Information-theoretic derivation of quantum theory via quantum logic

Alexei Grinbaum
LPHS, Nancy, France

Found. Phys. Lett. **18** (6) 563-572, 2005

Thesis

- Quantum theory is a general theory of information constrained by several important information-theoretic principles. It can be formally derived from the corresponding information-theoretic axiomatic system.

Historical context

- **Information-theoretic approach:**

  Wheeler (1978, 1988), Rovelli (1996), Steane (1998), Fuchs (2001), Brukner and Zeilinger (2002), Clifton, Bub and Halverson (2003), Jozsa (2004).

- **Axiomatic approach:**

  La théorie physique moderne manifeste une tendance certaine à rechercher une présentation *axiomatique*, sur le modèle des axiomatiques mathématiques. L'idéal axiomatique, emprunté à la géométrie, revient à définir tous les « objets » initiaux d'une théorie uniquement par des *relations*, nullement par des qualités substantielles.

  Jean Ullmo (1958)

Reconstruction

1. Physical principles
2. Mathematical axioms
3. Formal derivation

Rovelli

Quantum mechanics will cease to look puzzling only when we will be able to *derive* the formalism of the theory from a set of simple physical assertions ("postulates", "principles") about the world. Therefore, we should not try to append a reasonable interpretation to the quantum mechanical formalism, but rather to *derive* the formalism from a set of experimentally motivated postulates.

Mackey

M1 Function p is a probability measure. Mathematically, we have $p(x, f, \emptyset) = 0$, $p(x, f, \mathbb{R}) = 1$, and $p(x, f, M_1 \cup M_2 \cup M_3 \ldots) = \sum_{n=1}^{\infty} p(x, f, M_n)$ whenever the M_n are Borel sets that are disjoint in pairs.

M2 Two states, in order to be different, must assign different probability distributions to at least one observable; and two observables, in order to be different, must have different probability distributions in at least one state. Mathematically, if $p(x, f, M) = p(x', f, M)$ for all f in $\mathcal{S}$ and all M in $\mathfrak{B}$ then $x = x'$; and if $p(x, f, M) = p(x, f', M)$ for all x in $\mathcal{O}$ and all M in $\mathfrak{B}$ then $f = f'$.

M3 Let x be any member of $\mathcal{O}$ and let u be any real bounded Borel function on the real line. Then there exists y in $\mathcal{O}$ such that $p(y, f, M) = p(x, f, u^{-1}(M))$ for all f in $\mathcal{S}$ and all M in $\mathfrak{B}$.

M4 If $f_1, f_2, \ldots$ are members of $\mathcal{S}$ and $\lambda_1 + \lambda_2 + \ldots = 1$ where $0 \leq \lambda_n \leq 1$, then there exists f in $\mathcal{S}$ such that $p(x, f, M) = \sum_{n=1}^{\infty} \lambda_n p(x, f_n, M)$ for all x in $\mathcal{O}$ and M in $\mathfrak{B}$.

M5 Call *question* an observable e in $\mathcal{O}$ such that $p(e, f, \{0, 1\}) = 1$ for all f in $\mathcal{S}$. Questions e and e' are disjoint if $e \leq 1 - e'$. Then a question $\sum_{n=1}^{\infty} e_n$ exists for any sequence (e_n) of questions such that e_m and e_n are disjoint whenever $n \neq m$.

M6 If E is any compact, question-valued measure then there exists an observable x in $\mathcal{O}$ such that $\chi_M(E) = E(M)$ for all M in $\mathfrak{B}$, where χ_M is a characteristic function of M.

M7 The partially ordered set of all questions in quantum mechanics is isomorphic to the partially ordered set of all closed subspaces of a separable, infinite-dimensional Hilbert space.

M8 If e is any question different from 0 then there exists a *state* f in $\mathcal{S}$ such that $m_f(e) = 1$.

M9 For each sequence (f_n) of members of $\mathcal{S}$ and each sequence (λ_n) of non-negative real numbers whose sum is 1, one-parameter time evolution group $V_t : \mathcal{S} \mapsto \mathcal{S}$ acts as follows: $V_t\left(\sum_{n=1}^{\infty} \lambda_n f_n\right) = \sum_{n=1}^{\infty} \lambda_n V_t(f_n)$ for all $t \geq 0$; and for all x in $\mathcal{O}$, f in $\mathcal{S}$, and M in $\mathfrak{B}$, $t \to p(x, V_t(f), M)$ is continuous.

Hardy

H1 *Probabilities.* Relative frequencies (measured by taking the proportion of times a particular outcome is observed) tend to the same value for any case where a given measurement is performed on an ensemble of n systems prepared by some given preparation in the limit as n becomes infinite.

H2 *Simplicity.* K is determined by a function of N where $N = 1, 2, \ldots$ and where, for each given N, K takes the minimum value consistent with the axioms.

H3 *Subspaces.* A system whose state is constrained to belong to an M dimensional subspace behaves like a system of dimension M.

H4 *Composite systems.* A composite system consisting of subsystems A and B satisfies $N = N_A N_B$ and $K = K_A K_B$.

H5 *Continuity.* There exists a continuous reversible transformation on a system between any two pure states of that system.

CBH

CBH1 No superluminal information transfer via measurement.

CBH2 No broadcasting.

CBH3 No bit commitment.

Outline

I. **Philosophy of this approach**

II. **Quantum logical derivation of quantum theory**

Four points

1. Epistemological attitude: one is concerned with *theories.*

Science is the construction of theories. A theory is an *objective description of certain phenomena*, while selection criteria for phenomena and the understanding of objectivity of the description vary depending on the particular theory in question. All such phenomena, however, have *repeatable* traits. A theory is a description of repeatable traits of the observed phenomena with the goal of predicting these traits in unobserved (unknown, future) phenomena.

Four points

2. Universality: there is no repeatable trait of any phenomenon excluded from being described by a theory.

This is a claim of universal applicability of theoretic description. It does not imply the existence of a "theory of everything." Moreover, it excludes non-repeatable traits of phenomena from the scientific domain.

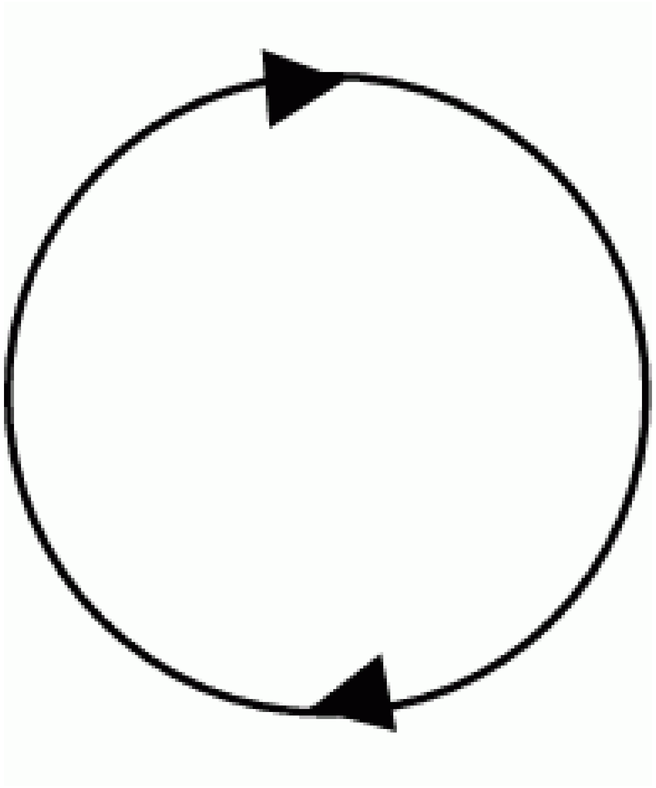
Four points

3. Various theories can be depicted in the loop form.

Theories can be classified by what they assume and by what they explain. Concepts assumed in one theory can be explained in another. This is a claim that, if depicted graphically and interconnected by their ‘end concepts,’ theories do not form a pyramid of infinite regress or another form; they form a circle, a loop.

Four points

3. The set of all theories can be depicted in the loop form.



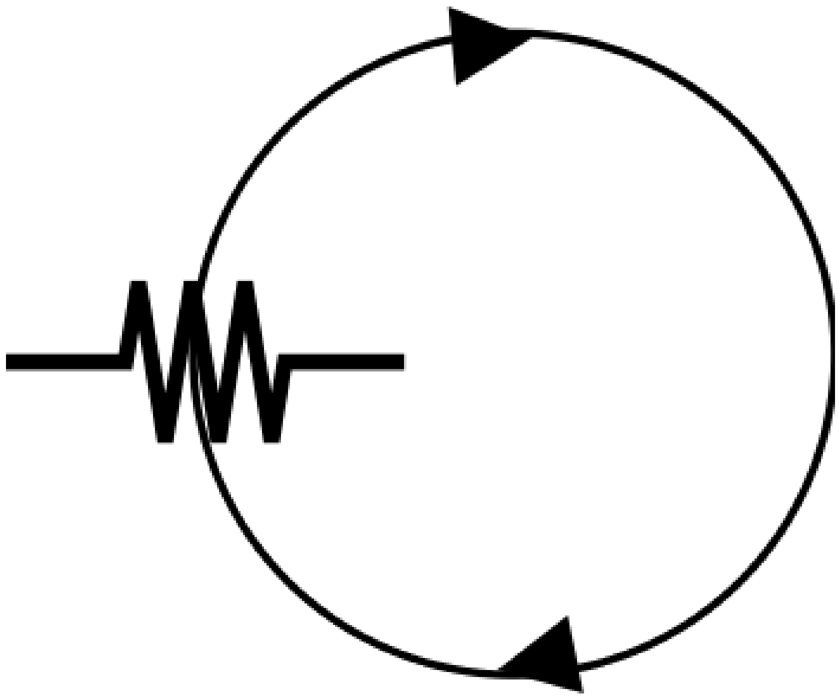
Four points

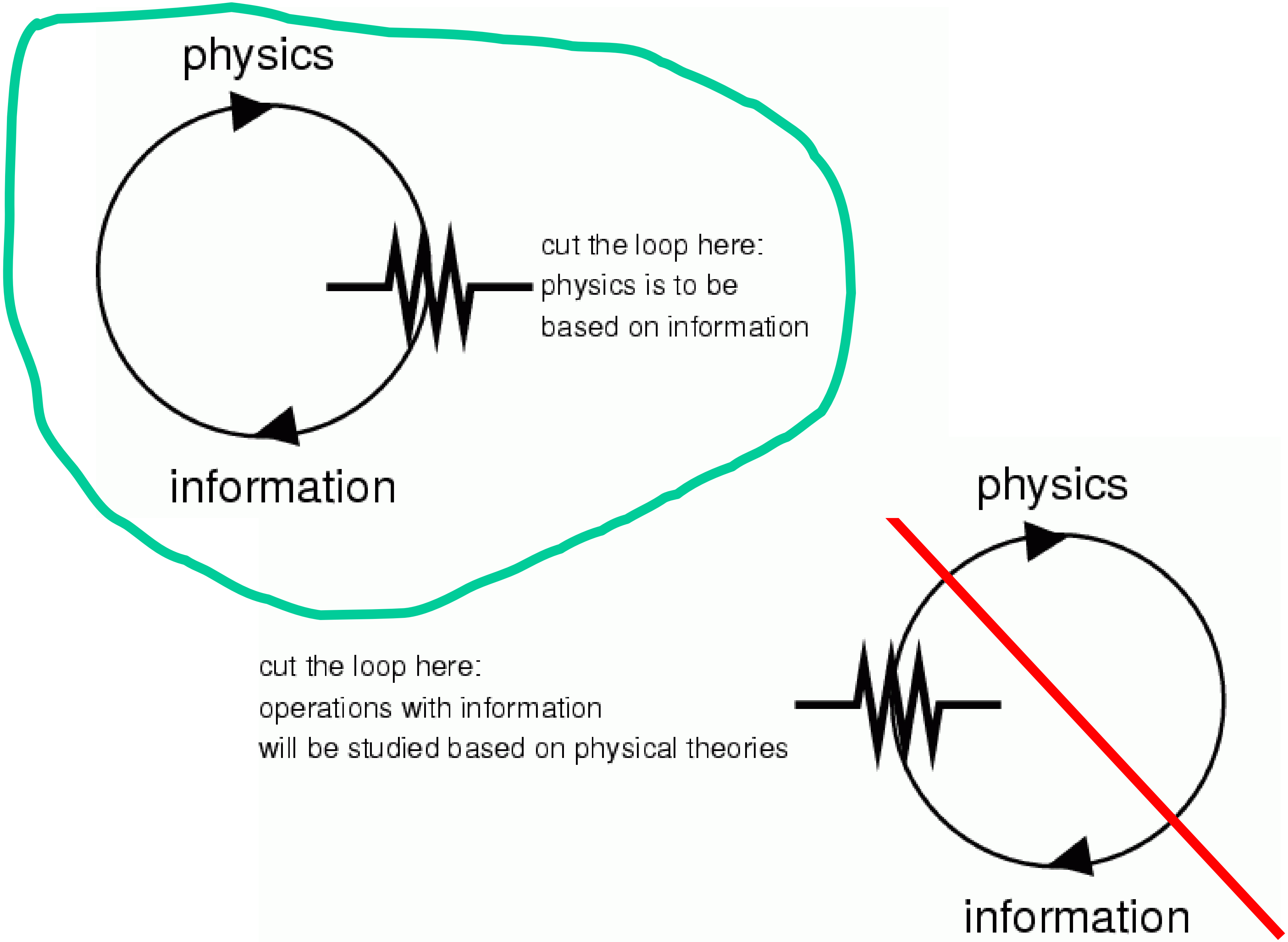
4. Construction of a particular theory requires at least one loop cut.

The cut separates *explanans* from *explanandum*, what is assumed from what is derived. A theory of the loop uncut is a logical circularity. Cutting the loop must be seen as a condition of possibility of the theoretical description as such.

Four points

4. Construction of a particular theory requires at least one loop cut.




physics
information
cut the loop here:
physics is to be
based on information
physics
information
cut the loop here:
operations with information
will be studied based on physical theories

Language

Fundamental notions	**Formal representation**
System	Systems S, O, P…
Information	Yes-no questions
Fact (act of bringing about information)	Answer to a yes-no question (given at time *t*)

Axioms

Axiom I: There is a maximum amount of **relevant** information that can be extracted from a system.

Axiom II: It is always possible to acquire new information about a system.

What do we need to reconstruct?

- Obtain Hilbert space and prove quantumness
- Obtain Born rule with the state space
- Obtain unitary dynamics

Quantum logical reconstruction of the Hilbert space

1. Definition of the lattice of yes-no questions.
2. Definition of orthogonal complement.
3. Definition of relevance and proof of orthomodularity.
4. Introduction of the space structure.
5. Lemmas about properties of the space.
6. Definition of the numeric field.
7. Construction of the Hilbert space.

Relevance: Motivation

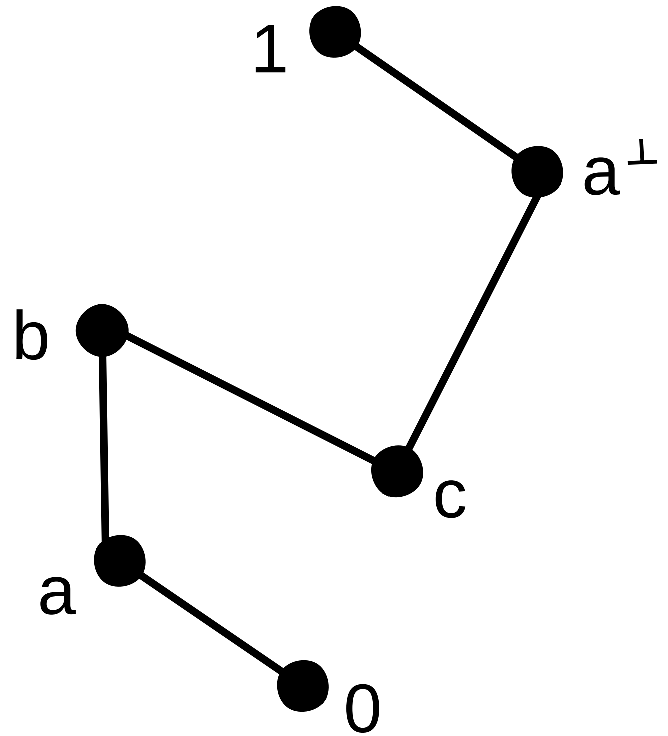
- Consider *b* such that it entails the negation of *a*: $b \rightarrow \neg a$
  - If the observer asks *a* and obtains an answer to *a*…

    …but then asks a *genuine* new question *b*, it means that the observer expects either a positive or a negative answer to b.

    - This, in turn, is only possible if information *a* is no more relevant; indeed, otherwise the observer would be bound to always obtain the negative answer to *b*.
      - We say that, by asking *b*, the observer renders *a* irrelevant.

Definition of Relevance

- Question *b* is called irrelevant with respect to question *a* if $b \wedge a^{\perp} \neq 0$.



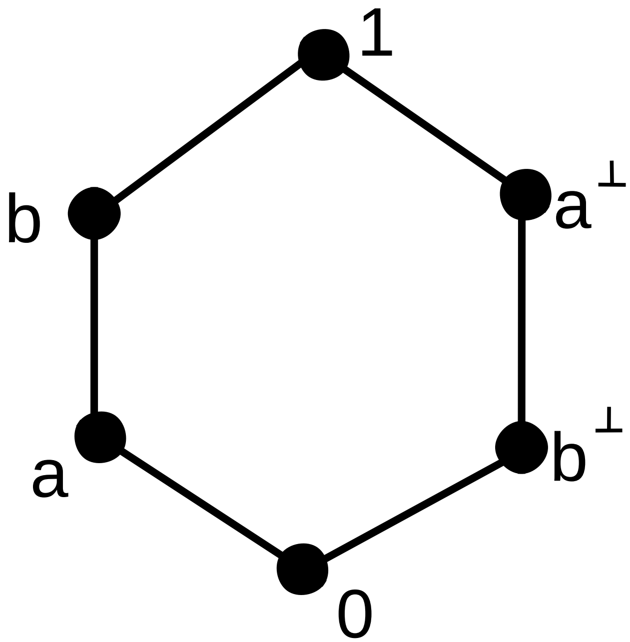
- Trivial in Hilbert lattices: $x \leq y$ are relevant with respect to y, all others irrelevant.
- Non-trivial if used to *derive* what in Hilbert lattices is *assumed.*

Non-trivial notion of relevance

- Question *b* is relevant with respect to question *a*

  and

- b ≥ a



Amount of information

- Two assumptions:

1. If relevance is not lost, the amount of information grows monotonously as new information comes in.
2. The lattice contains all possible information (yes-no questions). Thus, there are sufficiently many questions as to bring about any *a priori* allowed amount of information.

Proof of Orthomodularity

- By Axiom I there exists a finite upper bound of the amount of relevant information, call it *N*. Select an arbitrary question *a* and consider a question *ã* such that {a,ã} bring N bits of information. Then $a^{\perp} \wedge \tilde{a} = 0$.
- Lemma: An orthocomplemented lattice is orthomodular if and only if $a \leq b$ and $a^{\perp} \wedge b = 0$ imply $a = b$.
- Question *b* is relevant with respect to *a*; and question *ã* is relevant with respect to *b*.
- Consider {a,b,ã}. If $b > a$, this sequence preserves relevance and brings about strictly more than N bits of relevant information.
- From the contradiction follows $a = b$.

Kalmbach's theorem

Infinite-dimensional Hilbert space characterization theorem:

Let *H* be an infinite-dimensional vector space over real or complex numbers or quaternions. Let *L* be a complete orthomodular lattice of subspaces of *H* which satisfies:

(i) Every finite-dimensional subspace of *H* belongs to *L*.

(ii) For every element *U* of *L* and for every finite-dimensional subspace *V* of *H*, linear sum *U*+*V* belongs to *L*.

Then there exists an inner product *f* on *H* such that (*H*,*f*) is a Hilbert space with *L* as its lattice of closed subspaces.

Step 6: Definition of the numeric field

- Axiom VII: The underlying numeric field of *V* is one of the real or complex numbers or quaternions, and the involutory anti-automorphism (conjugation) is continuous.

- Substitutes: Solèr’s theorem assuming existence of an infinite orthonormal sequence of vectors. Also: Zieler, Holland, Landsman

Step 7: Construction of the Hilbert space

- **Theorem**:

  Let W(P) be an ensemble of yes-no questions that can be asked to a physical system and V a vector space over real or complex numbers or quaternions such that a lattice of its subspaces L is isomorphic to W(P).

  Then there exists an inner product f on V such that V together with f form a Hilbert space.

Quantumness

Axiom II: It is always possible to acquire new information about a system.

Criterion: Orthomodular lattice, in order to describe a quantum mechanical system, must be nondistributive.

- Lemma: all Boolean subalgebras of L(V) are proper.

- Corollary: W(P) is non-Boolean.

State space and the Born rule

- Axiom III

  "Intratheoretic non-contextuality":
  If information is obtained by an observer, then it is obtained independently of how the measurement was eventually conducted, i.e. independent of the measurement context.

  "No metainformation":
  If information *I* about a system has been brought about, then it happened without bringing in information *J* about the fact of bringing about information *I*.

- Gleason's theorem builds the state space.

Time and unitary dynamics

- Assume isomorphism between $W_t(P)$ at different time moments. In other words, time evolution commutes with orthogonal complementation, hence with relevance.
- Wigner's theorem: unitary or anti-unitary transformation $U(t_1,t_2)$: $W_{t_1}(P) \rightarrow W_{t_2}(P)$. Select unitary transformation only in virtue of the condition of continuity in the limit $t_2 \rightarrow t_1$.
- Stone's theorem: Hamiltonian description $U(t_2 - t_1) = \exp[-i(t_2 - t_1)H]$.

POVM description

- Twofold role of the observer:
Observer is at the same time a physical system (P-observer) and an informational agent (I-observer). Information-based physical theory must give an account of P-observer, while I-observer must remain meta-theoretic.
- Starting with an orthogonal projector description of measurement and factoring out P-observer, one obtains the general POVM description of measurement.

List of axioms

Information-theoretic axioms:

I. There is a maximum amount of relevant information that can be extracted from a system.

II. It is always possible to acquire new information about a system.

III. If information I about a system has been brought about, then it happened independently of information J about the fact of bringing about information I.

Supplementary assumptions:

IV. For any two yes-no questions there exists a yes-no question to which the answer is positive if and only if the answer to at least one of the initial question is positive.

V. For any two yes-no questions there exists a yes-no question to which the answer is positive if and only if the answer to both initial questions is positive.

VI. The lattice of questions is complete.

VII. The underlying field of the space of the theory is one of the numeric fields R, C or D and the involutory anti-automorphism in this field is continuous.

Open questions

1. Information-theoretic meaning of Axioms IV, V, and VI.
2. Information-theoretic meaning of Axiom VII or of Solèr's theorem.
3. Meaning of the probability function *f* used in Gleason's theorem.
4. Origin of assumptions concerning time evolution.
5. Problem of dimension of the Hilbert space.
6. Superselection rules.

I say that it is not illogical to think that the world is infinite. Those who judge it to be limited postulate that in remote places the corridors and stairways and hexagons can conceivably come to an end—which is absurd. Those who imagine it to be without limit forget that the possible number of books does have such a limit. I venture to suggest this solution to the ancient problem: *The Library is unlimited and cyclical*.

J.L. Borges *The Library of Babel*