Title: Foundations and Interpretation of Quantum Theory - Lecture 2

Date: Jan 14, 2010 02:30 PM

URL: http://pirsa.org/10010079

Abstract: After a review of the axiomatic formulation of quantum theory, the generalized operational structure of the theory will be introduced (including POVM measurements, sequential measurements, and CP maps). There will be an introduction to the orthodox (sometimes called Copenhagen) interpretation of quantum mechanics and the historical problems/issues/debates regarding that interpretation, in particular, the measurement problem and the EPR paradox, and a discussion of contemporary views on these topics. The majority of the course lectures will consist of guest lectures from international experts covering the various approaches to the interpretation of quantum theory (in particular, many-worlds, de Broglie-Bohm, consistent/decoherent histories, and statistical/epistemic interpretations, as time permits) and fundamental properties and tests of quantum theory (such as entanglement and experimental tests of Bell inequalities, contextuality, macroscopic quantum phenomena, and the problem of quantum gravity, as time permits).

1. Introduction and Motivation

2. Axioms for Quantum Theory
   - Ideal Preparations: Hilbert Space Vectors
   - Ideal Measurements: Self-adjoint Operators
   - Composite Systems: Tensor-Product Structure
   - Ideal Transformations 1: Unitary Operators
   - Ideal Transformations 2: Projections

3. Generalized Axioms for Quantum Theory
   - Generalized Preparations: Density Operators
   - Generalized Measurements: POVMs
   - Generalized Transformations: CP maps
   - Measurement as a Generalized Transformation
   - Composite Systems and Entanglement

1. Introduction and Motivation

2. Axioms for Quantum Theory
   - Ideal Preparations: Hilbert Space Vectors
   - Ideal Measurements: Self-adjoint Operators
   - Composite Systems: Tensor-Product Structure
   - Ideal Transformations 1: Unitary Operators
   - Ideal Transformations 2: Projections

3. Generalized Axioms for Quantum Theory
   - Generalized Preparations: Density Operators
   - Generalized Measurements: POVMs
   - Generalized Transformations: CP maps
   - Measurement as a Generalized Transformation
   - Composite Systems and Entanglement

Some Introductory Thoughts

- The purpose of this course is to gain a deeper understanding of what kind of theory quantum theory is, and to learn what it tells us about the world.

General states as mixtures of pure states

Suppose we want to describe a quantum system which is prepared according to one procedure, represented by state $|\psi_1\rangle$, with probability p_1 and according to a distinct procedure, represented by state $|\psi_2\rangle$, with probability p_2. How can we do this?

- If we are measuring the operator $A = \sum_a a\hat{P}_a$ which possesses non-degenerate eigenvalues $a \in \mathbb{R}$ associated with orthogonal eigenspaces $\hat{P}_a$, then the probability of obtaining outcome a given preparation ψ_1 is

$$\mathrm{Pr}(a|\psi_1) = \mathrm{Tr}(\hat{P}_a|\psi_1\rangle\langle\psi_1|),$$

and similarly for preparation 2.

- If we do not know which preparation took place then the net probability of finding outcome a is simply

$$\mathrm{Pr}(a) = p_1 \mathrm{Pr}(a|\psi_1) + p_2 \mathrm{Pr}(a|\psi_2).$$

By linearity of the trace we deduce that,

$$\mathrm{Pr}(a) = \mathrm{Tr}(\hat{P}_a \rho)$$

where

$$\rho = p_1 |\psi_1\rangle\langle\psi_1| + p_2 |\psi_2\rangle\langle\psi_2|$$

is non-negative operator called a *density operator* satisfying the normalization condition $\mathrm{Tr}(\rho) = 1$ (which ensures that probabilities are conserved).

General states as mixtures of pure states

In this way we can construct general quantum states from probabilistic mixtures (convex combinations) of pure states as follows:

- (i) Discrete case: $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ with $\sum_i p_i = 1$ and $p_i \geq 0$.
- (ii) Continuous case: $\rho = \int d\lambda p(\lambda) |\psi(\lambda)\rangle\langle\psi(\lambda)|$ for $\lambda \in \mathbb{R}$, with $\int d\lambda p(\lambda) = 1$ and $p(\lambda) \geq 0$.

General states from the partial trace

Suppose we have a quantum state (density operator) $\rho = \rho_{AB}$ on a composite Hilbert space $\mathcal{H}_{AB}$, where in general ρ need not correspond to a pure state $\rho = |\psi\rangle\langle\psi|$ but may be a probabilistic mixture of pure states. How does one describe the state of subsystem A alone (with a state ρ_A) or B alone (with a state ρ_B) ?

- The relationship between ρ_A and ρ_{AB} is generated by the *partial trace* operation:

$$\rho_A = \mathrm{Tr}_B(\rho_{AB}).$$

- The state ρ_A is called the *reduced state* associated with ρ_{AB}.

- This relationship can be deduced from physical consistency of *demanding* that

$$\langle \hat{A} \otimes \mathbb{1}_B \rangle = \langle \hat{A} \rangle$$

for all Hermitian operators $\hat{A}$ and states $\hat{\rho}_{AB}$. Hence,

$$(\rho_A)_{\ell\ell'} = \sum_k (\rho_{AB})_{\ell k \ell' k},$$

which gives us an explicit *matrix representation* of $\hat{\rho}_A$ in terms of the matrix elements of ρ_{AB} via the partial trace.

General states from the partial trace

Definition

The *partial trace* over a subsystem B of an operator O acting on the composite space $\mathcal{H}_{AB}$,

$$\hat{O}_A = \mathrm{Tr}_B[\hat{O}_{AB}],$$

can be defined in terms of the matrix representation,

$$(\hat{O}_A)_{\ell\ell'} = \langle \ell | \hat{O}_A | \ell' \rangle = \sum_k \langle \ell | \otimes \langle k | \hat{O}_{AB} | \ell' \rangle \otimes | k \rangle.$$

- It should be understood that the operation $\mathrm{Tr}_B(\cdot)$ takes as input *any* linear operator on $\mathcal{H}_{AB}$ (not necessarily a density operator) and generates a linear operator on $\mathcal{H}_A$.

Generalized states

Generalized Axiom 1: The physical configuration of a system positive semidefinite operator ρ subject to the normalization constraint $\mathrm{Tr}(\rho) = 1$.

- An operator P is positive semi-definite iff it is self-adjoint and satisfies $\langle u|P|u\rangle \geq 0$ for every vector u in the Hilbert space.
- A positive semidefinite operator (i.e., a *non-negative operator*) is often just called a *positive operator*.

Pure States vs Mixed States

For a state operator $\hat{\rho}$ subject to the normalization condition $\mathrm{Tr}(\hat{\rho}) = 1$ there are three equivalent definitions of *purity*:

i) $\hat{\rho}^2 = \hat{\rho}$, which means that ρ is projector.

ii) $\mathrm{Tr}(\hat{\rho}^2) = 1$.

iii) $\hat{\rho} = |\psi\rangle\langle\psi|$, defining a projector onto a one-dimensional subspace of $\mathcal{H}$.

Definition

If ρ can not be expressed in the form $\rho = |\psi\rangle\langle\psi|$ for any $\psi \in \mathcal{H}$, *i.e.*, if ρ is not a *pure state*, then it is called a *mixed state*.

Pure States vs Mixed States

General states obtained via partial trace are sometimes called *improper mixtures*, whereas the term *proper mixtures* refers to general states obtained from probabilistic mixing of pure states. These two conceptually distinct classes of mixed states are mathematically (and operationally) indistinguishable, as is evident from the following theorems:

Theorem

Any mixed state can be expressed as a convex combination of pure states.

Theorem

Any mixed state can be realized as the reduced state obtained from an (entangled) pure state on an extended Hilbert space.

Generalized Measurements

Recall from standard Axiom 2 that the primitives of a measurement, associated with a self-adjoint operator A, are the orthogonal projectors P_a (onto distinct, possibly degenerate, eigenspaces of A) in the spectral decomposition of A.
Any set of orthogonal projectors $\{P_a\}$, satisfying $\sum_a P_a = \mathbb{1}$ is called a *projector valued measure*, or PVM for short.

- We can construct measurements that have a more general structure than a PVM in two different ways.
  - First, we can build up a more general measurement by considering classical probabilistic mixtures of PVM measurements.
  - Second, we can consider what structure occurs when look at the "reduced measurement" obtained from different kinds of PVM measurement on an extended Hilbert space.

Mixtures of Projector Valued Measures

Consider two distinct PVMs, each given by a discrete set of D rank-one orthogonal projectors:

- $\{P_i\}$ with $i = 1, \ldots, D$ and $\{\tilde{P}_j\}$ with $j = 1, \ldots, D$, satisfying $\sum_i P_i = \mathbb{1}$ and $\sum_j \tilde{P}_j = \mathbb{1}$, where the orthogonality implies $P_i P_{i'} = P_i \delta_{ii'}$ and $\tilde{P}_j \tilde{P}_{j'} = \tilde{P}_j \delta_{jj'}$
- Note that in general the elements P_i and $\tilde{P}_j$ are non-orthogonal.

Mixtures of Projector Valued Measures

Suppose we have a device which performs the first PVM at random with probability p and the second with probability $1 - p$.

- Given a preparation ρ on a D-dimensional Hilbert space, from Axiom 2 we know that we can represent the probability of each of the $2D$ possible outcomes as follows:

$$\begin{aligned} \mathrm{Pr}(i) &= p\mathrm{Tr}(P_i\rho) \\ \mathrm{Pr}(j) &= (1-p)\mathrm{Tr}(\tilde{P}_j\rho). \end{aligned}$$

Mixtures of Projector Valued Measures

- Let $E_\nu = pP_i$ for $\nu = i$ and $E_\nu = (1-p)\tilde{P}_j$ for $\nu = D + j$.
- Then we can describe the probabilities of the $2D$ possible outcomes with the simple formula

$$\mathrm{Pr}(\nu) = \mathrm{Tr}(E_\nu \rho),$$

where these new operators satisfy:

$$\sum_\nu E_\nu = \mathbb{1}$$
$$E_\nu \geq \text{foreach}\nu.$$

- Note that when $p \in (0,1)$ the operators $\{E_\nu\}$ are not projectors.

PVMs on a Composite System

- The above measurement can be expressed in the form

$$\Pr(\nu) = \mathrm{Tr}[E_\nu \rho_A]$$

where $(E_\nu)_{ij} = \sum_{kl} (P_\nu)_{ik,jl} (\rho_B)_{lk}$ is an operator acting on $\mathcal{H}_A$.

- It is easy to see that the measurement operators E_ν satisfy

$$\sum_\nu E_\nu = \mathbb{1}_A$$
$$E_\nu \geq 0.$$

which are the same conditions we found for mixtures of PVMs.

PVMs on a Composite System

- Note that the measurement operators E_ν are not necessarily orthogonal (this is an important difference from a PVM) and hence that the number of elements in the set $\{E_\nu\}$ may be greater than the Hilbert space dimension.
- Indeed the measurement operators $\{E_\nu\}$ can also form a continuous set.

PVMs on an Extended Hilbert Space

Suppose instead now that we have a composite system represented by the state $\rho_A \otimes \rho_B$ and we perform a *joint measurement* of both systems.

- This is represented by a PVM $\{P_\nu\}$ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$, with the usual properties $P_\nu P'_\nu = P_\nu \delta_{\nu\nu'}$ and $\sum_\nu P_\nu = \mathbb{1}$, and where the Greek index run from 1 to $K \leq MN = \dim(\mathcal{H}_A)\dim(\mathcal{H}_B)$.
- The probability of outcome ν is

$$\Pr(\nu) = \mathrm{Tr}[P_\nu(\rho_A \otimes \rho_B)] = \sum_{ij=1}^{M}\sum_{kl=1}^{N}(P_\nu)_{ik,jl}(\rho_A)_{ji}(\rho_B)_{lk},$$

where

$$(P_\nu)_{ik,jl} = \langle i| \otimes \langle k|(P_\nu)|j\rangle \otimes |l\rangle$$

Mixtures of Projector Valued Measures

Suppose we have a device which performs the first PVM at random with probability p and the second with probability $1 - p$.

- Given a preparation ρ on a D-dimensional Hilbert space, from Axiom 2 we know that we can represent the probability of each of the $2D$ possible outcomes as follows:

$$\begin{aligned} \Pr(i) &= p\mathrm{Tr}(P_i \rho) \\ \Pr(j) &= (1-p)\mathrm{Tr}(\tilde{P}_j \rho). \end{aligned}$$

Generalized Measurements

Recall from standard Axiom 2 that the primitives of a measurement, associated with a self-adjoint operator A, are the orthogonal projectors P_a (onto distinct, possibly degenerate, eigenspaces of A) in the spectral decomposition of A.
Any set of orthogonal projectors $\{P_a\}$, satisfying $\sum_a P_a = \mathbb{1}$ is called a *projector valued measure*, or PVM for short.

- We can construct measurements that have a more general structure than a PVM in two different ways.
  - First, we can build up a more general measurement by considering classical probabilistic mixtures of PVM measurements.
  - Second, we can consider what structure occurs when look at the "reduced measurement" obtained from different kinds of PVM measurement on an extended Hilbert space.

Generalized Measurements

Recall from standard Axiom 2 that the primitives of a measurement, associated with a self-adjoint operator A, are the orthogonal projectors P_a (onto distinct, possibly degenerate, eigenspaces of A) in the spectral decomposition of A.
Any set of orthogonal projectors $\{P_a\}$, satisfying $\sum_a P_a = \mathbb{1}$ is called a *projector valued measure*, or PVM for short.

- We can construct measurements that have a more general structure than a PVM in two different ways.
  - First, we can build up a more general measurement by considering classical probabilistic mixtures of PVM measurements.
  - Second, we can consider what structure occurs when look at the "reduced measurement" obtained from different kinds of PVM measurement on an extended Hilbert space.

Mixtures of Projector Valued Measures

Suppose we have a device which performs the first PVM at random with probability p and the second with probability $1 - p$.

- Given a preparation ρ on a D-dimensional Hilbert space, from Axiom 2 we know that we can represent the probability of each of the $2D$ possible outcomes as follows:

$$\begin{aligned} \Pr(i) &= p\mathrm{Tr}(P_i\rho) \\ \Pr(j) &= (1-p)\mathrm{Tr}(\tilde{P}_j\rho). \end{aligned}$$

Mixtures of Projector Valued Measures

Suppose we have a device which performs the first PVM at random with probability p and the second with probability $1 - p$.

- Given a preparation ρ on a D-dimensional Hilbert space, from Axiom 2 we know that we can represent the probability of each of the $2D$ possible outcomes as follows:

$$\begin{aligned} \mathrm{Pr}(i) &= p\,\mathrm{Tr}(P_i\rho) \\ \mathrm{Pr}(j) &= (1-p)\,\mathrm{Tr}(\tilde{P}_j\rho). \end{aligned}$$

Mixtures of Projector Valued Measures

Suppose we have a device which performs the first PVM at random with probability p and the second with probability $1 - p$.

- Given a preparation ρ on a D-dimensional Hilbert space, from Axiom 2 we know that we can represent the probability of each of the $2D$ possible outcomes as follows:

$$\begin{aligned} \mathrm{Pr}(i) &= p\mathrm{Tr}(P_i \rho) \\ \mathrm{Pr}(j) &= (1-p)\mathrm{Tr}(\tilde{P}_j \rho). \end{aligned}$$

Mixtures of Projector Valued Measures

- Let $E_\nu = pP_i$ for $\nu = i$ and $E_\nu = (1-p)\tilde{P}_j$ for $\nu = D + j$.
- Then we can describe the probabilities of the $2D$ possible outcomes with the simple formula

$$\mathrm{Pr}(\nu) = \mathrm{Tr}(E_\nu \rho),$$

where these new operators satisfy:

$$\sum_\nu E_\nu = \mathbb{1}$$
$$E_\nu \geq \text{foreach}\nu.$$

- Note that when $p \in (0,1)$ the operators $\{E_\nu\}$ are not projectors.

PVMs on an Extended Hilbert Space

Suppose instead now that we have a composite system represented by the state $\rho_A \otimes \rho_B$ and we perform a *joint measurement* of both systems.

- This is represented by a PVM $\{P_\nu\}$ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$, with the usual properties $P_\nu P'_\nu = P_\nu \delta_{\nu\nu'}$ and $\sum_\nu P_\nu = \mathbb{1}$, and where the Greek index run from 1 to $K \leq MN = \dim(\mathcal{H}_A)\dim(\mathcal{H}_B)$.
- The probability of outcome ν is

$$\Pr(\nu) = \mathrm{Tr}[P_\nu(\rho_A \otimes \rho_B)] = \sum_{ij=1}^{M} \sum_{kl=1}^{N} (P_\nu)_{ik,jl} (\rho_A)_{ji} (\rho_B)_{lk},$$

where

$$(P_\nu)_{ik,jl} = \langle i| \otimes \langle k|(P_\nu)|j\rangle \otimes |l\rangle$$

PVMs on an Extended Hilbert Space

Suppose instead now that we have a composite system represented by the state $\rho_A \otimes \rho_B$ and we perform a *joint measurement* of both systems.

- This is represented by a PVM $\{P_\nu\}$ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$, with the usual properties $P_\nu P'_\nu = P_\nu \delta_{\nu\nu'}$ and $\sum_\nu P_\nu = \mathbb{1}$, and where the Greek index run from 1 to $K \leq MN = \dim(\mathcal{H}_A)\dim(\mathcal{H}_B)$.
- The probability of outcome ν is

$$\Pr(\nu) = \mathrm{Tr}[P_\nu(\rho_A \otimes \rho_B)] = \sum_{ij=1}^{M}\sum_{kl=1}^{N}(P_\nu)_{ik,jl}(\rho_A)_{ji}(\rho_B)_{lk},$$

where

$$(P_\nu)_{ik,jl} = \langle i| \otimes \langle k|(P_\nu)|j\rangle \otimes |l\rangle$$

PVMs on a Composite System

- The above measurement can be expressed in the form

$$\Pr(\nu) = \mathrm{Tr}[E_\nu \rho_A]$$

where $(E_\nu)_{ij} = \sum_{kl} (P_\nu)_{ik,jl} (\rho_B)_{lk}$ is an operator acting on $\mathcal{H}_A$.

- It is easy to see that the measurement operators E_ν satisfy

$$\sum_\nu E_\nu = 1_A$$
$$E_\nu \geq 0.$$

which are the same conditions we found for mixtures of PVMs.

PVMs on a Composite System

- Note that the measurement operators E_ν are not necessarily orthogonal (this is an important difference from a PVM) and hence that the number of elements in the set $\{E_\nu\}$ may be greater than the Hilbert space dimension.
- Indeed the measurement operators $\{E_\nu\}$ can also form a continuous set.

PVMs on a Composite System

Another important measurement paradigm is the following:

- In order to measure a property of system A, prepared in state ρ_A, we allow it to interact **in a controlled way** with another system B, which is initially prepared in some known state $\rho_B = |0\rangle_B \langle 0|_B$.
- We then perform a measurement on the system B alone.
- This paradigm models the important case of coupling the system to an *apparatus* which is, in turn, observed directly.

$$E_\nu \geq 0 \quad \forall \; \nu.$$

$$U \rho_A \otimes \rho_B U^\dagger$$

$$=$$

$$E_\nu \geq 0 \quad \forall \ \nu .$$

$$U \rho_A \otimes \rho_B U^\dagger$$

$$=$$

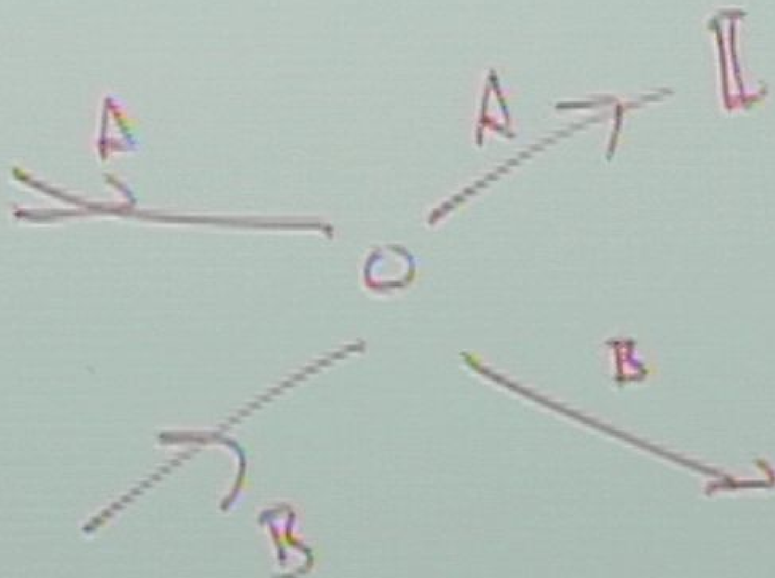
$$\rho'_{AB}$$

PVMs on a Composite System

- In this measurement method, not only do we gain information about the initial state of system A, but we can deduce also something about the state of system B *after the measurement*.
- That is, this paradigm provides a *filtering type-measurement* of system A, which is a method of *preparing a known state*.
- Note this paradigm is a model for the kind of measurement von Neumann considered (ie, the Compton experiment set-up) when he deduced the necessity of introducing the projection postulate as a *dynamical process* associated with measurement.

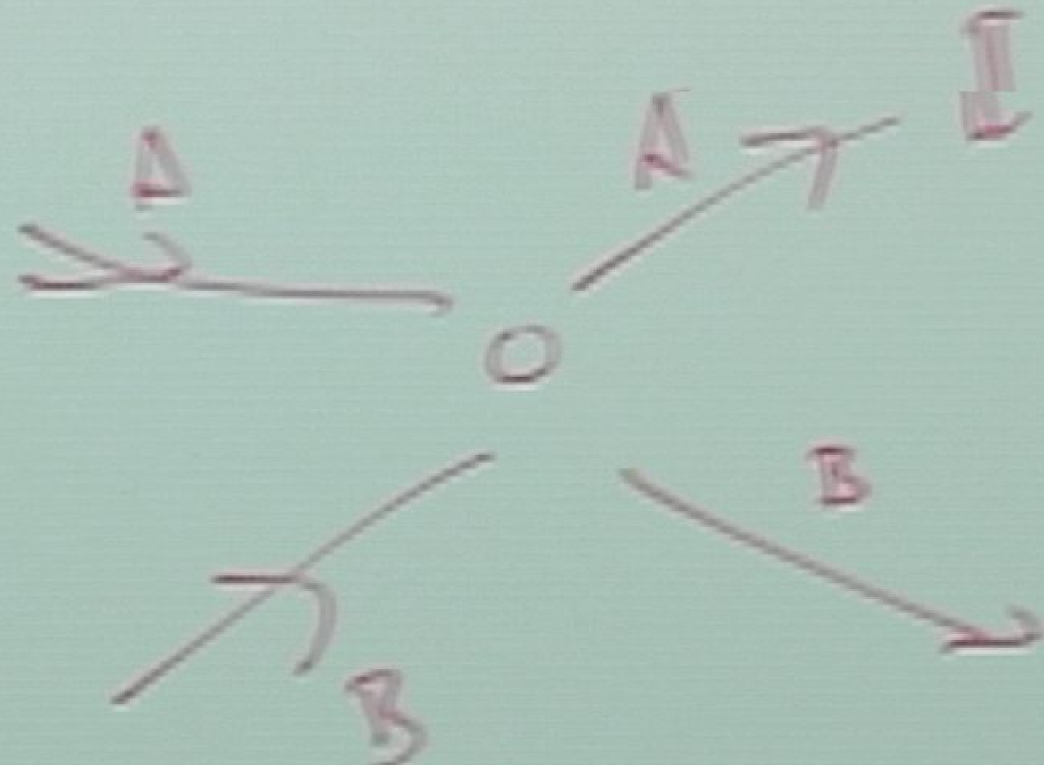
PVMs on a Composite System

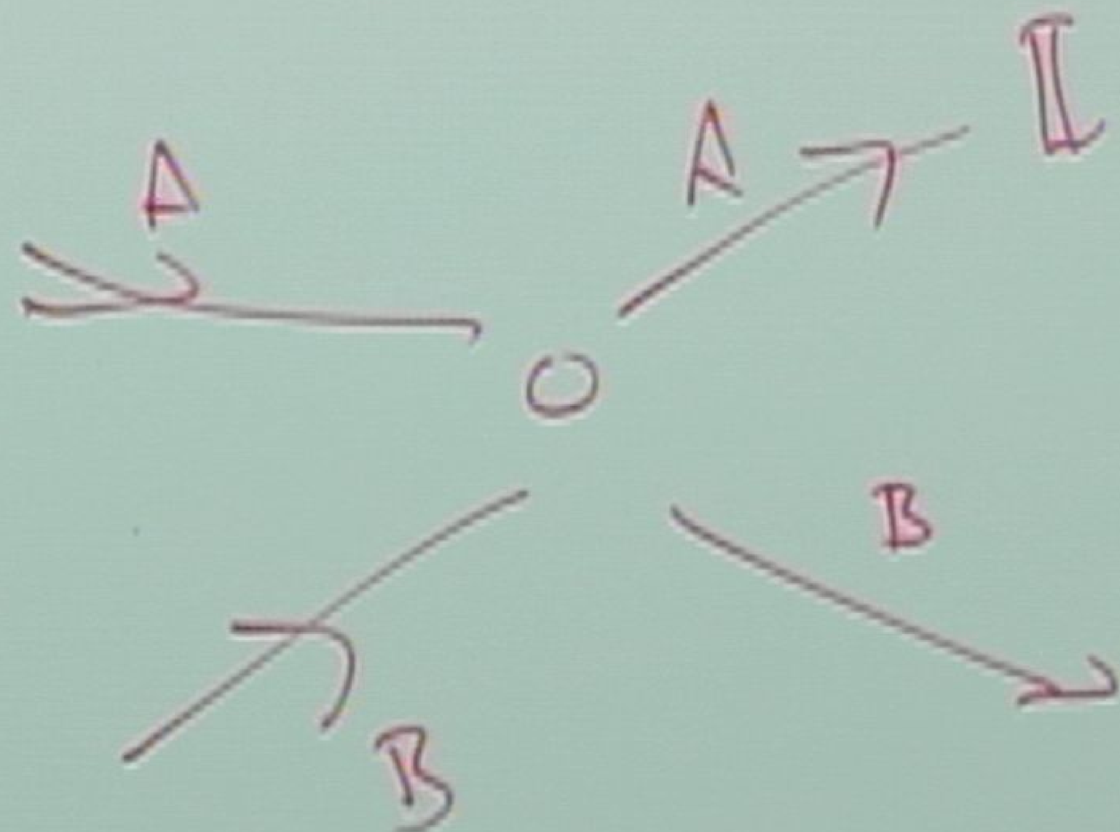
- In this measurement method, not only do we gain information about the initial state of system A, but we can deduce also something about the state of system B *after the measurement.*
- That is, this paradigm provides a *filtering type-measurement* of system A, which is a method of *preparing a known state.*
- Note this paradigm is a model for the kind of measurement von Neumann considered (ie, the Compton experiment set-up) when he deduced the necessity of introducing the projection postulate as a *dynamical process* associated with measurement.



$$E_\nu \geq 0 \quad \forall \nu$$

$U \rho \otimes \rho \, U^\dagger$


A
A
B
B
O


A
A
II
O
B
B

PVMs on a Composite System

How can we represent this process as a measurement operator acting on system A alone?

- Applying Axiom 2, we represent the **direct measurement** of the apparatus system with the PVM $\{P_m\}$, where $m = 1, \ldots, K$ with $K \leq N = \dim(\mathcal{H}_B)$.
- This measurement can be expressed on the joint system as the PVM $\{\mathbb{1}_A \otimes P_m\}$.

PVMs on a Composite System

How can we represent this process as a measurement operator acting on system A alone?

- Applying Axiom 2, we represent the **direct measurement** of the apparatus system with the PVM $\{P_m\}$, where $m = 1, \ldots, K$ with $K \leq N = \dim(\mathcal{H}_B)$.
- This measurement can be expressed on the joint system as the PVM $\{\mathbb{1}_A \otimes P_m\}$.

PVMs on a Composite System

- Let U be an arbitrary unitary operator that couples the two systems.
- Then the probability of outcome m is

$$\begin{aligned}\Pr(m) &= \mathrm{Tr}[(\mathbb{1}_A \otimes P_m)U(\rho_A \otimes \rho_B)U^\dagger] \\ &= \sum_i \langle i|_A \langle m|_B U|0\rangle_B \rho_A \langle 0|_B U^\dagger|i\rangle_A|m\rangle_B \\ &= \mathrm{Tr}[A_{m0}\rho_A A_{m0}^\dagger] = \mathrm{Tr}[E_m\rho_A],\end{aligned}$$

where we have used the cyclic property of the trace and defined $E_m \equiv A_{m0}^\dagger A_{m0}$.
- The operators E_m are positive (semi-definite) operators that act only on $\mathcal{H}_A$ and satisfy the properties: (i) $E_m \geq 0$ and (ii) $\sum_m E_m = \mathbb{1}_A$.
- These are the same conditions on the operators E_ν that we found previously.

PVMs on a Composite System

- Let U be an arbitrary unitary operator that couples the two systems.
- Then the probability of outcome m is

$$\begin{aligned}\Pr(m) &= \mathrm{Tr}[(\mathbb{1}_A \otimes P_m)U(\rho_A \otimes \rho_B)U^\dagger] \\ &= \sum_i \langle i|_A \langle m|_B U|0\rangle_B \rho_A \langle 0|_B U^\dagger |i\rangle_A |m\rangle_B \\ &= \mathrm{Tr}[A_{m0}\rho_A A_{m0}^\dagger] = \mathrm{Tr}[E_m \rho_A],\end{aligned}$$

  where we have used the cyclic property of the trace and defined $E_m \equiv A_{m0}^\dagger A_{m0}$.
- The operators E_m are positive (semi-definite) operators that act only on $\mathcal{H}_A$ and satisfy the properties: (i) $E_m \geq 0$ and (ii) $\sum_m E_m = \mathbb{1}_A$.
- These are the same conditions on the operators E_ν that we found previously.

Neumark's Theorem

We've seen three measurements paradigms which motivate the following definition and axiom:

Definition (Discrete POVM)

A discrete *positive operator valued measure* (POVM) is a set of operators $\{E_\nu\}$ satisfying:

(i) $E_\nu \geq 0$ for each $\nu \in \{1, 2, \dots\}$.

(ii) $\sum_\nu E_\nu = \mathbb{1}$.

Generalized Axiom 2 (Discrete Case): A measurement procedure with discrete outcomes is represented by a discrete POVM $\{E_\nu\}$, and the probability of observing outcome ν, given any preparation ρ, is

$$\Pr(\nu) = \mathrm{Tr}(E_\nu \rho).$$

Continuous Outcome POVMs

We can generalize the preceding to the case of continuous outcomes:

- Let Ω be a non-empty set and $\mathcal{F}$ be a σ-algebra of subsets of Ω so that $(\Omega, \mathcal{F})$ forms a measure space.
- Those unfamiliar with measure spaces can just think of Ω as a space of possible outcomes, e.g., the real line, and of $\mathcal{F}$ as the measurable subsets of Ω, e.g., arbitrary intervals on the real line.

Neumark's Theorem

We've seen three measurements paradigms which motivate the following definition and axiom:

Definition (Discrete POVM)

A discrete *positive operator valued measure* (POVM) is a set of operators $\{E_\nu\}$ satisfying:

(i) $E_\nu \geq 0$ for each $\nu \in \{1, 2, \dots\}$.

(ii) $\sum_\nu E_\nu = \mathbb{1}$.

Generalized Axiom 2 (Discrete Case): A measurement procedure with discrete outcomes is represented by a discrete POVM $\{E_\nu\}$, and the probability of observing outcome ν, given any preparation ρ, is

$$\Pr(\nu) = \mathrm{Tr}(E_\nu \rho).$$

Neumark's Theorem

We've seen three measurements paradigms which motivate the following definition and axiom:

Definition (Discrete POVM)

A discrete *positive operator valued measure* (POVM) is a set of operators $\{E_\nu\}$ satisfying:

(i) $E_\nu \geq 0$ for each $\nu \in \{1, 2, \ldots\}$.

(ii) $\sum_\nu E_\nu = \mathbb{1}$.

Generalized Axiom 2 (Discrete Case): A measurement procedure with discrete outcomes is represented by a discrete POVM $\{E_\nu\}$, and the probability of observing outcome ν, given any preparation ρ, is

$$\Pr(\nu) = \mathrm{Tr}(E_\nu \rho).$$

Neumark's Theorem

We've seen three measurements paradigms which motivate the following definition and axiom:

Definition (Discrete POVM)

A discrete *positive operator valued measure* (POVM) is a set of operators $\{E_\nu\}$ satisfying:

(i) $E_\nu \geq 0$ for each $\nu \in \{1, 2, \dots\}$.

(ii) $\sum_\nu E_\nu = \mathbb{1}$.

Generalized Axiom 2 (Discrete Case): A measurement procedure with discrete outcomes is represented by a discrete POVM $\{E_\nu\}$, and the probability of observing outcome ν, given any preparation ρ, is

$$\Pr(\nu) = \mathrm{Tr}(E_\nu \rho).$$

Continuous Outcome POVMs

We can generalize the preceding to the case of continuous outcomes:

- Let Ω be a non-empty set and $\mathcal{F}$ be a σ-algebra of subsets of Ω so that $(\Omega, \mathcal{F})$ forms a measure space.
- Those unfamiliar with measure spaces can just think of Ω as a space of possible outcomes, e.g., the real line, and of $\mathcal{F}$ as the measurable subsets of Ω, e.g., arbitrary intervals on the real line.

Continuous Outcome POVMs

We can generalize the preceding to the case of continuous outcomes:

- Let Ω be a non-empty set and $\mathcal{F}$ be a σ-algebra of subsets of Ω so that $(\Omega, \mathcal{F})$ forms a measure space.
- Those unfamiliar with measure spaces can just think of Ω as a space of possible outcomes, e.g., the real line, and of $\mathcal{F}$ as the measurable subsets of Ω, e.g., arbitrary intervals on the real line.

POVM - general definition

Definition

A *positive operator valued measure* (POVM) $E : \mathcal{F} \to \mathcal{L}(\mathcal{H})$ is defined by the properties:

(i) $E(X) \geq 0$ for all $X \in \mathcal{F}$

(ii) $E(\Omega) = \mathbb{1}$

(iii) $E(\bigcup_i X_i) = \sum_i E(X_i)$ for all disjoint sequences $\{X_i\} \subset \mathcal{F}$

POVM as a continuous PVM

If the POVM elements satisfy $E(X) = E(X)^2$ for all $X \in \mathcal{F}$ then the POVM reduces to a PVM: in which case the set Ω may be taken without loss of generality to be the real line $\mathbb{R}$ and the σ-algebra consists of the $\mathcal{B}(\mathbb{R})$, the Borel subsets of $\mathbb{R}$.
As a result we recover a continuous PVM as a one-parameter family of projection operators.
That is, in terms of the Borel sets we can define a PVM $E : \mathcal{B}(\mathbb{R}) \to \mathcal{L}(\mathcal{H})$ by the conditions:

(i) $E(X) = E^2(X)$ for all $X \in \mathcal{B}(\mathbb{R})$

(ii) $E(\mathbb{R}) = \mathbb{1}$

(iii) $E(\bigcup_i X_i) = \sum_i E(X_i)$ for all disjoint sequences $\{X_i\} \subset \mathcal{B}(\mathbb{R})$.

Note that i) implies $E(X \bigcap Y) = E(X)E(Y)$ for all $X, Y \in \mathcal{F}$ and also implies that $E(X) = E^\dagger(X)$.

Generalized Measurements

This gives a more general version of Axiom 2:
Generalized Axiom 2: Any measurement procedure can be represented by a POVM $E : \mathcal{F} \to \mathcal{L}(\mathcal{H})$, and tor any preparation ρ, the probability of observing an outcome $X \in \mathcal{F}$ is

$$\Pr(X) = \mathrm{Tr}(E(X)\rho).$$

You can think of outcome X as corresponding to a question like:
Was the position q found to be within the interval $X \subseteq \mathbb{R}$?

Non-uniqueness of purifications

So we have seen that physically realizable cases of generalized measurement correspond to a POVM measurement.

- But does every POVM measurement correspond to some physically realizable measurement, and, in particular, to some realizable PVM measurement?

Non-uniqueness of purifications

So we have seen that physically realizable cases of generalized measurement correspond to a POVM measurement.

- But does every POVM measurement correspond to some physically realizable measurement, and, in particular, to some realizable PVM measurement?

Neumark's Theorem

The answer to this question is given by *Neumark's theorem* (actually a simplified version of it):

Theorem (Neumark)

For any POVM $\{E\}$ acting on a Hilbert space $\mathcal{H}_A$ there exists a PVM $\{P\}$ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$ and a state $|\phi\rangle\langle\phi|$ acting on $\mathcal{H}_B$ such that

$$\mathrm{Tr}[(\rho \otimes |\phi\rangle\langle\phi|)P(X)] = \mathrm{Tr}[E(X)\rho]$$

for any state ρ acting on $\mathcal{H}_A$ and any $X \in \mathcal{F}$. The PVM can always be expressed in the form $U^\dagger(\mathbb{1}_A \otimes P)U$, i.e., the PVM P acts only on $\mathcal{H}_B$.

Neumark's Theorem

Recall in the case of generalized preparations, which were given by density operators, the sets of proper and improper mixtures were mathematically equivalent (and hence operationally indistinguishable).
This is **not** the case for POVM measurements.

- That is, we can define proper POVMs as those obtained from convex combinations of PVMs.
- Similarly, we can define improper POVMs as those obtained from a PVM measurement on an extended Hilbert space.
- From Neumark's theorem we know that proper POVMs must be a subset of improper POVMs. However, they are a strict subset.
- This means that operationally implementing some POVM measurements requires access to (and control over) a larger Hilbert space.

Example of improper POVM

Here is a simple example of an *improper* POVM:

Example

Consider the *trine* given by the set of three projectors $|\chi_\nu\rangle\langle\chi_\nu|$ acting on $\mathbb{C}^2$ defined by:

$$(\sigma \cdot \mathbf{n}_\nu)\chi_\nu = \chi_\nu$$

where n_1, n_2 and n_3 denote three unit vectors making angles of 120 degrees with each other. Let $E_\nu = (2/3)|\chi_\nu\rangle\langle\chi_\nu|$.

The trine is the smallest possible POVM that is not a PVM.
It also can not be generated, mathematically or operationally, from taking convex combinations of PVMs.

Example of improper POVM

Here is a simple example of an *improper* POVM:

Example

Consider the *trine* given by the set of three projectors $|\chi_\nu\rangle\langle\chi_\nu|$ acting on $\mathbb{C}^2$ defined by:

$$(\sigma \cdot \mathbf{n}_\nu)\chi_\nu = \chi_\nu$$

where n_1, n_2 and n_3 denote three unit vectors making angles of 120 degrees with each other. Let $E_\nu = (2/3)|\chi_\nu\rangle\langle\chi_\nu|$.

The trine is the smallest possible POVM that is not a PVM.
It also can not be generated, mathematically or operationally, from taking convex combinations of PVMs.

Example of improper POVM

Here is a simple example of an *improper* POVM:

Example

Consider the *trine* given by the set of three projectors $|\chi_\nu\rangle\langle\chi_\nu|$ acting on $\mathbb{C}^2$ defined by:

$$(\sigma \cdot \mathbf{n}_\nu)\chi_\nu = \chi_\nu$$

where n_1, n_2 and n_3 denote three unit vectors making angles of 120 degrees with each other. Let $E_\nu = (2/3)|\chi_\nu\rangle\langle\chi_\nu|$.

The trine is the smallest possible POVM that is not a PVM.
It also can not be generated, mathematically or operationally, from taking convex combinations of PVMs.

Example of improper POVM

Here is a simple example of an *improper* POVM:

Example

Consider the *trine* given by the set of three projectors $|\chi_\nu\rangle\langle\chi_\nu|$ acting on $\mathbb{C}^2$ defined by:

$$(\sigma \cdot \mathbf{n}_\nu)\chi_\nu = \chi_\nu$$

where n_1, n_2 and n_3 denote three unit vectors making angles of 120 degrees with each other. Let $E_\nu = (2/3)|\chi_\nu\rangle\langle\chi_\nu|$.

The trine is the smallest possible POVM that is not a PVM.
It also can not be generated, mathematically or operationally, from taking convex combinations of PVMs.

Generalized Transformations

As with measurements and states, there are two ways to construct *generalized transformations*:

- By taking convex combinations of unitary transformations.
- By considering a unitary acting on an extended Hilbert space and then tracing out the ancillary system.

Generalized Transformations

As with measurements and states, there are two ways to construct *generalized transformations*:

- By taking convex combinations of unitary transformations.
- By considering a unitary acting on an extended Hilbert space and then tracing out the ancillary system.

Mixtures of Unitary Operators

Consider a procedure whereby we subject a preparation ρ to transformation U_j with probability p_j.

- The effective transformation is then given by a convex combination of unitary operators
$$\Lambda(\rho) = \sum_j p_j U_j \rho U_j^\dagger .$$
- Clearly this map is in general non-unitary, but it always preserves the trace of the input state. Specifically, if $\rho' = \Lambda(\rho)$, from the linearity of the trace we see that
$$\mathrm{Tr}\rho' = \sum_j p_j \mathrm{Tr}(U_j \rho_j U_j^\dagger) = 1.$$
- By convexity, the output state will remain a positive (semi-definite) operator. Hence the map Λ is called *positive*.

Unitary acting on Extended Hilbert Space

Consider the effect of a unitary operator on extended Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ acting on an *uncorrelated* initial state

$$\rho_A(t) = \Lambda_t(\rho_A(0)) \equiv \mathrm{Tr}_B[U(t)\rho_A(0) \otimes |0\rangle_B\langle 0|_B U^\dagger(t)] = \sum_k A_k \rho(0) A_k^\dagger$$

where $A_k = \langle k|U(t)|0\rangle$ is a linear operator acting on $\mathcal{H}_A$.

- By linearity, a decomposition of the same form is obtained also in the case that the initial environment state is an arbitrary mixed state ρ_B.
- The requirement that the initial state is uncorrelated is strictly stronger than the requirement that the state be separable.

Unitary acting on Extended Hilbert Space

- Clearly this map preserves the trace of the output state:

$$\mathrm{Tr}[\sum_k A_k \rho_A A_k^\dagger] = \mathrm{Tr}[U\rho_A(0) \otimes |0\rangle_B \langle 0|_B U^\dagger] = 1.$$

- Because this holds for any ρ_A, from the cyclic property of the trace we deduce that

$$\sum_k A_k^\dagger A_k = \mathbb{1}_A.$$

Hence it is easy to see from the properties of the partial trace that this map also guarantees the positivity of the reduced state.

Kraus Decomposition

Definition

The expression

$$\Lambda(\rho) = \sum_k A_k \rho(0) A_k^\dagger$$

subject to the constraint

$$\sum_k A_k^\dagger A_k = \mathbb{1}$$

is called a *Kraus decomposition* or an *operator-sum decomposition* of the map Λ, and the set of (bounded) linear operators $\{A_k\}$ are called Kraus operators.

For a map Λ constructed from a mixture of unitary operators, one choice for the Kraus operators is the unitary operators weighted by the appropriate probabilities.

Generalized Transformations

Definition

Any linear map Λ taking linear operators to linear operators is called a *superoperator*.

Definition

Any superoperator Λ representing a dynamical transformation on the space of quantum states is called a *quantum dynamical map*.

Remark: For some physicists these terms are used interchangeably.

Generalized Transformations

Any quantum dynamical map Λ_t describing the evolution of a quantum state over a time t

$$\rho(0) \rightarrow \rho(t) = \Lambda_t(\rho(0))$$

that is constructed by either of the above methods satisfies the following properties:

(i) Convex Linear: $p_1\rho_1(t) + p_2\rho_2(t) = \Lambda_t(p_1\rho_1(0) + p_2\rho_2(0))$ where $\rho_i(t) = \Lambda_t(\rho_i(0))$ and $p_i \geq 0$.

(ii) Completely positive: $\rho_{AB}(t) = \Lambda \otimes \mathbb{1}_B(\rho_{AB}(0))$ is positive if $\rho_{AB}(0)$ is positive – this guarantees that probabilities are positive (and hence real) - note that it is stronger than positivity because it guarantees that probabilities must be positive even when the map is acting on part of an extended system *provided that the initial state is uncorrelated between the two systems.*

(iii) Trace preserving: $\mathrm{Tr}(\rho(t)) = \mathrm{Tr}(\rho(0))$ – this guarantees that

Unitary acting on Extended Hilbert Space

Consider the effect of a unitary operator on extended Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ acting on an *uncorrelated* initial state

$$\rho_A(t) = \Lambda_t(\rho_A(0)) \equiv \mathrm{Tr}_B[U(t)\rho_A(0) \otimes |0\rangle_B\langle 0|_B U^\dagger(t)] = \sum_k A_k \rho(0) A_k^\dagger$$

where $A_k = \langle k|U(t)|0\rangle$ is a linear operator acting on $\mathcal{H}_A$.

- By linearity, a decomposition of the same form is obtained also in the case that the initial environment state is an arbitrary mixed state ρ_B.
- The requirement that the initial state is uncorrelated is strictly stronger than the requirement that the state be separable.

Generalized Transformations

Definition

Any linear map Λ taking linear operators to linear operators is called a *superoperator*.

Definition

Any superoperator Λ representing a dynamical transformation on the space of quantum states is called a *quantum dynamical map*.

Remark: For some physicists these terms are used interchangeably.

Generalized Transformations

Any quantum dynamical map Λ_t describing the evolution of a quantum state over a time t

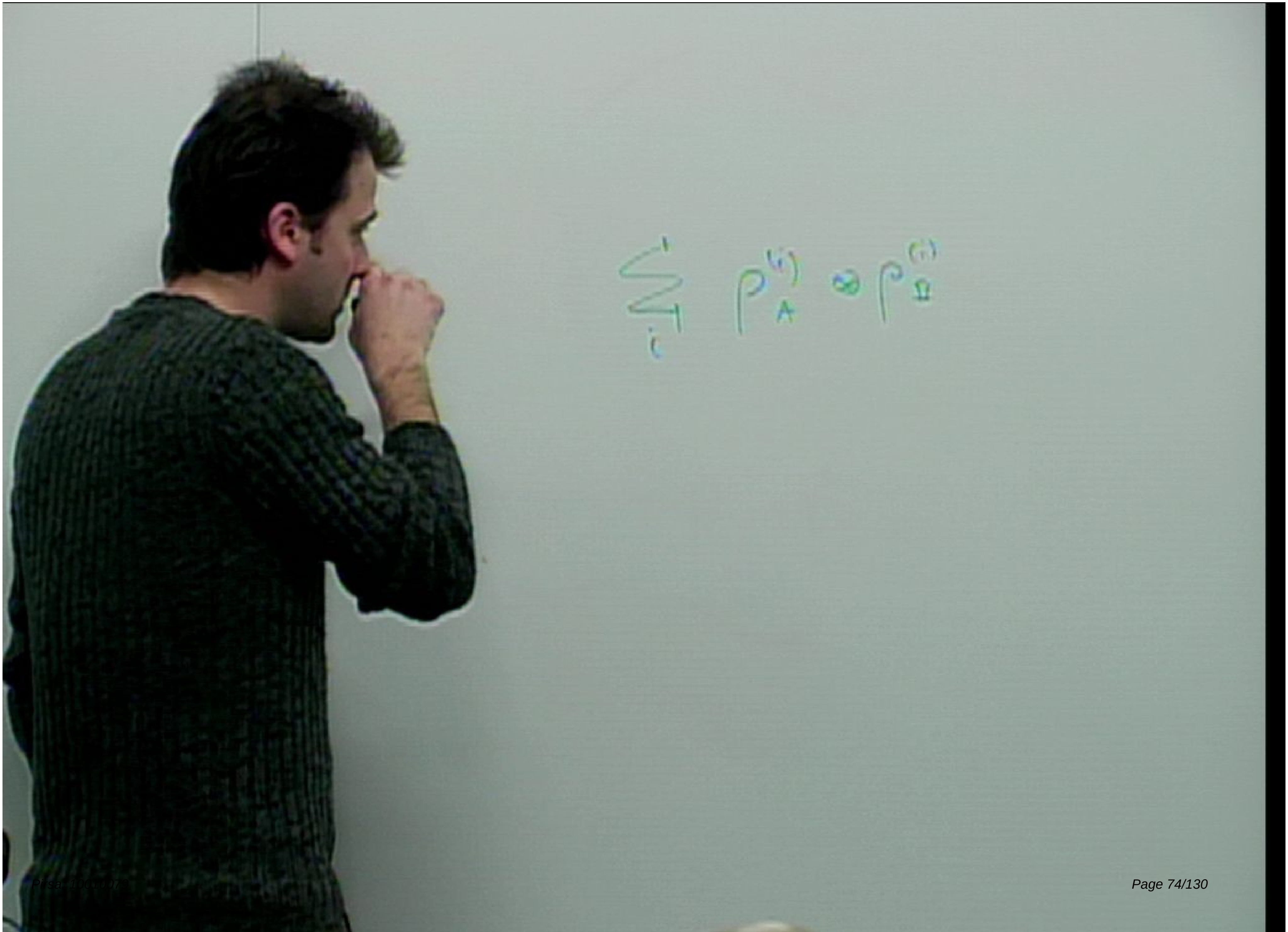
$$\rho(0) \rightarrow \rho(t) = \Lambda_t(\rho(0))$$

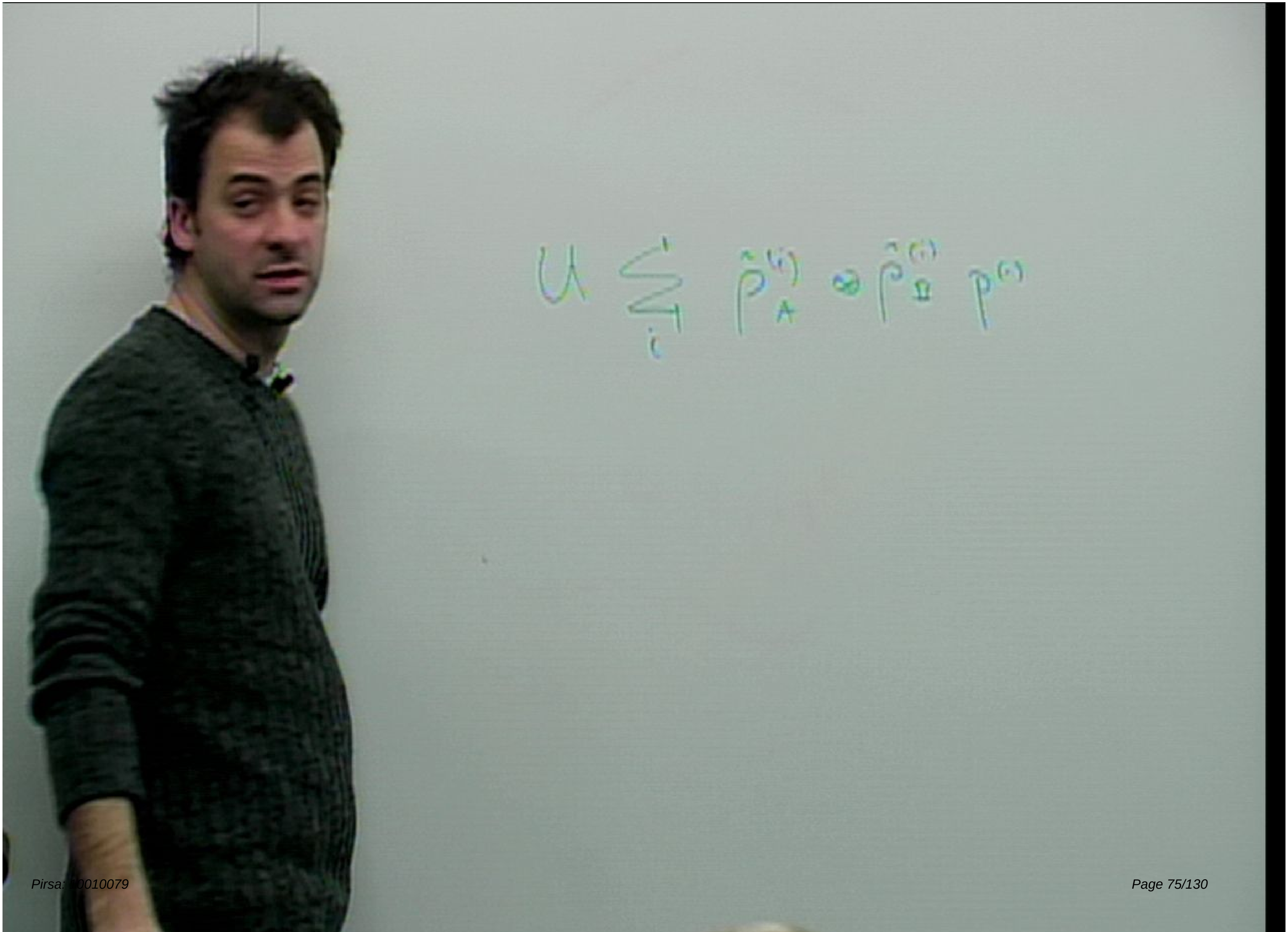
that is constructed by either of the above methods satisfies the following properties:

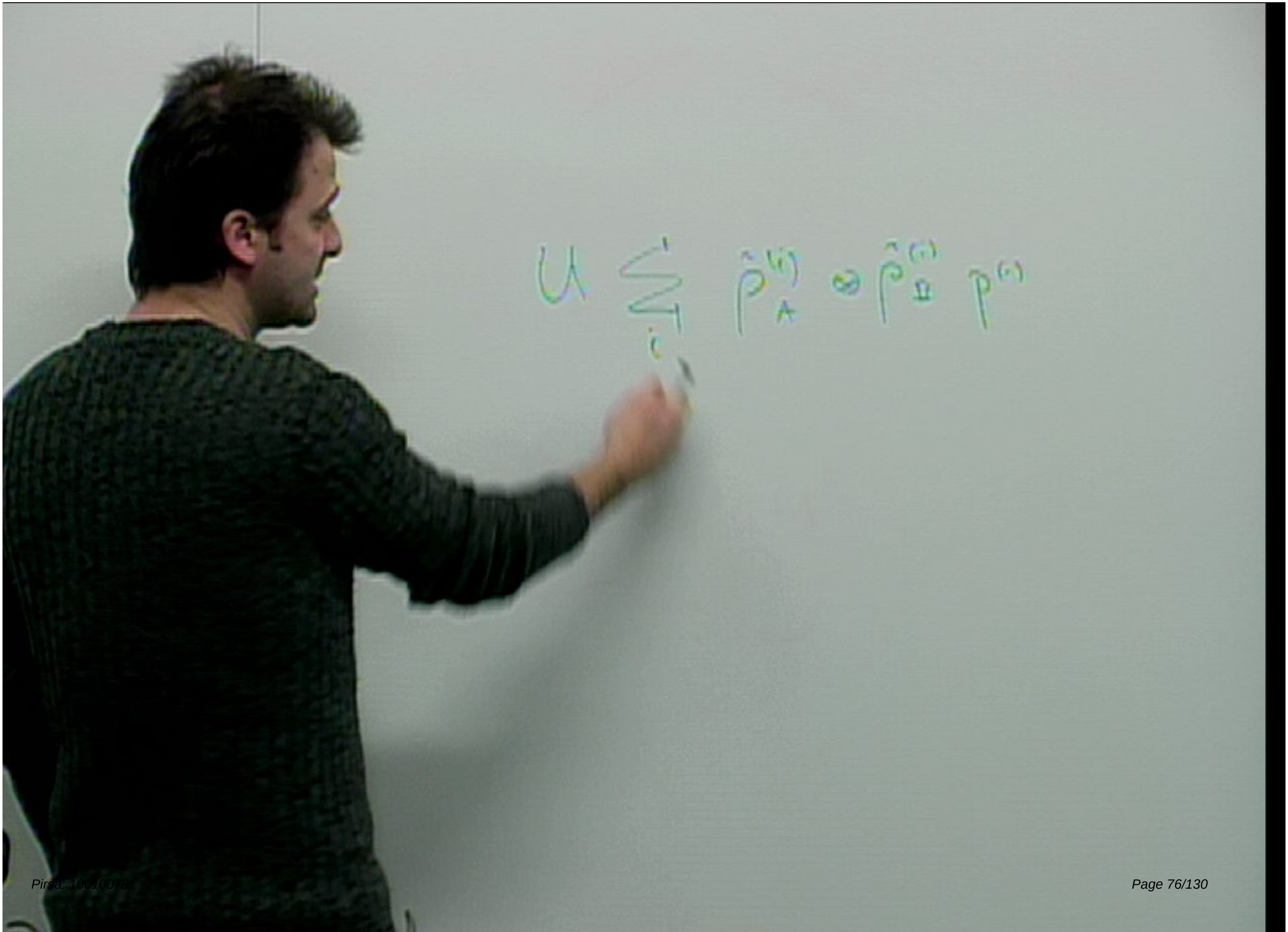
(i) Convex Linear: $p_1\rho_1(t) + p_2\rho_2(t) = \Lambda_t(p_1\rho_1(0) + p_2\rho_2(0))$ where $\rho_i(t) = \Lambda_t(\rho_i(0))$ and $p_i \geq 0$.

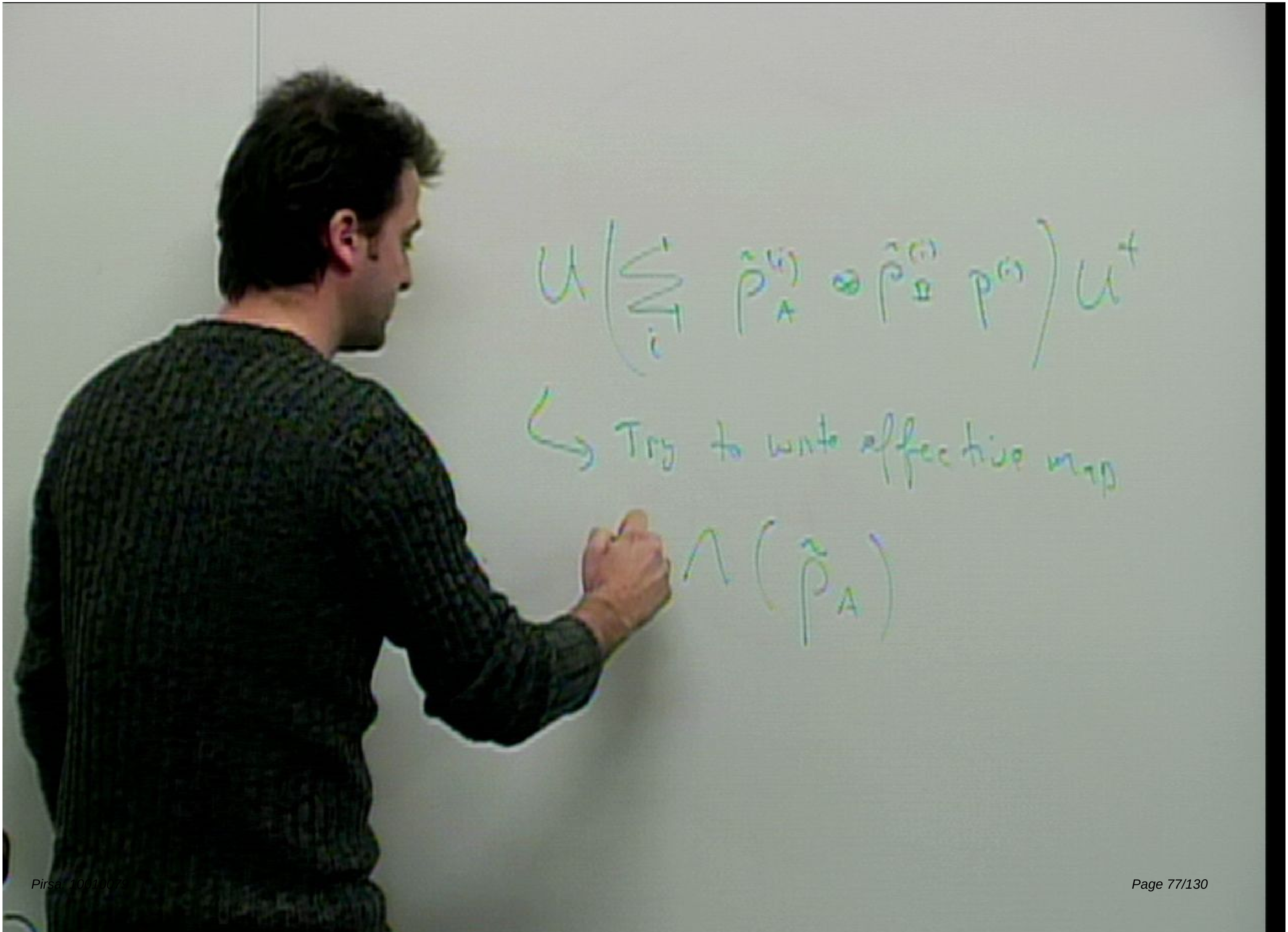
(ii) Completely positive: $\rho_{AB}(t) = \Lambda \otimes \mathbb{1}_B(\rho_{AB}(0))$ is positive if $\rho_{AB}(0)$ is positive – this guarantees that probabilities are positive (and hence real) - note that it is stronger than positivity because it guarantees that probabilities must be positive even when the map is acting on part of an extended system *provided that the initial state is uncorrelated between the two systems.*

(iii) Trace preserving: $\mathrm{Tr}(\rho(t)) = \mathrm{Tr}(\rho(0))$ – this guarantees that


$\sum_i \rho_A^{(i)} \otimes \rho_B^{(i)}$


$\mathcal{U} \sum_i \tilde{\rho}_A^{(i)} \otimes \tilde{\rho}_B^{(i)} \; p^{(i)}$




$U\left(\sum_i \tilde{\rho}_A^{(i)} \otimes \tilde{\rho}_\Omega^{(i)} p^{(i)}\right)U^\dagger$
Try to write effective map
$\Lambda(\tilde{\rho}_A)$

$$U\left(\sum_i \tilde{\rho}_A^{(i)} \otimes \tilde{\rho}_\Omega^{(i)}\, p^{(i)}\right)U^\dagger$$

↳ Try to write effective map

$$\tilde{\rho}_A' = \Lambda\left(\tilde{\rho}_A\right)$$

Generalized Transformations

- The complete positivity condition implies that when the map acts on part of composite system it still produces a positive operator.
- This requirement implies positivity but the converse does not hold.
- The partial transpose map is an example of a map that is positive but not completely positive.

Generalized Transformations

- The complete positivity condition implies that when the map acts on part of composite system it still produces a positive operator.
- This requirement implies positivity but the converse does not hold.
- The partial transpose map is an example of a map that is positive but not completely positive.

Generalized Transformations

Definition

A *completely positive map* (CP map) is a superoperator satisfying conditions (i) and (ii).

Definition

A *completely positive trace-preserving map* (CPTP map) is a superoperator satisfying conditions (i)-(iii).

Generalized Transformations

Any quantum dynamical map Λ_t describing the evolution of a quantum state over a time t

$$\rho(0) \rightarrow \rho(t) = \Lambda_t(\rho(0))$$

that is constructed by either of the above methods satisfies the following properties:

(i) Convex Linear: $p_1\rho_1(t) + p_2\rho_2(t) = \Lambda_t(p_1\rho_1(0) + p_2\rho_2(0))$ where $\rho_i(t) = \Lambda_t(\rho_i(0))$ and $p_i \geq 0$.

(ii) Completely positive: $\rho_{AB}(t) = \Lambda \otimes \mathbb{1}_B(\rho_{AB}(0))$ is positive if $\rho_{AB}(0)$ is positive – this guarantees that probabilities are positive (and hence real) - note that it is stronger than positivity because it guarantees that probabilities must be positive even when the map is acting on part of an extended system *provided that the initial state is uncorrelated between the two systems.*

(iii) Trace preserving: $\mathrm{Tr}(\rho(t)) = \mathrm{Tr}(\rho(0))$ – this guarantees that

Generalized Transformations

Definition

A *completely positive map* (CP map) is a superoperator satisfying conditions (i) and (ii).

Definition

A *completely positive trace-preserving map* (CPTP map) is a superoperator satisfying conditions (i)-(iii).

Kraus Representation Theorem

While the properties deduced from our derivation of quantum dynamical maps in terms of unitary operators *implied* the properties for our definition of a CPTP map, it turns out that any map satisfying these properties can also be identified with a unitary operator on an extended space. This is made explicit by the following representation theorem due to Kraus and the associated dilation theorem due to Stinespring:

Theorem (Kraus Representation Theorem)

A superoperator Λ is a CPTP map iff it admits an operator-sum decomposition.

Theorem (Stinespring Dilation Theorem)

Any CPTP map can be expressed as the reduced action of a unitary operator acting on an extended Hilbert space, where the initial state in the ancilla Hilbert space is uncorrelated with the initial system state.

Generalized Transformations

Generalized Axiom 4: Over any finite time, the dynamical transformation of a quantum system is described by a completely positive trace-preserving map.

- For finite dimensional systems, the maximum number of operators $\{A_k\}$ required to represent any CP map acting on $\mathcal{L}(\mathbb{C}^D)$ is D^2.
- Stinespring's representation is unique up to unitary transformations on the ancilla system.
- For finite dimensional state spaces the theorem also comes with a bound on the dimension of the ancilla system.

Proper vs Improper

We can also define a notion of proper and improper CPTP maps depending on whether they can be decomposed as a convex combination of unitary operators.

Definition

A CPTP map is called *unital* if it maps the identity operator to the identity operator.

- If the CPTP map is unital this implies the condition $\sum_k A_k A_k^\dagger = \mathbb{1}$ on the Kraus operators.
- Any proper CPTP must be unital.
- A simple example of a non-unital map is the spontaneous decay of an atom.

Hence it is clear that proper CPTP maps form a strict subset of all CPTP

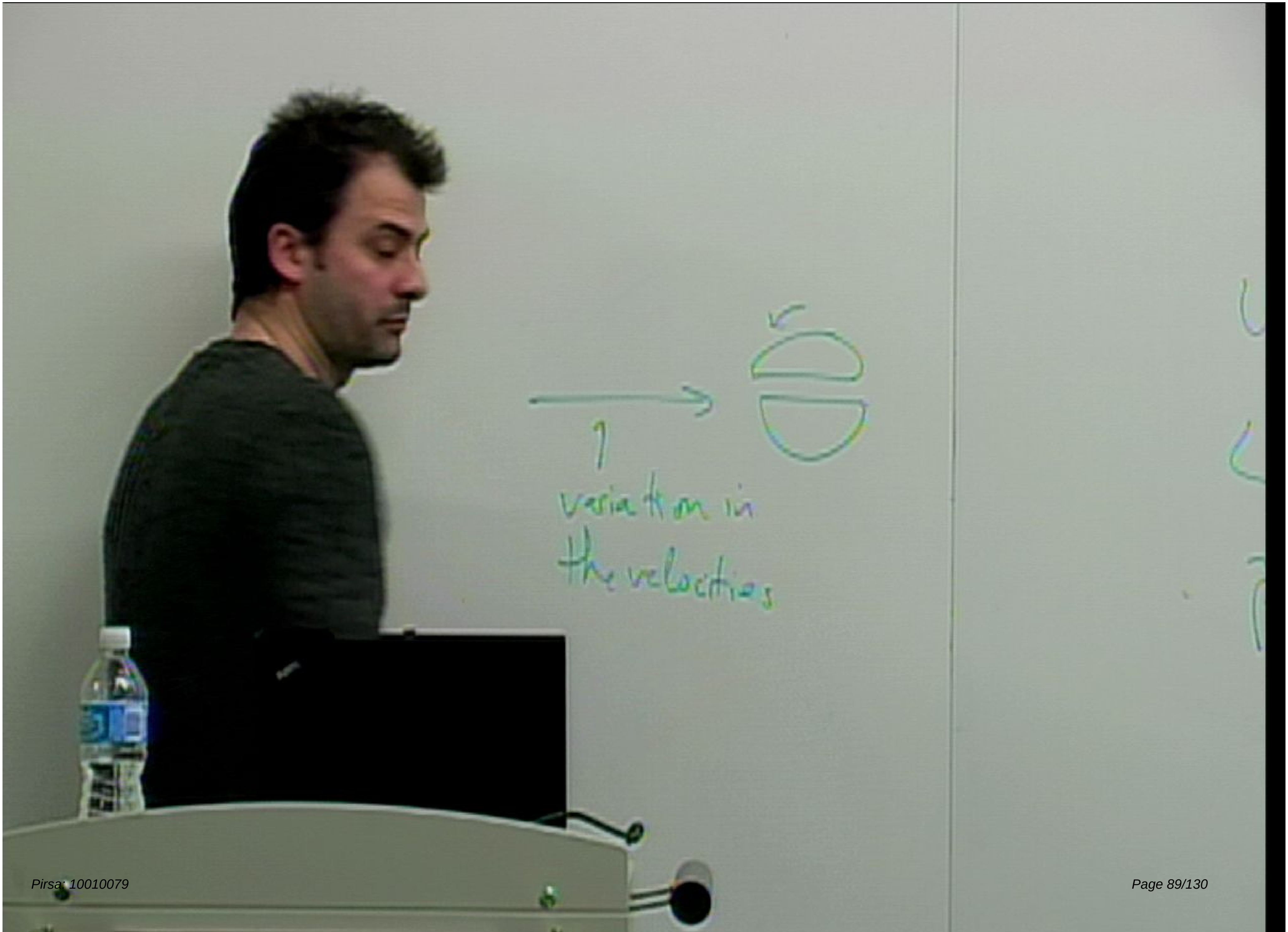
Generalized Transformations and Decoherence

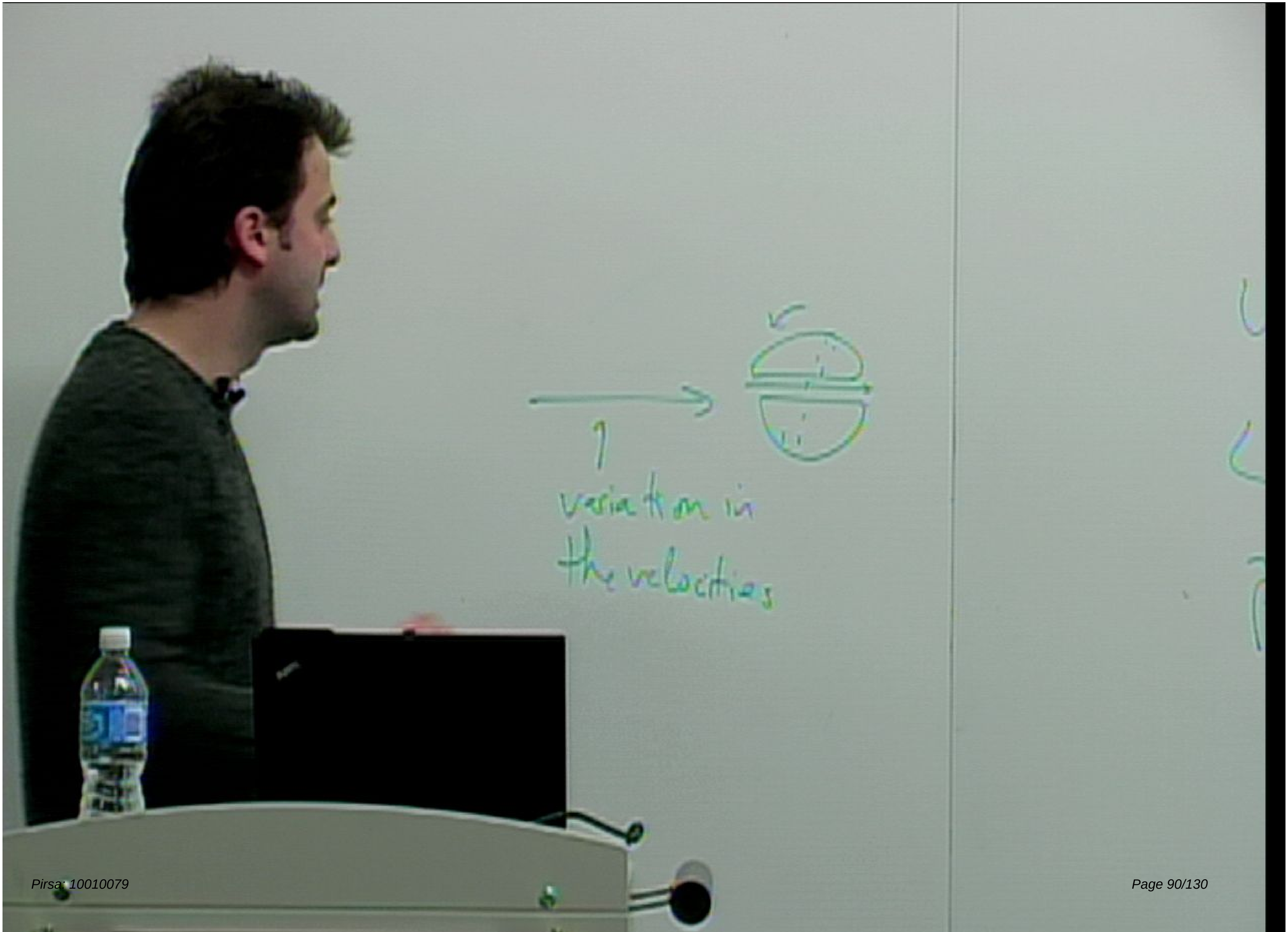
- The non-unitary transformations that occur in the form of CPTP maps often produce *decoherence*.
- Broadly speaking, *decoherence* is a dynamical process whereby the purity of the system state decreases.
- Often decoherence occurs due to coupling of the system of interest to an ancillary quantum system (often the uncontrolled "environment"), which is subsequently traced over (via the partial trace).
- In the quantum optics community, decoherence sometimes refers specifically to *de-phasing* - the attenuation of off-diagonal terms in some fixed basis without any changes to the diagonal terms (ie, the populations of each of the basis states).

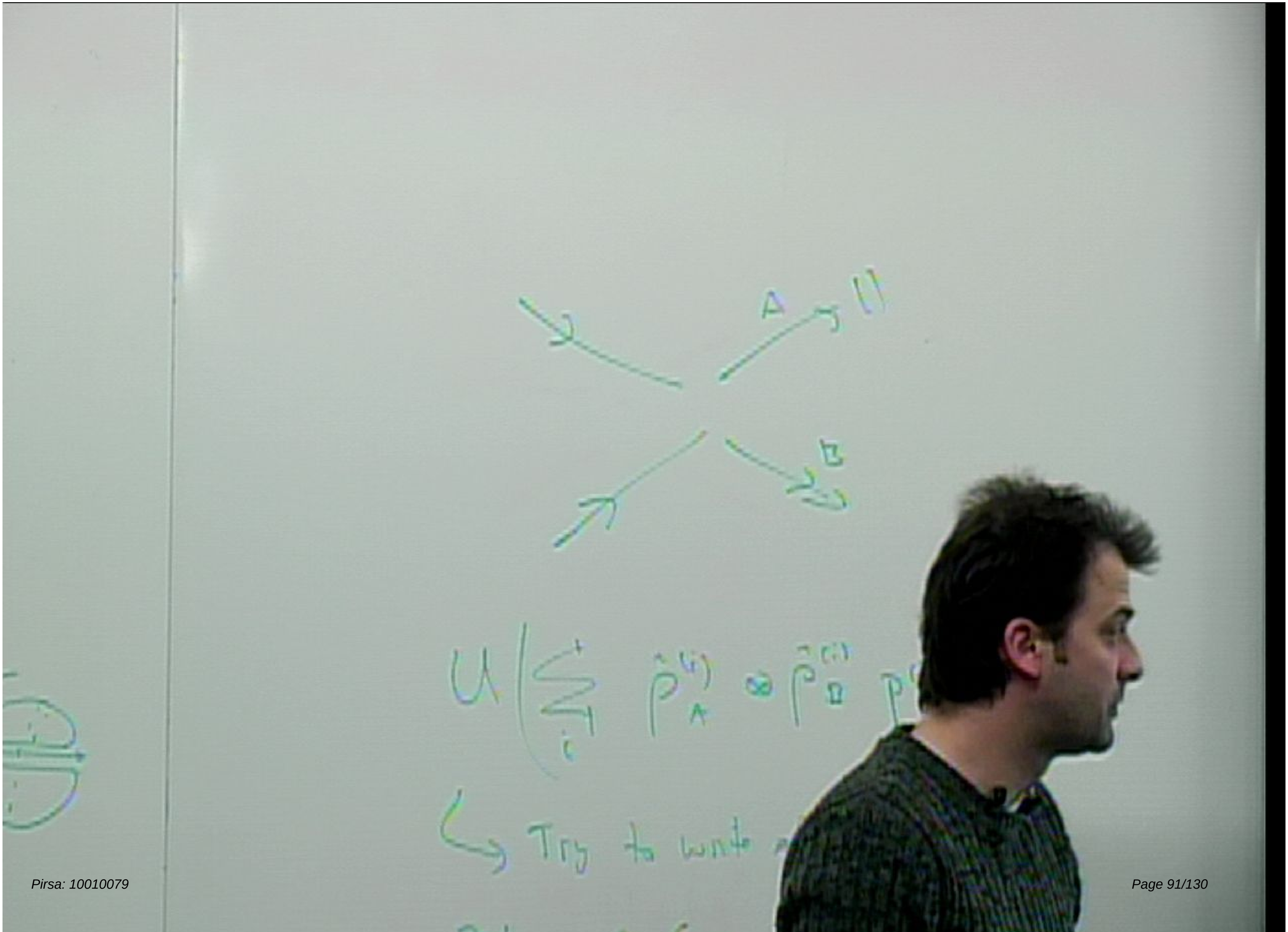
Generalized Transformation under Filtering Measurements

In the context of the standard axioms, it was necessary to postulate a second kind of transformation to describe the quantum state after an ideal "filtering" measurement.

- A ideal "filtering" type measurement was one where the outcome associated with the measurement is verifiable under repeated sequential measurements -
- However, some measurements are "destructive" and do not have this property, eg, absorbing a photon to measure its momentum. For such measurements the Born rule still applies, but the state update rule is not given by the projection postulate.


variation in
the velocities


variation in
the velocities


A
B
Try to write

Luders' Rule

Consider the ideal measurement of an observable $A = \sum_n a_n P_n$ (with eigenvalues a_n and associated projectors P_n) applied to a given preparation ρ that yields the outcome a_m.

- The post-measurement state ρ_m, conditional upon the outcome a_m, is determined by *Luders' rule*,

$$\rho \to \rho_m = \frac{P_m \rho P_m}{\mathrm{Tr}(P_m \rho)}.$$

  where the factor in the denominator is required for normalization.

- If the observed eigenvalue is non-degenerate, then P_m is rank-one, and the state $\rho_m = P_m = |\psi_m\rangle\langle\psi_m|$ will be pure - in this special case Luders' rule reduces to the projection postulate considered by von Neumann.

Luders' Rule

Consider the ideal measurement of an observable $A = \sum_n a_n P_n$ (with eigenvalues a_n and associated projectors P_n) applied to a given preparation ρ that yields the outcome a_m.

- The post-measurement state ρ_m, conditional upon the outcome a_m, is determined by *Luders' rule*,

$$\rho \to \rho_m = \frac{P_m \rho P_m}{\mathrm{Tr}(P_m \rho)}.$$

  where the factor in the denominator is required for normalization.

- If the observed eigenvalue is non-degenerate, then P_m is rank-one, and the state $\rho_m = P_m = |\psi_m\rangle\langle\psi_m|$ will be pure - in this special case Luders' rule reduces to the projection postulate considered by von Neumann.

Luders' Rule

Consider the ideal measurement of an observable $A = \sum_n a_n P_n$ (with eigenvalues a_n and associated projectors P_n) applied to a given preparation ρ that yields the outcome a_m.

- The post-measurement state ρ_m, conditional upon the outcome a_m, is determined by *Luders' rule*,

$$\rho \rightarrow \rho_m = \frac{P_m \rho P_m}{\mathrm{Tr}(P_m \rho)}.$$

  where the factor in the denominator is required for normalization.
- If the observed eigenvalue is non-degenerate, then P_m is rank-one, and the state $\rho_m = P_m = |\psi_m\rangle\langle\psi_m|$ will be pure - in this special case Luders' rule reduces to the projection postulate considered by von Neumann.

Luders' Rule

Of course it always possible to describe the system after measurement without post-selecting based on the outcome (ie, without conditioning on the observed eigenvalue), either by ignoring the outcome or else because the outcome may not be observable in practice.

- In this case we can accurately describe the post-measurement state by simply constructing the weighted classical mixture over the set of possible post-selected states,

$$\rho \to \rho' = \sum_m \mathrm{Tr}(P_m \rho) \frac{P_m \rho P_m}{\mathrm{Tr}(P_m \rho)} = \sum_m \mathrm{Tr}(P_m \rho P_m).$$

- The lack of post-selection usually leads to a loss of purity. For example, if the input state ρ is a pure state which is a coherent superposition over two eigenspaces P_m and $P_{m'}$, then the output state ρ' will be a mixture over those two eigenspaces.
- In this way measurement without post-selection leads to a form of

Luders' Rule

Of course it always possible to describe the system after measurement without post-selecting based on the outcome (ie, without conditioning on the observed eigenvalue), either by ignoring the outcome or else because the outcome may not be observable in practice.

- In this case we can accurately describe the post-measurement state by simply constructing the weighted classical mixture over the set of possible post-selected states,

$$\rho \to \rho' = \sum_m \mathrm{Tr}(P_m \rho) \frac{P_m \rho P_m}{\mathrm{Tr}(P_m \rho)} = \sum_m \mathrm{Tr}(P_m \rho P_m).$$

- The lack of post-selection usually leads to a loss of purity. For example, if the input state ρ is a pure state which is a coherent superposition over two eigenspaces P_m and $P_{m'}$, then the output state ρ' will be a mixture over those two eigenspaces.
- In this way measurement without post-selection leads to a form of

State Update Rule for Generalized Measurements

Recall from Neumark's theorem that any POVM measurement $\{E_k\}$ on a Hilbert space $\mathcal{H}_A$ can be expressed as a PVM measurement $\{P_k\}$ acting on an extended Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ with a state $|\phi\rangle\langle\phi|$ representing the initial state of $\mathcal{H}_B$. Specifically,

$$\mathrm{Tr}[(\rho \otimes |\phi\rangle\langle\phi|) P_k] = \mathrm{Tr}[\rho E_k]$$

for any state ρ acting on $\mathcal{H}_A$ and any outcome k.

- Here for simplicity of the analysis we will assume that the POVM has discrete outcomes.

State Update Rule for Generalized Measurements

Hence, in order to describe the post-selected quantum state after measurement under a POVM, we can simply apply the Luders rule to the description given by Neumark's theorem.
Specifically, we have

$$\rho \to \rho_k = \frac{\mathrm{Tr}_B(P_k(\rho \otimes |\phi\rangle\langle\phi|)P_k)}{\mathrm{Tr}[P_k \rho \otimes |\phi\rangle\langle\phi| P_k]}$$

where the denominator $\mathrm{Tr}[(\rho \otimes |\phi\rangle\langle\phi|)P_k]$, the probability of outcome k, is included for normalization of the conditional state.
After some algebra it is possible to show that ρ_k takes the form,

$$\rho_k = \frac{M_k \rho M_k^\dagger}{\mathrm{Tr}[M_k \rho M_k^\dagger]}$$

for some measurement operator M_k such that $E_k = M_k^\dagger M_k$.

State Update Rule for Generalized Measurements

Hence, in order to describe the post-selected quantum state after measurement under a POVM, we can simply apply the Luders rule to the description given by Neumark's theorem.
Specifically, we have

$$\rho \rightarrow \rho_k = \frac{\mathrm{Tr}_B(P_k(\rho \otimes |\phi\rangle\langle\phi|)P_k)}{\mathrm{Tr}[P_k\rho \otimes |\phi\rangle\langle\phi|P_k]}$$

where the denominator $\mathrm{Tr}[(\rho \otimes |\phi\rangle\langle\phi|)P_k]$, the probability of outcome k, is included for normalization of the conditional state.
After some algebra it is possible to show that ρ_k takes the form,

$$\rho_k = \frac{M_k\rho M_k^\dagger}{\mathrm{Tr}[M_k\rho M_k^\dagger]}$$

for some measurement operator M_k such that $E_k = M_k^\dagger M_k$.

State Update Rule for Generalized Measurements

The post-measurement state that describes the situation in which we *do not condition upon an observed outcome* takes the form:

$$\rho_k = \sum_k (\mathrm{Tr}[M_k \rho M_k^\dagger]) \frac{M_k \rho M_k^\dagger}{\mathrm{Tr}[M_k \rho M_k^\dagger]} = \sum_k M_k \rho M_k^\dagger.$$

State Update Rule for Generalized Measurements

It is worth emphasizing that the measurement operators $\{M_k\}$ (or, equivalently, the PVM $\{P_k\}$ and ancilla state $|0\rangle$) are not uniquely determined by the POVM $\{E_k\}$ because different measurement procedures in the extended Hilbert space can result in the same POVM acting on the system Hilbert space.

Hence the post-measurement state ρ_k is not uniquely determined by specification of the POVM; it is however uniquely determined by specifying a particular implementation of the measurement in the extended Hilbert space.

Expressing Generalized Measurement as a Completely Positive Map

If we consider the *un-normalized* expression for the post-measurement state, conditioned on outcome k, we have

$$\rho_k = M_k \rho M_k^\dagger.$$

Hence the state update rule *with post-selection* has the form of a *non-trace preserving* completely positive map - just interpret the M_k as Kraus operators and observe that $M_k^\dagger M_k \leq \mathbb{1}$.
Of course, the final state *without post-selection* takes the form

$$\rho_k = \sum_k M_k \rho M_k^\dagger,$$

which is just a CPTP map because $\sum_k M_k^\dagger M_k = \mathbb{1}$.

Expressing Generalized Measurement as a Completely Positive Map

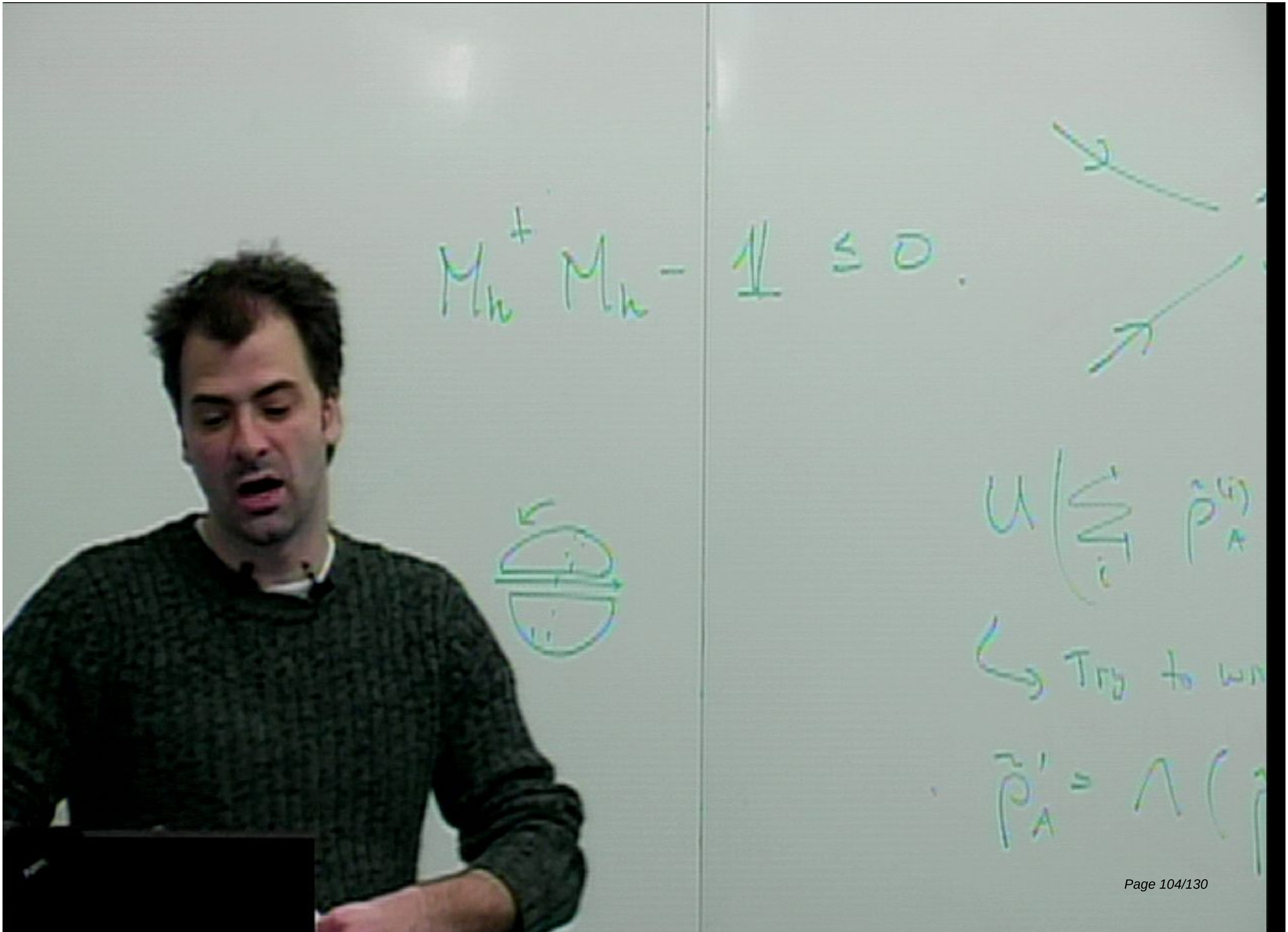
If we consider the *un-normalized* expression for the post-measurement state, conditioned on outcome k, we have

$$\rho_k = M_k \rho M_k^\dagger.$$

Hence the state update rule *with post-selection* has the form of a *non-trace preserving* completely positive map - just interpret the M_k as Kraus operators and observe that $M_k^\dagger M_k \leq \mathbb{1}$.
Of course, the final state *without post-selection* takes the form

$$\rho_k = \sum_k M_k \rho M_k^\dagger,$$

which is just a CPTP map because $\sum_k M_k^\dagger M_k = \mathbb{1}$.


$M_k^\dagger M_k - \mathbb{1} \leq 0.$
Try to

Expressing Generalized Measurement as a Completely Positive Map

If we consider the *un-normalized* expression for the post-measurement state, conditioned on outcome k, we have

$$\rho_k = M_k \rho M_k^\dagger.$$

Hence the state update rule *with post-selection* has the form of a *non-trace preserving* completely positive map - just interpret the M_k as Kraus operators and observe that $M_k^\dagger M_k \leq \mathbb{1}$.
Of course, the final state *without post-selection* takes the form

$$\rho_k = \sum_k M_k \rho M_k^\dagger,$$

which is just a CPTP map because $\sum_k M_k^\dagger M_k = \mathbb{1}$.

Expressing Generalized Measurement as a Completely Positive Map

Hence if we think of measurement as a transformation, it can be included in the same mathematical formalism that accounts for the generalized transformations generated by unitary evolution (on some extended Hilbert space) by simply *weakening the requirement* that the transformation must be trace-preserving.

Generalized Axiom 4 (revised): Over any finite time, the dynamical transformation of quantum system is described by a completely positive map.

Composite Systems and Entanglement

Axiom 3 remains unchanged in the generalized formalism. But it is worthwhile reviewing some basic properties of *entanglement*.

- An arbitrary pure state $|\psi_{AB}\rangle \in \mathcal{H}_{AB}$ has the form:

$$|\psi_{AB}\rangle = \sum_{k,\ell} \psi_{k\ell}|a_k\rangle \otimes |b_\ell\rangle, \quad \psi_{k\ell} = (\langle a_k| \otimes \langle b_\ell|)|\psi_{AB}\rangle \in \mathbb{C}.$$

- An arbitrary state operator, $\hat{\rho}$, acting on $\mathcal{H}_{AB}$ has the form:

$$\hat{\rho} = \sum_{\ell,k} \sum_{\ell',k'} \rho_{\ell\ell' kk'}|a_\ell\rangle\langle a_{\ell'}| \otimes |b_k\rangle\langle b_{k'}|,$$

where

$$\hat{\rho}_{\ell\ell' kk'} = \langle a_\ell| \otimes \langle b_k|\hat{\rho}|a_{\ell'}\rangle \otimes |b_{k'}\rangle.$$

Composite Systems and Entanglement

Axiom 3 remains unchanged in the generalized formalism. But it is worthwhile reviewing some basic properties of *entanglement*.

- An arbitrary pure state $|\psi_{AB}\rangle \in \mathcal{H}_{AB}$ has the form:

$$|\psi_{AB}\rangle = \sum_{k,\ell} \psi_{k\ell} |a_k\rangle \otimes |b_\ell\rangle, \quad \psi_{k\ell} = (\langle a_k| \otimes \langle b_\ell|)|\psi_{AB}\rangle \in \mathbb{C}.$$

- An arbitrary state operator, $\hat{\rho}$, acting on $\mathcal{H}_{AB}$ has the form:

$$\hat{\rho} = \sum_{\ell,k} \sum_{\ell',k'} \rho_{\ell\ell' kk'} |a_\ell\rangle\langle a_{\ell'}| \otimes |b_k\rangle\langle b_{k'}|,$$

where

$$\hat{\rho}_{\ell\ell' kk'} = \langle a_\ell| \otimes \langle b_k|\hat{\rho}|a_{\ell'}\rangle \otimes |b_{k'}\rangle.$$

Composite Systems and Entanglement

Axiom 3 remains unchanged in the generalized formalism. But it is worthwhile reviewing some basic properties of *entanglement*.

- An arbitrary pure state $|\psi_{AB}\rangle \in \mathcal{H}_{AB}$ has the form:

$$|\psi_{AB}\rangle = \sum_{k,\ell} \psi_{k\ell} |a_k\rangle \otimes |b_\ell\rangle, \quad \psi_{k\ell} = (\langle a_k| \otimes \langle b_\ell|)|\psi_{AB}\rangle \in \mathbb{C}.$$

- An arbitrary state operator, $\hat{\rho}$, acting on $\mathcal{H}_{AB}$ has the form:

$$\hat{\rho} = \sum_{\ell,k} \sum_{\ell',k'} \rho_{\ell\ell' kk'} |a_\ell\rangle\langle a_{\ell'}| \otimes |b_k\rangle\langle b_{k'}|,$$

where

$$\hat{\rho}_{\ell\ell' kk'} = \langle a_\ell| \otimes \langle b_k|\hat{\rho}|a_{\ell'}\rangle \otimes |b_{k'}\rangle.$$

Composite Systems and Entanglement

Definition

A pure state $|\chi\rangle \in \mathcal{H}_{AB}$ that *can* be expressed as $|\chi\rangle = |\alpha\rangle \otimes |\beta\rangle$ for some $|\alpha\rangle \in \mathcal{H}_A$, $|\beta\rangle \in \mathcal{H}_B$, is called a *product state*, or a *factorable state*; otherwise it is called *entangled*.

Definition

A general state $\hat{\rho}$ acting on $\mathcal{H}_{AB}$ is called *separable* if and only if

$$\hat{\rho}_{AB} = \sum_i p_i \hat{\rho}_{iA} \otimes \hat{\rho}_{iB},$$

i.e., iff it can be expressed as a statistical mixture of states of the form $\hat{\rho}_{iA} \otimes \hat{\rho}_{iB}$; otherwise it is called *entangled*.

Composite Systems and Entanglement

Definition

A pure state $|\chi\rangle \in \mathcal{H}_{AB}$ that *can* be expressed as $|\chi\rangle = |\alpha\rangle \otimes |\beta\rangle$ for some $|\alpha\rangle \in \mathcal{H}_A$, $|\beta\rangle \in \mathcal{H}_B$, is called a *product state*, or a *factorable state*; otherwise it is called *entangled*.

Definition

A general state $\hat{\rho}$ acting on $\mathcal{H}_{AB}$ is called *separable* if and only if

$$\hat{\rho}_{AB} = \sum_i p_i \hat{\rho}_{iA} \otimes \hat{\rho}_{iB},$$

i.e., iff it can be expressed as a statistical mixture of states of the form $\hat{\rho}_{iA} \otimes \hat{\rho}_{iB}$; otherwise it is called *entangled*.

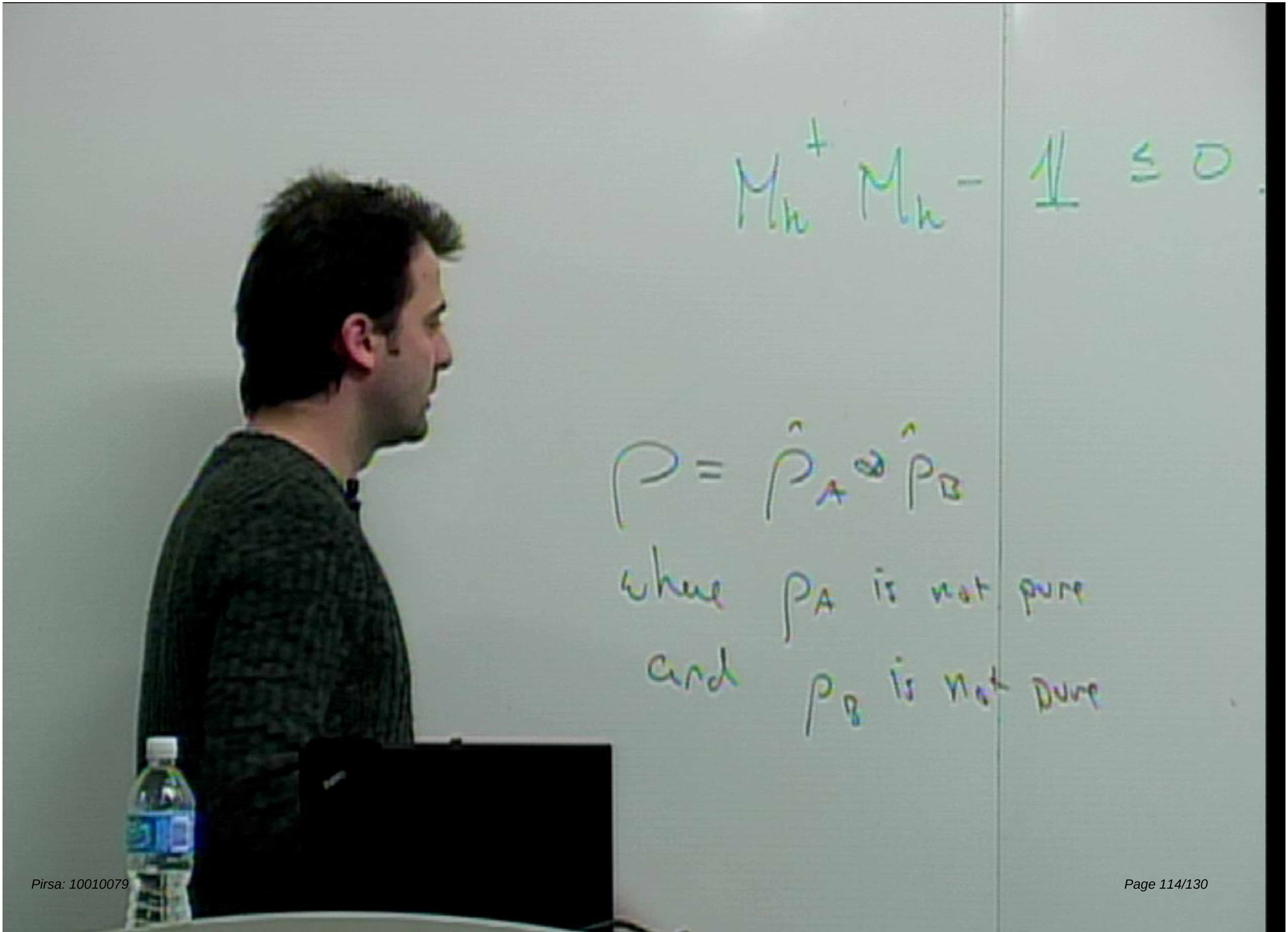
Composite Systems and Entanglement

- An entangled pure state has the property that independent measurements on A and B exhibit correlations in the *outcomes*. Entangled pure states have *in a well defined sense* (due to Bell-type theorems) stronger correlations than any classical state.
- Note that any classical correlation between A and B can be modeled by a separable state, ie, a mixed state can exhibit correlations without being entangled.

Criteria for Pure-state Entanglement

How can we tell when a pure composite state is entangled?

- If the composite system state is pure then we can compute whether it is entangled by calculating the purity of the reduced state of either system A or system B: a pure composite system state is entangled iff $\mathrm{Tr}(\rho_A^2) < 1$.
- This criterion will not work if the composite system state is a mixed state, because even a separable mixed state will produce a mixed reduced state.
- Note also that $\mathrm{Tr}(\rho_A^2) = \mathrm{Tr}(\rho_B^2)$ whenever the reduced states are obtained from a pure composite system state.


M_h^+ M_h - 1 ≤ 0
ρ = ρ̂_A ⊗ ρ̂_B
where ρ_A is not pure
and ρ_B is not pure

Criteria for Pure-state Entanglement

How can we tell when a pure composite state is entangled?

- If the composite system state is pure then we can compute whether it is entangled by calculating the purity of the reduced state of either system A or system B: a pure composite system state is entangled iff $\mathrm{Tr}(\rho_A^2) < 1$.
- This criterion will not work if the composite system state is a mixed state, because even a separable mixed state will produce a mixed reduced state.
- Note also that $\mathrm{Tr}(\rho_A^2) = \mathrm{Tr}(\rho_B^2)$ whenever the reduced states are obtained from a pure composite system state.

Criteria for Pure-state Entanglement

How entangled is an entangled state?

- We can quantify the amount of entanglement of a pure state using the von Neumann entropy $S(\rho_A) = -\mathrm{Tr}[\rho_A \log_2(\rho_A)]$ of either reduced state.
- Note that $\log_2(\rho_A)$ can be defined in terms of the spectral decomposition of ρ_A, and that $\lambda_i \log_2 \lambda_i = 0$ if $\lambda_i = 0$.

Criteria for Pure-state Entanglement

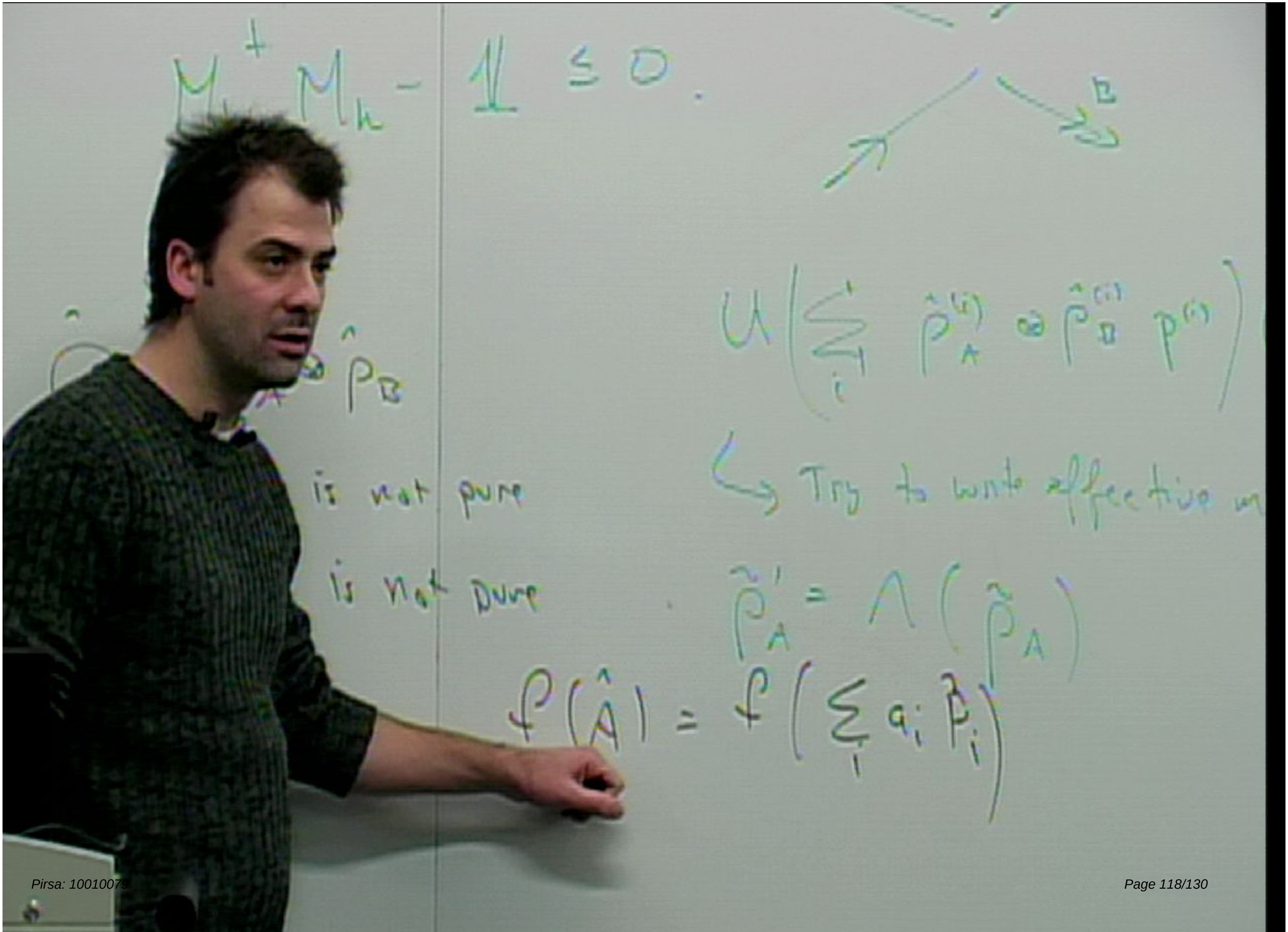
The von Neumann entropy gives a measure of the amount ignorance one has about a system (as encoded by the state one is using to describe the system). It has the following properties:

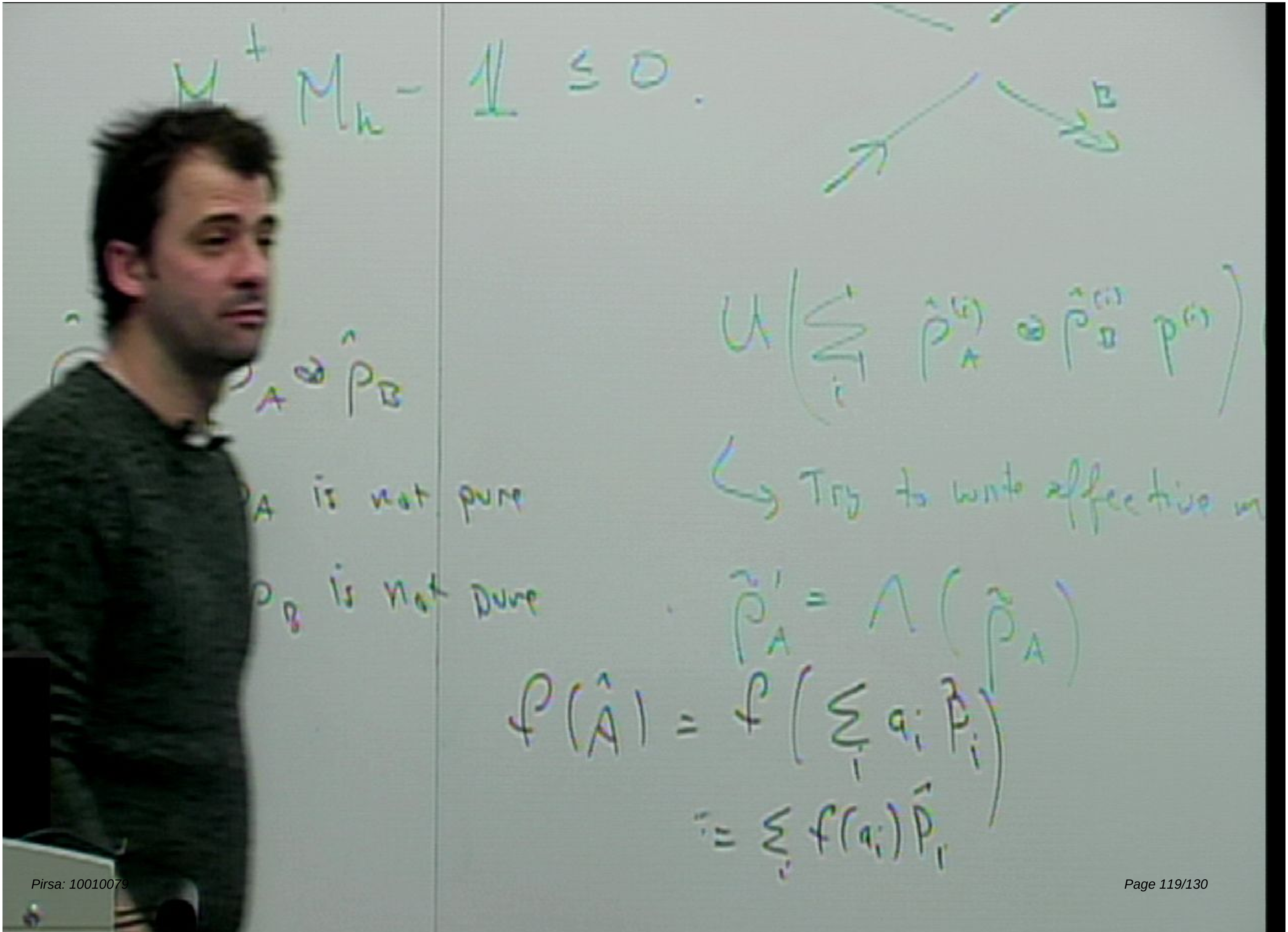
i) $S(\rho) = 0$ iff ρ is pure.

ii) For ρ acting on $\mathcal{H} = \mathbb{C}^D$, $0 \leq S(\rho) \leq \log(D)$, where the upper bound is saturated iff the state is completely mixed, ie, $\rho = \mathbb{1}/D$.

iii) If the composite system state is pure then $S(\rho_A) = S(\rho_B)$.

iv) Concavity: $S(\sum_i p_i \rho_i) \geq \sum_i p_i S(\rho_i)$. (Intuitively, ignorance about the mixture must be greater than the average of the ignorance associated with each of the component states.)


is not pure
is not pure
Try to write effective



M† Mₕ − 𝟙 ≤ 0.
ρ_A ⊗ ρ_B
ρ_A is not pure
ρ_B is not pure
U(Σᵢ ρ_A^(i) ⊗ ρ_B^(i) p^(i))
Try to write effective m
ρ'_A = Λ(ρ_A)
f(Â) = f(Σᵢ aᵢ Pᵢ)
:= Σᵢ f(aᵢ) Pᵢ

Criteria for Pure-state Entanglement

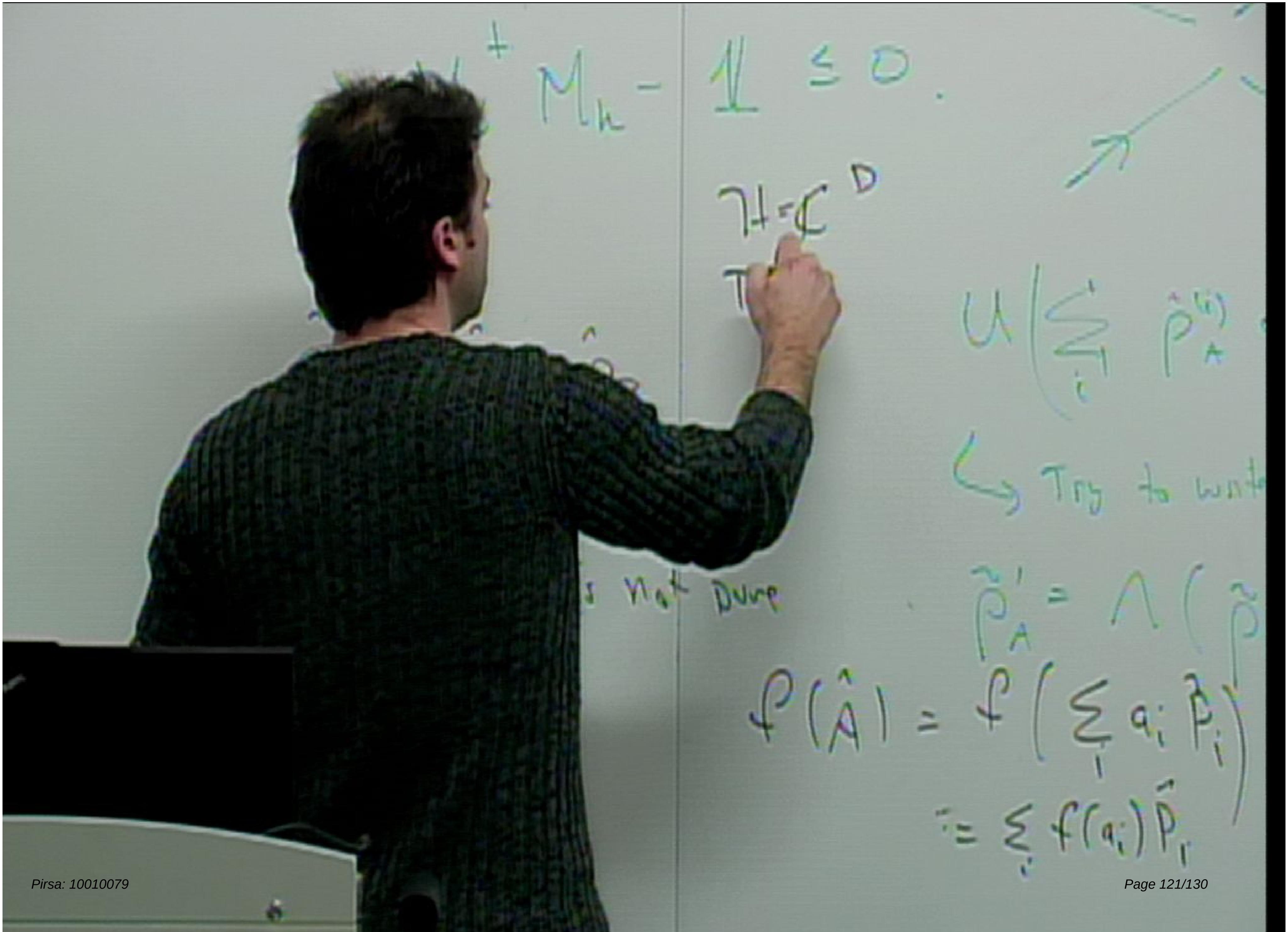
The von Neumann entropy gives a measure of the amount ignorance one has about a system (as encoded by the state one is using to describe the system). It has the following properties:

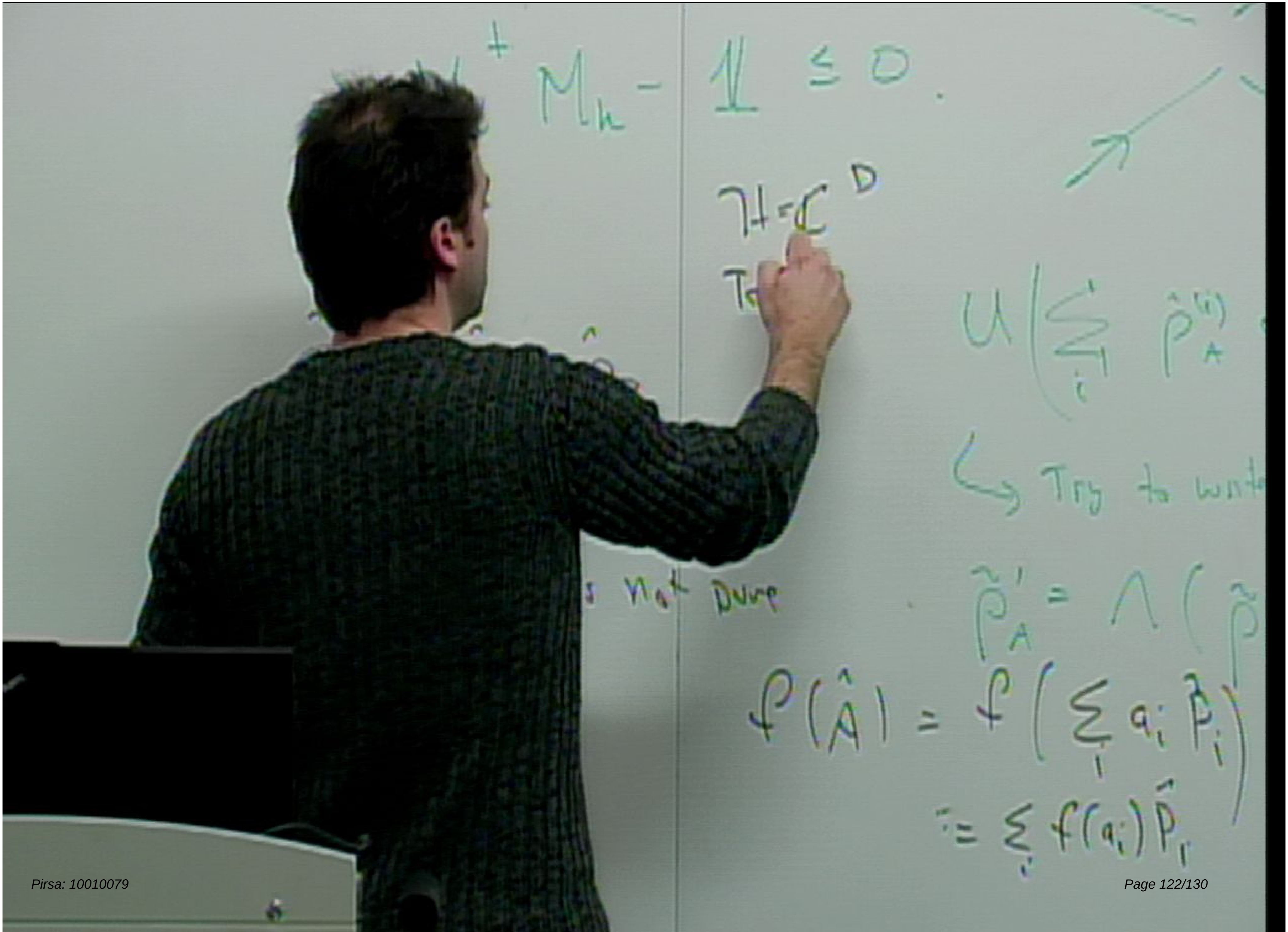
i) $S(\rho) = 0$ iff ρ is pure.

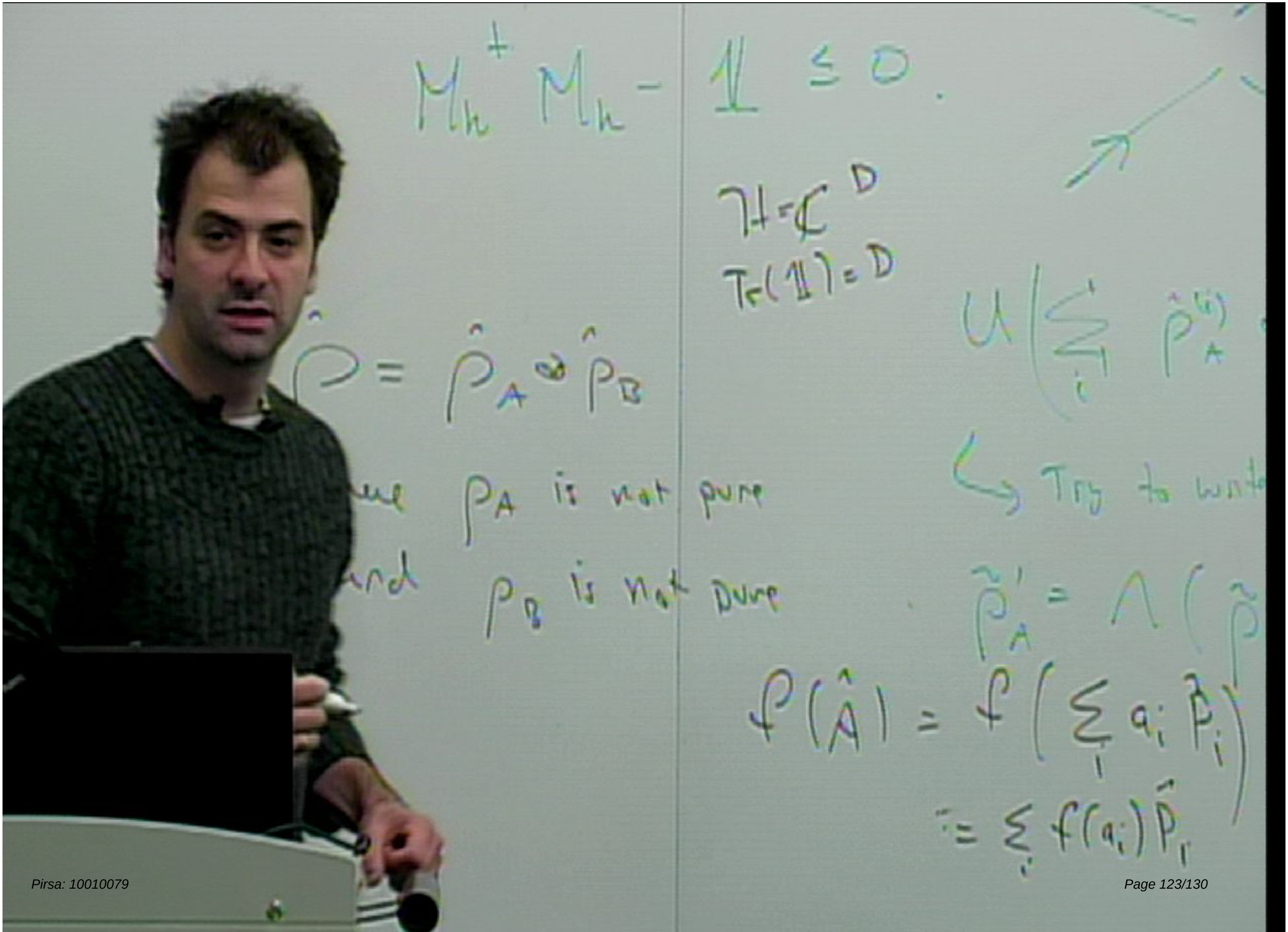
ii) For ρ acting on $\mathcal{H} = \mathbb{C}^D$, $0 \leq S(\rho) \leq \log(D)$, where the upper bound is saturated iff the state is completely mixed, ie, $\rho = \mathbb{1}/D$.

iii) If the composite system state is pure then $S(\rho_A) = S(\rho_B)$.

iv) Concavity: $S(\sum_i p_i \rho_i) \geq \sum_i p_i S(\rho_i)$. (Intuitively, ignorance about the mixture must be greater than the average of the ignorance associated with each of the component states.)


$f(\hat{A}) = f\left(\sum_i a_i \hat{P}_i\right)$
$= \sum_i f(a_i) \hat{P}_i$




$M_k^\dagger M_k - \mathbb{1} \leq 0$
$\mathcal{H} = \mathbb{C}^D$
$Tr(\mathbb{1}) = D$
$\hat{\rho} = \hat{\rho}_A \otimes \hat{\rho}_B$
ρ_A is not pure
and ρ_B is not pure
Try to write
$f(\hat{A}) = f\left(\sum_i a_i \hat{P}_i\right)$
$= \sum_i f(a_i) \hat{P}_i$

$1 - \mathbb{1} \leq 0.$

$\ldots D$

not pure

not pure

$f(\hat{A}) =$

$|i\rangle, \quad i = 1, \ldots, D$

Let $\hat{\rho} = |i\rangle\langle i|$

with prob. $\frac{1}{D}$

$\mathbb{1}_h - \mathbb{1} \leq 0$.

$\ldots D$

$\otimes \hat{\rho}_B$

is not pure

is not pure

$f(\hat{A}) =$

$|i\rangle, \quad i = 1, \ldots, D$

Let $\hat{\rho} = |i\rangle\langle i|$

with prob. $\frac{1}{D}$

Effective description

$$\hat{\rho} = \sum_{i=1}^{D} \frac{1}{D} |i\rangle\langle i|$$

$$= \frac{\mathbb{1}}{D}$$

Criteria for Pure-state Entanglement

Another useful characterization of entanglement is from the Schmidt decomposition:

Definition

Schmidt decomposition: Given a pure state in $\mathcal{H}_{AB}$, there exist ON bases $\{|i_A\rangle\}$ and $\{|i_B\rangle\}$ of $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively (called *Schmidt bases*) such that

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle |i_B\rangle$$

where λ_i are non-negative real numbers (Schmidt coefficients) satisfying $\sum_i \lambda_i^2 = 1$.

Criteria for Pure-state Entanglement

Another useful characterization of entanglement is from the Schmidt decomposition:

Definition

Schmidt decomposition: Given a pure state in $\mathcal{H}_{AB}$, there exist ON bases $\{|i_A\rangle\}$ and $\{|i_B\rangle\}$ of $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively (called *Schmidt bases*) such that

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle |i_B\rangle$$

where λ_i are non-negative real numbers (Schmidt coefficients) satisfying $\sum_i \lambda_i^2 = 1$.

Schmidt Number

- The *Schmidt number* is the number of non-zero coefficients λ_i in the Schmidt decomposition. Clearly the maximum possible Schmidt number is less than or equal to the smaller of $\dim(\mathcal{H}_A)$ and $\dim(\mathcal{H}_B)$.
- Clearly a state is entangled iff the Schmidt number is greater than 1.
- If $N = \dim(\mathcal{H}_A) = \dim(\mathcal{H}_B)$, a state is called *maximally entangled* if the Schmidt number is N. This holds for any state of the form

$$|\psi\rangle = \sum_{n=1}^{N} \frac{\exp(ia_n)}{\sqrt{N}} |n_A\rangle |n_B\rangle$$

where $a_n \in \mathbb{R}$.

Usefulness of the Schmidt decomposition

- The Schmidt decomposition gives an easy way to calculate the reduced density operators: $\rho_A = \sum_i \lambda_i^2 |i_A\rangle\langle i_A|$ and $\rho_B = \sum_i \lambda_i^2 |i_B\rangle\langle i_B|$.
- Note that the eigenvalues of the subsystem states are identical - this holds whenever the composite state is pure - and hence this justifies the earlier claim that $\mathrm{Tr}(\rho_A^2) = \mathrm{Tr}(\rho_B^2)$ whenever the joint state is pure.

Usefulness of the Schmidt decomposition

- The Schmidt decomposition gives an easy way to calculate the reduced density operators: $\rho_A = \sum_i \lambda_i^2 |i_A\rangle\langle i_A|$ and $\rho_B = \sum_i \lambda_i^2 |i_B\rangle\langle i_B|$.
- Note that the eigenvalues of the subsystem states are identical - this holds whenever the composite state is pure - and hence this justifies the earlier claim that $\mathrm{Tr}(\rho_A^2) = \mathrm{Tr}(\rho_B^2)$ whenever the joint state is pure.