Additivity properties of Quantum Maps

Author: Ana Kontrec

Supervisor: prof. dr hab. Karol Życzkowski

Krakow, June 2014
## Contents

Introduction 2

Chapter 1. Quantum maps 3
   1.1. Basic notions of quantum mechanics 3
   1.2. Properties of quantum maps 4
   1.3. Environmental and operator-sum representations 5
   1.4. Bloch sphere representation 7
   1.5. Examples of single qubit maps 8

Chapter 2. Measures of performance for quantum channels 11
   2.1. Capacity of a channel 11
   2.2. Minimum output entropy 13
   2.3. Maximal output $p$-norm 14
   2.4. Additivity conjectures 15

Chapter 3. Results about additivity 16
   3.1. Depolarizing channel and generalizations 16
   3.2. Weyl-covariant channels 19

Summary and open questions 26

Bibliography 27
Introduction

Quantum information theory is an interdisciplinary field of science, incorporating mathematical techniques and quantum mechanical framework into information-processing tasks. In the last two decades great progress has been made, thus providing a novel perspective on quantum theory and generating new questions in mathematics.

In the first chapter we present the background required to understand and work with quantum maps, along with some classical results and examples. We discuss in more detail the single qubit maps. This may seem like a very special case, but they are of utmost importance in quantum information processing - since the qubit (much like a bit, in the classical realm) is the basic indivisible unit of quantum information.

After that, we introduce some measures of performance for quantum maps (also called quantum channels), that is, different ways of measuring how well a given channel does the job of transporting information. Although the statement of the problem appears to belong purely to information science, many interesting mathematical questions arise.

Additivity conjectures for capacity of quantum channels state that the capacity cannot be increased by using entangled inputs - that is, inputs which possess non-classical correlations. The fact that they have been proven to be false illustrates the deep difference between quantum and classical information, and quantum and classical world in general. We present several examples of quantum maps for which additivity does hold, and establish additivity for certain classes of Weyl-covariant maps.
CHAPTER 1

Quantum maps

1.1. Basic notions of quantum mechanics

1.1.1. States. To any isolated physical system there is an associated Hilbert space, called the state space of the system. Pure states of the system are given by unit norm vectors in the Hilbert space $\mathcal{H}$. If the state space is finite dimensional, then it is simply the vector space $\mathbb{C}^d$ equipped with the standard scalar product.

In quantum mechanics, the Dirac bra-ket notation is used: a vector in the Hilbert space $\mathcal{H}$ is denoted by the ket $|\psi\rangle$, and its dual vector is the bra $\langle \psi |$.

The set of all states consists of both pure and mixed states. If we write the pure states as projectors $\rho = |\psi\rangle \langle \psi |$, we can define the mixed states as convex combinations of $k$ pure states:

$$\rho = \sum_{j=1}^{k} a_j |\psi_j\rangle \langle \psi_j |,$$

where $\sum_{j=1}^{k} a_j = 1$. The operator $\rho$ is called a density matrix. The term “state” is used both to denote a vector and the associated density matrix.

A density matrix $\rho \in M_d$ has the following properties:

- it is Hermitian and positive semidefinite: $\rho \geq 0$,
- it has unit trace: $\text{Tr}(\rho) = 1$.

1.1.2. Composite systems and separability. The state space of a composite physical system is the tensor product of the state spaces of the component physical systems. That is, if two systems $A$ and $B$ have associated state spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, then the state space of the composite system is given by $\mathcal{H}_A \otimes \mathcal{H}_B$.

A pure state $\rho_{AB} = |\psi_{AB}\rangle \langle \psi_{AB} |$ of the composite system, for which $|\psi_{AB}\rangle$ can be expressed in the product form:

$$|\psi_{AB}\rangle = |\phi_A\rangle \otimes |\varphi_B\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$$
is called a *separable* state. A pure state which cannot be expressed in the above form is said to be *entangled*.

A mixed state $\rho_{AB}$ is called *separable* if it can be represented as a convex sum of product states,

$$\rho = \sum_{j=1}^{k} a_j \rho_j^A \otimes \rho_j^B,$$

where $\rho_j^A$ acts in $\mathcal{H}_A$ and $\rho_j^B$ acts in $\mathcal{H}_B$, and $\sum_{j=1}^{k} a_j = 1$. A mixed state which is not separable is called *entangled*.

### 1.2. Properties of quantum maps

According to the postulates of quantum mechanics, the dynamics of *closed* quantum systems (those which do not interact with their environment) is given by unitary transformations. However, since real-world systems are rarely perfectly isolated, this leads us to the study of *open* quantum systems - those which do interact with the outside world. In the quantum information processing context, this interaction manifests as noise in the system.

The dynamics of open quantum systems is described using the mathematical formalism of quantum operations. Most generally, a quantum operation is a map $\Phi : M_n \to M_m$ which transforms density matrices to density matrices. In order for this transformation to be physically realizable, it has to satisfy certain conditions:

(i): Linearity: $\Phi(a \rho_1 + a \rho_2) = a \Phi(\rho_1) + b \Phi(\rho_2)$

(ii): Trace-preserving (TP): $\text{Tr}(\Phi(\rho)) = \text{Tr}(\rho)$

(iii): Positivity: if $\rho \geq 0$, then $\Phi(\rho) \geq 0$.

These criteria ensure that the image under $\Phi$ of a quantum state is indeed a quantum state. However, this is not enough: since any quantum state $\rho \in \mathcal{H}_A$ can be extended to the tensor product $\rho \otimes \sigma \in \mathcal{H}_A \otimes \mathcal{H}_{\text{env}}$ (by adding an ancilla), we have to check that the map $\Phi$ takes density matrices of this joint system again to density matrices, i.e. that the map $\Phi \otimes I$ is positive. This gives us the last requirement for the map $\Phi$:

(iv): Completely positive (CP): a linear map $\Phi : M_n \to M_m$ is called **completely positive** if

$$\Phi \otimes I_k : M_n \otimes M_k \to M_m \otimes M_k$$

is positivity preserving for every $k \geq 1$. 
1.3. Environmental and operator-sum representations

**Definition 1.** A linear completely positive trace-preserving map (CPTP) \( \Phi : M_n \rightarrow M_m \) between quantum states is called a **quantum operation** (or a quantum channel).

**Example 2.** An example of a map which is positive but not completely positive is the transpose map. Suppose we have a pair of qubits in the entangled state
\[
|\psi^+\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}.
\]
After applying the transpose \( T \) on the first qubit and identity on the second, the density matrix of the system, \( \rho = |\psi^+\rangle \langle \psi^+| \), becomes
\[
(1.2.1) \quad \rho' = (T \otimes I)(\rho) = \frac{1}{2} \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix},
\]
which has eigenvalues \( \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \text{ and } -\frac{1}{2} \), and hence is not positive definite and thus not a density matrix.

1.3. Environmental and operator-sum representations

Another way to describe the dynamics of open quantum systems is to extend the principal system with an auxiliary system (which represents the environment), so that they form together a closed system which evolves unitarily. We will assume that the combined system starts out in a product state, \( \rho \otimes \rho_{\text{env}} \).

After adding the ancilla, undergoing an unitary transformation, and removing the ancilla, the initial state \( \rho \) of the principal system is changed to
\[
(1.3.1) \quad \Phi(\rho) = \text{Tr}_{\text{env}} \left[ U (\rho \otimes \rho_{\text{env}}) U^* \right],
\]
where \( \text{Tr}_A X := \sum_i \langle e_i^B | X | e_i^B \rangle \) is an orthonormal basis of the subsystem \( \mathcal{H}_B \) defined for any operator \( X \) acting on the composite Hilbert space \( \mathcal{H}_A \otimes \mathcal{H}_B \), denotes the **partial trace** over the subsystem \( \mathcal{H}_A \). The following theorem connects the above description with the one given in the previous section:

**Theorem 3.** *(Stinespring dilation theorem, [1])*  
Let \( \Phi : \mathcal{S}(\mathcal{H}) \rightarrow \mathcal{S}(\mathcal{H}) \) be a linear completely positive trace preserving map. Then there exists a Hilbert space \( \mathcal{H}_{\text{env}} \) and a unitary operator \( U \) acting on \( \mathcal{H} \otimes \mathcal{H}_{\text{env}} \) such that, for all \( \rho \in \mathcal{S}(\mathcal{H}) \),
\[
(1.3.2) \quad \Phi(\rho) = \text{Tr}_{\text{env}} \left[ U (\rho \otimes \rho_{\text{env}}) U^* \right],
\]
where \( \rho_{\text{env}} \) is some fixed state in \( \mathcal{H}_{\text{env}} \) (which can be chosen to be pure).

The ancilla space can be chosen such that \( \dim \mathcal{H}_{\text{env}} \leq (\dim \mathcal{H})^2 \).

Quantum maps can be represented in an elegant form known as the operator-sum or Kraus representation [2]. It is basically a restatement of (1.3.2) in terms of operators acting on the principal system’s Hilbert space alone.

**Theorem 4. (Operator-sum or Kraus representation, [2])**

A linear map \( \Phi : M_n \to M_m \) is completely positive if and only if it is of the form

\[
\Phi(\rho) = \sum_k A_k \rho A_k^*,
\]

where \( \{A_k\} \) is a finite set of operators acting on the Hilbert space of the system, and \( A^* \) denotes the Hermitian adjoint of \( A \).

The map will be trace preserving if and only if

\[
\sum_k A_k^* A_k = I.
\]

**Proof.** Without loss of generality we can assume that the initial state of the environment is given by a pure state \( \rho_{\text{env}} = |e_0\rangle \langle e_0| \). Let \( \{|e_k\}\) be an orthonormal basis of the Hilbert space of the environment, \( \langle e_j| e_k \rangle = \delta_{jk} \) for \( j, k = 1, \ldots, d \). Then (1.3.2) can be rewritten as

\[
\Phi(\rho) = \sum_k \langle e_k| [U (\rho \otimes |e_0\rangle \langle e_0|) U^*] |e_k \rangle
\]

\[
= \sum_k A_k \rho A_k^*,
\]

where \( A_k := \langle e_k| U |e_0\rangle \) is an operator acting on the Hilbert space of the principal system.

If \( \Phi \) is trace-preserving, then it follows that

\[
1 = \text{Tr} \Phi(\rho)
\]

\[
= \text{Tr} \left( \sum_k A_k \rho A_k^* \right) = \sum_k \text{Tr} (A_k \rho A_k^*)
\]

\[
= \text{Tr} \left( \sum_k A_k^* A_k \rho \right).
\]
Since the last equality must be true for all density matrices $\rho$, it must hold that
\[ \sum_k A_k^* A_k = I. \]
\[ \square \]

As the Kraus representation describes dynamics of a system without having to explicitly consider properties of the environment, it simplifies calculations and provides a useful tool in situations when we are only interested in the principal system.

### 1.4. Bloch sphere representation

The action of quantum map acting on a single qubit can be visualized using the **Bloch sphere representation** - a geometric representation of the state of a qubit.

An arbitrary hermitian matrix $\rho \in M_2$ with $\text{Tr}\rho = 1$ can be parametrized as
\[
\rho = \frac{1}{2} \begin{bmatrix} 1 + z & x - iy \\ x + iy & 1 - z \end{bmatrix}.
\]

Let
\[
\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}
\]
be the **Pauli matrices**. The parametrization (1.4.1) is usually rewritten in terms of the Pauli matrices $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ as
\[
\rho = \frac{1 + \vec{s} \cdot \vec{\sigma}}{2},
\]
where $\vec{s} \in \mathbb{R}^3$ is known as the **Bloch vector** for the state $\rho$. Since $\rho$ must be positive definite, it follows that $\| \vec{s} \| \leq 1$.

A state $\rho$ is pure if and only if $\| \vec{s} \| = 1$, since for a pure state it holds that $\rho^2 = \rho$. Let $|\psi\rangle = a |0\rangle + b |1\rangle$ be a state of a qubit. Setting $a = \cos \frac{\theta}{2}$ and $b = e^{i\phi} \sin \frac{\theta}{2}$, we see that a pure state $\rho = |\psi\rangle \langle \psi|$ of a qubit can be viewed as a point on the surface of a unit sphere - the Bloch sphere. The Bloch vector $\vec{s}$ is given by
\[
\vec{s} = (\cos \phi \sin \theta, \sin \phi \cos \theta, \cos \theta).
\]

Any mixed state corresponds to a point inside the Bloch sphere. If $\rho = \frac{1}{2} I$ (completely mixed state), then $\vec{s} = 0$, and hence it corresponds to the origin of the Bloch sphere.
Definition 5. A quantum map $\Phi$ is said to be **unital** if it preserves the identity, i.e. if $\Phi(I) = I$.

1.5. Examples of single qubit maps

1.5.1. Bit-flip map. This map flips the state of the qubit from $|0\rangle$ to $|1\rangle$ with probability $1 - p$ and leaves it unchanged with probability $p$:

$$
\Phi(\rho) = (1 - p) \sigma_x \rho \sigma_x + pp.
$$

The Kraus operators are

$$
A_1 = \sqrt{1 - p} I, \quad A_2 = \sqrt{p} \sigma_x.
$$

Under the action of the bit-flip map, the $x$-axis of the Bloch sphere stays untouched while the $yz$-plane is compressed uniformly by a factor of $1 - 2p$ [3, 4].

1.5.2. Phase-flip map. Leaves the qubit unchanged with probability $p$ and causes a phase flip with probability $1 - p$:

$$
\Phi(\rho) = (1 - p) \sigma_z \rho \sigma_z + pp.
$$

The corresponding Kraus operators are

$$
A_1 = \sqrt{1 - p} I, \quad A_2 = \sqrt{p} \sigma_z.
$$

This map compresses the $xy$-plane by a factor of $(1 - 2p)$ while leaving the $z$-axis unchanged [3, 4].
1.5.3. **Depolarizing map.** Depolarizing map leaves the qubit unaffected with probability $1 - \alpha$ and replaces the state of the qubit with a completely mixed state $\frac{I}{2}$ with probability $\alpha$:

\[
\Delta_\alpha (\rho) := (1 - \alpha) \rho + \alpha \frac{I}{2}.
\]

For arbitrary state $\rho$, it holds that

\[
\frac{I}{2} = \rho + \sigma_x \rho \sigma_x + \sigma_y \rho \sigma_y + \sigma_z \rho \sigma_z
\]

and hence the equation (1.5.3) can be rewritten as

\[
\Delta_\alpha (\rho) = \left(1 - \frac{3}{4} \alpha\right) \rho + \frac{\alpha}{4} (\sigma_x \rho \sigma_x + \sigma_y \rho \sigma_y + \sigma_z \rho \sigma_z).
\]

There are four Kraus operators:

\[
A_1 = \sqrt{1 - \frac{3}{4} \alpha} I, \quad A_2 = \frac{\sqrt{\alpha}}{2} \sigma_x, \quad A_3 = \frac{\sqrt{\alpha}}{2} \sigma_y, \quad A_4 = \frac{\sqrt{\alpha}}{2} \sigma_z.
\]

The Bloch sphere is contracted uniformly, depending on the parameter $\alpha$ [3].
1.5.4. Phase-damping map. Phase-damping map is a model for decoherence of quantum states. It acts on a state by suppressing the off-diagonal elements and keeping the diagonal ones unchanged [3].

**Definition 6.** Let $\mathcal{B} = \{|\psi_i\rangle\}_{i=1}^{d}$ be an orthonormal basis for $\mathcal{H}$ (dim $\mathcal{H} = d$), and let $E_i = |\psi_i\rangle \langle \psi_i|$. The phase-damping channel corresponding to $\mathcal{B}$ is the map:

\begin{equation}
\Phi_{\lambda}(\rho) = \lambda \rho + (1 - \lambda) \sum_{i=1}^{d} E_i \rho E_i.
\end{equation}

We say that a phase-damping channel $\Phi_{\lambda}$ is **uniform** [5] if every vector $|\psi_i\rangle$ of the basis $\mathcal{B}$, $|\psi_i\rangle = (v_1, ..., v_d) \in \mathbb{C}^d$, has the property that $|v_i| = |v_j|$, for all $i, j = 1, ..., d$. 
CHAPTER 2

Measures of performance for quantum channels

2.1. Capacity of a channel

The capacity of a channel (quantum or classical) is defined as the maximal amount of information that can be reliably transmitted, per channel use. However, since there is a lot of flexibility in the use of a quantum channel (unlike a classical channel, which can be used in one way only), it has various distinct capacities.

2.1.1. Classical channel. In the Shannon’s model ([6]), the input $X$ and the output $Y$ of a classical channel are modeled as random variables (with probability distributions $\{\pi_i\}$ and $\{q_j\}$, respectively).

**Definition 7.** The classical channel is a linear map $\mathcal{N}$ between probability distributions $\pi = (\pi_1, \pi_2, \ldots)$ and $q = (q_1, q_1, \ldots)$. It acts independently on each input letter, according to the fixed stochastic (transition) matrix $\{p_{ij}\}$, (where $p_{ij} = P(Y = j \mid X = i)$):

$$\mathcal{N}(\pi) = q, \quad \mathcal{N}(\pi)_j = \sum_i p_{ij} \pi_i.$$ 

We define the following quantity:

**Definition 8.** Let $X \sim p(x)$ and $Y \sim p(y)$ be a pair of random variables. The mutual information of $X$ and $Y$ is given by

$$I(X : Y) = \sum_{x,y} p(x, y) \log \left( \frac{p(x, y)}{p(x)p(y)} \right).$$

It is a measure of the amount of information that one random variable contains about another random variable.
2.1. CAPACITY OF A CHANNEL

Theorem 9. (Shannon, 1948. [6]) The channel capacity of a discrete memoryless channel \( N \) is given by

\[
C_{\text{class}} (N) = \max_{\{\pi(x)\}} I (X : Y),
\]

where \( I (X : Y) \) is the mutual information of the random variables \( X \) and \( Y \), and the maximum is taken over all possible input distributions \( \{\pi(x)\} \).

2.1.2. Shannon capacity of a quantum channel. This capacity of a quantum channel is obtained by viewing it as a particular realization of a classical channel: only product states are allowed at the input and single measurements at the output. Measurements are defined by a positive-operator valued measure (POVM), that is a collection of positive semidefinite matrices \( \{E_j\} \) such that \( \sum_{j=1}^k E_j = 1 \).

Definition 10. The Shannon capacity of a quantum channel is defined as

\[
C_{\text{Shan}} (\Phi) := \sup_{\{\rho_i, E_j\}} C_{\text{class}} (N),
\]

where \( N \) is the classical channel corresponding to the stochastic transition matrix \( p_{ij} = \Tr E_j \Phi (\rho_i) \), and the maximum is taken over all input states and output measurements.

If we allow both entangled inputs and collective measurements, it can be shown [7] that there exist channels satisfying the superadditivity property

\[
C_{\text{Shan}} (\Phi \otimes \Phi) > 2C_{\text{Shan}} (\Phi).
\]

2.1.3. Holevo capacity.

Definition 11. Define the Holevo capacity of the map \( \Phi \) as

\[
\chi (\Phi) := \sup_{\{p_i, \rho_i\}} \left[ S \left( \Phi \left( \sum_i p_i \rho_i \right) \right) - \sum_i p_i S \left( \Phi (\rho_i) \right) \right],
\]

where \( S (\rho) = -\Tr_2 \rho \log \rho \) is the von Neumann entropy.

Following the convention used in quantum information theory, we shall use logarithm of base two throughout this thesis.

Suppose that a sender uses only product states at the input, but the receiver is allowed to perform collective measurements over multiple channel uses. The capacity of the channel in this case is called the product-state classical capacity \( (C^{(1)}) \).
Holevo proved the following bound [7]:

\[ C_{Sh} (\Phi) \leq \chi (\Phi). \]

**Theorem 12.** (HOLEVO-SCHUMACHER-WESTMORELAND, [8, 9]) The product-state classical capacity \( C^{(1)} \) of the channel \( \Phi \) is given by

\[ C^{(1)} (\Phi) = \chi (\Phi). \]

A natural question arises: can one increase the classical capacity of a quantum channel by using entangled input states? The capacity of a channel in this case \( (C_{ult}) \) is given by the regularized Holevo capacity:

\[ C_{ult} (\Phi) = \lim_{n \to \infty} \frac{1}{n} \chi (\Phi^{\otimes n}). \]

If Holevo capacity is additive (this will be discussed later), then \( \chi (\Phi^{\otimes n}) = n \chi (\Phi) \), which implies that \( C_{ult} (\Phi) = \chi (\Phi). \)

### 2.2. Minimum output entropy

**Definition 13.** The **minimum output entropy** of a quantum map \( \Phi \) is defined as

\[ S_{\min} (\Phi) = \inf_{\rho} S (\Phi (\rho)), \]

where the infimum is taken over all input states.

In the \( d \)-dimensional space, the von Neumann entropy \( S \) takes values in the interval \([0, \log_d d]\). For a state \( \rho \) it holds that:

- \( S (\rho) = 0 \iff \rho \) is a pure state
- \( S (\rho) = \log_d d \iff \rho = \frac{1}{d} I \) is the maximally mixed state, proportional to identity.

Hence we see that in two dimensions, the states which have minimum output entropy are those which are mapped to points on the ellipsoid (the image of the Bloch sphere) which are closest to the Bloch sphere.
two unique states of minimum entropy

a whole “waistband” of states of minimum entropy

2.3. Maximal output $p$-norm

The following notion has been introduced by Amosov, Holevo and Werner [10] as a way to characterize the noisiness of a channel:

**Definition 14.** The **maximal output $p$-norm** for a map $\Phi$ is defined as

$$
\nu_p(\Phi) = \sup_{\rho} \|\Phi(\rho)\|_p,
$$

where $\|A\|_p = (\text{Tr} A^p)^{\frac{1}{p}}$ is the Schatten $p$-norm of a positive matrix (defined for $p \geq 1$).

Since the entropy of a state is the negative of the derivative of the Schatten $p$-norm at $p = 1$, it follows that for any channel $\Phi$,

$$
\frac{d}{dp} \nu_p(\Phi) \bigg|_{p=1} = -S_{min}(\Phi).
$$

(2.3.1)
2.4. Additivity conjectures

The additivity conjecture for Holevo capacity stated [11] that for any two quantum channels $\Phi$ and $\Omega$, it holds that

\[ \chi(\Phi \otimes \Omega) = \chi(\Phi) + \chi(\Omega). \]

This would imply that for a tensor product of a channel one has $\chi(\Phi^{\otimes n}) = n\chi(\Phi)$, and hence

\[ C_{ult}(\Phi) = \chi(\Phi). \]

If this conjecture were true, it would have important operational consequences, namely: the classical capacity of a quantum channel is achieved by using product states, i.e. using entangled inputs does not increase the capacity.

As a way to approach the additivity problem for Holevo capacity, a similar conjecture was stated for the minimum output entropy:

\[ S_{\text{min}}(\Phi \otimes \Omega) = S_{\text{min}}(\Phi) + S_{\text{min}}(\Omega). \]

The following theorem proves the equivalence of these conjectures in a global sense (i.e. on the class of all quantum maps). However, they are not necessarily equivalent for particular channels.

Theorem 15. (Shor, 2003., [12]) The following conjectures are equivalent:

1. Additivity of the Holevo capacity of a quantum channel;
2. Additivity of the minimal entropy output $S_{\text{min}}(\Phi)$ of a quantum channel, where

\[ S_{\text{min}}(\Phi) = \inf_{\rho} S(\Phi(\rho)); \]

3. Additivity of the entanglement of formation $E_F(\rho)$ of a bipartite state $\rho$, where

\[ E_F(\rho) = \min_{\xi} \sum_i p_i S(\text{Tr}_B(|\rho_i\rangle\langle \rho_i|)). \]

In 2008, Hastings proved [13] (non-constructively) the existence of quantum channels for which the minimal output entropy is not additive. Hence all of the above conjectures are false - and therefore we can increase the capacity of some quantum channels by using entangled inputs.
CHAPTER 3

Results about additivity

Although the additivity conjectures have been proven to be false in general, it remains an open question to find classes of quantum maps for which they do hold. Some of the most important examples of such maps include:

- unital qubit channels \([14]\), for which \(\Phi (I) = I\)

**Theorem 16.** (King, 2001. [14]) Let \(\Phi\) be a unital qubit channel. Then for any channel \(\Omega\), it holds that

\[
\begin{align*}
\chi (\Phi \otimes \Omega) &= \chi (\Phi) + \chi (\Omega), \\
S_{\text{min}} (\Phi \otimes \Omega) &= S_{\text{min}} (\Phi) + S_{\text{min}} (\Omega), \\
\nu_p (\Phi \otimes \Omega) &= \nu_p (\Phi) \nu_p (\Omega), \quad \forall p \geq 1.
\end{align*}
\]

- \(d\)-dimensional depolarizing channel (King, 2002. [5])
- some generalizations of the depolarizing channel (Datta and Ruskai, 2005. [15])
- entanglement-breaking channel (Shor, 2002. [16])

In this chapter we discuss the theorems of King and of Datta and Ruskai about depolarizing channel(s) and introduce Weyl-covariant channels, following the paper [17].

3.1. Depolarizing channel and generalizations

The depolarizing channel is a simple and highly symmetric example of a unital quantum map. In two dimensions, it contracts the Bloch sphere uniformly, depending on a single parameter. A natural generalization in the \(d\)-dimensional case is the map \(\Delta_\alpha\), which maps a state \(\rho\) into a linear combination of itself and the \(d \times d\) identity matrix:

\[
\Delta_\alpha (\rho) = \alpha \rho + (1 - \alpha) \frac{I}{d}.
\]

The condition of complete positivity requires that \(-\frac{1}{d^2 - 1} \leq \alpha \leq 1\).
Since the output $\Delta_{\alpha} (|\varphi\rangle \langle \varphi|)$ for any pure state $|\varphi\rangle \langle \varphi|$ has eigenvalues $\alpha + \frac{1-a}{d}$ (with multiplicity 1) and $\frac{1-a}{d}$ (with multiplicity $d-1$), the expressions for the three measures of performance can be easily computed. Their values are

(3.1.2) \[ S_{\min} (\Delta_{\alpha}) = - \left( a + \frac{1-a}{d} \right) \log \left( a + \frac{1-a}{d} \right) - (d-1) \left( \frac{1-a}{d} \right) \log \left( \frac{1-a}{d} \right), \]

(3.1.3) \[ \nu_p (\Delta_{\alpha}) = \left[ \left( \alpha + \frac{1-a}{d} \right)^p + (d-1) \left( \frac{1-a}{d} \right)^p \right]^{\frac{1}{p}}, \]

(3.1.4) \[ \chi (\Delta_{\alpha}) = \log_2 d - S_{\min} (\Delta_{\alpha}). \]

The additivity conjectures for the depolarizing channel have been proven by King [5]:

**Theorem 17.** (King, 2002. [5]) Let $\Delta$ be the $d$-dimensional depolarizing channel given by $\Delta_{\alpha} (\rho) = \alpha \rho + (1 - \alpha) \frac{I}{d}$. Then for any channel $\Omega$, it holds that

\[
\chi (\Delta \otimes \Omega) = \chi (\Delta) + \chi (\Omega), \]
\[
S_{\min} (\Delta \otimes \Omega) = S_{\min} (\Delta) + S_{\min} (\Omega), \]
\[
\nu_p (\Delta \otimes \Omega) = \nu_p (\Delta) \nu_p (\Omega), \quad \forall p \geq 1.
\]

The main idea of the proof is to express the depolarizing channel as a convex combination of uniform phase damping channels:

(3.1.5) \[ \Delta (\rho) = \sum_{n=1}^{2d^2(d+1)} c_n U_n^\dagger \Phi^{(n)} (\rho) U_n. \]

Since the $p$-norm is convex, to prove the bound

\[ \| (\Delta \otimes \Omega) (\rho) \|_p \leq \nu_p (\Delta) \nu_p (\Omega) \]

it is enough to prove it for each term separately, i.e. that for every $n$,

\[ \| (\Phi^{(n)} \otimes \Omega) (\rho) \|_p \leq \nu_p (\Delta) \nu_p (\Omega). \]

The following theorem proves multiplicativity of the maximal $p$-norm and additivity of the minimal output entropy for a generalization of the depolarizing channel:
Theorem 18. (Datta, Ruskai, 2005. [15]) Let $\Phi$ be a channel given by

\begin{equation}
\Phi (\rho) = \sum_{k} a_k V_k \rho V_k^\dagger + (1 - \alpha) \frac{I}{d},
\end{equation}

with $0 < \alpha = \sum_k a_k < 1$ and $V_k$ unitary operators which all have a common eigenvector $|\varphi\rangle$.

Then for any channel $\Omega$, it holds that

\begin{align*}
S (\Phi (|\varphi\rangle \langle \varphi|)) &= S_{\min} (\Phi) = S_{\min} (\Delta_\alpha), \\
S_{\min} (\Phi \otimes \Omega) &= S_{\min} (\Phi) + S_{\min} (\Omega), \\
\|\Phi (|\varphi\rangle \langle \varphi|)\|_p &= \nu_p (\Phi) = \nu_p (\Delta_\alpha), \quad \forall p \geq 1, \\
\nu_p (\Phi \otimes \Omega) &= \nu_p (\Phi) \nu_p (\Omega).
\end{align*}
3.2. Weyl-covariant channels

Let \( \{ e_k \}_{k=0}^{d-1} \) be an orthonormal basis of the Hilbert space \( \mathcal{H} \) of dimension \( d \). Consider the additive cyclic group \( \mathbb{Z}_d \) and the map

\[
z = (x, y) \mapsto W_z = U^x V^y,\]

where \( x, y \in \mathbb{Z}_d \) and \( U \) and \( V \) are unitary operators such that

\[
U |e_k\rangle = |e_{k+1(\text{mod } d)}\rangle,
\]

\[
V |e_k\rangle = \exp \left( \frac{2\pi i k}{d} \right) |e_k\rangle.
\]

The operators \( W_z \) are called discrete Weyl operators [17].

Discrete Weyl operators satisfy commutation relations:

\[
W_z W_{z'} = \exp \left( i \langle x', y \rangle - \langle y', x \rangle \right) W_{z'} W_z,
\]

\[
W_z W_{z'} = \exp \left( i \langle y, x' \rangle \right) W_{z+z'},
\]

where \( \langle y, x \rangle = \frac{2\pi y x}{d} \).

Let \( \{ e_k \}_{k=0}^{d-1} \) be the standard basis. We can represent the operators \( W_z = W_{(x,y)} \) with the following matrix (where \( \exp \left( \frac{2\pi i}{d} (t - x) y \right) \) are scaling factors and \( e_{t-x} \) are vectors of the basis):

\[
\begin{bmatrix}
\exp \left( \frac{2\pi i}{d} (1 - x) y \right) e_{1-x} \\
\vdots \\
\exp \left( \frac{2\pi i}{d} (t - x) y \right) e_{t-x} \\
\vdots \\
\exp \left( \frac{2\pi i}{d} (d - x) y \right) e_{d-x}
\end{bmatrix}
\]

**Definition 19.** A channel of the form

\[
\Phi (\rho) = \sum_{z \in \mathbb{Z}} a_z W_z \rho W_z^*,
\]

where \( Z := \mathbb{Z}_d \oplus \mathbb{Z}_d \), is called a Weyl-covariant channel.

Cortese [18] considered general Weyl-covariant channels of the form

\[
\Phi (\rho) = \sum_{z \in \mathbb{Z}} a_z W_z \rho W_z^*,
\]
where \( \sum_{z \in \mathbb{Z}} a_z = \sum_{x,y=0}^{d-1} a_{(x,y)} = 1 \), and found that the closed-form expression for the one-shot Holevo capacity is the same as the expression for the depolarizing channel:

**Theorem 20.** (Cortese, 2002. [18]) Let \( \Phi (\rho) = \sum_{z \in \mathbb{Z}} a_z W_z \rho W^*_z \), where \( \sum_{z \in \mathbb{Z}} a_z = 1 \), be the \( d \)-dimensional Weyl-covariant channel. Then the channel \( \Phi \) has the **one-shot** Holevo capacity of the form

\[
\chi (\Phi) = \log_2 d - S_{\min} (\Phi).
\]

It was conjectured in [15] that the channels of the form (3.2.4) behave similarly to unital qubit channels and hence have the same additivity properties (3.0.3).

### 3.2.1. Weyl-covariant channels associated to degenerate subgroups.

Consider a subset \( G_\psi \) of \( \mathbb{Z} = \mathbb{Z}_d \oplus \mathbb{Z}_d \) such that

\[
G_\psi = \{ z : |\psi\rangle \text{ is an eigenvector of } W_z \}.
\]

**Fact 21.** \( G_\psi \) is a subgroup of \( \mathbb{Z} \) and \( |G_\psi| \leq d \).

A subset \( F \) of \( \mathbb{Z} \) will be called **degenerate** if \( \langle z', Jz \rangle = 0 \) for every \( z, z' \in \mathbb{Z} \), where \( \langle z', Jz \rangle := \langle x', y \rangle - \langle y', x \rangle \).

If \( F \) is degenerate, then the operators \( \{ W_z : z \in F \} \) all commute and hence they have common eigenvectors. It follows that \( F \subseteq G_\psi \) and \( |F| \leq d \). If equality holds, then \( F \) is called a maximal degenerate subgroup of \( \mathbb{Z} \).

**Example 22.** The cyclic subgroup generated by an element \( z = (x, y) \in \mathbb{Z} \),

\[
G(z) := \{ kz : k = 0, 1, ..., d - 1 \}
\]

is degenerate. It is a maximal degenerate subgroup of \( \mathbb{Z} \) (i.e. \( |G(z)| = d \)) when \( x, y, d \) have no common divisor (in particular when \( d \) is prime).

In the paper [17], Fukuda and Holevo analysed a channel of the form

\[
(3.2.5) \quad \Phi (\rho) = \frac{a}{d} \sum_{z \in G} W_z \rho W^*_z + b \rho + (1 - a - b) I, \quad G
\]

where \( G \) is a maximal degenerate subgroup of \( \mathbb{Z} \).

**Lemma 23.** If \( G \) is a cyclic subgroup of order \( d \), then for the channel \( \Psi (\rho) = \frac{1}{d} \sum_{z \in G} W_z \rho W^*_z \) it holds that

\[
(3.2.6) \quad \Psi (\rho) = \sum_{j=1}^{d} |h_j\rangle \langle h_j| \rho |h_j\rangle \langle h_j|,
\]
where \( \{|h_j\}\) is the orthonormal basis of the commuting operators \( \{W_z : z \in G\} \).

**Proof.** Let \( G = \{jz_0 : k = 0, 1, \ldots, d-1\} \). We have

\[
\Psi(\rho) = \frac{1}{d} \sum_{j=0}^{d-1} W_{jz_0} \rho W_{jz_0}^* = \frac{1}{d} \sum_{j=0}^{d-1} (W_{jz_0})^j \rho (W_{jz_0}^*)^j.
\]

Let \( W_{jz_0} |h_j\rangle = c_j |h_j\rangle \), where \( c_j = \exp\left(\frac{2\pi i}{d} \alpha_j\right) \) and \( \alpha_j = j + \alpha_0, j = 0, 1, \ldots, d-1 \). It follows that

\[
\Psi(|h_m\rangle \langle h_n|) = \frac{1}{d} \sum_{j=0}^{d-1} (W_{jz_0})^j |h_m\rangle \langle h_n| (W_{jz_0}^*)^j
\]

\[= \frac{1}{d} \sum_{j=0}^{d-1} \exp\left(\frac{2\pi i}{d} (m-n) k\right) |h_m\rangle \langle h_n|
\]

\[= \begin{cases} |h_m\rangle \langle h_n|, & m = n \\ 0, & m \neq n \end{cases}.
\]

\[\square\]

### 3.2.2. A generalization.

In this section we present the original results obtained in this thesis.

Let us consider a generalized shifted Weyl channel of the form

\[
(3.2.7) \quad \Phi_{G_1, \ldots, G_k}(\rho) = \sum_{i=1}^{k} \alpha_i \frac{1}{d} \sum_{z \in G_i} W_z \rho W_z^* + \ldots + \alpha_k \frac{1}{d} \sum_{z \in G_k} W_z \rho W_z^* + b \rho + c I_d,
\]

where \( G_1, \ldots, G_k \) are maximal cyclic subgroups of \( Z = \mathbb{Z}_d \oplus \mathbb{Z}_d \), and \( 1 \leq k \leq d+1 \). Then from the Lemma 23, it follows that:

**Corollary 24.** The map \( \Phi_{G_1, \ldots, G_k}(\rho) \) can be rewritten as

\[
\Phi_{G_1, \ldots, G_k}(\rho) = \sum_{i=1}^{k} \left( \alpha_i \sum_{j=1}^{d} \left( a_i' |h_j^i\rangle \langle h_j^i| \rho |h_j^i\rangle \langle h_j^i| + (1 - a_i') I \right) \right),
\]

where \( B_i := \{|h_j^i\rangle\}, i = 1, \ldots, k \) are the \( k \) orthonormal bases of the commuting operators \( \{W_z : z \in G_i\} \) and \( \sum_i \alpha_i = 1 \). In particular, for every \( i \), the operators \( \{W_z : z \in G_i\} \) have all eigenvectors in common.
Theorem 25. (additivity for generalized shifted Weyl channels) For a generalized shifted Weyl channels \( \Phi^{G_1,\ldots,G_k} \) and an arbitrary CPTP map \( \Omega \), one has

\[
\nu_p(\Phi^{G_1,\ldots,G_k} \otimes \Omega) = \nu_p(\Phi^{G_1,\ldots,G_k}) \nu_p(\Omega).
\]

Proof. Since the left side of (3.2.8) is at least as big as the right side, it suffices to prove that for every state \( \rho \),

\[
\left\| (\Phi^{G_1,\ldots,G_k} \otimes \Omega)(\rho) \right\|_p \leq \nu_p(\Phi^{G_1,\ldots,G_k}) \nu_p(\Omega).
\]

Denote with \( \Psi^{(i)} \) the map \( \Psi^{(i)}(\rho) := \sum_{z \in G_i} \alpha_z^i W_z \rho W_z^\dagger + (1 - \alpha)^I \), for every \( i = 1,\ldots,k \). We can express the map \( \Phi^{G_1,\ldots,G_k} \) as a convex combination of \( \Psi^{(i)} \), i.e.

\[
\Phi^{G_1,\ldots,G_k}(\rho) = \sum_{i=1}^k \beta_i \Psi^{(i)}(\rho).
\]

For every \( G_i \), the map \( \Psi^{(i)} \) has the property that the corresponding Weyl operators \( W_z \) have all eigenvectors in common. Hence from the theorem of Datta and Ruskai (see [15]) it follows that

\[
\nu_p(\Psi^{(i)} \otimes \Omega) = \nu_p(\Psi^{(i)}) \nu_p(\Omega) \quad \text{and} \quad \nu_p(\Psi^{(i)}) = \nu_p(\Delta \alpha), \quad \forall p \geq 1.
\]

Therefore, by the convexity of the \( p \)-norm, we have that

\[
\left\| (\Phi^{G_1,\ldots,G_k} \otimes \Omega)(\rho) \right\|_p \leq \sum_{i=1}^k \beta_i \left\| \Psi^{(i)} \otimes \Omega \right\|_p
\]

\[
\leq \sum_{i=1}^k \beta_i \nu_p(\Psi^{(i)}) \nu_p(\Omega)
\]

\[
\leq \sum_{i=1}^k \beta_i \nu_p(\Phi^{G_1,\ldots,G_k}) \nu_p(\Omega) = \nu_p(\Phi^{G_1,\ldots,G_k}) \nu_p(\Omega),
\]

and the claim follows. \( \square \)

3.2.3. Weyl operators and mutually unbiased bases.

Definition 26. Let \( \{|e_1\rangle,\ldots,|e_d\rangle\} \) and \( \{|f_1\rangle,\ldots,|f_d\rangle\} \) be two orthonormal bases in the Hilbert space \( \mathbb{C}^d \) of dimension \( d \). We say that the bases \( \{|e_i\rangle\} \) and \( \{|f_j\rangle\} \) are mutually unbiased [19] if

\[
\left| \langle e_i | f_j \rangle \right|^2 = \frac{1}{d}, \quad \forall i,j \in \{1,\ldots,d\}.
\]

The problem of finding mutually unbiased bases in a given finite dimensional Hilbert space has important consequences for quantum information theory, in particular for quantum cryptography. The following results have been established:
• if $d$ is a power of a prime number, $d = p^n$, then one can find a set of $d + 1$ mutually unbiased bases \[20, 21]\;
• let $d = p_1^{n_1} \ldots p_k^{n_k}$ be the prime number decomposition of $d$, with $p_1^{n_1} < ... < p_k^{n_k}$. Then the number of mutually unbiased bases (MUBs) that can be constructed satisfies \[22\]

\[p_1^{n_1} + 1 \leq \#\text{MUBs} \leq d + 1.\]

One method of constructing mutually unbiased bases is by using Weyl operators. In his 1928 book \[23\], Hermann Weyl stated the following:

**Claim 27.** Suppose that $X$ and $Z$ are unitary operators acting on a Hilbert space $\mathbb{C}^d$, satisfying

\[XZ = \omega ZX,\]

where $\omega \in \mathbb{C}^d$, $|\omega| = 1$ (\(\omega\) is called a phase vector). If \(\omega\) is a primitive root of unity (that is, if $\omega = \frac{2\pi k}{d} \neq 1$ for $k < d$), then the eigenbases of $X$ and $Z$ are mutually unbiased.

Due to the commutation relations (3.2.2), it follows immediately that the eigenbases of Weyl operators, as defined in (3.2.1), are mutually unbiased when the dimension of the Hilbert space is a prime number. If it is not a prime, then the phase factor $\omega$ may not be a primitive root of unity.

**Example 28.** Let us return to the generalized shifted Weyl channel (3.2.7) and calculate it explicitly in dimension $d = 3$.

The maximal cyclic subgroups $G_1, G_2, G_3, G_4$ are

\[G_1 = \{(0, 0), (0, 1), (0, 2)\}, \quad G_2 = \{(0, 0), (1, 0), (2, 0)\}, \quad G_3 = \{(0, 0), (1, 1), (2, 2)\}, \quad G_4 = \{(0, 0), (2, 1), (1, 2)\}.\]

The Weyl operators corresponding to $G_1$ are

\[W_{(0,0)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad W_{(0,1)} = \begin{bmatrix} e^{\frac{2\pi i}{3}} & 0 & 0 \\ 0 & e^{\frac{4\pi i}{3}} & 0 \\ 0 & 0 & e^{2\pi i} \end{bmatrix}, \quad W_{(0,2)} = \begin{bmatrix} e^{\frac{4\pi i}{3}} & 0 & 0 \\ 0 & e^{2\pi i} & 0 \\ 0 & 0 & e^{4\pi i} \end{bmatrix};\]
corresponding to \( G_2 \) are
\[
W_{(0,0)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad W_{(1,0)} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad W_{(2,0)} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix};
\]
corresponding to \( G_3 \) are
\[
W_{(0,0)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad W_{(1,1)} = \begin{bmatrix} 0 & 0 & 1 \\ e^{\frac{2\pi i}{3}} & 0 & 0 \\ 0 & e^{\frac{2\pi i}{3}} & 0 \end{bmatrix}, \quad W_{(2,2)} = \begin{bmatrix} 0 & e^{-\frac{4\pi i}{3}} & 0 \\ 0 & 0 & 1 \\ e^{\frac{4\pi i}{3}} & 0 & 0 \end{bmatrix};
\]
corresponding to \( G_4 \) are
\[
W_{(0,0)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad W_{(2,1)} = \begin{bmatrix} 0 & e^{-\frac{2\pi i}{3}} & 0 \\ 0 & 0 & 1 \\ e^{\frac{2\pi i}{3}} & 0 & 0 \end{bmatrix}, \quad W_{(1,2)} = \begin{bmatrix} 0 & 0 & 1 \\ e^{\frac{4\pi i}{3}} & 0 & 0 \\ 0 & 0 & e^{\frac{8\pi i}{3}} \end{bmatrix}.
\]

The eigenbases \( B_1, B_2, B_3, B_4 \) corresponding to \( G_1, G_2, G_3, G_4 \) are
\[
B_1 = \{(1,0,0), (0,1,0), (0,0,1)\}
\]
\[
B_2 = \left\{ (1,1,1), \left( -\frac{1}{2}i \left( -i + \sqrt{3} \right), \frac{1}{2}i \left( i + \sqrt{3} \right), 1 \right), \left( \frac{1}{2}i \left( i + \sqrt{3} \right), -\frac{1}{2}i \left( -i + \sqrt{3} \right), 1 \right) \right\}
\]
\[
B_3 = \left\{ \left( 1, -1 \right) \frac{2}{3}, 1 \right), \left( -\sqrt{-1}, -\sqrt{-1}, 1 \right), \left( -1 \frac{2}{3}, 1, 1 \right) \right\}
\]
\[
B_4 = \left\{ \left( -1 \frac{2}{3}, -1 \frac{2}{3}, 1 \right), \left( 1, -\sqrt{-1}, 1 \right), \left( -\sqrt{-1}, 1, 1 \right) \right\}.
\]

The above set of mutually unbiased bases also has the property that the vectors belonging to the bases (except for \( B_1 \)) are uniform in the following sense [5]: if \( |\psi_i\rangle = (v_1, v_2, v_3) \in \mathbb{C}^3 \), then \(|v_i| = |v_j|\), for all \( i, j = 1, 2, 3 \). This property holds in general, and it follows from the definition of mutually unbiased bases.

Using this property of the eigenbases corresponding to maximal cyclic subgroups \( G_1, ..., G_k \), we can rewrite the generalized shifted Weyl channel \( \Phi^{G_1, ..., G_k} \) (3.2.7) (after choosing suitable coefficients) as a convex combination of uniform phase-damping channels

\[
\sum_{i=1}^{k} \left( \alpha_i' \sum_{j=1}^{d} \alpha_i'' \langle h^i_j \rangle \langle h^i_j \rangle \rho \langle h^i_j \rangle \langle h^i_j \rangle + (1 - \alpha_i'') \rho \right),
\]

(3.2.9)
Note that the representation (3.2.9) closely resembles the decomposition of the depolarizing channel into uniform phase-damping channels in King’s proof of additivity (3.1.5), and further illustrates the similarity in their behaviour.
Summary and open questions

The aim of this thesis was to provide a concise introduction to the problem of additivity of Holevo capacity and related measures of performance for quantum maps, including minimal output entropy and maximal $p$-norm. Mathematical, quantum-mechanical and information-theoretical background was presented, along with concepts and theorems important for the field. We discussed Weyl-covariant channels for which it was conjectured that the additivity holds [15] and established additivity under some technical assumptions.

The additivity conjectures emerged from the attempt to find a closed-form expression for the classical capacity of an arbitrary quantum channel. Hastings was able to prove [13] the existence of quantum channels in high dimensions for which the minimal output entropy and hence the capacity is not additive, using a non-constructive, randomized approach. This result confirmed that there is a deep structural difference between classical and quantum information. The picture to have in mind for the classical case is the following: the amount of information one can send using two classical telephone wires is the same as the sum of the amounts for each of the wires used separately. This does not hold for quantum channels - the amount of information using two quantum channels in principle can be larger than the sum of amounts for separate channels. This is due to the unique quantum phenomenon of entanglement.

Current directions of research include finding explicit counterexamples [24], obtaining useful bounds for the capacity and using techniques from convex geometry to investigate counterexamples [25]. It is known that the capacity of unital qubit channels is additive; however, the question of whether additivity holds for non-unital qubit channels remains open. A different line of research for which there has recently been progress in solving the additivity problem are bosonic Gaussian channels [26].
Bibliography

[23] H. Weyl, Gruppentheorie und Quantenmechanik, Hirzel, Leipzig (1928)
[27] N. Datta, online lecture notes: http://www.qi.damtp.cam.ac.uk/node/223