Kochen-Specker Theorem

Introduction and Proof

Jan-Philipp Christ

LMU Munich

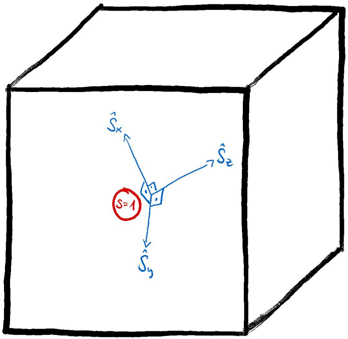
May 30, 2023

Overview

Measuring the squared spin components of a particle

Spin-1-Particle in a box

- Setup a Spin-1-Particle (e.g. atomic carbon $1s^2 2s^2 2p^2$ in triplet ground state) in a box
- Spin Operator $\underline{\hat{S}} = (\hat{S}_x, \hat{S}_y, \hat{S}_z)^T$ with well known commutation relations: $[\hat{S}_i, \hat{S}_j] = i\hbar\varepsilon_{ijk}\hat{S}_k,\ i,j \in \{x,y,z\}$
- $\hat{S}_i$ has three eigenvalues $s_i = -1\hbar, 0, 1\hbar$ for $i \in \{x,y,z\}$ (from now on $\hbar = 1$)



- $\hat{S}_i, \hat{S}_j, i \neq j$ not compatible observables
- $\left\{\hat{S}_i^2\right\}_{i\in\{x,y,z\}}$ are compatible observables
  Proof:

$$
\begin{aligned}
[\hat{S}_i^2, \hat{S}_j^2] &= \hat{S}_i\hat{S}_j[\hat{S}_i, \hat{S}_j] + \hat{S}_i[\hat{S}_i, \hat{S}_j]\hat{S}_j + \hat{S}_j[\hat{S}_i, \hat{S}_j]\hat{S}_i + [\hat{S}_i, \hat{S}_j]\hat{S}_j\hat{S}_i \\
&= i(\underbrace{\hat{S}_i\hat{S}_j\varepsilon_{ijk}\hat{S}_k + \hat{S}_i\varepsilon_{ijk}\hat{S}_k\hat{S}_j}_{=0(j\leftrightarrow k)} + \underbrace{\hat{S}_j\varepsilon_{ijk}\hat{S}_k\hat{S}_i + \varepsilon_{ijk}\hat{S}_k\hat{S}_j\hat{S}_i}_{=0(j\leftrightarrow k)}) \\
&= 0
\end{aligned}
$$

- measure $\hat{S}_i^2$ along any given axis
- $\hat{S}_i^2$ has two eigenvalues $s_i^2 = 0, 1$ for $i \in \{x, y, z\}$
- $\hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2 = s(s+1)\big|_{s=1} = 2$

$\Rightarrow$ measuring the squared spin components in three perpendicular directions yields (101) and permutations thereof

BUT:

Consequence of Kochen-Specker Theorem

The measurement outcome CANNOT be the result from detecting (hypothetically) predetermined values of the squared spin components

Definition 101-function (cp. [CK08]):

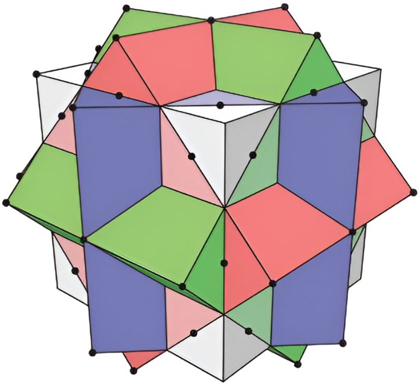
1. assigns measurement outcome to an axis
2. Opposite directions give the same answer for measuring the squared spin components.
3. Two perpendicular directions cannot both be 0.
4. Three perpendicular directions cannot all be 1.

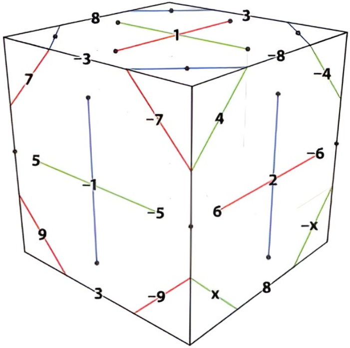
We will see:

Kochen-Specker Theorem (our early version)

$\nexists$ 101-function for arbitrary directions

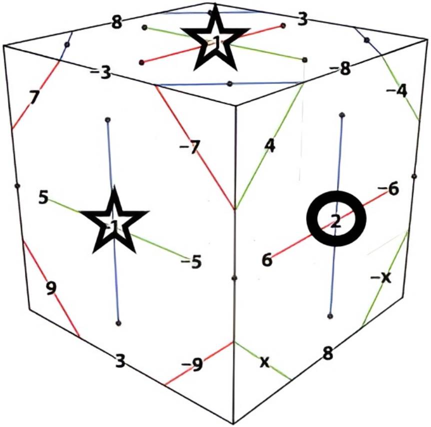
Proof of the KS Theorem
(following [CK08] and [Per91])



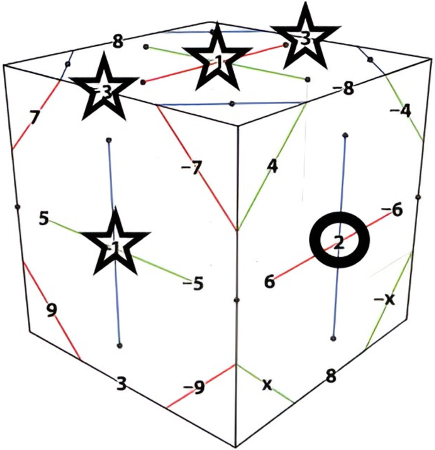


images from [CK08], p. 3, upscaled

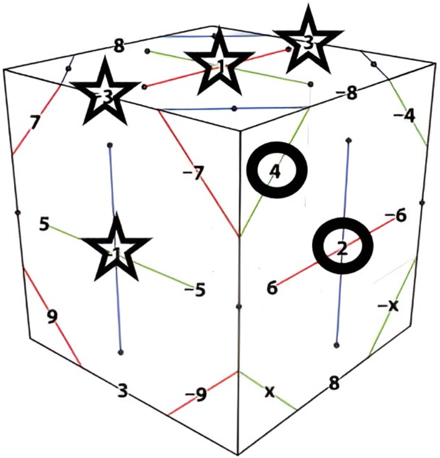
- star node: 101-function takes value 1 on it; circle node: 101-function takes value 0 on it
- axis defined by point goes to center of cube, each coordinate triple (a, b, c) we are interested in defines three orthogonal axes.
- $(1, -1, 2)$ is orthogonal triple, w.l.o.g. assign circle to 2



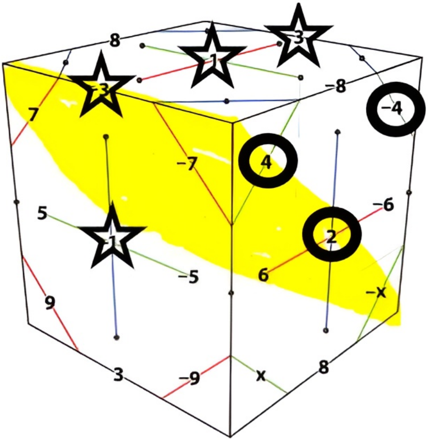
- $(2, 3, -3)$ is orthogonal triple $\Rightarrow \pm 3$ star



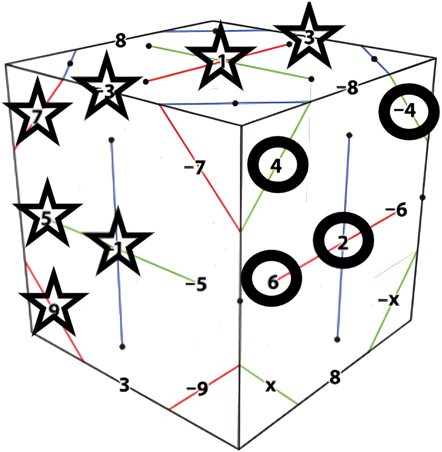
- $(4, -x, 3)$ is orthogonal triple
- w.l.o.g. 4 is circle (proof can proceed analogously if we say -4 is circle)



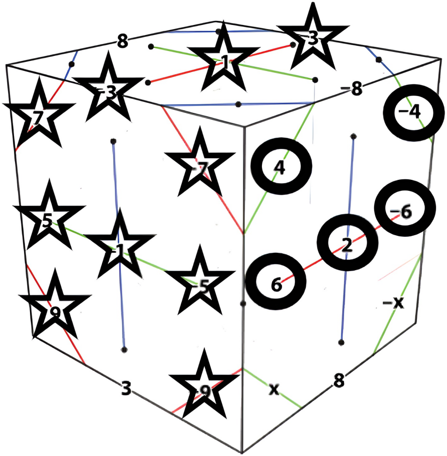
- $(-4, x, -3)$ fixes that either -4 OR x is circle because -3 is star
- if -4 is star: reflect around yellow plane
  $\rightarrow$ -4 and x are interchanged every other circle or star node is left invariant
  $\Rightarrow$ we may set -4 to circle



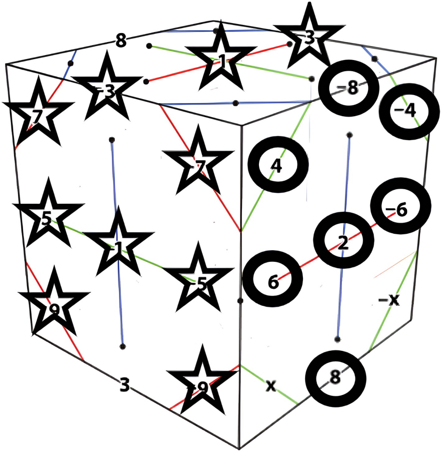
- 5 is orth. to 4 $\Rightarrow$ 5 star
- $(1,5,6)$ is orth. triple $\Rightarrow$ 6 circle
- $(6,7,9)$ is orth. triple $\Rightarrow$ 7 and 9 star



- -5 is orth. to -4 ⇒ -5 star
- $(1, -5, -6)$ is orth. triple ⇒ -6 circle
- $(-6, -7, -9)$ is orth. triple ⇒ -7 and -9 star



- $(8, -7, 9)$ is orth. triple $\Rightarrow$ 8 has to be circle
- $(-8, 7, 9)$ is orth. triple $\Rightarrow$ -8 has to be circle
- BUT: 8 is orthogonal to -8 $\Rightarrow$ $\lightning$



Tracing the origin of the contradiction

What assumptions were (secretly) made?

(NA1) The squared spin components are well defined for arbitrary orthogonal axes

(NA2) (measurement) value of the sum of the squared spin components = Sum of the (measurement) values of the squared spin components

(NA3) (measurement) value of the squared spin component = Square of the (measurement) value of the spin component

These lead directly to...

The Kochen-Specker Theorem

(our semi-final version, cited from [Hel22])

KS Theorem

Let $\mathcal{H}$ be a Hilbert space of QM state vectors of dimension $x \geq 3$. There is a set $\mathcal{M}$ of observables on $\mathcal{H}$, containing y elements, such that the following two assumptions are contradictory:

(KS1) All y members of $\mathcal{M}$ simultaneously have values, i.e. are unambiguously mapped onto real numbers (designated, for observables $A, B, C, ... \in \mathcal{M}$ by $\nu(A), \nu(B), \nu(C), ... \in \mathbb{R}$

(KS2) Values of all observables in $\mathcal{M}$ conform to the following constraints:

(a) $A, B, C \in \mathcal{M}$ pairwise compatible with $C = A + B \Rightarrow \nu(C) = \nu(A) + \nu(B)$

(b) $A, B, C \in \mathcal{M}$ pairwise compatible with $C = A \cdot B \Rightarrow \nu(C) = \nu(A) \cdot \nu(B)$

Remarks:

- There is no such statement for $\dim\mathcal{H} = x < 3$
- If proven for $x = 3$, the theorem follows for $\dim\mathcal{H} > 3$ because $\mathcal{H}$ has at least one subspace of dimension three in which the statement holds.
- If we do not assume a correspondence between operators on $\mathcal{H}$ and observables in $\mathcal{M}$, (KS2) could be understood as a definition of the addition and multiplication of observables.
- (KS2)$\Rightarrow$(NA2) and (KS2)$\Rightarrow$(NA3)
- (KS1)$\Rightarrow$(NA1)

$\Rightarrow$ The non-existence of 101-functions proves the KS Theorem in full strength.

For future reference:

(VD) All observables defined for a QM system have definite values at all times.

Obviously (KS1) $\Longleftrightarrow$ (VD)

For the term 'measurement' to be well-defined in the way we used it before we need another assumption:

(NC) If a QM system possesses a property (value of an observable), then it does so independently of any measurement context, i.e. independently of *how* that value is eventually measured.

The FUNC Principle

Valuation Functions in Classical Theory

(following [Flo13])

Physical quantities

Any physical quantity A is represented by function from the state space S to the reals

$$f_A : S \to \mathbb{R}$$

We call $f_A(s_i)$ the value of A for a state $s_i \in S$.

Remark:

- Let A be a physical quantity which is described by two distinct functions f_A, f'_A (from the state space to the reals). Then $\exists s \in S : f_A(s) \neq f'_A(s) \Rightarrow$ the value of A is not well-defined for a state s ↯ to classical theory
  $\Rightarrow$ for each A there corresponds exactly one function f_A

Classical valuation functions

$\forall s_i \in S$ define the valuation function V_{s_i} as the function which assigns an observable A from the set observables $\mathcal{O}$ the value of A for the state:

$$V_{s_i} : \mathcal{O} \to \mathbb{R},\ A \mapsto V_{s_i}(A) := f_A(s_i)$$

The functional composistion principle

in classical theory

The classical FUNC principle

For all functions $h : \mathbb{R} \to \mathbb{R}$, physical quantities $A \in \mathcal{O}$ and states $s_i \in S$:

$$V_{s_i}(h(A)) = h(V_{s_i}(A))$$

where $h(A) := (h \circ f_A)$

Example:
Let $h : \mathbb{R} \to \mathbb{R}, x \mapsto x^2$ and $A := E$ be the energy of a state s_i. The classical FUNC principle now tells us: *"the value of the energy squared is equal to the square of the value of the energy"*

Valuation Functions in Quantum Theory

- if a state $|\psi\rangle$ is not an eigenstate of $A \in \mathcal{O}$, it makes not sense to extend the classical definition for valuation functions to quantum systems

The quantum FUNC principle

A valuation function for quantum theory is a map $V_{|\psi\rangle} : \mathcal{O} \to \mathbb{R}$ which satisfies the following two conditions:

(*i*) $V_{|\psi\rangle}(\hat{A}) \in \sigma(\hat{A}) \subseteq \mathbb{R}$ represents the value of the operator $\hat{A}$ given a state $|\psi\rangle \in \mathcal{H}$. $\sigma(\hat{A})$ is the spectrum of $\hat{A} \in \mathcal{O}$

(*ii*) FUNC:

$$V_{|\psi\rangle}(h(\hat{A})) = h(V_{|\psi\rangle}(\hat{A}))\ \forall h : \mathbb{R} \to \mathbb{R}$$

Remark:

- In some contexts it makes sense to define $V_{|\psi\rangle} : \mathcal{O} \to P([\mathbb{R}])$ where $P([\mathbb{R}])$ is the set of all probability distributions.

What is meant by $h(\hat{A})$?

1. If $|\psi\rangle \in \mathcal{H}$ and $\hat{A}|\psi\rangle = a|\psi\rangle$ for some $a \in \mathbb{R}$:

$$h(\hat{A})|\psi\rangle := h(a)|\psi\rangle$$

2. $\hat{A} = \sum_{a\in\sigma(\hat{A})} a\hat{P}_a \to h(\hat{A}) := \sum_{a\in\sigma(\hat{A})} h(a)\hat{P}_a$

Remarks:

(a) $\hat{P}_a$ is the projector on the eigenspace of eigenvalue $a \in \sigma(\hat{A})$

(b) We assume discrete $\sigma(\hat{A})$. Generalization for general $\sigma(\hat{A}) \subseteq \mathbb{R}$ possible, but not trivial ($\to$ Projection-valued measures)

The statistical functional composition principle

The STAT FUNC Principle

Let $\hat{A}$ be a self-adjoint operator representing an observable A and let $f : \mathbb{R} \to \mathbb{R}$ be a function. Then for all $a \in \mathbb{R}$:

$$\text{prob}\left[V_{|\psi\rangle}(f(\hat{A})) = a\right] = \text{prob}\left[f(V_{|\psi\rangle}(\hat{A})) = a\right]$$

- Note that we do not require (ii) from the definition of the quantum FUNC principle for our $V_{|\psi\rangle}$ here.
- Gleasons Theorem [Gle57]: On a Hilbert space $\mathcal{H}$ with $\dim\mathcal{H} \geq 3$, the only probability measures are of the form

$$\mu(\hat{P}_\alpha) := \text{Tr}\left(\hat{P}_\alpha \hat{\rho}\right)$$

  where $\hat{P}_\alpha$ is a projection operator and $\hat{\rho}$ is the operator which characterizes the system's state (”density matrix”).

$$\Rightarrow \text{prob}\left[V_{|\psi\rangle}(\hat{A}) = a\right] = \text{Tr}(\hat{P}_{|a\rangle} \cdot \hat{P}_{|\psi\rangle})$$

Proving the STAT FUNC Principle I

- Characteristic function: $\chi_r : \mathbb{R} \to \{0,1\}, t \mapsto \chi_r(t) = \begin{cases} 1, r = t \\ 0, r \neq t \end{cases}$
- $\chi_r(\hat{A}) := \sum_{a \in \sigma(\hat{A})} \chi_r(a)\hat{P}_a = \begin{cases} \hat{P}_r, r \in \sigma(\hat{A}) \\ 0, \text{else} \end{cases}$

  $\Rightarrow \text{prob}\left[V_{|\psi\rangle}(\hat{A}) = a\right] = \text{Tr}(\hat{P}_{|a\rangle} \cdot \hat{P}_{|\psi\rangle}) = \text{Tr}(\chi_a(\hat{A}) \cdot \hat{P}_{|\psi\rangle})$
- $\chi_r(f(\hat{A})) = \left\{\begin{matrix} \hat{P}_r, f(r) \in \sigma(f(\hat{A})) \\ 0, \text{else} \end{matrix}\right\} = \chi_{f^{-1}(r)}(\hat{A})$

Proving the STAT FUNC Principle II

$$\begin{aligned}
\text{prob}\left[V_{|\psi\rangle}(f(\hat{A})) = a\right] &= \text{Tr}\left(\chi_{f^{-1}(a)}(\hat{A}) \cdot \hat{P}_{|\psi\rangle}\right) \\
&= \text{Tr}\left(\hat{P}_{f^{-1}(a)} \cdot \hat{P}_{|\psi\rangle}\right) \\
&= \text{prob}\left[V_{|\psi\rangle}(\hat{A}) = f^{-1}(a)\right] \\
&= \text{prob}\left[f(V_{|\psi\rangle}(\hat{A})) = a\right]
\end{aligned}$$

$\square$

The FUNC Principle: Proof I

Value realism

$\forall \alpha \in [0,1]$ and for all operators $\hat{A}$ s.t. $\alpha = \text{prob}\left[V(\hat{A}) = \beta \in \mathbb{R}\right]$ there corresponds an observable A with value β

$$\left.\begin{array}{l}\text{(NC)}\\ \text{(VD)}\\ \text{(VR)}\\ \text{STAT FUNC}\end{array}\right\} \Rightarrow \text{FUNC}$$

The FUNC Principle: Proof II

Proof:

- Consider an observable B represented by the (self-adjoint) operator $\hat{B}$ and a state $|\psi\rangle$.
- From (VD): $\exists b \in \mathbb{R}: \; V_{|\psi\rangle}(\hat{B}) = b$
- For $f : \mathbb{R} \to \mathbb{R}$ we obtain $f(V_{|\psi\rangle}(\hat{B})) = f(b) =: a \in \mathbb{R}$
- From STAT FUNC:
  prob $\left[f(V_{|\psi\rangle}(\hat{B})) = a\right] = \text{prob}\left[f(V(_{|\psi\rangle}(\hat{B})) = a\right] \Rightarrow \exists$ self-adjoint $f(\hat{B})$
- From (VR): The observable which corresponds to $f(\hat{B})$ exists and has value $a \Rightarrow f(V_{|\psi\rangle}(\hat{B})) = V_{|\psi\rangle}(f(\hat{B}))$.
- From (NC): a is unique $\square$

Consequences of the FUNC Principle I

The sum (KS2a) and product rule (KS2b) follow from FUNC:

Proof of (KS2a):

- From functional analysis: $\forall \hat{A}, \hat{B}$ compatible and $\exists f, g : \mathbb{R} \to \mathbb{R}, \hat{C}$ s.t. $\hat{A} = f(\hat{C}), \hat{B} = g(\hat{C})$
- Let $h := f + g \Rightarrow \hat{A} + \hat{B} = h(\hat{C})$
- $V(\hat{A} + \hat{B}) = V(h(\hat{C})) = h(V(\hat{C})) = f(V(\hat{C})) + g(V(\hat{C})) = V(f(\hat{C})) + V(g(\hat{C})) = V(\hat{A}) + V(\hat{B})$

(KS2a) follows similiarly.

Consequences of the FUNC Principle II

From the sum and product rule we immediately get:

- The identity operator has to take value 1
- The zero operator has to take value 0
- Projectors have to take either value 0 or 1

Consequences of the FUNC Principle III

Alternate form of KS Theorem

Let $\mathcal{H}$ be a Hilbert space of QM state vectors of dimension $x \geq 3$. There is a set $\mathcal{M}$ of observables on $\mathcal{H}$, containing y elements, such that the following two assumptions are contradictory:

(KS1') All y members of $\mathcal{M}$ simultaneously have values, i.e. are unambiguously mapped onto real numbers (designated, for observables $A, B, C, ... \in \mathcal{M}$ by $\nu(A), \nu(B), \nu(C), ... \in \mathbb{R}$.

(KS2') Values of all observables obey the FUNC principle.

Consequences of KS Theorem

Consequences of KS Theorem

If our reasoning to this point was correct (i.e. the axioms of QM themselves are flawed!), (KS1$^{(')}$) or (KS2$^{(')}$) has to be incorrect.

$$\rightsquigarrow \text{John :)}$$

Bibliography I

[CK08] John Conway and Simon Kochen. *The Strong Free Will Theorem*. 2008. arXiv: 0807.3286 [quant-ph].

[Per91] A Peres. "Two simple proofs of the Kochen-Specker theorem". In: *Journal of Physics A: Mathematical and General* 24.4 (1991), p. L175. DOI: 10.1088/0305-4470/24/4/003. URL: https://dx.doi.org/10.1088/0305-4470/24/4/003.

[Hel22] Carsten Held. "The Kochen-Specker Theorem". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta and Uri Nodelman. Fall 2022. Metaphysics Research Lab, Stanford University, 2022.

[Flo13] Cecilia Flori. *A First Course in Topos Quantum Theory*. Vol. 868. Jan. 2013. ISBN: 978-3-642-35712-1. DOI: 10.1007/978-3-642-35713-8.

Bibliography II

[Gle57] Andrew M. Gleason. “Measures on the Closed Subspaces of a Hilbert Space”. In: *Journal of Mathematics and Mechanics* 6.6 (1957), pp. 885–893. ISSN: 00959057, 19435274. URL: `http://www.jstor.org/stable/24900629` (visited on 04/29/2023).