On quantum computation theory

van Dam, W.K.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
This chapter contains a standard introduction to quantum information theory. Topics that will be discussed are: the Hilbert space formalism for quantum states, unitary transformations and the probability rules for measurement outcomes. Also the theory of mixed states, density matrices and completely positive operators is discussed.

1.1 Modeling Information

The term ‘bit’ stands for ‘binary digit’, which reflects the fact that it can be described and implemented by a two-level system. Conventionally, these two levels are indicated by the labels “zero” and “one”, or “0” and “1”. If we want to capture more than two possibilities, more bits are needed: with \( k \) bits we have \( 2^k \) different labels.

The abstraction from \( k \) two-level systems to the set \( \{0, 1\}^k \) of size \( 2^k \) takes us away from the physical details of the implementation of a piece of memory in a computer, and instead focuses on a more mathematical description of information. This ‘physics independent’ approach to standard information theory has been extremely successful in the past decades: it enables a general understanding of computational and communicational processes that is applicable to all the different ways of implementing these processes. It is for this reason that the Turing machine model of computation gives an accurate description of both the mechanical computer suggested by Charles Babbage and the latest Silicon based Pentium IV processors, despite their obvious physical differences. This does not mean that Turing’s model ignores the physical reality of building a computer, on the contrary. The observation that it would be unphysical to assume an infinite or unbounded precision in the components of a computer is expressed by Turing’s rule that per time-step only a fixed, finite amount of computational work can be done.\[99\] The proper analysis of algorithms in the theory of computational complexity relies critically on the exclusion of computational models that are not realistic. Such models often give the wrong impression that certain complicated tasks are easy. (A good example of this is the result that the factorization of integers can be done in
polynomial time if we assume that addition, multiplication and division of arbitrary big numbers can be done in constant time. (See Chapter 4.5.4, Exercise 40 in [63] and [88].) There is, however, also a danger with this axiomatization of the physical assumptions in information theory: believing that the assumptions are true. This is what happened with the traditional view on information; forgotten were the implicit classical assumptions that ignore the possibilities of quantum mechanics. The realization that quantum physics describes a world where information behaves differently than in classical theory led to the blossoming of several fields—quantum information, quantum computing, quantum communication, et cetera. In this thesis we will focus on the differences in query complexity between classical and a quantum computation (Chapters 3–5), the possibility of ‘self-testing’ a quantum computer (Chapter 6) and a definition of quantum Kolmogorov complexity (Chapter 7). Before doing so, it is necessary to define what we mean by quantum information and computation.

1.2 Quantu m Information

At the heart of quantum mechanical information theory lies the superposition principle. Where a classical bit is either in the state “zero” or “one”, a quantum bit is allowed to be in a superposition of the two states. A qubit with the label $q$ is therefore described in Dirac’s bra-ket notation by the linear combination:

$$|q\rangle = \alpha|\text{"zero"}\rangle + \beta|\text{"one"}\rangle,$$

where for the complex valued amplitudes $\alpha, \beta \in \mathbb{C}$, the normalization restriction $|\alpha|^2 + |\beta|^2 = 1$ applies. Here $|\alpha|$ denotes the norm of $\alpha$: if $\alpha = a + bi$, then $|\alpha| := \sqrt{a^2 + b^2}$. Alternatively we can write $|\alpha| := \sqrt{\alpha\alpha^*}$, where $\alpha^*$ is the complex conjugate $a - bi$ of the complex value $\alpha = a + bi$. In this formalism, the state space of a single qubit is built up by the unit vectors in the two-dimensional Hilbert space $\mathcal{H}_2$. For $k$ qubits, there are $2^k$ basis states and hence the corresponding superposition is a linear combination of all $2^k$ possible strings of $k$ bits:

$$|q_1 \cdots q_k\rangle = \sum_{i \in \{0,1\}^k} \alpha_i |i\rangle.$$

Again it is required that the amplitudes $\alpha_i$ obey the normalization condition: $\sum_i |\alpha_i|^2 = 1$. (In Section 1.4 we will see the reason behind this stipulation.) The state space of $k$ qubits is the $k$-fold tensor product of the state space of a single qubit. This space is identical with a single $2^k$-dimensional Hilbert space:

$$|q_1 \cdots q_k\rangle \in \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_2 = \mathcal{H}_{2^k}.$$

For our purposes we will only use finite sets of quantum bits, so there is no need to look at infinite-dimensional Hilbert spaces.
1.3 Time Evolution of Quantum Bits

Quantum mechanics only allows transformations of states that are linear and respect the normalization restriction. When acting on an $n$-dimensional Hilbert space, these are the $n \times n$ complex valued rotation matrices that are norm preserving: the unitary matrices of $U(n)$. It is easy to show that this corresponds exactly to the requirement that the inverse of $U$ is the conjugate transpose $U^*$ of the matrix.

The effect of a unitary transformation $U$ on a state $x$ is exactly described by the corresponding rotation of the vector $|x\rangle$ in the appropriate Hilbert space. For this reason, "$U$" stands both for the quantum mechanical transformation as well as for the unitary rotation:

$$|U(x)\rangle = U|x\rangle = U \left( \sum_i \alpha_i |i\rangle \right) = \sum_i \alpha_i U|i\rangle = \sum_i \alpha_i \sum_j U_{ji} |j\rangle,$$

where $U_{ji}$ denotes the matrix element of $U$ positioned at the $j$-th row and the $i$-th column. It follows from the associativity of matrix multiplication that the effect of two consecutive transformation $U$ and $W$ is the same as the single transformation $(W \cdot U)$. Just as matrix multiplication does not commute, so does the order of a sequence of unitary transformations matter: in general $WU \neq UW$. We can restate this in a more intuitive way by saying that it makes a difference if we first do $U$ and then $W$, or the other way around. A typical example of this phenomenon is given by the matrices

$$W = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad U = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

with clearly $WU \neq UW$.

1.4 Measurements

When measuring the state $|x\rangle = \sum_i \alpha_i |i\rangle$, the probability of observing the outcome "$i$" equals $|\alpha_i|^2$. This explains the normalization restriction on the amplitudes: the different probabilities have to add up to one. But what exactly is a 'measurement' and an 'observation', and how do we describe this mathematically? These are thorny issues that this thesis will leave untouched. Here we will only give a formal description of the measurement process and a short explanation of why this is such a problematic part of quantum mechanics.

The possible outcomes "$i$" of $x$ correspond to a set of orthogonal vectors $\{|m_i\rangle\}$ of the measuring device. This device can be our own eye or some kind of machine, but the crucial point is that 'measuring $x$' implies 'interacting with $x$'. The effect on $x$ of such a measurement is that the state collapses according to the outcome "$m_i$" of our observation. This is described by the transformation:

$$\sum_i \alpha_i |i\rangle \quad \text{outcome} \quad m_i \quad |i\rangle.$$
The above described collapse is a non-unitary transformation. This is typical when we try to describe the behavior of $x$ as it interacts with a system that lies outside of the state. (We say that $x$ is an ‘open system’.) When we view $x$ and the measurement device together during the observation, the evolution becomes unitary again. Our current example is then described by the transformation:

$$\sum_i \alpha_i |i\rangle \otimes |\text{measurement device}\rangle \rightarrow \sum_i \alpha_i |i\rangle |\text{outcome } m_i\rangle.$$  

The problem with this last description is that it no longer specifies the specific outcome “$i$” that we seem to observe. It is here where the debate on the measurement problem starts and our discussion ends.

For the purposes of this thesis it is more convenient to use the terminology of the collapsing quantum state. We will therefore describe the effect of a measurement as in Equation 1.2 for practical reasons. (This does not imply that the author really thinks that there is such a collapse, but these issues are outside the scope of this text. They concern the interpretation of quantum mechanics, which is irrelevant for the purposes of this thesis.)

We just described the traditional ‘Von Neumann measurement’ where we observe the state $x$ in a canonical basis spanned by the basis vectors $i$. Other, more subtle, measurement procedures are also possible by choosing an in- or over-complete basis. We will postpone the description of these two options to the point when we discuss the density matrix formalism, which is more suitable for the general theory of interacting quantum mechanical systems.

### 1.5 Limitations of Dirac’s Notation

The brake t notation that we discussed above is tailor-made for the description of closed quantum mechanical systems. By this we mean the evolution of states that do not interact with an exterior environment. When we also want to consider the behavior of open systems, the ket-notation becomes less suitable. This was already obvious in the discussion of the measurement procedure where we had to expand the set of unitary operations with a probabilistic procedure that ‘collapses’ the quantum state to one of the basis states. One cannot help but feel uncomfortable about this sudden change of rules: is it not possible to deal with open and closed quantum systems in the same way? Luckily, we find in the formalism of density matrices a positive answer to this question.

### 1.6 Density Matrices

An $n$-dimensional pure state $x$ can be expressed as a normalized vector $|x\rangle$ in the Hilbert space $\mathcal{H}_n$. The complex conjugate $|x\rangle^*$ of this vector is the bra $\langle x|$.
1.6. Density Matrices

Element of the adjoint space $\mathcal{H}_n^*$. By taking the direct product between the ket $|x\rangle$ and the bra $\langle x|$, we thus obtain an $n \times n$ complex valued, Hermitian matrix: the density matrix of $x$.

As an example, for the state $|x\rangle = \sum \alpha_i |i\rangle$, the density matrix is:

$$|x\rangle\langle x| = \left( \sum_i \alpha_i |i\rangle \right) \left( \sum_j \alpha_j^* \langle j| \right) = \sum_{i,j} \alpha_i \alpha_j^* |i\rangle \langle j|.$$  

In the case of a single qubit with the ket description $|q\rangle = \alpha|0\rangle + \beta|1\rangle$, this leads to the $2 \times 2$ matrix in the standard basis

$$|q\rangle\langle q| = \begin{bmatrix} |\alpha|^2 & \alpha\beta^* \\ \alpha^*\beta & |\beta|^2 \end{bmatrix}.$$  

From now on, the density matrix of the state $x$ will be denoted by the same symbol $x$, and the fact that a matrix is a density matrix will be indicated by its square brackets.

The great advantage of this formalism is that it also allows the description of an ensemble of pure quantum states. If we have such a state $\rho$, which is a probabilistic mixture of the pure states $|x_t\rangle$ with probabilities $p_t$, then the matrix $\rho$ is the weighted linear combination of the corresponding pure states matrices,

$$\rho = \sum_t p_t \cdot |x_t\rangle\langle x_t|,$$

with $p_t \geq 0$ and $\sum_t p_t = 1$.

Every density matrix that can be written as such a convex combination of pure states is a legal, or ‘allowed’, state, where allowed means: “allowed by the laws of quantum physics”. It follows from linear algebra that this restriction coincides with the requirement that the matrix is a Hermitian, positive semidefinite matrix with unit trace.

The spectral decomposition of a proper density matrix $\rho$ is done in terms of its eigenvalues $\lambda_t$ and eigenvectors $|\omega_t\rangle$, by the equality

$$\rho = \sum_t \lambda_t |\omega_t\rangle\langle \omega_t|.$$  \hspace{1cm} (1.3)

This shows that we can interpret $\rho$ as the mixture $\{(\lambda_t, |\omega_t\rangle)\}_t$, where the states $\omega_t$ are pure and mutually orthogonal.

The above decomposition gives a convenient way of assigning a mixture to a given density matrix. It is important to realize, however, that a density matrix corresponds to a whole family of possible mixtures. Take, for example, the ensembles $\{(\frac{1}{2}, |0\rangle), (\frac{1}{2}, |1\rangle)\}$ and $\{(\frac{1}{2}, \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)), (\frac{1}{2}, \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle))\}$, which have the same density matrix:

$$\frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$
We shall see that this implies that these two mixtures are indistinguishable from each other; it is therefore more accurate and less confusing to consider them as equivalent mixtures.

The density matrix of a qubit $\rho$ in the standard basis is always of the form

$$
\rho(p, \alpha) = \begin{pmatrix}
p & \alpha^* \\
\alpha & 1-p
\end{pmatrix},
$$

with the probability $p$ between 0 and 1 and the ‘off-diagonal term’ $|\alpha|^2 \leq p(1-p)$. If $|\alpha|^2 = p(1-p)$ then $\rho$ is a pure state with $|\rho\rangle = \sqrt{p}|0\rangle + \sqrt{1-p}|1\rangle$ (or $|\rho\rangle = |1\rangle$ if $p = 0$); otherwise the qubit $\rho$ corresponds to a mixture.

1.7 Separated Systems

We need the formalism of density matrices to be able to describe the evolution of an open system. By ‘open’ we mean that there is a possible interaction between the quantum mechanical state and its environment. An example of such a situation was already mentioned when we saw how a qubit changed into a probabilistic mixture after it interacted with a measurement device outside the qubit system. An important operation in this context is the ‘tracing out’ operation that describes how we can ignore a part of a quantum system.

**Definition 1 (Partial trace)** Let $\mathcal{H}_{AB}$ be the combination of the two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, with the respective bases $\{|a_i\rangle\}$ and $\{|b_j\rangle\}$. The partial trace $\operatorname{tr}_B$ of a state $\rho$ in $\mathcal{H}_{AB}$ is defined by

$$
\operatorname{tr}_B(\rho) := \sum_j \langle b_j | \rho^{AB} | b_j \rangle,
$$

where $\langle x|\rho|y\rangle$ expresses the inner product of the row vector $\langle x|$, the matrix $\rho$ and the column vector $|y\rangle$.

When we are dealing with a general state $\rho$ and we want to describe its content for the subsystem $\mathcal{H}_A$, we indicate this by the notation “$\rho^A$“. Hence in terms of the above definition we would write $\rho^A := \operatorname{tr}_B(\rho^{AB})$. Conversely, we also have $\rho^B := \operatorname{tr}_A(\rho^{AB})$.

1.8 Von Neumann Entropy and Fidelity

The eigenvalues $\lambda$, of a density matrix are always nonnegative and sum up to one. If we decompose a mixture into a linear combination of orthogonal pure states, then the $\lambda$’s will correspond to the probabilities of the respective eigenvectors. (See Equation 1.3.) Although the eigenvectors of a density matrix are not always unique, its eigenvalues are. This allows us to unambiguously define the *Von Neumann entropy* $S(\rho)$ of a state,
1.8. Von Neumann Entropy and Fidelity

which reflects how ‘mixed’ or random $\rho$ is. As a result, pure states will have zero entropy.

**Definition 2 (Von Neumann entropy)** The Von Neumann entropy of a mixed state $\rho$ is defined as

$$S(\rho) = S \left( \sum_i p_i |\phi_i\rangle \langle \phi_i| \right) := - \sum_i p_i \log p_i,$$

where $\sum_i p_i |\phi_i\rangle \langle \phi_i|$ is a spectral decomposition of $\rho$ in its eigenvectors.

If we understand the logarithm of the matrix $\rho$ to be the standard Taylor expansion:

$$(\rho - I) - \frac{1}{2}(\rho - I)^2 + \frac{1}{3}(\rho - I)^3 - \cdots,$$

then the above definition can also be written as $S(\rho) := - \text{tr}(\rho \log_2 \rho)$. It should be clear that the Von Neumann entropy equals the Shannon entropy of the eigenvalues of the density matrix $\rho$.

A source $E = \{ (\rho_i, p_i) \}$ has an associated Von Neumann entropy $S(\rho)$ of the average state $\rho = \sum_i p_i \rho_i$. Schumacher’s noiseless coding theorem [83] shows how to obtain an encoding with average letter-length $S(\rho)$ for a source of pure states, where the fidelity of the encoding goes to 1 as the number of letters emitted by the source goes to infinity. (A survey can be found in Preskill’s lecture notes [78, page 190], Nielsen’s thesis [73, Chapter 7], or the standard book by Nielsen and Chuang [74].)

How close two mixed states $\rho$ and $\sigma$ are, can be expressed by the fidelity between the two density matrices. This notion generalizes the inner product between two Hilbert space vectors for pure states. The matrix $\rho$ represents a pure state if and only if $\rho^2 = \rho$, in which case we can also say $\sqrt{\rho} = \rho$. In general, the square root of a mixed state is defined by

$$\sqrt{\rho} = \sqrt{\sum_i p_i |\phi_i\rangle \langle \phi_i|} := \sum_i \sqrt{p_i} |\phi_i\rangle \langle \phi_i|.$$

We will use this root in the following definition.

**Definition 3 (Fidelity)** The fidelity $F(\rho, \sigma)$ between two density matrices $\rho$ and $\sigma$ is defined by

$$F(\rho, \sigma) := \text{tr} \left( \sqrt{\sqrt{\rho} \cdot \sigma \cdot \sqrt{\rho}} \right). \quad (1.4)$$

For pure states $\phi$ and $\psi$, the above definition coincides again with the familiar $|\langle \phi | \psi \rangle|$ (although some authors use the square of this value). If $F(\rho, \sigma) = 1$, then $\rho = \sigma$, and vice versa.
1.9 Operations on Mixed States

A unitary transformation $U$ maps the pure state $|x\rangle$ to the new pure state $U|x\rangle$. The latter can be written as the density matrix $U|x\rangle\langle x|U^*$. In the language of density matrices, the corresponding transformation $U$ is therefore calculated by ‘sandwiching’ the matrix $x$ between $U$ and its conjugate $U^*$:

$$U (|x\rangle\langle x|) := U|x\rangle\langle x|U^*.$$ 

If we have a mixed state $\rho$, then $U$ acts linearly on the eigenvectors of $\rho$. The following equation shows us that this calculation can be done without having to decompose $\rho$, and that our sandwich expression therefore also holds for mixed states:

$$U (\rho) = U \left( \sum_{t} \lambda_{t} |\omega_{t}\rangle\langle\omega_{t}| \right) = \sum_{t} \lambda_{t} \cdot U (|\omega_{t}\rangle\langle\omega_{t}|) \Delta U^{*} = U \left( \sum_{t} \lambda_{t} \cdot |\omega_{t}\rangle\langle\omega_{t}| \right) U^{*} = U \cdot \rho \cdot U^{*}.$$ 

It is clear that the positive eigenvalues $\lambda_{t}$ of $\rho$ remain unchanged, and that $U$ only rotates the eigenvectors $|\omega_{t}\rangle$ to the new eigenstates $U|\omega_{t}\rangle$.

Unitary operations are an example of completely-positive, trace preserving maps: every positive semidefinite matrix is mapped to (another) positive semidefinite matrix, and the trace of the matrix remains unaltered. Complete-positivity, in combination with the preservation of the trace, assures us that the result of a transformation will be a proper state if we started with a proper one.

Besides the unitary functions, there are other transformations that are possible in quantum mechanics. Just as mixed states are composed of pure states, so can a positive map be a linear combination of matrix multiplications similar to the ones we discussed above. An example of such a non-unitary mapping is the mapping $P$, corresponding to a measurement of a qubit in the standard basis $\{0, 1\}$. This function consists of two ‘projectors’ $P_0 = |0\rangle\langle 0|$ and $P_1 = |1\rangle\langle 1|$ that transform a qubit $\rho$ into a probabilistic
mixture of the states 0 and 1. Explicitly:

\[
P(\rho) = P\left(\begin{bmatrix} p & \alpha^* \\ \alpha & 1 - p \end{bmatrix}\right)
= P_0\left(\begin{bmatrix} p & \alpha^* \\ \alpha & 1 - p \end{bmatrix}\right) + P_1\left(\begin{bmatrix} p & \alpha^* \\ \alpha & 1 - p \end{bmatrix}\right)
= \begin{bmatrix} p & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 - p \end{bmatrix}
= \begin{bmatrix} p & 0 \\ 0 & 1 - p \end{bmatrix}.
\]

We see that the eigenvalues of the new density matrix are \( p \) and \( 1 - p \) with the corresponding eigenvectors \(|0\rangle\langle 0|\) and \(|1\rangle\langle 1|\). In general, the eigenvalues of \( \rho \) will change under this transformation and hence there is no unitary operation that can establish the above mapping. In the next section we will give a formal description of all transformations, such as the above \( P \), that are allowed by quantum physics.

### 1.10 Operator Sum Representation

The following requirements for an operator \( E \) are necessary and sufficient for \( E \) to be a proper quantum mechanical transformation:

1. The mapping \( E \) can be written as a set of matrices \( \{E_i\}_i \) with which it maps a state \( \rho \) to the linear combination \( \sum_i E_i \cdot \rho \cdot E_i^* \).

2. The set of operators \( \{E_i\} \) has to obey the identity restriction \( \sum_k E_k^* \cdot E_k = I \). (Note the change of order of \( E \) and \( E^* \) in the multiplication.)

These two requirements exactly describe the set of completely-positive, trace preserving maps. Complete-positivity means that we require both \( E \) as well its trivial extensions \( E \otimes I \) to higher dimensions to be positive. This is a stronger condition than positivity. An example of a positive but not completely-positive map is the partial transpose \( T \), which is defined by \( T(\rho) = \rho^T \).

We have properly extended the set of unitary transformations and measurements by the above 'operator sum' formalism. An example of this is the mapping that erases a qubit and replaces it with the value zero. This non-unitary function is the combination of two operators

\[
E = \left\{ \left( \begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array} \right), \left( \begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array} \right) \right\},
\]
and has the same effect on every qubit $\rho$, namely

$$E(\rho) = E\left(\begin{bmatrix} p & \alpha^* \\ \alpha & 1 - p \end{bmatrix}\right)$$

$$= \left(\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array}\right) \left(\begin{array}{cc} p & \alpha^* \\ \alpha & 1 - p \end{array}\right) \left(\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array}\right) + \left(\begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array}\right) \left(\begin{array}{cc} p & \alpha^* \\ \alpha & 1 - p \end{array}\right) \left(\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array}\right)$$

$$= \left[\begin{array}{cc} p & 0 \\ 0 & 0 \end{array}\right] + \left[\begin{array}{cc} 1 - p & 0 \\ 0 & 0 \end{array}\right]$$

$$= |0\rangle\langle 0|.$$

We previously argued that a measurement has a non-unitary effect on a state because we ignored its interaction with an outside system (the measurement device). This lesson holds for all allowed transformations:

Every completely-positive, trace preserving transformation $E$ of a system $\mathcal{H}_A$ can be viewed as a part of unitary mapping $U_E$ on a bigger system $\mathcal{H}_A \otimes \mathcal{H}_B$. That $E$ by itself appears to be non-unitary is due to the fact that we ignore the space $\mathcal{H}_B$.

It can be shown that for the extension of the system it is sufficient to assume that the dimension of the appended space $\mathcal{H}_B$ is twice as large as that of $\mathcal{H}_A$, and that its initial state is $|0\cdots0\rangle$. Hence, for every allowed quantum mechanical transformation $E$ that acts on an $n$-dimensional system, there exists a unitary matrix $U_E \in U(n^3)$ such that

$$E(x) = \operatorname{tr}_B [U_E(x \otimes |0^B\cdots0^B\rangle\langle 0^B\cdots0^B|)U_E^*)]$$

for all $x$. This is, in more general terms, the difference that we encountered between the Equations 1.2 and 1.3. The non-unitary 'collapse' associated with an observation, or any other kind of interaction, is again a unitary transformation when we incorporate the measurement device into the description of the event.

The converse of the earlier statement also holds: every mapping that can be written as a traced-out, unitary transformation on a larger Hilbert space is a completely-positive, trace preserving mapping.

In the literature on quantum information theory the linear functions on density matrices are sometimes called 'super operators'. We thus have the following definition.

**Definition 4 (Completely positive super operator/CPSO)** A transformation $E$ is a completely positive super operator, or CPSO, if only if $E$ is linear, trace-preserving, and completely positive.

The reader is referred to the standard book by Asher Peres[77] or the article by Benjamin Schumacher[84] for a more extended and rigorous treatment of this 'operator sum representation'.