{\mathcal{H}}$ holds for any set of representing Kraus operators. The operator $\sum_{k=1}^n V_k \otimes e_{k1}$ is easily seen to be a partial isometry on the Hilbert space $\hat{\mathcal{H}} := \mathcal{H} \otimes \mathbb{C}^n$, where $e_1, \dots, e_n$ is an orthonormal base for $\mathbb{C}^n$, therefore it can be extended to a unitary U on $\hat{\mathcal{H}}$. Straightforward computation shows that $\Phi(A) = \mathrm{Tr}_2\, U\,(A \otimes |e_1\rangle\langle e_1|)\, U^*$ holds for all $A \in \mathcal{A}$, which, together with Example A.4.9 yields the following:

A.4.12. Theorem. (Lindblad representation)
Let $\mathcal{A}$ be a C^*-subalgebra of $\mathcal{B}(\mathcal{H})$ for some finite dimensional Hilbert space $\mathcal{H}$. A map $\Phi : \mathcal{A} \to \mathcal{A}$ is stochastic, if and only if there exists a Hilbert space $\mathcal{K}$, a unitary $U : \mathcal{H} \otimes \mathcal{K} \to \mathcal{H} \otimes \mathcal{K}$ and a state ρ on $\mathcal{K}$ such that

$$\Phi(A) = \mathrm{Tr}_2\, U\,(A \otimes \rho)\, U^*$$

holds for all $A \in \mathcal{A}$.

A.5 Properties of unital 2-positive maps

Let $\alpha : \mathcal{A} \to \mathcal{B}$ be a unital 2-positive map. The element $\hat{x} := \begin{bmatrix} \mathbb{1} & x \\ x^* & x^*x \end{bmatrix}$ is positive in $\mathcal{M}_2 \otimes \mathcal{A}$, therefore $(\mathrm{id}_2 \otimes \alpha)\,(\hat{x}) = \begin{bmatrix} \mathbb{1} & \alpha(x) \\ \alpha(x)^* & \alpha(x^*x) \end{bmatrix}$ is also positive, which in turn is equivalent to

$$\alpha(x)^*\, \alpha(x) \leq \alpha(x^*x). \tag{A.12}$$

Inequality (A.12) is called **2-positivity inequality** (sometimes also referred to as **Schwarz inequality**). Easy algebra shows that the set

$$\mathcal{A}_\alpha := \{x \in \mathcal{A} \,:\, \alpha(x)^* \, \alpha(x) = \alpha(x^* x) \;\text{ and }\; \alpha(x)\, \alpha(x)^* = \alpha(x x^*)\} \tag{A.13}$$

is a C^*-algebra, and the restriction of α to $\mathcal{A}_\alpha$ is a C^*-algebra morphism. Moreover, the stronger multiplicativity property

$$\alpha(xy) = \alpha(x)\alpha(y) \qquad\qquad x \in \mathcal{A}_\alpha,\; y \in \mathcal{A}; \tag{A.14}$$

holds.

The fixed point set $\mathcal{F}_\alpha$ of a 2-positive unital map α is in general a linear subspace of $\mathcal{A}$, that is also closed under the $*$-operation. However, if there exists a faithful invariant state φ of α then $\mathcal{F}_\alpha$ is also an algebra. Indeed, let $x \in \mathcal{F}_\alpha$; then by the Schwarz inequality $\alpha(x^* x) - \alpha(x)^* \alpha(x)$ is positive,

$$\varphi\left(\alpha(x^* x) - \alpha(x)^* \alpha(x)\right) = \varphi\left(\alpha(x^* x) - x^* x\right) = 0,$$

and faithfulness of φ implies $\alpha(x^* x) - \alpha(x)^* \alpha(x) = 0$. Taking now linear combinations of fixed points, and repeating the above argument, we obtain the desired statement. A sideresult is

$$\mathcal{F}_\alpha \subset \mathcal{A}_\alpha\,. \tag{A.15}$$

Suppose now that α is a unital 2-positive map from $\mathcal{B}(\mathcal{H})$ to $\mathcal{B}(\mathcal{H})$ for some finite dimensional Hilbert space $\mathcal{H}$, and φ is a faithful α-invariant state. Then

$$\langle A, B\rangle_\varphi := \varphi(A^* B)\,, \qquad\qquad A, B \in \mathcal{B}(\mathcal{H});$$

is easily seen to give an inner product on $\mathcal{B}(\mathcal{H})$, and Schwarz inequality implies that α is a contraction in the norm induced by it. Hence von Neumann's mean ergodic theorem can be applied, and gives the existence of the limit

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} \alpha^k =: E, \tag{A.16}$$

where E is the projection onto the fixed point algebra of α, self-adjoint in the scalar-product $\langle\,.\,,\,.\,\rangle_\varphi$. Moreover, it is a so-called **conditional expectation**:

A.5.1. Definition. Let $\mathcal{A}$ be a C^*-algebra, and $\mathcal{A}_0$ a C^*-subalgebra of it. A map $E : \mathcal{A} \to \mathcal{A}_0$ is called a conditional expectation onto $\mathcal{A}_0$, if

(i) E is positivity-preserving and unital;

(ii) $E(xy) = xE(y)$ for all $x \in \mathcal{A}_0$ and $y \in \mathcal{A}$.

The conditional expectation E is said to be φ-preserving for a state φ if $\varphi \circ E = \varphi$.

Note that condition (ii) with the choice $b := \mathbb{1}$ implies $E(a) = a \;\; \forall a \in \mathcal{A}_0$. The first criterion obviously holds for E defined as in (A.16); it is even 2-positive. The second condition follows from (A.14) and (A.15).

Now let D be the density of the faithful state φ. The group of unitaries $t \mapsto \mathrm{Ad}_{D^{it}}$ is called the **modular group** of the state φ.

A.5.2. Theorem. (Takesaki, [106])
The modular group of φ leaves a subalgebra $\mathcal{A}_0$ invariant (i.e. $D^{it}\,\mathcal{A}_0\,D^{-it} \subset \mathcal{A}_0$) if and only if there exists a φ-preserving conditional expectation from $\mathcal{A}$ to $\mathcal{A}_0$.

A.5.3. Corollary. The modular group of a faithful α-invariant state leaves the fixed point algebra of α invariant.

Note that since $\mathrm{Ad}_{D^{it}}$ is a unitary on $\mathcal{B}(\mathcal{H})$ with respect to the Hilbert-Schmidt inner product, the invariance property can actually be written as $D^{it}\,\mathcal{A}_0\,D^{-it} = \mathcal{A}_0$.

Note that the conditional expectation in (A.16) is independent of the invariant state φ, we only used the existence of φ. Therefore Theorem 9.3 in [83] gives

$$D_1^{it} D_2^{-it} \in \mathcal{A}_0 \qquad\qquad \forall t \in \mathbb{R} \tag{A.17}$$

for any two faithful α-invariant states φ_1 and φ_2, with densities D_1 and D_2.

References

An excellent introduction to the theory of C^*-algebras is [27]; it covers the material of sections A.1, A.2 and A.3. For a more detailed summary of positive maps we refer to [3] and [81]. Extensive material about modular theory can be found in [105], [83] and [21].

Appendix B

In the standard formalism of quantum mechanics states of a system are described by unit vectors (or, more precisely, rays) of the Hilbert space of the system, the time evolution is given by a unitary group and measurements are described by complete sets of projections. When dealing with open systems, however, these notions are not sufficient anymore. Here we present a few examples that indicate how some more general notions appear naturally in certain situations. For the rest, let $\mathcal{H}_S$ denote the Hilbert space of the system in interest and $\mathcal{H}_E$ the Hilbert space of an environment. For technical convenience we assume all Hilbert spaces to be finite dimensional.

B.1.1. Example. Let the composite system be in the state $\psi \in \mathcal{H}_S \otimes \mathcal{H}_E$. When performing a measurement only on the system S, the measurement operators have the form $P_x \otimes \mathbb{1}_E$ for some projections P_x on $\mathcal{H}_S$. The outcome probabilities are

$$p_x = \langle \psi, (P_x \otimes \mathbb{1}_E)\, \psi \rangle = \operatorname{Tr} \rho_1 \, P_x \,,$$

where $\rho_1 = \operatorname{Tr}_2 |\psi\rangle\langle\psi|$, i.e.

$$(\rho_1)_{ij} = \sum_k \langle e_i \otimes f_k, \psi \rangle \langle \psi \, e_j \otimes f_k \rangle$$

for an orthonormal base $\{e_i\}$ and $\{f_j\}$ of $\mathcal{H}_E$ and $\mathcal{H}_S$. Hence the outcome probabilities of any measurement on system S only are fully described by the density operator ρ_1.

B.1.2. Example. Another situation when it is convenient to work within the density matrix formalism is when one does not know the exact state of the system, only that it is one of the states $\psi_1, \dots, \psi_r$, with probabilities

$p_1, \ldots, p_r$. Performing a measurement with projections P_x, the outcome probabilities are

$$\begin{aligned} \text{Prob(outcome}{=}x) &= \sum_{k=1}^{r} \text{Prob(outcome}{=}x \,|\, \text{the true state is } \psi_k) \\ &\qquad \cdot \text{Prob(the true state is } \psi_k) \\ &= \sum_{k=1}^{r} \langle \psi_k, P_x \, \psi_k \rangle \, p_k = \operatorname{Tr} \rho \, P_x \, , \end{aligned}$$

where $\rho = \sum_{k=1}^{r} p_x \, |\psi_k\rangle\langle\psi_k|$ is a density operator.

B.1.3. Example. Suppose the system and the environment are independent i.e. their state vector is a product $\psi = \psi_1 \otimes \psi_2$. If a measurement is performed on the joint system with measurement operators $P_x = \sum_k A_{x,k} \otimes B_{x,k}$ then the outcome probabilities are

$$p_x = \langle \psi_1 \otimes \psi_2, \, \left(\sum_k A_{x,k} \otimes B_{x,k} \right) \, \psi_1 \otimes \psi_2 \rangle = \langle \psi_1, \, M_x \, \psi_1 \rangle \, ,$$

with

$$M_x = \sum_k A_{x,k} \, \langle \psi_2, \, B_{x,k} \, \psi_2 \rangle \, .$$

Hence any measurement can be described by the state vector ψ_1 and a set of positive semi-definite operators $\{M_x\}$, satisfying $\sum_x M_x = \mathbb{1}_E$. Such a set of operators is called a **positive operator valued measurement** (POVM).

B.1.4. Example. Let the system and the environment be independent at time $t = 0$, in the state $\psi_1 \otimes \psi_2$, and suppose their joint evolution is governed by a Hamiltonian that contains interaction terms between S and E. The state of the joint system after time t is $\psi(t) = U(t) \, \psi_1 \otimes \psi_2$, where $U(t)$ is the unitary group generated by the Hamiltonian. Due to the interaction, ψ_t is not a product state anymore. This implies that the restricted state $\rho_1(t)$ of the system is not a vector state, hence the evolution $\Phi_t : \psi_1 \mapsto \rho_1(t)$ cannot be described by unitary operators. However, Φ_t is a completely positive trace-preserving map for every t (see Example A.4.9).

Motivated by the above examples, we will work with the following set of postulates:

(P1) To every physical system there is a Hilbert space $\mathcal{H}$ attached.

(P2) Possible measurements on the system are described by POVM's. All measurement operators generate a C^*-algebra $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$, called the observable algebra of the system.

(P3) States of the system are described by positive, normalized functionals φ on the observable algebra. If the system is in state φ then the outcome probabilities of a measurement $\{M_x\}$ are $p_x = \varphi(M_x)$.

(P4) Any change of the state of the system is described by a CPTP map.

As a closing remark, we show how discrete classical systems can be treated in this formalism.

B.1.5. Example. Let the state space of a classical system be the finite set $\mathcal{X}$. The corresponding Hilbert space is $\mathcal{F}(\mathcal{X})$, the set of all complex-valued functions on $\mathcal{X}$, with the inner product $\langle f, g\rangle := \sum_x \overline{f}(x)\, g(x)$. The observable algebra is the algebra of multiplication operators, which can be identified by $\mathcal{F}(\mathcal{X})$. Hence a measurement is described by a set of functions $\{m_i\}$ such that $m_i(x) \geq 0$ and $\sum_i m_i(x) = 1\ \forall x$. States of the system are described by probability measures μ on $\mathcal{X}$; the outcome probabilities of a measurement are $p_i = \sum_x \mu(x)\, m_i(x)$. CPTP maps are in one-to-one correspondence with stochastic matrices $(T_{xy})_{x,y\in\mathcal{X}}$; the change of the state is then $\mu \mapsto \mu T$, where $(\mu T)(y) = \sum_x \mu(x)\, T_{xy}$.

A detailed introduction into the above formalism can be found e.g. in [81] or [3].

Appendix C

C.1 Ground state of the XY model

Using the notations of section 5.3, the local Hamiltonians of the XX-model have the form

$$H_N := \frac{1}{2} \sum_{k=-N}^{N-1} \left(\hat{\sigma}_x^k \hat{\sigma}_x^{k+1} + \hat{\sigma}_y^k \hat{\sigma}_y^{k+1} \right) + h \sum_{k=-N}^{N} \hat{\sigma}_z^k \tag{C.1}$$

on the local algebra $\hat{\mathcal{A}}_{[-N,N]}$.

By the help of the identities (5.31) and (5.33), we can rewrite the local Hamiltonians in terms of fermionic creation and annihilation operators:

$$H_N = \sum_{k=-N}^{N-1} \left(a_k^* a_{k+1} + a_{k+1}^* a_k \right) - 2h \sum_{k=-N}^{N} a_k^* a_k + h(2N+1)I. \tag{C.2}$$

Note that the last term, $h(2N+1)I$ doesn't influence the ground state, so we can simply leave it. The resulting operator, H_N' is the second-quantized form of the one-particle operator

$$A_N = \sum_{k=-N}^{N-1} \left(|\delta_k\rangle\langle\delta_{k+1}| + |\delta_{k+1}\rangle\langle\delta_k| \right) - 2h \sum_{k=-N}^{N} |\delta_k\rangle\langle\delta_k|,$$

acting on the space $\mathcal{H}_N := \operatorname{span}\{\delta_{-N}, \ldots, \delta_N\}$.

The quasi-free state φ_N with symbol Q_N is easily seen to be a ground state of H_N', where Q_N is the eigenprojection of A_N, corresponding to the negative part of its spectrum. The density of this pure state is $e_1 \wedge \ldots \wedge e_r$, where $e_1, \ldots, e_r$ are the eigenvectors of A_N, corresponding to negative eigenvalues.

Note that H_N is an even element, thus φ_N is an even state, therefore the thermodynamical limit exists on any odd local element b, and gives zero:

$$\lim_{N\to\infty} \varphi_N(b) = \lim_{N\to+\infty} 0 = 0.$$

Hence in the study of the thermodynamical limit we can restrict our attention to the states restricted to the even subalgebras $(\hat{\mathcal{A}}_{[-N,N]})_+$, or equivalently we can study the behavior of the corresponding even states on the algebras $\mathcal{A}_{[-N,N]}$.

Let P_N denote the natural embedding of Q_N into $l^2(\mathbb{Z})$; then we have

$$\varphi_N\left(c_{i_1}^* \dots c_{i_n}^* c_{j_m} \dots c_{j_1}\right) = \delta_{m,n} \det\{\langle \delta_{j_k}, P_N\ \delta_{i_l}\rangle\}_{k,l=1}^n, \tag{C.3}$$

for sufficiently large N (i.e. $N \geq \max\{|i_1|, \dots, |i_n|,\ |j_1|, \dots, |j_n|\}$). If we could show that the sequence of operators $(P_N)_{N\in\mathbb{N}}$ converges in the weak operator topology (that is, the matrix elements $(P_N)_{t,s} = \langle \delta_t, P_N \delta_t\rangle$ converge) then we have convergence in formula (C.3), and therefore the existence of the thermodynamical limit is shown.

The matrix of A_N is of the form

$$\begin{bmatrix} -2h & 1 & 0 & 0 & \dots & 0 \\ 1 & -2h & 1 & 0 & \dots & 0 \\ 0 & 1 & -2h & 1 & \ddots & \vdots \\ 0 & 0 & 1 & -2h & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & 0 & \dots & 0 & 1 & -2h \end{bmatrix}.$$

The eigenvalues and corresponding eigenvectors of this matrix are

$$\lambda_k = -2h + 2\cos\left(k\frac{\pi}{2N+2}\right); \qquad k = 1, \dots, 2N+1;$$

$$v_k = \sum_{s=-N}^{N} \sin\left(k(s+N+1)\frac{\pi}{2N+2}\right)\delta_s;$$

(see e.g. [75], Example (7.2.5)).

If $h \geq 1$ then all eigenvalues are negative, therefore P_N is the identity on $\mathcal{H}_N$, and zero on the orthocomplement, which sequence obviously converges to the identity of $l^2(\mathbb{Z})$. On the other hand, if $h \leq -1$ then all the P_N's are zero, and so the weak limit is also the zero operator. The resulting states on $\mathcal{A}$ are the Fock state for $h \leq -1$ and the anti-Fock state for $h \geq 1$; the corresponding states on the spin chain are $|\uparrow\rangle\langle\uparrow|^{\otimes\infty}$ (all spins up), and $|\downarrow\rangle\langle\downarrow|^{\otimes\infty}$ (all spins down). Note that we get the same states if we leave the interaction term from the Hamiltonian, that is, for $|h| \geq 1$ the interaction between the spins doesn't conribute to the ground state; the direction of the outer magnetic field determines it completely.

When $h \in (-1,1)$ then $\lambda_k < 0$ if and only if

$$k > \arccos(h)\frac{2N+2}{\pi},$$

and therefore the matrix elements of P_N are

$$(P_N)_{t,s} = \sum_{k=k(N,h)}^{2N+1} \frac{1}{\|v_k\|^2} \sin\left(\frac{k(t+N+1)\,\pi}{2N+2}\right) \sin\left(\frac{k(s+N+1)\,\pi}{2N+2}\right), \tag{C.4}$$

where $k(N,h)$ is the smallest integer that is strictly bigger than $\arccos(h)\frac{2N+2}{\pi}$.

For the normalization factor $\|v_k\|^2$ we have

$$\begin{aligned}
\|v_k\|^2 &= \sum_{s=1}^{2N+1} \sin^2\left(ks\frac{\pi}{2N+2}\right) \\
&= \frac{1}{2}\sum_{s=1}^{2N+1}\left(1-\cos\left(ks\frac{\pi}{N+1}\right)\right) \\
&= N+\frac{1}{2}-\frac{1}{2}\cos(k\pi)\sin\left(\frac{2N+1}{2N+2}k\pi\right)/\sin\left(\frac{k\pi}{2N+2}\right) \\
&= N+1.
\end{aligned}$$

In the third equality we used the identity

$$\sum_{m=1}^{n}\cos(mx) = \cos\left(\frac{n+1}{2}x\right)\sin\left(\frac{n}{2}x\right) big/\sin\left(\frac{1}{2}x\right).$$

Using the identity $\sin(x)\sin(y) = \frac{1}{2}(\cos(x-y)-\cos(x+y))$ we can rewrite formula (C.4) as

$$
\begin{aligned}
(P_N)_{t,s} &= \frac{1}{\pi}\frac{\pi}{2N+2}\sum_{k=k(N,h)}^{2N+1}\cos\left(k(t-s)\frac{\pi}{2N+2}\right) \\
&+ \frac{1}{2N+2}\sum_{k=k(N,h)}^{2N+1}(-1)^k\cos\left(k(t+s)\frac{\pi}{2N+2}\right).
\end{aligned}
$$

The first term in the sum is a Riemannian approximation of the integral $\int_{\arccos h}^{\pi}\frac{1}{\pi}\cos\left((t-s)\,x\right)\ \mathrm{d}x$, and therefore converges to

$$
\begin{cases}
\dfrac{\pi-\arccos h}{\pi} & \text{when } t=s; \\[2ex]
-\dfrac{1}{\pi}\dfrac{\sin\left((t-s)\arccos h\right)}{t-s} & \text{when } t\neq s.
\end{cases}
$$

We denote the second term by Σ, and show that it converges to zero when N goes to infinity. We show it for $k(N,h)$ even; the odd case can be obtained by a trivial modification of the argument. Using the identity $\cos x-\cos y = 2\sin\left(\frac{x+y}{2}\right)\sin\left(\frac{y-x}{2}\right)$ we obtain

$$
\begin{aligned}
\Sigma &= -\frac{1}{2N+2}\sum_{l=k(N,h)/2}^{N}\left[\cos\left(\frac{(2l+1)(t+s)\,\pi}{2N+2}\right)-\cos\left(\frac{2l(t+s)\,\pi}{2N+2}\right)\right] \\
&= \frac{1}{2N+2}\sum_{l=k(N,h)/2}^{N}\sin\left(\frac{4l+1}{2}(t+s)\frac{\pi}{2N+2}\right)\sin\left((t+s)\frac{\pi}{2N+2}\right),
\end{aligned}
$$

therefore

$$
|\Sigma| \leq \frac{N}{2N+2}\left|\sin\left((t+s)\frac{\pi}{2N+2}\right)\right| \xrightarrow[N\to+\infty]{} 0.
$$

$$
\lim_{N\to+\infty}(P_N)_{t,s} =
\begin{cases}
\dfrac{\pi-\arccos h}{\pi} & \text{when } t=s; \\[2ex]
-\dfrac{1}{\pi}\dfrac{\sin\left((t-s)\arccos h\right)}{t-s} & \text{when } t\neq s.
\end{cases}
\tag{C.5}
$$

These are exactly the matrix elements of the symbol Q of the translation-invariant pure quasi-free state ω_Q with defining set $[\arccos h; 2\pi - \arccos h]$ (see section 5.2.2). The weak convergence means that

$$\lim_{N\to+\infty} \varphi_N \left(c^*_{i_1} \dots c^*_{i_n} c_{j_m} \dots c_{j_1}\right) = \omega_Q \left(c^*_{i_1} \dots c^*_{i_n} c_{j_m} \dots c_{j_1}\right),$$

and so the thermodynamical limit of the ground states is $\hat{\omega}_Q$.

C.2

Suppose T is an element of $\mathcal{A}$. The unitary on the one-particle Hilbert space, implementing the parity automorphism is $-I$, therefore

$$\begin{aligned}
\alpha(T) &= (-I)_{\mathrm{F}}\ T\ (-I)_{\mathrm{F}} = (-I)_{\mathrm{F}}\ U_{\mathrm{F}}\ (-I)_{\mathrm{F}} \\
&= \left(\bigoplus_{n\in\mathbb{N}} \wedge^n(-I)\right)\left(\bigoplus_{n\in\mathbb{N}} \wedge^n U\right)\left(\bigoplus_{n\in\mathbb{N}} \wedge^n(-I)\right) \\
&= \bigoplus_{n\in\mathbb{N}} [\wedge^n(-I)]\,[\wedge^n(U)]\,[\wedge^n(-I)] = \bigoplus_{n\in\mathbb{N}} \wedge^n\,[(-I)U(-I)] \\
&= T,
\end{aligned}$$

where U is given in formula (5.30). That is, T is an even element.

Let $\mathcal{A}_0$ denote the linear span of monomials of the form $a^*_{i_1} \dots a^*_{i_n} a_{j_m} \dots a_{j_1}$. Since $\mathcal{A}_0$ is dense in $\mathcal{A}$, we have a sequence t_n in $\mathcal{A}_0$, converging to T. We can assume that t_n is even, since $\lim_{n\to+\infty} \frac{1}{2}\left(t_n + \alpha(t_n)\right) = \frac{1}{2}\left(T + \alpha(T)\right) = T$. Now we fix an N such that $\|t_N - T\| < 1$ holds. Since t_N is a finite linear combination of monomials, there are only finitely many a_k's appearing in it, thus we can pick $K < 0$ such that a_K doesn't appear in any of the monomials of t_N. Since t_N is even, all its monomials is a product of an even number of a_k's, therefore a_K commutes with all of them, and so also with t_N, thus

$$\|[T, a_K]\| = \|[T, a_K] - [t_N, a_K]\| \le 2\,\|t_N - T\| < 2.$$

On the other hand, by (5.29) we have

$$[T, a_K] = 2Ta_K,$$

and so

$$\|[T, a_K]\|^2 = \|2Ta_K\|^2 = 4\,\|a^*_K T^* T a_K\| = 4\,\|a^*_K a_K\| = 4\,\|a_K\|^2 = 4,$$

giving a contradiction.

Samenvatting

In de fysica gebruikt men statistische methodes omwille van de onvolledige informatie die een waarnemer heeft over het te bestuderen systeem. Men kan er zich over verbazen dat er een precieze wiskundige theorie kan opgebouwd worden die de 'onvolledigheid van informatie' kwantificeert in het kader van stochastische modellen: dit is de informatietheorie, ontwikkeld door Shannon [101].

Het belangrijkste begrip in informatietheorie is de entropie van een eindige kansverdeling $\{p_i\}$, gedefinieerd door

$$S\left(\{p_i\}\right) := -\sum_i p_i \log p_i \,. \tag{C.6}$$

De wiskundige uitdrukking voor de entropie van een kwantumsysteem in de toestand beschreven door de dichtheidsoperator ρ werd afgeleid door John von Neumann, aan de hand van een Gedankenexperiment in verband met het gedrag van een ideaal gas, in zijn basiswerk [78]. De uitdrukking

$$S(\rho) = -\mathrm{Tr}\,\rho \log \rho \tag{C.7}$$

die hij bekwam, is een directe uitbreiding van de entropie van Shannon en valt samen met de Shannonentropie van de eigenwaardenverdeling van ρ (we veronderstellen hier een eindigdimensionaal systeem om technische moeilijkheden te vermijden).

De gelijkheid van de thermodynamische en informatietheoretische entropie berust zeker niet op een eenvoudig toeval, maar wijst op een diep verband tussen de twee vakgebieden. De theorie van ensembles kan bv. opgebouwd worden door een maximale onzekerheid over de toestand van het systeem aan te nemen, in de zin dat de informationele entropie gemaximaliseerd wordt, rekening houdend met de beschikbare informatie over het systeem. Zo kan men het microcanonisch ensemble verkrijgen als de kansverdeling die de entropie maximaliseert onder voorwaarde dat de energie constant gehouden wordt. Anderzijds bestaat informatie niet onafhankelijk van het medium waardoor ze gedragen wordt ('Informatie is fysisch', in de woorden van Rolf Landauer), vandaar dat de mogelijkheden om informatie te manipuleren, te verzenden of op te slaan, bepaald worden door de wetten van de natuurkunde.

De kwantuminformatietheorie ontstond in de jaren '70 (vroeg werk is o.a. [41, 68, 45, 49, 50, 57, 52, 28]) en ze heeft de laatste tien jaar een snelle uitbreiding ondergaan. De toegenomen interesse is te danken aan enkele spectaculaire resultaten zoals kwantumalgoritmes die aanzienlijk sneller werken dan hun klassieke tegenhangers [81], dense coding [13, 74], kwantumteleportatie [14] en kwantumcryptografie [11]. Terwijl de theoretische resultaten over kwantumcomputing een doorbraak in de computerwetenschappen beloven, behoort de fysische realisatie van een kwantumcomputer wellicht tot een nog verre toekomst. Daarentegen bevinden kwantumcryptografie en kwantumteleportatie zich in een stadium waarin experimentele realisatie mogelijk is [12, 20, 19, 39, 79]. Los van de mogelijke praktische toepassingen, ligt het belang van de kwantuminformatietheorie in de hernieuwde interesse die ze opgewekt heeft in de studie van de fundamenten van de kwantummechanica en van correlaties in kwantummechanische veeldeeltjessystemen.

De interpretatie van het formalisme van de kwantummechanica is altijd al een punt van discussie geweest en de vraag is nog lang niet afgehandeld [38]. Het lijkt echter duidelijk dat kwantummechanica een intrinsiek probabilistische theorie is en dat de niet-klassieke eigenschappen van kwantumsystemen fundamenteel verwant zijn aan de niet-Kolmogoroviaanse eigenschappen van de onderliggende kanstheorie. De belangrijkste verschillen met de klassieke theorie komen voort uit de beduidend meer ingewikkelde algebraïsche structuur van kwantumtoestanden: in tegenstelling tot het klassieke geval is het niet voldoende om een verzameling gewichten toe te kennen om een toestand van een eindig systeem te beschrijven, maar moet ook een bijbehorende verzameling eigenvectoren gegeven worden, zodat er veel meer vrijheidsgraden geïntroduceerd worden. Deze rijke algebraïsche structuur heeft als gevolg dat een zuivere toestand van een samengesteld kwantumsysteem niet noodzakelijk het product is van zuivere toestanden van de deelsystemen. Dit is de eenvoudigste vorm van het fenomeen 'entanglement', ook wel verstrengeling genoemd.

Terwijl entanglement het leven bemoeilijkt, of juist interessanter maakt, zoals men verkiest, blijkt het ook een fenomeen te zijn dat aan de basis ligt van vele nieuwigheden in de informatieverwerking die door de kwantuminformatietheorie voorspeld worden. Bovendien wijzen experimentele resultaten erop dat macroscopisch meetbare grootheden, zoals magnetische susceptibiliteit, door entanglement beïnvloed worden [40]. Vandaar dat een goed theoretisch begrip van entanglement belangrijk lijkt om het fysisch gedrag van samengestelde kwantumsytemen te doorgronden. Het eerste waar men naar op zoek kan gaan is een goede maat voor de hoeveelheid van entanglement. Het blijkt dat enkele eenvoudige criteria de entropie van de marginalen van de toestand als de enige aanvaardbare maat voor entanglement in een zuivere tweepartijentoestand selecteren

$$\mathcal{E}(|\psi\rangle) = S\left(\mathrm{Tr}_2\left(|\psi\rangle\langle\psi|\right)\right) = S\left(\mathrm{Tr}_1\left(|\psi\rangle\langle\psi|\right)\right) . \tag{C.8}$$

Dit voorbeeld geeft aan dat de entropie niet enkel bruikbaar is als thermodynamische grootheid of maat voor de hoeveelheid informatie, maar ook als een nuttig instrument bij de studie van de algebraïsche structuur van meerpartijentoestanden. Zo kan bv. het zuiver en entangled zijn van een tweepartijentoestand uitgedrukt worden als $S(\rho_{\mathrm{AB}}) = 0$ en $S(\rho_{\mathrm{A}}) > 0$. De onafhankelijkheid van subsystemen is equivalent met het verzadigd zijn van de ongelijkheid $S(\rho_{\mathrm{AB}}) \leq S(\rho_{\mathrm{A}}) + S(\rho_{\mathrm{B}})$, dewelke op haar beurt equivalent is met het opsplitsen van de dichtheidsoperator in een eenvoudig tensorproduct van haar marginalen. Verder blijkt de verzadiging van de ongelijkheid $S(\rho_{\mathrm{ABC}}) + S(\rho_{\mathrm{B}}) \leq S(\rho_{\mathrm{AB}}) + S(\rho_{\mathrm{BC}})$ equivalent te zijn met het Markoviaans zijn, een eigenschap die weer gekarakteriseerd wordt door een specifieke algebraïsche structuur van de driepartijentoestand ρ_{ABC} [44]. In dit proefschrift bestuderen we het gedrag van entropie in enkele concrete fysische modellen, verder de rol van entropie in kwantuminformatieverwerking, te samen met de informatie die de entropie van subsystemen ons kan geven over de algebraïsche structuur van meerpartijenkwantumtoestanden.

Kwantumkanalen

Het wiskundig model voor een kwantumkanaal met ruis bestaat uit een stochastische afbeelding Φ die toestanden van een input-Hilbertruimte $\mathcal{H}_{\mathrm{in}}$ afbeeldt op toestanden van een output-Hilbertruimte $\mathcal{H}_{\mathrm{out}}$. Om klassieke informatie te verzenden door het kanaal codeert de zender de verzameling mogelijke boodschappen $\mathcal{X}$ in kwantumtoestanden $\{\rho_x : x \in \mathcal{X}\}$ van het inputsysteem. Wanneer een boodschap x verzonden wordt, voert de ontvanger een meting $\{M_x : x \in \mathcal{X}\}$ uit op het output-systeem en decodeert hij het bericht als $y \in \mathcal{X}$ indien het resultaat van de meting y bedraagt. Er treedt een fout op wanneer $y \neq x$, de waarschijnlijkheid hiervoor is $\mathrm{Tr}\,\Phi(\rho_x)\,M_y$. Een kanaal kan meerdere maal gebruikt worden: bij n-voudig gebruik worden de in- en output-Hilbertruimten vervangen door $\mathcal{H}_{\mathrm{in}}^{\otimes n}$ en $\mathcal{H}_{\mathrm{out}}^{\otimes n}$, de verzameling te verzenden berichten is $\mathcal{X}_n$. Merk op dat terwijl het product $\Phi^{\otimes n}$ als model gebruikt wordt bij meervoudig gebruik van een kanaal, zowel de coderende kwantumtoestanden als de metingsoperatoren entangled kunnen zijn.

De klassieke capaciteit voor informatieoverdracht van een kanaal is het maximaal aantal verzendbare berichten per gebruik van het kanaal, wanneer de kans op verkeerd decoderen asymptotisch dient de verdwijnen. Holevo, Schumacher en Westmoreland [53, 100] hebben aangetoond dat deze capaciteit gelijk is aan

$$\overline{\chi}\,(\Phi) := \lim_{n\to\infty} \frac{1}{n}\,\chi\left(\Phi^{\otimes n}\right)\,, \tag{C.9}$$

waar

$$\begin{aligned}\chi\left(\Phi^{\otimes n}\right) \;:=\; \sup\Big\{ S\Big(\sum_{x\in\mathcal{X}_n} p_x\,\Phi^{\otimes n}\left(\rho_x\right)\Big) - \sum_{x\in\mathcal{X}_n} p_x\,S\left(\Phi^{\otimes n}\left(\rho_x\right)\right)\;:\;\{\rho_x\} \\ \text{toestanden op } \mathcal{H}_{\mathrm{in}}^{\otimes n};\; \{p_x\} \text{ kansverdeling op } \mathcal{X}_n\Big\}.\end{aligned}$$

De vraag naar de additiviteit van een kwantumkanaal is of $\overline{\chi}\,(\Phi)$ vervangen kan worden door $\chi\,(\Phi)$, met andere woorden: of $\chi\,(\Phi^{\otimes n}) = n\chi\,(\Phi)$, of meer algemeen: geldt

$$\chi(\Phi_1 \otimes \Phi_2) = \chi(\Phi_1) + \chi(\Phi_2) \tag{C.10}$$

voor twee willekeurige kanalen Φ_1 en Φ_2? Het is duidelijk dat additiviteit de evaluatie van de capaciteit van kanalen sterk zou vereenvoudigen. Een andere manier om de vraag te formuleren is of men de informatiedragende capaciteit van een kanaal kan vergroten door verstrengelde dichtheidsoperatoren te gebruiken bij het coderen (het is bekend dat men de capaciteit van bepaalde kanalen kan opdrijven door het gebruik van verstrengelde meetoperatoren). Bovendien heeft men onlangs aangetoond [8, 72, 103, 104] dat additiviteit van de kanaalcapaciteit equivalent is met andere open vragen zoals de additiviteit van 'entanglement of formation', een probleem van centraal belang in de theorie van entanglement. Ook is de vraag equivalent met de addiviteit van de minimale entropie-output

$$S_{\min}(\Phi) := \min\big\{\; S\big(\Phi(|\psi\rangle\langle\psi|)\big)\;:\;\psi \in \mathcal{H}_{\mathrm{in}},\;\|\psi\| = 1\big\}\,, \tag{C.11}$$

die de de hoeveelheid ruis op een kwantumkanaal meet aan de hand van de minimale hoeveelheid onzekerheid die veroorzaakt wordt op een zuivere input.

In Hoofdstuk 3 bestuderen we de zgn. covariante kanalen, die voldoen aan

$$\Phi\left(U^* \rho U\right) = \alpha(U)^* \, \Phi(\rho) \, \alpha(U) \tag{C.12}$$

voor een bepaalde representatie α van $SU(\mathrm{d_{in}})$. Nadat we de structuur van zulke kanalen bestudeerd hebben voor willekeurige representaties α behandelen we de toegevoegde representatie $\alpha(U) = \overline{U}$. We tonen we aan dat de verzameling covariante kanalen samenvalt met de verzameling 'transpose depolarising channels'

$$\Lambda_t(A) := tA^T + (1-t)\,(\mathrm{Tr}\,A)\,\frac{1}{d}\,\mathbb{1}\,, \qquad -\frac{1}{\mathrm{d}-1} \leq t \leq \frac{1}{\mathrm{d}+1}, \tag{C.13}$$

waar $A^{\,T}$ de getransponeerde is van A in dezelfde basis waarin de U's toegevoegd worden. Tenslotte bewijzen we de additiviteit

$$S_{\min}(\Lambda_t \otimes \Lambda_t) = 2S_{\min}(\Lambda_t) \tag{C.14}$$

voor het parameterbereik $-2/(\mathrm{d}^2-2) \leq t$.

Asymptotisch gedrag van entropie op de spinketen

Stel dat φ een translatie-invariante toestand is op de ééndimensionale spinketen. Men kan hier bijvoorbeeld denken aan de thermodynamische limiet van de locale Gibbstoestanden voor een translatie-invariante locale interactie. Voor de beperking φ_N tot N opeenvolgende spins, kunnen we de ε-typische deelruimte $\mathcal{H}_{N,\varepsilon}$ definiëren als de deelruimte met minimale dimensie zó dat het spoor van de beperking van φ_N tot $\mathcal{H}_{N,\varepsilon}$ groter is dan $1-\varepsilon$. Met andere woorden: voor ε 'klein genoeg', verschilt de drager van φ_N niet beduidend van $\mathcal{H}_{N,\varepsilon}$. Typische deelruimten zijn belangrijk bij numerieke berekeningen in de veronderstelling dat hun dimensies t_N trager toenemen dan die van de hele dragers. Men kan verwachten dat $S(\varphi_N)$ en t_N verwant zijn aan elkaar: een grote t_N wijst op een verspreid spectrum, net zoals een grote entropie, terwijl een goed geconcentreerd spectrum een kleine t_N en een kleine entropie tot gevolg heeft. Men heeft inderdaad aangetoond [46, 15] dat voor een grote klasse van toestanden met positieve gemiddelde entropie $s(\varphi) := \lim_{N\to\infty} \frac{1}{N} S(\varphi_N)$, de grootte van de typische deelruimten toeneemt volgens

$$t_N \sim e^{S(\varphi_N)}, \tag{C.15}$$

waarbij $S(\varphi_N)$ op lineaire wijze toeneemt: $S(\varphi_N) \sim N\, s(\varphi)$.

Voor zuivere toestanden, zoals de thermodynamische limiet van de grondtoestand van een locale translatie-invariante interactie, is de gemiddelde entropie nul in het algemeen en is er geen exact verband gekend tussen het asymptotisch gedrag van de entropie en de grootte van de typische deelruimten. De entropieën $S(\varphi_N)$ kan men beschouwen als een maat voor de entanglement tussen een eindig deel van de keten en de rest. Een divergerende reeks entropieën wijst in bepaalde modellen op kritisch gedrag [110].

In hoofdstuk 7 bestuderen we het asymptotisch gedrag van de entropie voor de klasse van zuivere quasi-vrije toestanden, gekarakteriseerd door een deelverzameling K van het interval $[0, 2\pi]$. We laten zien dat wanneer K een unie is van een eindig aantal disjuncte intervallen (bv. de grondtoestand van het XX-model, die overeenkomt met één enkel interval), de entropie toeneemt als

$$c_1 \log N \leq S_N \leq c_2 \log^2 N, \tag{C.16}$$

waar c_1 en c_2 constanten zijn. Numerieke berekeningen tonen een logaritmisch gedrag en bovendien lijkt de grootte van de typische deelruimten lineair toe te nemen, wat er op wijst dat (1.10) waarschijnlijk geldig blijft in dit geval.

Door gebruik te maken van toestanden met een meer ingewikkelde karakteriserende verzameling K, kan men sneller asymptotisch gedrag bekomen. Als K een verzameling van het Cantortype is, waarbij in de n-de stap van de constructie 2^{n-1} intervallen met lengte aq^n verwijderd worden, dan kan het asymptotisch gedrag

$$S_N \sim N^\alpha \tag{C.17}$$

aangetoond worden, met $\alpha = -\frac{\log 2}{\log q}$.

Kwantumdatacompressie

De problemen van de kwantuminformatietheorie worden in twee verwante categorieën ingedeeld. De eerste soort vraag die men zich kan stellen gaat over de efficiëntie van informatieverwerking wanneer een kwantummechanisch medium gebruikt wordt om klassieke informatie te verzenden; problemen verwant aan de klassieke capaciteit van een kwantumkanaal vallen hieronder. Anderzijds kan men vragen wat er gebeurt met de informatie die gedragen wordt door een kwantumtoestand, wanneer deze toestand kwantumoperaties ondergaat. Het 'Schumacher source coding' theorema [99] toont aan dat de entropie van een kwantumtoestand kan beschouwd worden als een maat voor de informatie-inhoud, net zoals in het klassieke geval. Dit wil zeggen dat de entropie aangeeft hoever we een kwantumtoestand kunnen comprimeren wanneer we een groot aantal copiën beschouwen en de decompressie alleen asymptotisch perfect hoeft te zijn (wanneer het aantal copiën naar oneindig gaat).

In hoofdstuk 6.1 laten we zien dat in meer algemene systemen de gemiddelde entropie dezelfde rol speelt als de asymptotische comprimeerbaarheid, namelijk voor ergodische toestanden van de spinketen. We kunnen ze dan ook weer interpreteren als de gemiddelde informatie-inhoud van de toestand. Schumachers theorema kadert in een model waar de klassieke informatie in een kwantummechanisch systeem opgeslagen wordt en waar nadien de kwantummechanische toestand gecomprimeerd wordt. De opeenvolgende letters van een lange boodschap worden elk om beurt in een kwantumtoestand gecodeerd van een systeem met Hilbertruimte $\mathcal{H}$, zodat de hele boodschap vertaald wordt in een producttoestand van $\mathcal{H}^{\otimes n}$. Als we aannemen dat er een kansverdeling, de 'taalstatistiek', bestaat op de verzameling van de mogelijke boodschappen dan kunnen we ook het statistische gemiddelde van de gecodeerde boodschappen beschouwen. De taalstatistiek is een productmaat indien er geen correlaties tussen opeenvolgende letters bestaan, in dit geval zal de kansmaat na coderen ook een productmaat zijn. Het theorema van Schumacher kadert precies deze situatie. In meer realistische situaties bestaan er wel degelijk correlaties in de taalstatistiek en daarom verkrijgt men ook niet langer een productmaat op de gecodeerde boodschappen. In 6.1 bespreken we zulk een model van een kwantuminformatiebron, waaruit een meer algemene (ergodische) toestand voorkomt waarop het datacompressietheorema van toepassing is.

Monotoniciteit van de relatieve entropie

Positieve gemiddelde entropie betekent een asymptotisch lineaire toename van de reeks entropieën S_N voor een translatie-invariante toestand op de spinketen. De toename is exact lineair als en slechts als $S_{N+1} - S_N = S_2 - S_1$ voor alle N, m.a.w. als de sterke subadditiviteitsongelijkheid (SSA), die geldt voor elke translatie-invariante toestand van de spinketen, verzadigd wordt. Sterke subadditiviteit is een karakteriserende eigenschap van driepartijentoestanden. Voor de karakterisatie van lineaire entropietoename dient men dus het geval te bestuderen waarbij de SSA-gelijkheid bereikt wordt voor een algemene driepartijentoestand. Deze vraag werd reeds behandeld in o.a. [91, 96] en in [44] werd een nodige en voldoende voorwaarde gegeven voor de algebraïsche structuur van toestanden die aan de SSA-gelijkheid voldoen. Hier behandelen we een meer algemeen probleem: de algebraïsche karakterisatie van toestanden die de monotoniciteitsongelijkheid van de relatieve entropie verzadigen. Hieruit volgen de resultaten van [44] als een speciaal geval.

De relatieve entropie van een toestand ρ t.o.v. een toestand σ wordt gegeven door

$$S(\rho, \sigma) := \mathrm{Tr}\, \rho\, (\log \rho - \log \sigma)\ , \tag{C.18}$$

en ze kan geïnterpreteerd worden als een maat voor de onderscheidbaarheid van twee toestanden door metingen op een groot aantal copiën van het systeem [46, 18]. Via deze interpretatie wordt op heuristische wijze duidelijk dat kwantumoperaties de relatieve entropie zouden moeten verminderen. Dit noemt men de monotoniciteitseigenschap van de relatieve entropie:

$$S\left(\Phi\left(\rho\right), \Phi\left(\sigma\right)\right) \leq S(\rho, \sigma)\,. \tag{C.19}$$

Alle gangbare entropieongelijkheden, o.a. (SSA), kunnen uit de monotoniciteitseigenschap afgeleid worden, vandaar dat een karakterisatie van de verzadiging ervan tot een karakterisatie leidt van andere entropieongelijkheden.

In [44] werd ook aangetoond dat verzadiging van (SSA) equivalent is met het Markoviaans zijn van een driepartijentoestand, waar het Markoviaans zijn van kwantumtoestanden een algebraïsche uitbreiding is van het begrip Markoviaans voor klassieke maten. Deze karakterisatie kan uitgebreid worden naar toestanden op de spinketen met een exact lineaire entropietoename. De structurele karakterisatie van Markoviaanse driepartijentoestanden zorgt ervoor dat Markoviaanse toestanden op de spinketen gespecificeerd kunnen worden aan de hand van een klassieke Markoviaanse maat en een eindige verzameling kwantumtoestanden. Dit levert op zijn beurt een expliciete uitdrukking voor de gemiddelde entropie van zulke toestanden.

Opbouw van het proefschrift

Omwille van de volledigheid hebben we twee inleidende **Hoofdstukken (2 en 5)** en twee **Appendices (A en B)** bijgevoegd. Het materiaal dat in deze onderdelen besproken wordt, kan men terugvinden in de literatuur, echter wel verspreid over meerdere boeken en artikels uit verschillende vakgebieden, elk met hun eigen doelstellingen en terminologie. Vandaar dat het noodzakelijk leek om deze materie samen te vatten in een gestructureerd geheel, aangepast aan de doelstellingen van het proefschrift. In onze bespreking gebruiken we steeds het C^*-algebraïsche formalisme van de kwantummechanica en dit om twee redenen. Enerzijds omdat het ons toelaat om klassieke en kwantummechanische systemen met eenzelfde formalisme te beschrijven, anderzijds omdat we zo op meer eenvoudige wijze een beschrijving kunnen geven van oneindig uitgestrekte systemen, zoals de kwantumspinketen of de CAR-algebra. Hiertoe vatten we enkele basisbegrippen en resultaten uit de theorie van C^*-algebra's samen in **Appendix A**, waarbij we vooral aandacht besteden aan eindigdimensionale systemen. De argumentatie in **Hoofdstuk** 4 steunt op enkele gevorderde resultaten uit de theorie van algebra's van operatoren, die we eveneens bespreken, hoofdzakelijk in **Appendix A.5**. **Appendix A** verschaft ook de noodzakelijke terminologie om de postulaten van de kwantummechanica weer te geven in **Appendix B**, zoals we ze verder gebruiken. In **Appendix B** tonen we ook hoe deze terminologie verwant is aan de gebruikelijke manier van formuleren in de standaardwerken over kwantummechanica. **Appendix C** bevat enkele berekeningen die weggelaten zijn uit de hoofdtekst maar het proefschrift meer volledig maken.

De overige vier **Hoofdstukken** 3, 4, 6 en 7 bevatten de resultaten die verkregen werden in het onderzoek dat tot dit proefschrift geleid heeft. Ze bestaan telkens uit een wat uitgebreidere en vernieuwde versie van de artikels [36], [77], [93] en [35]. In **Hoofdstuk 4** geven we een meer doorzichtige versie van de structurele karakterisatie van 'saturating triples', o.a. de structuur van de stochastische afbeelding Φ. **Sectie 4.1** geeft een overzicht van de toepassingen van het structurele karakterisatietheorema, vooral gebaseerd op de resultaten van [44]. Recente resultaten bieden de mogelijkheid om het 'source coding' theorema in een sterkere vorm te gieten in **Hoofdstuk 6** dan in het oorspronkelijk werk [93], met slechts een kleine aanpassing van de oorspronkelijke argumentatie. Het model voor een kwantuminformatiebron is eveneens extra ten opzichte van [93] en het gaat hier, voor zover de kennis van de auteur reikt, om een nieuw idee. Tenslotte wordt in **Hoofdstuk 7** het resultaat van een numerieke berekening i.v.m. de grootte van typische deelruimten toegevoegd aan de inhoud van [35].

Bibliography

[1] L. Accardi, C. Cecchini: Conditional expectations in von Neumann algebras and a theorem of Takesaki; *J. Funct. Anal.* **45**, 245–273, (1982)

[2] L. Accardi, V. Liebscher: Markovian KMS-states for one-dimensional spin chains; *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **2**, 645–661, (1999)

[3] R. Alicki, M. Fannes : Quantum Dynamical Systems; Oxford University Press, (2001)

[4] R. Alicki and M. Fannes: Note on multiple additivity of minimal Renyi entropy output of the Werner-Holevo channels; *Open systems and Information Dynamics* **11**, 1–4, (2004)

[5] G.G. Amosov, A.S. Holevo and R.F. Werner: On some additivity problems in quantum information theory; *Problems in Information Transmission* **36**, 305–313, (2000)

[6] H. Araki: On the XY-model on two-sided infinite chain; *Publications of the Research Institute for Mathematical Sciences of Kyoto* **20**, 277–296, (1984)

[7] A. Aspect, P. Grangier, G. Roger: Experimental Realization of the Einstein-Podolsky-Rosen-Bohm *Gedankenexperiment*: A New Violation of Bell's Inequalities; *Phys. Rev. Lett.* **49**, 91–94, (1982)

[8] K.M.R. Audenaert and S.L. Braunstein: On strong superadditivity of the entanglement of formation; *Communications in Mathematical Physics* **246**, 443–452, (2004)

[9] H. Barnum, E. Knill: Reversing quantum dynamics with near optimal quantum and classical fidelity; *J. Math. Phys.* **43**, 2097–2106, (2002)

[10] E. Barouch, B.M. McCoy: Statistical mechanics of the XY-model II-Spin correlation functions.; *Phys. Rev. A* **3**, 786–804, (1971)

[11] C.H. Bennett, G. Brassard: Quantum cryptography: Public key distribution and coin tossing; *Proc. of the IEEE Inf. Conf. on Computers, Systems and Signal Processing*, 175–179, (1984)

[12] C.H. Bennett, F. Bessette, G. Brassard, L. Salvail and J. Smolin: Experimental Quantum Cryptography; *J. of Cryptology* **5**, (1992)

[13] C.H.Bennett, S.J. Wiesner: Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states; *Phys. Rev. Lett.* **69**, 2881–2884, (1992)

[14] C.H. Bennett, G. Brassard, C. Crépau, R. Jozsa, A. Peres and W. Wootters: Teleporting an unknown quantum state via dual classical and EPR channel; *Phys. Rev. Lett.* **70**, 1895–1899, (1993)

[15] I. Bjelakovic, T. Krueger, R. Siegmund-Schultze, A. Szkola: The Shannon-McMillan Theorem for Ergodic Quantum Lattice Systems; *Invent. Math.* **155**, 203–222, (2004)

[16] T. Bjelakovic, T. Krueger, R. Siegmund-Schultze, A. Szkola: Chained Typical Subspaces - a Quantum Version of Breiman's Theorem; *quant-ph/0301177*

[17] I. Bjelakovic, A. Szkola: The Data Compression Theorem for Ergodic Quantum Information Sources; *quant-ph/0301043*; to appear in *Quantum Information Processing*

[18] I. Bjelakovic, R. Siegmund-Schultze: An ergodic theorem for the quantum relative entropy; *quant-ph/0306094*

[19] D. Boschi, S. Branca, F.D. Martini, L. Hardy and S. Popescu: Experimental realization of teleporting an unknown pure quantum state via dual classical and Einstein-Podolsky-Rosen channels; *Phys. Rev. Lett.* **80**, 1121–1125, (1998)

[20] D. Bouwmeester, J.W. Pan, K. Mattle, M. Eibl, H. Weinfurter and A. Zeilinger: Experimental quantum teleportation; *Nature* **390**, 575–579, (1997)

[21] O. Bratteli, D.W. Robinson: Operator algebras and quantum statistical mechanics I.; Springer, (1979)

[22] O. Bratteli, D.W. Robinson: Operator algebras and quantum statistical mechanics II.; Springer, (1981)

[23] M.D. Choi: Completely positive linear maps on complex matrices; *Linear Algebra and its Applications* **10**, 285–290, (1975)

[24] I. Csiszár, J. Körner: Information theory. Coding theorems for discrete memoryless systems; Akadémiai Kiadó, Budapest, 1981

[25] N. Datta, A.S. Holevo and Y.M. Suhov: A quantum channel with additive minimum output entropy; *quant-ph/0403072*

[26] N. Datta, A.S. Holevo and Y.M. Suhov: Additivity for transpose depolarizing channels; *quant-ph/0412034*

[27] Kenneth R. Davidson: C^*-algebras by example; American Mathematical Society, (2001)

[28] E.B. Davies: Quantum communication systems; *IEEE Trans. Inform. Theory* **23**, 530–534, (1977)

[29] M.J. Donald, M. Horodecki, O. Rudolph: The uniqueness theorem for entanglement measures; *J. Math. Phys.* **43**, 4252–4272, (2002)

[30] D.E. Evans, Y. Kawahigashi: Quantum Symmetries on Operator Algebras; Oxford University Press, Oxford, (1998)

[31] M. Fannes: A continuity property of the entropy density for spin lattice systems; *Communications in Mathematical Physics* **31**, 291–294.

[32] M. Fannes: Canonical commutation and anticommutation relations; *Quantum Probability Communications XI*, World Scientific, Singapore, 171–198, (2003)

[33] M. Fannes, B. Nachtergaele, R.F. Werner: Finitely Correlated States on Quantum Spin Chains; *Communications in Mathematical Physics* **144**, 443–490, (1992)

[34] M. Fannes, B. Nachtergaele, R.F. Werner: Finitely correlated pure states; *Journal of Functional Analysis* **120**, 511–534, (1994)

[35] M. Fannes, B. Haegeman, M. Mosonyi: Entropy growth of shift-invariant states on a quantum spin chain; *Journal of Mathematical Physics* **44**, 6005–6019, (2003)

[36] M. Fannes, B. Haegeman, M. Mosonyi, D. Vanpeteghem: Additivity of minimal entropy output for a class of covariant channels; *quant-ph/0410195*

[37] R. Feynman: Quantum mechanical computers; *Optics News* **11**, 11–20, (1985)

[38] C.A. Fuchs: Quantum Mechanics as Quantum Information (and only a little bit more); *quant-ph/0205039*

[39] A. Furusawa, J.L. Sorensen, S.L. Braunstein, C.A. Fuchs, H.J. Kimble and E.S. Polzik: Unconditional quantum teleportation; *Science* **282**, 706–709, (1998)

[40] S. Ghosh, T.F. Rosenbaum, G. Aeppli, S.N. Coppersmith: Entangled quantum state of magnetic dipoles; *Nature* **425**, 48–51, (2003)

[41] J.P. Gordon: Quantum effects in communication systems; *Proc. IRE* **50**, 1898-1908, (1962)

[42] W.T. Grandy, Jr.: Foundations of Statistical Mechanics. Volume I: Equilibrium Theory; D. Reidel, Dordrecht, 1987

[43] U. Grenander, G. Szegö: Toeplitz Forms and their Applications; University of California Press, (1958)

[44] P. Hayden, R. Jozsa, D. Petz, A. Winter: Structure of states which satisfy strong subadditivity of quantum entropy with equality; *Commun. Math. Phys.* **246**, 359–374, (2004)

[45] C.W. Helstrom, J. Liu, J.P. Gordon: Quantum-mechanical communication theory; *Proc. IEEE* **58**, 1578–1598, (1970)

[46] F. Hiai, D. Petz: The proper formula for relative entropy and its asymptotics in quantum probability; *Commun. Math. Phys.* **143**, 99-114, (1991)

[47] F. Hiai, D. Petz: Entropy density for algebraic states; *J. Functional Anal.* **125**, 287–308, (1994)

[48] F. Hiai, D. Petz: Entropy densities for Gibbs states of quantum spin systems; *Rev. Math. Phys.* **5**, 693–712, (1994)

[49] A.S. Holevo: Towards the mathematical theory of quantum communication channels *Probl. Inform. Transm.* **8**, 63–71, (1972) (Russian)

[50] A.S. Holevo: Information theoretic aspects of quantum measurement *Probl. Inform. Transm.* **9**, 31–42, (1973) (Russian)

[51] A.S. Holevo: Some estimates for the amount of information transmittable by a quantum communication channel; *Problems Inf. Transmission*, **9**, 177–183, (1973) (Russian)

[52] A.S. Holevo: Problems in the mathematical theory of quantum communication channels; *Rep. Math. Phys.* **12**, 273–278, (1977)

[53] A.S. Holevo: The Capacity of the Quantum Channel with General Signal States; *IEEE Trans. Inform. Theory* **44**, 269–273, (1998)

[54] A.S. Holevo: Quantum coding theorems; *Russian Math. Surveys,* **53**, 1295–1331, (1998)

[55] A.S. Holevo: Remarks on the classical capacity of quantum channel; *quant-ph/0212025*

[56] M. Horodecki, P. Horodecki, R. Horodecki: Mixed-state entanglement and quantum communication; in *Quantum Information: An introduction to Basic Theoretical Concepts and Experiments* edited by G. Alber, T. Beth, M. Horodecki, P. Horodecki, R. Horodecki,M. Rotteler, H. Weinfurter, R. Werner and A. Zeilinger; Springer Tracts in Modern Physics, (2001)

[57] R. Ingarden: Quantum information theory; *Rep. Math. Phys.* **10**, 43–72, (1976)

[58] A. Jamiolkowski: Linear transformations which preserve trace and positive semidefiniteness of operators; *Rep. Math. Phys.* **3**, 275–278, (1972)

[59] B.Q. Jin, V.E. Korepin: Quantum spin chain, Toeplitz determinants and Fisher–Hartwig conjecture, *Journal of Statistical Physics* **116**, 79–95, (2004)

[60] H.F. Jones: Groups, Representations and Physics, Adam Hilger, Bristol, (1990)

[61] *Journal of Mathematical Physics*, vol. 43, issue 9, Special Issue: Quantum Information Theory; (2002)

[62] L.B.W. Jolley: Summation of series; 2nd edition, Dover Publications, Inc. (1961)

[63] R. Jozsa, B. Schumacher: A new proof of the quantum noiseless coding theorem; *Journal of Modern Optics* **41**, 2343–2349, (1994)

[64] C.King: Additivity for unital qubit channels; *Journal of Mathematical Physics* **43**, 4641–4653, (2002)

[65] C. King: The capacity of the depolarizing channel; *IEEE Transactions on Information Theory* **49**, 221–229, (2003)

[66] C. King: Maximal p-norms of entanglement breaking channels; *Quantum Information and Computation* **3**, 186–190, (2003)

[67] M. Koashi N. Imoto: Operations that do not disturb partially known quantum states; *Phys. Rev. A* **66**, 022318, (2002)

[68] L.B. Levitin: On the quantum measure of information; *Proc. 4th All-Union Conference on Information Theory*, 111–115, Taschkent, (1969); English version: *Ann. Fond. L. de Broglie* **21**, 345–348, (1996)

[69] E. H. Lieb, M.B. Ruskai: Proof of the strong subadditivity of quantum mechanical entropy; *J. Math. Phys.* **14**, 1938–1941, (1973)

[70] E. Lieb, P. Schultz, D. Mattis: Two soluble models of an antiferromagnetic chain; *Annals of Physics* **16**, 407–466, (1961)

[71] G. Lindblad: Completely Positive Maps and Entropy Inequalities; *Commun. Math. Phys.* **40**, 147–151, (1975)

[72] K. Matsumoto, T. Shimono and A. Winter: Remarks on additivity of the Holevo channel capacity and of the entanglement of formation; *Communications in Mathematical Physics* **246**, 427–442 (2004)

[73] K. Matsumoto and F. Yura: Entanglement cost of antisymmetric states and additivity of capacity of some quantum channels; *Journal of Physics A* **37**, L167–L171, (2004)

[74] K. Mattle, H. Weinfurter, P.G. Kwiat and A. Zeilinger: Dense coding in experimental quantum communication; *Phys. Rev. Lett.* **76**, 4656–4659, (1996)

[75] C.D. Meyer: Matrix analysis and applied linear algebra; Society for Industrial and Applied Mathematics, (2000)

[76] G. Mitchison, R. Jozsa: Towards a geometrical interpretation of quantum information compression; *quant-ph/0309177*

[77] M. Mosonyi, D. Petz: Structure of Sufficient Coarse-Grainings; *Letters in Mathematical Physics* **68**, 19–30, (2004)

[78] John von Neumann: Mathematische Grundlagen der Quantenmechanik; Springer, Berlin, (1932); English translation: Mathematical Foundations of Quantum Mechanics by Robert T. Beyer; Princeton University Press, (1955)

[79] M.A. Nielsen, E. Knill, R. Laflamme: Complete quantum teleportation using nuclear magnetic resonance; *Nature* **396**, 52–55, (1998)

[80] M.A. Nielsen: Probability distributions consistent with a mixed state; *Physical Review A* **61**, (2000)

[81] M.A. Nielsen, I.L. Chuang: Quantum Computation and Quantum Information; Cambridge University Press, (2000)

[82] H. Ohno: Translation-invariant quantum Markov states; to be published in *Interdisc. Inf. Sci.*

[83] M. Ohya, D. Petz: Quantum Entropy and Its Use; Springer-Verlag, Heidelberg, (1993)

[84] M. Ohya, D. Petz, N. Watanabe: On capacities of quantum channels; *Prob. Math. Stat.* **17**, 179–196, (1997)

[85] A. Peres: Separability criterion for density matrices; *Physical Review Letters* **77**, 1413–1415, (1996)

[86] I. Peschel, X. Wang, M. Kaulke, K. Hallberg (Eds.): Density Matrix Renormalization, LNP **528**, Springer, Berlin, Heidelberg, (1999)

[87] D. Petz: A dual in von Neumann algebras; *Quart. J. Math. Oxford* **35**, 475–483, (1984)

[88] D. Petz: Sufficiency of channels over von Neumann algebras; *Quart. J. Math. Oxford* **39**, 907–1008, (1988)

[89] D. Petz: Characterization of sufficient observation channels; in *Mathematical Methods in Statistical Mechanics*, 167–178, Leuven University Press, (1989)

[90] D. Petz: Entropy of Markov states; *Riv. di Math. Pura ed Appl.* **14**, 33–42, (1994)

[91] D. Petz: Monotonicity of quantum relative entropy revisited; *Rev. Math. Physics.* **15**, 79–91, (2003)

[92] D. Petz: Entropy density in quantum statistical mechanics and information theory; in *Contributions in Probability*, ed. C. Cecchini, 221–226, Forum, Udine, (1996)

[93] D. Petz, M. Mosonyi: Stationary quantum source coding; *Journal of Mathematical Physics*, volume 42, number 10, 4257–4264, (2001)

[94] P. Pfeuty: The one dimensional Ising model with a transverse field; *Annals of Physics* **57**, 79–90, (1970)

[95] M. Reed, B. Simon: Methods of modern mathematical physics, I: Functional analysis; Academic Press, (1972)

[96] M. B. Ruskai: Inequalities for quantum entropy: A review with conditions with equality; *J. Math. Phys.* **43**, 4358–4375, (2002)

[97] Erwin Schrödinger: Discussion of probability distributions between separated systems; *Proc. Camb. Phil. Soc.* **31**, (1935)

[98] Erwin Schrödinger: Probability relations between separated systems; *Proc. Camb. Phil. Soc.* **32**, (1936)

[99] B. Schumacher: Quantum coding; *Phys. Rev. A* **51**, 2738–2747, (1995)

[100] B. Schumacher, M.D. Westmoreland: Sending classical information via noisy quantum channels; *Physical Review A* **56**, 131–138, (1997)

[101] C.E. Shannon: A mathematical theory of communication; *The Bell System Technical Journal* **27**, (1948)

[102] P.W. Shor: Additivity of the classical capacity of entanglement-breaking quantum channels; *Journal of Mathematical Physics* **43**, 4334–4340, (2002)

[103] P.W. Shor: Capacities of quantum channels and how to find them; *Mathematical Programming* **97**, 311–335, (2003)

[104] P.W. Shor: Equivalence of additivity questions in quantum information theory; *Communications in Mathematical Physics* **246**, 453–472 (2004)

[105] S. Strătilă: Modular theory in operator algebras; Abacuss Press, Tunbridge Wells, 1981.

[106] M. Takesaki: Conditional expectation in von Neumann algebra; *Journal of Functional Analysis* **9**, 306–321, (1972)

[107] A. Uhlmann: The 'transition probability' in the state space of a $*$-algebra; *Rep. Math. Phys* **9**, 273–279, (1976)

[108] A. Uhlmann: Relative Entropy and the Wigner-Yanase-Dyson-Lieb Concavity in an Interpolation Theory; *Commun. Math. Phys.* **54**, 21–32, (1977)

[109] V. Vedral, M.B. Plenio: Entanglement Measures and Purification Procedures; *Phys. Rev. A* **57**, 1619–1633, (1998)

[110] G. Vidal, J.I. Latorre, E. Rico and A. Kitaev: Entanglement in quantum critical phenomena *Phys. Rev. Lett.* **90**, 227902, (2003)

[111] K.G.H. Vollbrecht and R.F. Werner: Entanglement measures under symmetry; *Phys. Rev. A* **64**, 062307, (2001)

[112] R.F. Werner, M.M. Wolf: Bell inequalities and entanglement; *Quantum Information & Computation* **1**, (2001)