What we learned from proving a quantum postulate redundant

Gabriele Carcassi

Physics Department

University of Michigan

Assumptions of Physics

The paper

Four Postulates of Quantum Mechanics Are Three

Gabriele Carcassi, Lorenzo Maccone, and Christine A. Aidala
Phys. Rev. Lett. **126**, 110402 – Published 16 March 2021

Gabriele Carcassi Christine A. Aidala Lorenzo Maccone

University of Michigan Università di Pavia

Plan

- The setup
    - Postulates, how to remove them and the nature of composite systems
- The proof
    - Projective spaces, their bridge between probabilistic events and quantum states, the fundamental theorem of projective geometry and the universal property of the tensor product
- The commentary
    - The anti-linearity debacle, the lack of tensor product in Hilbert spaces and the wrong math

THE SET-UP

Physics		Math
State of a quantum system	⇨	Ray in a Hilbert space
Quantities and measurements	⇨	Hermitian operators and Born rule
Composite quantum system	⇨	Tensor product
Time evolution	⇨	Schrödinger equation

Physics	Math
State of a quantum system	Ray in a Hilbert space
Quantities and measurements	Hermitian operators and Born rule
Composite quantum system	Tensor product
Time evolution	Schrödinger equation

Recipe for removing a postulate

- Identify basic physical requirements a composite system must have to be meaningful
- Translate those requirements into mathematical definitions
- Show the use of the tensor product to model a composite quantum system follows mathematically from those definitions and the other postulates

⇒ Postulate is no longer necessary: the physics is enough to constrain the math

Requirement one: preparation independence

- **R1**: Two systems are said **independent** if the preparation of one does not affect the preparation of the other

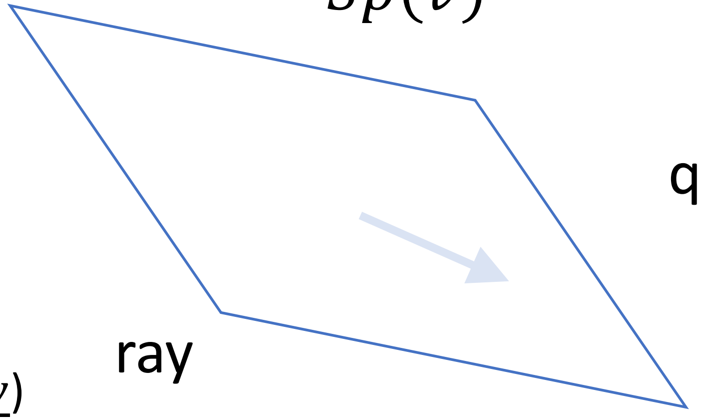
Ultimately, the physics of QM is expressed in probabilistic terms, so let us formalize independence in terms of probability

- I.1/I.2: Let $\underline{\mathcal{A}}$ and $\underline{\mathcal{B}}$ be the state spaces for two quantum systems A and B. Two states $(\underline{a}, \underline{b}) \in \underline{\mathcal{A}} \times \underline{\mathcal{B}}$ are **compatible** if the event/proposition $\underline{a} \wedge \underline{b}$ (i.e. system A is in state $\underline{a}$ and system B is in state $\underline{b}$) is possible (i.e. it does not correspond to the empty set in the σ-algebra). Two systems are **independent** if all pairs $(\underline{a}, \underline{b}) \in \underline{\mathcal{A}} \times \underline{\mathcal{B}}$ are compatible.

Projective space $\underline{\mathcal{H}}$

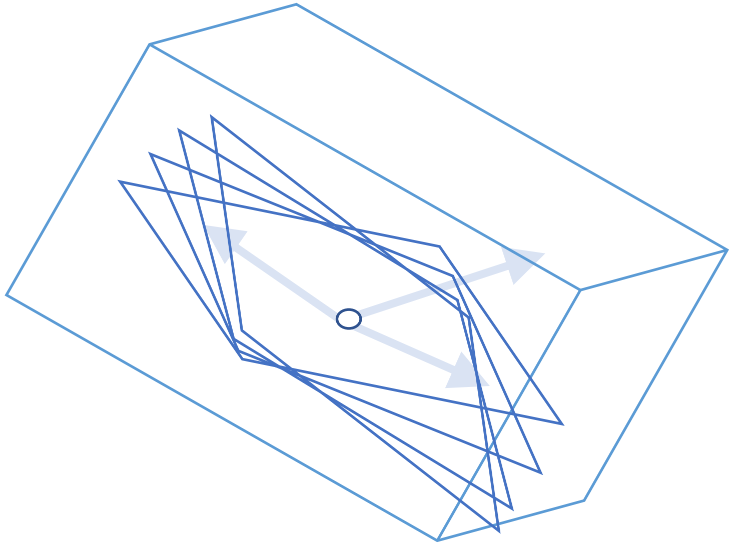
$\underline{v} = \{kv \mid k \in \mathbb{C}\}$
$= Sp(v)$

$\underline{Sp(v_1, v_2, v_3)}$



ray

quantum state



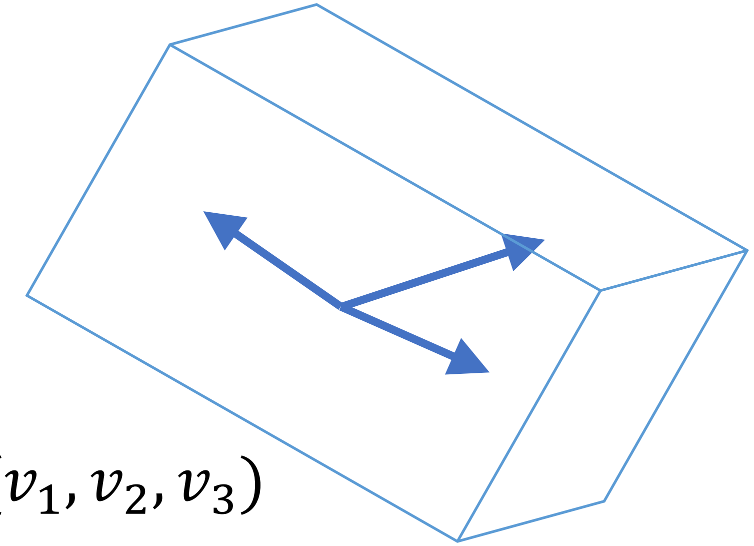
$P(v|w) = \frac{|\langle v,w \rangle|^2}{\langle v,v \rangle \langle w,w \rangle} = P(\underline{v}|\underline{w})$

Hilbert space $\mathcal{H}$

~~quantum state~~

v

vector



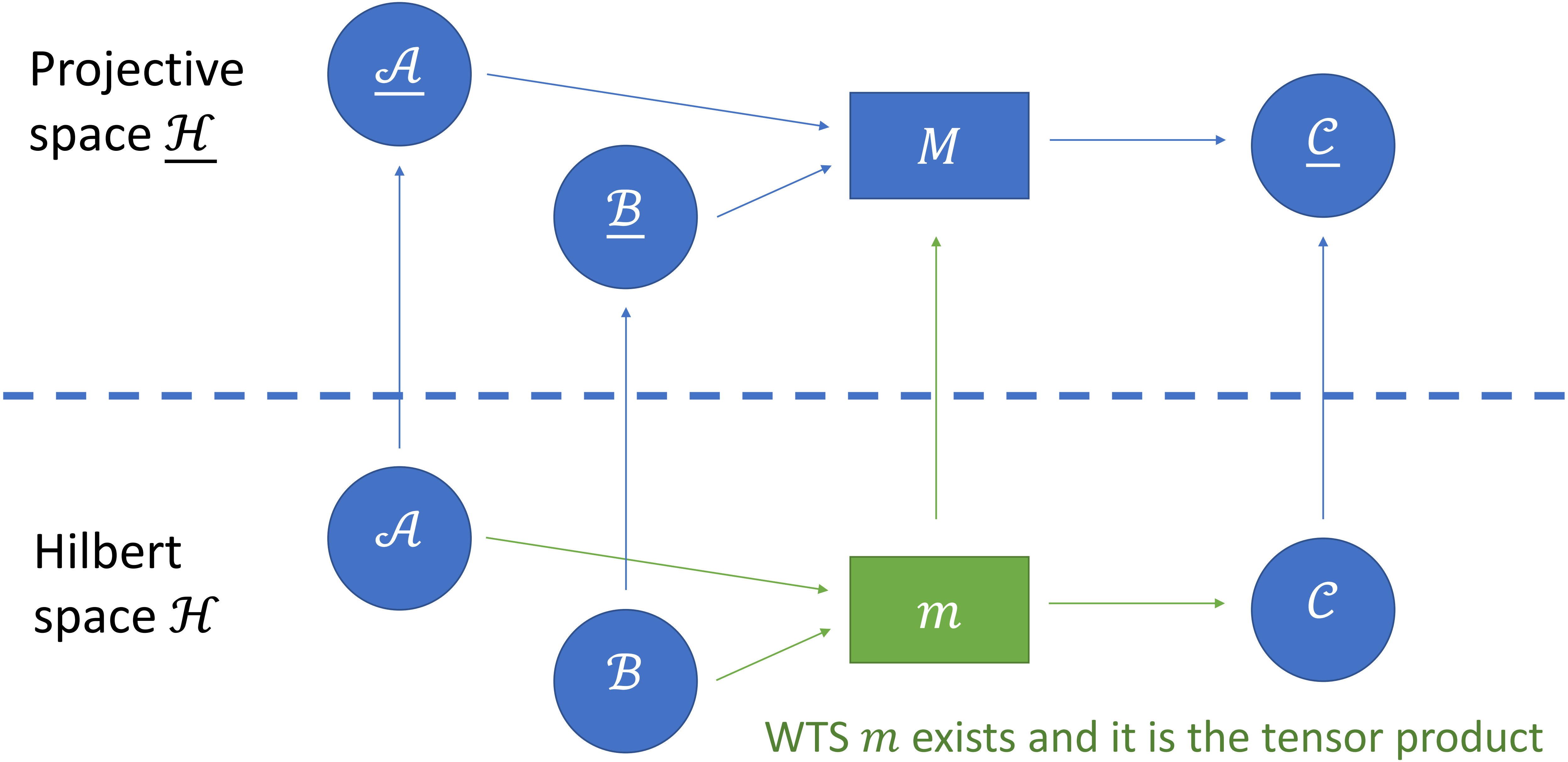
$Sp(v_1, v_2, v_3)$

Requirement two: composite system

- **R2**: Given two systems A and B, their composite system C is the simple collection of those and only those systems (the smallest system that contains both)

We break this into two:

- I.4.1: C is made of A and B…
  - Whenever we prepare A and B independently, we have prepared C. Formally, let $\underline{\mathcal{C}}$ be the state space of the composite of two quantum systems A and B. There exists a map $M: \underline{\mathcal{A}} \times \underline{\mathcal{B}} \to \underline{\mathcal{C}}$ such that $\underline{a} \wedge \underline{b}$ and $M(\underline{a}, \underline{b})$ corresponds to the same event.
- I.4.2: … and only A and B
  - Given any state of C, measuring A and B independently leads to a pair of respective states with non-zero probability. Formally, for every $\underline{c} \in \underline{\mathcal{C}}$, we can find at least a pair $\left(\underline{a}, \underline{b}\right) \in \underline{\mathcal{A}} \times \underline{\mathcal{B}}$ such that $P\left(\underline{a} \wedge \underline{b} \middle| \underline{c}\right) \neq 0$.



WTS m exists and it is the tensor product

Goal: tensor product

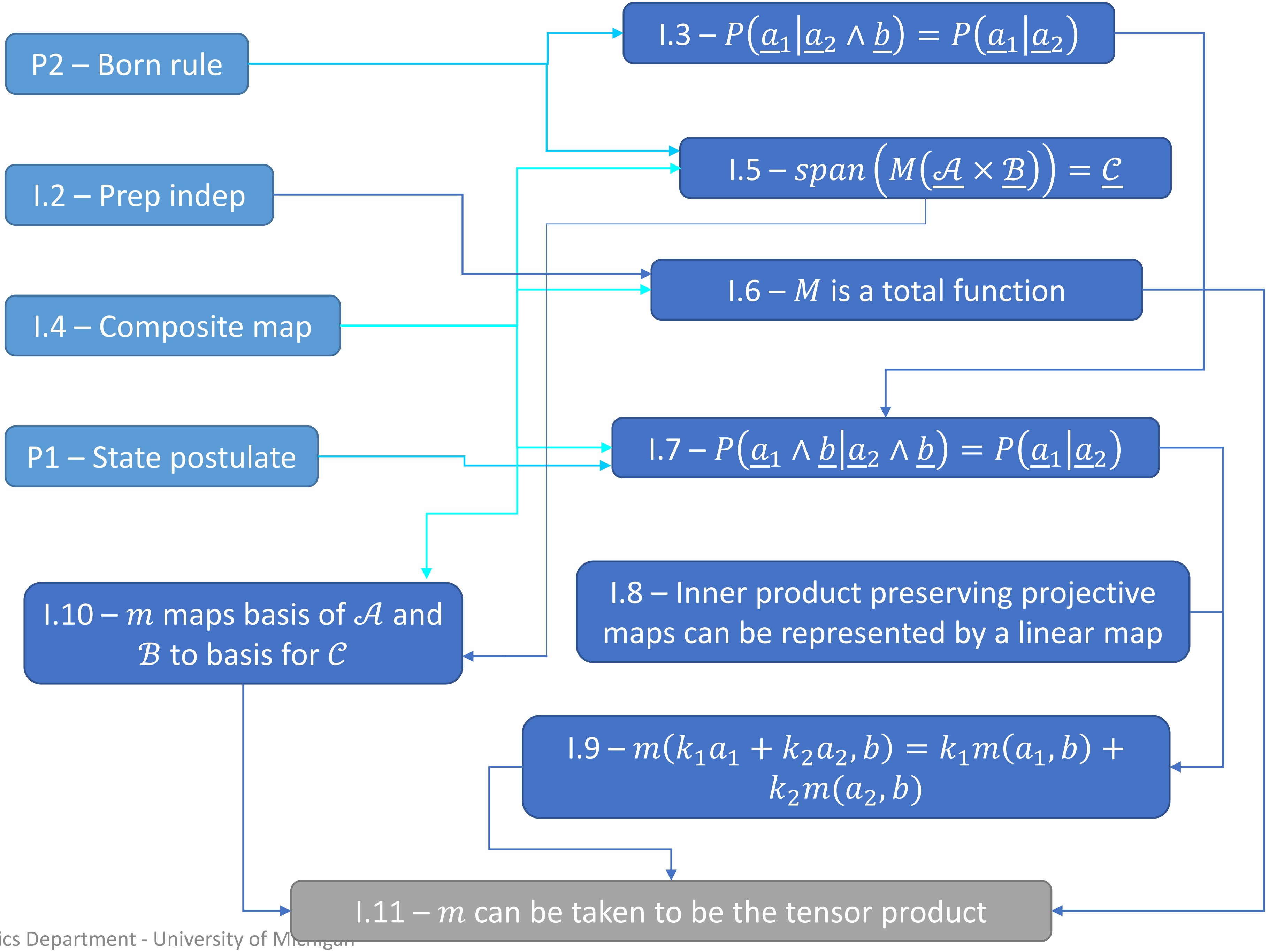
- **G**: The Hilbert space of the composite system of two independent quantum systems is represented by the tensor product of the Hilbert spaces of the component systems

- I.11: There exists a bilinear map $m: \mathcal{A} \times \mathcal{B} \to \mathcal{C}$ such that $\underline{m(a,b)} = M\left(\underline{a}, \underline{b}\right)$ and that map can be taken to be, without loss of generality, the tensor product

THE PROOF

Not the simplest thing

I'll try to cover the main points



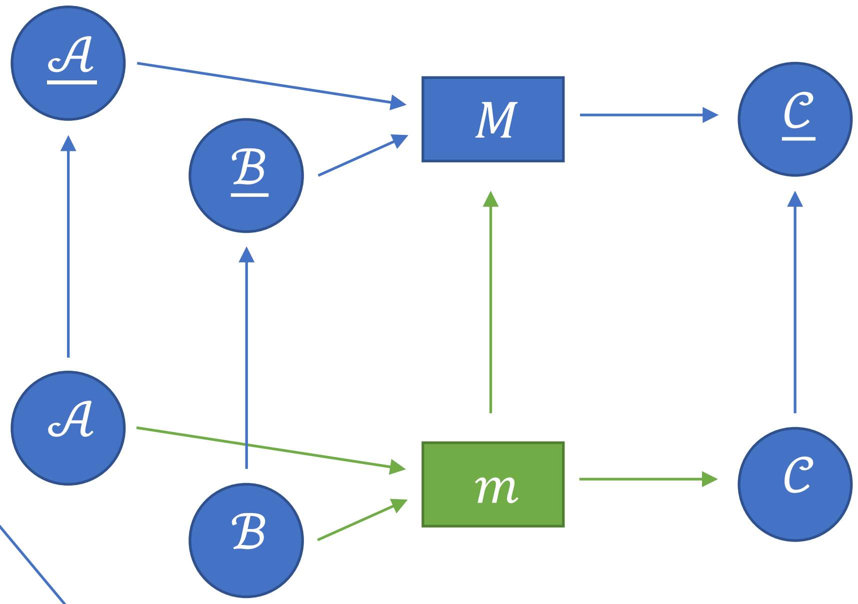
Outline

We break up the final **G**oal into 3 intermediate conditions (**H**ypotheses):

- **H1**: M is total
  - it is defined on all pairs $(\underline{a}, \underline{b})$
- **H2**: Show that if m exists, it must be bilinear

$$m(k_1 a_1 + k_2 a_2, b) = k_1 m(a_1, b) + k_2 m(a_2, b)$$
$$m(a, k_1 b_1 + k_2 b_2) = k_1 m(a, b_1) + k_2 m(a, b_2)$$

- **H3**: M is span-surjective
  - $\underline{Sp\left(M(\underline{\mathcal{A}} \times \underline{\mathcal{B}})\right)} = \underline{\mathcal{C}}$
- **G**: m exists and can be taken to be the tensor product



These are the harder bits

Labels used in the paper

I.6 **H1**: M is total

- Preparation independence (1.2 **R1**) tells us that all events $\underline{a} \wedge \underline{b}$ are possible
- The definition of composite system (1.4 **R2**) tells us that $M(\underline{a}, \underline{b})$ is equivalent to $\underline{a} \wedge \underline{b}$
- If $\underline{a} \wedge \underline{b}$ is not possible, the function M would not be defined for that pair: M would be a partial function
- Assuming preparation independence, M is defined on all pairs and is a total function

- M is total really means we have preparation independence
- Physically, if we don't have preparation independence (e.g. super-selection rules) we will not have the tensor product

I.5 **H3**: M is span-surjective

- Consider the span of the image of M: $\underline{Sp\left(M(\underline{\mathcal{A}} \times \underline{\mathcal{B}})\right)}$
- It's a subspace of $\underline{\mathcal{C}}$. Does it cover the full space?
- Suppose we have $\underline{c} \in \underline{\mathcal{C}}$ that is not in the span of the image of M
- Then $\underline{c}$ is perpendicular to all elements of the image (i.e. linearly independent)
- Therefore $P\left(M(\underline{a}, \underline{b})\middle|\underline{c}\right) = P\left(\underline{a} \wedge \underline{b}\middle|\underline{c}\right) = 0$ for all $\left(\underline{a}, \underline{b}\right) \in \underline{\mathcal{A}} \times \underline{\mathcal{B}}$
- This violates the requirement for the composite system (I.4.2 **R2**): we prepare the composite but we never find the parts
- M is span-surjective

- M is span-surjective means that the composite doesn't have anything else
- Mathematically, any state of the composite is a superposition of independent pairs of the individual systems

The road to bilinearity

Projective space $\underline{\mathcal{H}}$

Hilbert space $\mathcal{H}$

$$
\begin{tikzcd}
\underline{\mathcal{A}} \arrow[r] & F \arrow[r] & \underline{\mathcal{C}} \\
\mathcal{A} \arrow[r] \arrow[u] & f \arrow[r] \arrow[u] & \mathcal{C} \arrow[u]
\end{tikzcd}
$$

Colinear: preserves subgroup structure

$$G_1 \subseteq G_2 \Leftrightarrow F(G_1) \subseteq F(G_2)$$

Linear: $f(k_1 a_1 + k_2 a_2) = k_1 f(a_1) + k_2(a_2)$

The road to bilinearity

Projective space $\underline{\mathcal{H}}$

Hilbert space $\mathcal{H}$

$$\begin{tikzcd}
\underline{\mathcal{A}} \arrow[r] & M(\cdot, \underline{b}) \arrow[r] & \underline{\mathcal{C}} \\
\mathcal{A} \arrow[r] \arrow[u] & m(\cdot, b) \arrow[r] \arrow[u] & \mathcal{C} \arrow[u]
\end{tikzcd}$$

We first need to show that $M(\cdot, \underline{b})$ is colinear

Colinearity of $M(\cdot,\underline{b})$

- The Born rule (implicitly) tells us that a measurement on A depends only on the preparation of A: $P(\underline{a}_1|\underline{a}_2 \wedge \underline{b}) = \frac{|\langle a_1,a_2\rangle|^2}{\langle a_1,a_1\rangle\langle a_2,a_2\rangle} = P(\underline{a}_1|\underline{a}_2)$
- $P(\underline{a}_1 \wedge \underline{b}|\underline{a}_2 \wedge \underline{b}) = P(\underline{b}|\underline{a}_2 \wedge \underline{b})P(\underline{a}_1|\underline{a}_2 \wedge \underline{b} \wedge \underline{b}) = P(\underline{b}|\underline{b})P(\underline{a}_1|\underline{a}_2)$
- $P(\underline{a}_1 \wedge \underline{b}|\underline{a}_2 \wedge \underline{b}) = P(M(\underline{a}_1,\underline{b})|M(\underline{a}_2,\underline{b})) = P(\underline{a}_1|\underline{a}_2)$

- The map $M(\cdot,\underline{b})$ preserves the probability, therefore orthogonality and therefore the subgroup structure
- The map $M(\cdot,\underline{b})$ is colinear

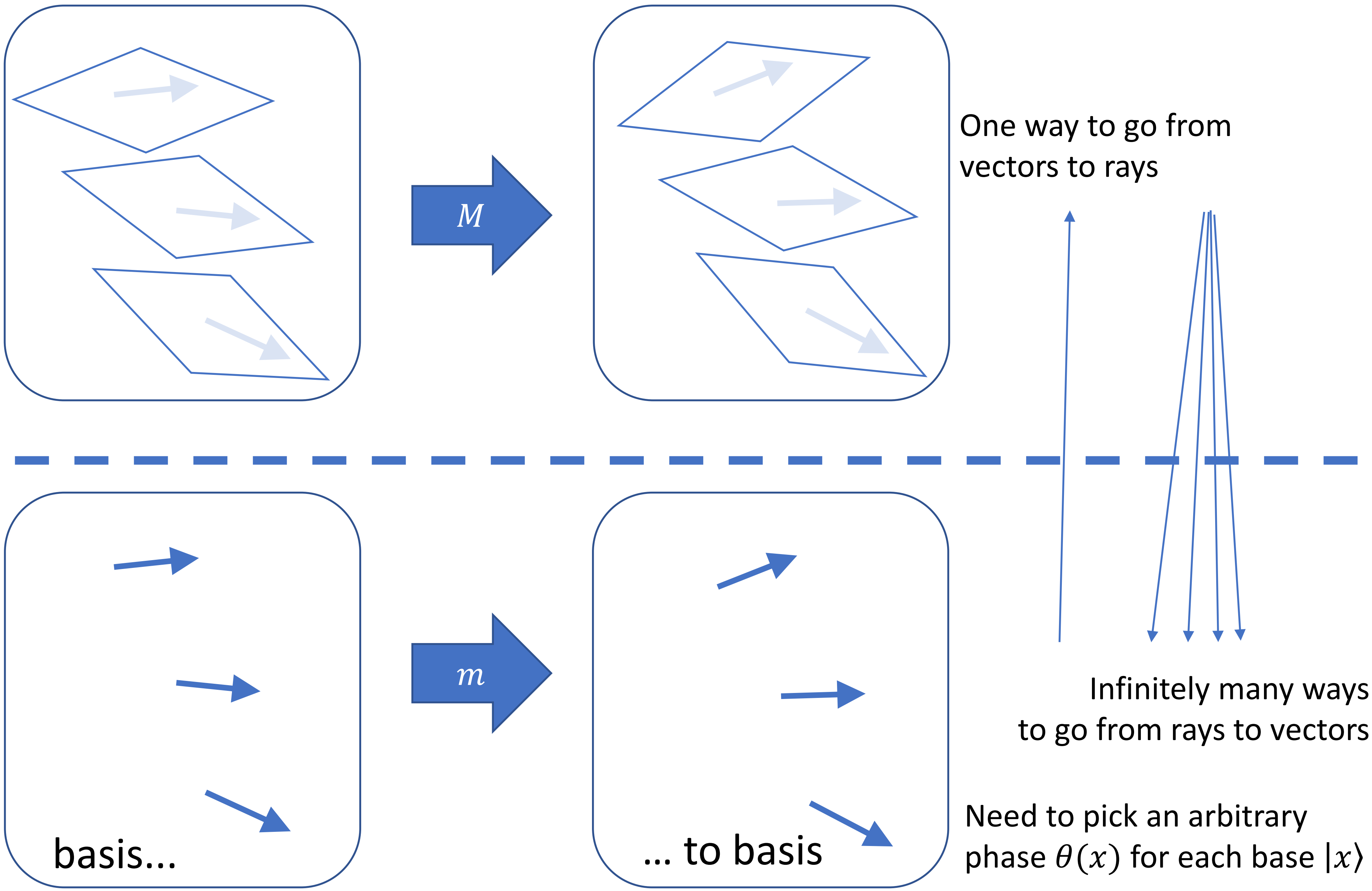
Fundamental theorem of projective geometry

- This theorem allows us to go from colinear maps in the projective space to linear maps in the Hilbert space:
  every colinear map M on the projective space induces a map m on the Hilbert space that is either linear or anti-linear (i.e. linear in the complex conjugate)
  - Technically, we use an adaptation of the fundamental theorem of projective geometry
  - The general result states that for every colinear function M between the projective spaces we can find a semi-linear transformation m on the vector spaces
  - Because we have $P\left(M\left(\underline{a}_1, \underline{b}\right) \middle| M\left(\underline{a}_2, \underline{b}\right)\right) = P\left(\underline{a}_1 \middle| \underline{a}_2\right)$, the transformation is either linear (i.e. $\langle m(a_1, b) | m(a_2, b) \rangle = \langle a_1 | a_2 \rangle$) or anti-linear (i.e. $\langle m(a_1, b) | m(a_2, b) \rangle = \langle a_2 | a_1 \rangle$)

- Note: there are infinitely many $m(\cdot, b)$ that induce $M\left(\cdot, \underline{b}\right)$, but we pick those that are linear (or anti-linear)

Projective space $\underline{\mathcal{H}}$

Hilbert space $\mathcal{H}$



Fixing the representation

- When going from the rays to the vectors, one picks a "gauge" $\theta(x)$
  - The gauge changes the representation, but not the probability:
$$\int \psi^\dagger(x)\phi(x)dx = \int e^{-\imath\theta(x)}\psi^\dagger(x)\phi(x)e^{\imath\theta(x)}dx$$
- In the proof, we use this freedom to construct the linear map: we fix "the same" gauge

- Linearity vs anti-linearity is also a choice of representation
  - We formally switch $\langle\psi|\phi\rangle$ with $\langle\phi|\psi\rangle$ in all of QM and all predictions (i.e. probabilities and eigenvalues of Hermitian operators) do not change
- If the map is anti-linear, we can transform to the linear case
- We will assume the map is linear without loss of generality

I.9 H2: m is bilinear

- Without loss of generality, we can say that if m exists it must be linear when fixing either side:

$$m(k_1 a_1 + k_2 a_2, b) = k_1 m(a_1, b) + k_2 m(a_2, b)$$
$$m(a, k_1 b_1 + k_2 b_2) = k_1 m(a, b_1) + k_2 m(a, b_2)$$

- We have all the ingredients we needed

Universal property of the tensor product

Any bilinear map factors uniquely through the tensor product

$$\begin{tikzcd} \mathcal{A} \times \mathcal{B} \arrow[r, "T"] \arrow[dr, "m"'] & \mathcal{A} \otimes \mathcal{B} \arrow[d, "I"] \\ & \mathcal{C} \end{tikzcd}$$

Note: we typically use the same symbol $\otimes$ for the operation on the spaces (i.e. $\mathcal{A} \otimes \mathcal{B}$) and the map on vectors (i.e. $a \otimes b$). Here $T(a, b)$ indicates the map on vectors.

For any bilinear map m there exists a unique linear map I such that $m = I \circ T$

Final proof

Because m has to be bilinear (I.9 **H2**), we can find a corresponding I

Because M was span surjective (I.5 **H3**), the basis of $\mathcal{C}$ cannot be "bigger" than $\mathcal{A} \otimes \mathcal{B}$ nothing but A and B

Because M was total (I.6 **H3**), I cannot send to zero any element of $\mathcal{A} \otimes \mathcal{B}$, so the basis of $\mathcal{A} \otimes \mathcal{B}$ cannot be "bigger" than $\mathcal{C}$

preparation independence

$$\begin{tikzcd}
\mathcal{A} \times \mathcal{B} \arrow[r, "T"] \arrow[dr, "m"'] & \mathcal{A} \otimes \mathcal{B} \arrow[d, "I"] \\
 & \mathcal{C}
\end{tikzcd}$$

I is an isomorphism: $\mathcal{C} \cong \mathcal{A} \otimes \mathcal{B}$

Postulate removed

- We showed that we can recover the tensor product for the composite system based on very narrow physically motivated requirements (preparation independence and the composite made of only the parts)
- Could we use something else apart from the tensor product? Yes! We could use other maps that introduce arbitrary gauges and phase flips. But why should we make our life complicated, since we can always pick a representation that behaves nicely?
- Now we know **exactly**, at both a physical level and a mathematical level, why we use the tensor product for composite systems in quantum mechanics

THE COMMENTARY

The commentary

- Note how the starting points are simple, yet the derivation is not
- There are two details that have been sources of confusion, let us go through them quickly

The anti-linear debacle

- Some take the anti-linear case to be physically distinct (e.g. related to time reversal)

T. Matolcsi, Tensor product of Hilbert lattices and free orthodistributive product of orthomodular lattices, Acta Sci. Math. (Szeged), **37**, 263 (1975).

Theorem 1. *Let H_1 and H_2 be Hilbert spaces*, $\dim H_1 \geqq 3$, $\dim H_2 \geqq 3$. *If the Hilbert spaces are complex, then there exist exactly two* $\boxed{\textit{(non-equivalent)}}$ *tensor products of $P(H_1)$ and $P(H_2)$ satisfying the condition of fullness. They are given by*

$$\text{(i)}\quad H = H_1 \otimes H_2, \quad u_1(M_1) = M_1 \otimes H_2, \quad u_2(M_2) = H_1 \otimes M_2;$$

$$\text{(ii)}\quad H = \overline{H}_1 \otimes H_2, \quad u_1(M_1) = \overline{M}_1 \otimes H_2, \quad u_2(M_2) = \overline{H}_1 \otimes M_2,$$

where $\otimes$ denotes the usual tensor products of Hilbert spaces.

The anti-linear debacle

- The fact that the conjugate representation is physically equivalent was something known to the founders of quantum mechanics

as well as by (3). We thus arrive at one of the two equations

$$\nabla^2\psi - \frac{8\pi^2}{h^2}V\psi \mp \frac{4\pi i}{h}\frac{\partial\psi}{\partial t} = 0. \tag{4''}$$

We will require the complex wave function ψ to satisfy one of these two equations. Since the conjugate complex function $\bar{\psi}$ will then satisfy the *other* equation, we may take the real part of ψ as the real wave function (if we require it). In the case of a conservative system

E. Schrödinger, Annalen der Physik **102**, 81 (1926); English translation in E. Schrödinger, *Collected papers on Wave Mechanics* (Blackie & Son, London, 1928).

E. Wigner, On Unitary Representations of the Inhomogeneous Lorentz Group, Annals Math. **40**, 149 (1939).

It follows from the second condition[5] that there either exists a unitary operator S by which the wave functions $\Phi^{(2)}$ of the second representation can be obtained from the corresponding wave functions $\Phi^{(1)}$ of the first representation

$$\Phi^{(2)} = S\Phi^{(1)} \tag{4}$$

or that this is true for the conjugate imaginary of $\Phi^{(2)}$. Although, in the latter case, the two representations are still equivalent physically, we shall, in keeping with the mathematical convention, not call them equivalent.

Maybe we should stop doing that?!?!?

No tensor product on Hilbert spaces

- Another objection comes from the use of the universal property of the tensor product
- The objection is that, in the category of Hilbert spaces, the universal property of the tensor product yields nothing: there is no tensor product (according to category theory) https://www-users.cse.umn.edu/~garrett/m/v/nonexistence_tensors.pdf
- In the proof, we use the universal property on linear spaces (not Hilbert spaces) so there is no issue

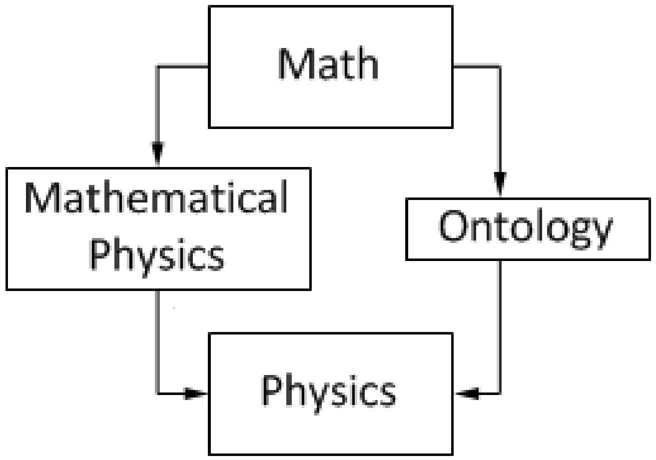
Do we have the “right” math?

- Conceptually simple physical premises lead to complicated math
- The anti-linear case is not part of the same mathematical category, and causes confusion
- The category of Hilbert spaces does not yield a “correct” categorical tensor product
- There is a mismatch between the physical content of the theory and the mathematical structures we use to represent it
- Should we, in physics, perhaps stop simply using the tools the mathematicians create for themselves, and maybe start developing some that have a tighter connection to the physics (though still mathematically sound)?

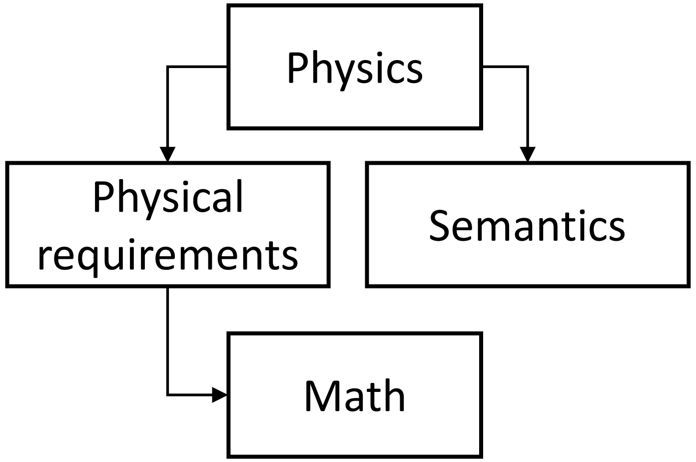
Math/physics relationship

Quantum mechanics (like other modern theories) starts by setting the mathematical structure

In general, there is a prevalent attitude that mathematics comes "before" the physics (through an interpretation)



From Wikipedia "Mathematical Physics"



To do this properly, we have to have to understand EXACTLY what each mathematical construct (Hilbert space, differential geometry, manifolds, real numbers, topology, etc…) physically represents and under what assumptions

Assumptions of Physics

- This is the approach we follow in our broader project Assumptions of Physics (see https://assumptionsofphysics.org/):
    - Identify a specific physical requirement (e.g. scientific theory must be grounded in experimental verifiability)
    - Encode that requirement in the math (e.g. the lattice of statements must be generated by a countable set of verifiable statements)
    - We prove results (e.g. the set of physically distinguishable cases form a T_0 second countable topological space, they can't exceed the cardinality of the continuum, causal relationships are topologically continuous functions …)
- To do this, we need to coalesce ideas from different fields of physics (classical, quantum, thermo, stat mech, …), mathematics (including foundations), computer science and philosophy of science

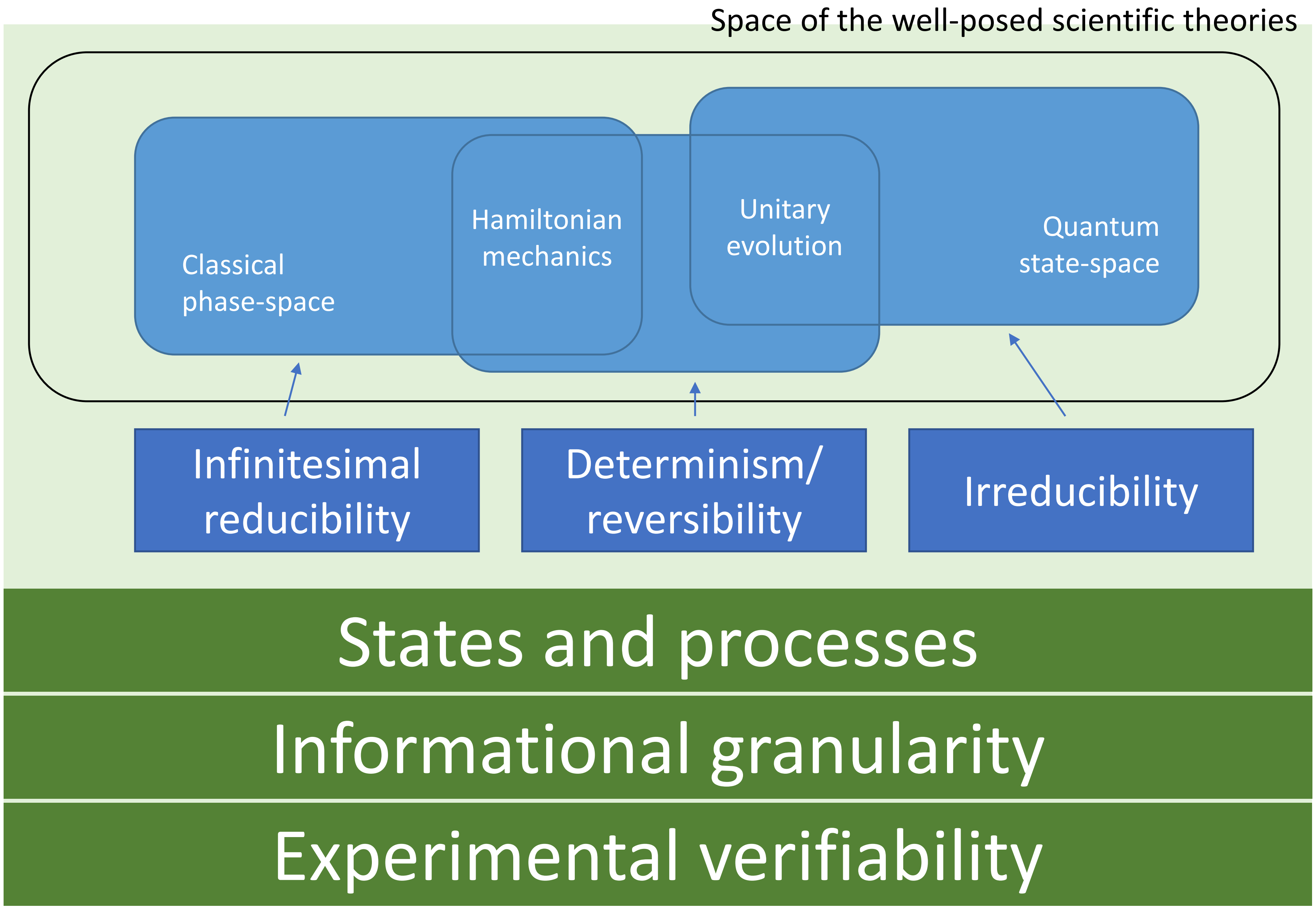
Physical theories

Specializations of the general theory under the different assumptions

Assumptions

General theory

Basic requirements and definitions valid in all theories



SUPPLEMENTAL

Example of colinear but non-linear map

- Let $\mathcal{H}$ be a two dimensional Hilbert space. Let $e_1, e_2 \in \mathcal{H}$ be a basis. Define the map $m : \mathcal{H} \to \mathcal{H}$ such that

$$m(c_1 e_1 + c_2 e_2) = c_1 e_1 + c_2 e^{\imath\theta \frac{|c_2|}{\sqrt{|c_1|^2+|c_2|^2}}} e_2$$

cosine of the angle across basis

- The map is colinear (maps rays to rays):

$$m(kv) = m\big(k(c_1 e_1 + c_2 e_2)\big) = m\big((kc_1)e_1 + (kc_2)e_2\big)$$
$$= kc_1 e_1 + kc_2 e^{\imath\theta \frac{|kc_2|}{\sqrt{|kc_1|^2+|kc_2|^2}}} e_2 = kc_1 e_1 + kc_2 e^{\imath\theta \frac{|k||c_2|}{|k|\sqrt{|c_1|^2+|c_2|^2}}} e_2 = km(v)$$

- The map is not linear (linear only if $\theta = 0$):

$$m(e_1) = e_1 \quad m(e_2) = e^{\imath\theta} e_2 \quad m(e_1 + e_2) = e_1 + e^{\imath\theta/\sqrt{2}} e_2$$

- If we don't fix the "correct" phase at the basis, a continuous map will change the phase gradually as we go from one basis vector to the other; the phase shift will depend on the angle between the basis, creating the non-linearity

Anti-linear

$$\langle \psi | \phi \rangle \mapsto \langle \phi | \psi \rangle$$

$$\langle \psi | O\phi \rangle \mapsto \langle O\phi | \psi \rangle = \langle \phi | O^\dagger \psi \rangle$$

Self-adjoint: $O = O^\dagger$

$$O \mapsto O$$

Skew-adjoint: $O = -O^\dagger$

$$O \mapsto -O$$

Time reversal

$E \mapsto E$ $\quad$ $t \mapsto -t$

$Q \mapsto Q$ $\quad$ $P \mapsto -P$

$\theta \mapsto \theta$ $\quad$ $S \mapsto -S$

Self-adjoint