

## Negative and Complex Probability in Quantum Information

Vasil Penchev<sup>1</sup>

**Abstract:** “Negative probability” in practice. Quantum Communication: Very small phase space regions turn out to be thermodynamically analogical to those of superconductors. Macro-bodies or signals might exist in coherent or entangled state. Such physical objects having unusual properties could be the basis of quantum communication channels or even normal physical ones ... Questions and a few answers about negative probability: Why does it appear in quantum mechanics? It appears in phase-space formulated quantum mechanics; next, in quantum correlations ... and for wave-particle dualism. Its meaning:- mathematically: a ratio of two measures (of sets), which *are not collinear*; physically: the ratio of the measurements of two physical quantities, which *are not simultaneously measurable*. The main innovation is in the mapping between phase and Hilbert space, since both are sums. Phase space is a sum of cells, and Hilbert space is a sum of qubits. The mapping is reduced to the mapping of a cell into a qubit and vice versa. Negative probability helps quantum mechanics to be represented quasi-statistically by quasi-probabilistic distributions. Pure states of negative probability cannot exist, but they, where the conditions for their expression exists, decrease the sum probability of the integrally positive regions of the distributions. They reflect the immediate interaction (interference) of probabilities common in quantum mechanics.

Key words: negative probability, quantum correlation, phase space, transformation between phase and Hilbert space, entanglement, Bell’s inequalities

---

<sup>1</sup>Institute of Philosophical Research, Bulgarian Academy of Science; [vasildinev@gmail.com](mailto:vasildinev@gmail.com);  
<http://vasil7penchev.wordpress.com>, <http://my.opera.com/vasil%20penchev>,  
<http://www.esnips.com/web/vasilpenchevsnews>, <http://www.philosophybulgaria.org>

The main highlights are:

1. Negative or complex probability appears where the measure of two “parts” or of a “part” and the “whole” forms any angle. It can happen when the probability sum is geometric in form.

2. Negative or complex probability cannot be excluded from consideration in quantum mechanics since any quantum object consists of two “parts”: wave and particle .

3. This form of probability is why the effects known as “relative rotated measures” are observed in any separated quantum object as well as in the systems of quantum objects. They reflect the immediate interaction of probabilities without any “hidden” parameters.

**I. Why does negative probability appear in quantum mechanics?**

Negative probability appears in phase-space formulated quantum mechanics; also, in quantum correlations and also in wave-particle dualism. Why?

One may give a few answers to that question.

The definition of negative probability expressed mathematically is the following: a ratio of two measures (of sets) which *are not collinear*; or physically: a ratio of the measurement of two physical quantities, which *are not simultaneously measurable*.

An revealing example is Heisenberg’s uncertainty measure:  $\Delta x \Delta p^x \geq \frac{h}{4\pi}$ . Its expression as a negative probability would mean that the real axis of the measurement either of  $x$ , or of  $p^x$  has been rotated between the coupled measurements; where  $\Delta x = |x_1 - x_2|$  and  $\Delta p^x = |p_1^x - p_2^x|$ :

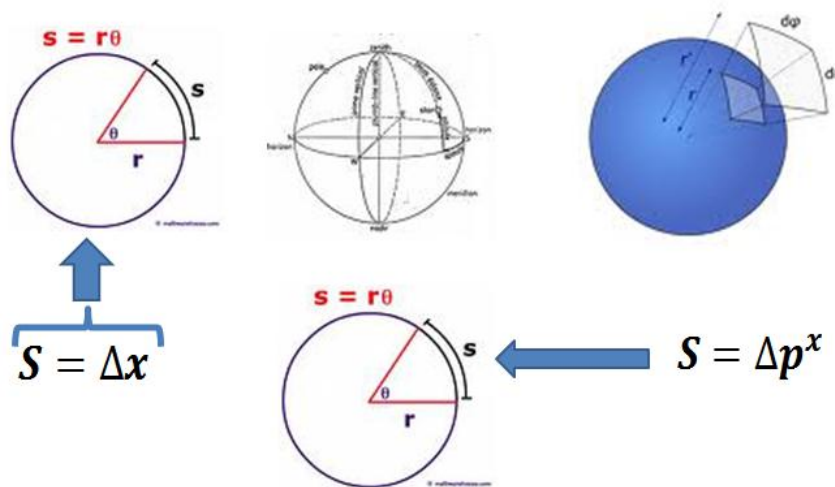


Fig. 1. The mechanism of Heisenberg’s uncertainty

**II. The “appearance” of negative probability in the phase-space formalism of quantum mechanics** can be represented schematically, using a correspondence with an ordinary Hilbert-space formalism of the phase-space or by a similar correspondence between their “atoms”, i.e. a phase-space cell and a Hilbert-space qubit.

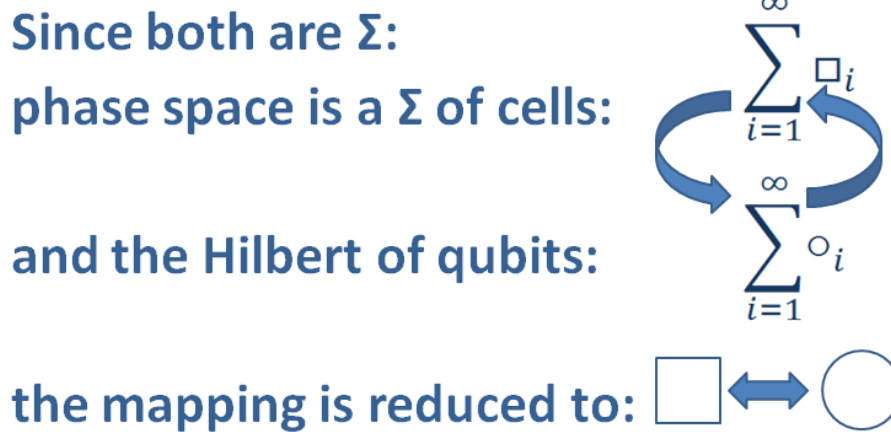


Fig. 2. The mapping between phase and Hilbert space reduced to the transformation between a cell and a qubit

The question is how the transformation between the two-dimensional cell of phase space and the qubit of Hilbert space works:

$$\Psi(q) = \sum_{i=1}^{\infty} C_i \circ_i \Leftrightarrow \sum_{i=1}^{\infty} A_i \square_i = P(x, p)$$

$$c_i \in \mathbb{C}, \text{ but } 1 \leq P(\circ_i) = \frac{|C_i|^2}{|\sum_1^{\infty} C_i|^2} \geq 0$$

$$A_i \in \mathbb{R}, \text{ but } 1 \leq P(\square_i) = \frac{A_i}{\sum_1^{\infty} A_i} \geq -1$$

If  $x$  and  $p$  is simultaneously measurable, then a cell can be transformed into a qubit:

$$P(x, p) = \frac{1}{\pi\hbar} \int_{-\infty}^{\infty} \psi^*(x + y)\psi(x - y)e^{2ipy/\hbar}$$

‘ $\Psi \rightarrow P$ ’ is the Wigner function (1932) in a contemporary view. Wigner function in its original form is:

$$P(x_1, \dots, x_n; p_1, \dots, p_n) = \left(\frac{1}{h\pi}\right)^n \int_{-\infty}^{\infty} \dots \int dy_1 \dots dy_n \psi(x_1 + y_1 \dots x_n + y_n)^* \psi(x_1 - y_1 \dots x_n - y_n) e^{2i(p_1 y_1 + \dots + p_n y_n)/h}$$

(Wigner 1932: 750).

Wigner transformation in general:

$$\mathbb{H} \ni \widehat{G} \rightarrow g(x, p) \in \mathbb{P}\hbar$$

$$g(x, p) = \int_{-\infty}^{\infty} ds e^{ips/\hbar} \langle x - s/2 | \widehat{G} | x + s/2 \rangle$$

Reverse Wigner transformation (Weyl transformation) in general:

$$\mathbb{P}\hbar \ni g(x, p) \rightarrow \widehat{G} \in \mathbb{H}$$

$$\langle x | \widehat{G} | y \rangle = \int_{-\infty}^{\infty} \frac{dp}{h} e^{ip(x-y)/\hbar} g\left(\frac{x+y}{2}, p\right)$$

Hermann Weyl's paper (1927) is historically first. He considered abstractly and mathematically the transformation  $\mathbb{H} \ni (q) \Leftrightarrow P(x, p) \in \mathbb{P}\hbar$ , but did not interpret  $P(x, p)$  as probability, and did not discuss the fact that it can obtain negative values.

The original Weyl transformation (1927: 116-117):

$$f(p, q) = \iint_{-\infty}^{+\infty} e(p\sigma + q\tau) \xi(\sigma, \tau) d\sigma d\tau$$

$$F = \iint_{-\infty}^{+\infty} e(P\sigma + Q\tau) \xi(\sigma, \tau) d\sigma d\tau$$

It follows:

$$F[f] = \frac{1}{2\pi} \iint \iint f(q, p) e[(P - p)\sigma + (Q - q)\tau] dq dp d\sigma d\tau$$

$F[f]$  turns out to be partly analogical to Dirac's  $\delta$ -functions (Schwartz distributions).

Now, we are going to consider the time-frequency (= time-energy) reformulation of Wigner function derived by [Ville \(1948\)](#):

“The transmission of communication signals is accomplished by means of a transmission of energy, generally of electromagnetic or of acoustic energy. ... It is not energy itself which is of interest, but rather the changes in this energy in the course of time. The more complicated the function which represents, as a function of time, the change in voltage, current, pressure, or any other carrier, the greater is the amount of the information carried by the transmitted energy”<sup>2</sup> ([Ville 1948: 63](#)).

When thinking about quantum information, it is especially important to look at all physical processes as informational ones; not simply transmitting, but also being themselves

---

<sup>2</sup> The English translation of the citations of Ville's paper is according to the translation by I. Selin, “Theory and Applications of the Notion of Complex Signal” – <http://www.rand.org/pubs/translations/2008/T92.pdf> .

signals. The necessary condition for such a viewpoint is the reciprocity of time and energy implied by the Heisenberg uncertainty measure (coordinates and momentums being formally linked to these quantities, and consequently, by the quantum/ discrete character of mechanical motion.

We may discuss  $\Psi$ -function as a special kind of complex signal modulated by dual processes:

“Complex signals ... may be considered as the result of the modulation of their envelope by a carrier is itself frequency modulated” ([Ville 1948: 67](#)).

Very important for the interpretation of negative probability is the following corollary made by Ville:

“Any signal modulated by a sufficiently high frequency may be considered complex” ([Ville 1948: 68](#)).

Consequently, any “high energetic“, physical object-like macro-body may be considered as such a complex signal modulated by dual processes. A similar corollary may be found described as follows:

*„For any signal  $s(t)$ , the function*

$$\bar{\Psi}(t) = s(t)e^{j\omega_0 t} \quad \omega_0 > 0$$

*which in general is not complex approaches the complex signal  $\Psi(t)$  (associated with  $s(t) \cos \omega_0 t$ ) as  $\omega_0$  increases”* ([Ville 1948: 68](#)).

We are inclined to consider  $\Psi$ -function in quantum mechanics “analogically” (cf., digitally) namely as the sum of an infinite series of energetic constants. Any signal of high enough frequency, and one, which may consequently be “discussed” as complex, dual-modulated, is in fact aperiodic, enough to represent  $\mathbf{s}(t)$ .

A crucial part of it happens, or maybe is better to say, is happening in a given time interval, which we may denote as “now and here”.

Consequently, setting time-frequency analysis, on the one hand, and  $\Psi$ -function, on the other, a closer connection with the former originates from the “material particle” of classical physics, but a much clearer meaning within much wider limits is needed for that scenario: those of quantum generalization.

Under quantum generalization a point-particle still remains localized, but inherently partitioned. By means of the “duality” it smoothly turns into a wave, and this originates from an isolated and defined part of the localized point-particle, in turn affecting the totality.

This quantum status, the whole as a set of its possibly non-additive parts, and completely in a Gibbsian manner, is already thought equivalent to the unity of corresponding probable states of the whole, or “worlds”. Eventually they correlate to each other.

Naturally, the question is: If the signal is a complex one, then where is the boundary of the two modulations; how is it to be distributed in the “space”; and by what is it determined? Or in other words: Where is the boundary between the signal part that is transmitted by amplitude and amplitude modulation; and its alternative part that is transmitted by frequency, phase and frequency-phase modulation?

The answer à la Ville is the following:

“The proposition which we shall come to use is itself an immediate consequence of the fact that a complex signal is characterized by the peculiarity of having a spectrum whose amplitude is zero for negative frequencies. Now, modulating  $s(t)$  by  $e^{j\omega_0 t}$  amounts to causing the spectrum of  $s(t)$  to be translated by the amount  $\omega_0$ . For a large value of  $\omega_0$ , the spectrum lies entirely in the region of positive frequencies, and  $s(t)e^{j\omega_0 t}$  becomes complex” (Ville 1948: 68-69).

Using wave-particle “spectacles” reveals a “corresponding” perspective. Ville’s answer has an obvious and simple interpretation: Amplitude and AM codes the particle properties of a quantum object while frequency and phase (FM and PM) code its wave properties. That hints at Cohen’s generalization of Wigner’s function:

“We now ask what the analytic signal procedure has done in terms of choosing the particular amplitude and phase, that is, what is special about the amplitude and phase to make an analytic signal? Generally speaking, the answer is that the spectral content of the amplitude is lower than the spectral content of  $e^{j\varphi(t)}$ ” (Cohen 1995: 35).

So, we can restrict the spectrum of amplitude within a fixed frequency interval, corresponding to the (introduced above) “now”, by the following two steps:

The first is: correspondingly linear dependence; and the second is: arbitrary time-frequency dependence: –

“ $A(t)e^{j\omega_0 t}$  is analytic if the spectrum of  $A(t)$  is contained within  $(-\omega_0, \omega_0)$ ” (Cohen 1995: 36).

“ $A(t)e^{j\varphi(t)}$  is analytic if the spectrum of  $A(t)$  is contained within  $(-\omega_1, \omega_1)$  and the spectrum of  $e^{j\varphi(t)}$  is zero for  $\omega \leq \omega_1$ ” (Cohen 1995: 36).

The division specific for each complex signal between a part transmitted by amplitude and a part transmitted by frequency and phase, suggests a new idea and an even more

decisive generalization. Resulting from this generalization is the following interpretation: We have arrived at a clear division between ‘finite’ and ‘infinite’ in an arithmetical, set-theoretical and meta-mathematical sense; or between ‘syntactic’ and ‘semantic’ in a logical view.

For example, the size of amplitude, treated as a signal and thus as coded information, could be thought of as the Gödel coding of any finite syntax, or the axiomatic-deductive kernel of a certain set of tautologies, i.e. as logic: Whether in an absolute or in a relative meaning, i.e. as the “logic of A”, it can be thought of as the logic of a certain thematic domain.

A bit more detailed development of these ideas follows:

After sketching the generalization of the Wigner function<sup>3</sup>, in which it is interpreted also as a signal and after applying a time-frequency analysis of it, we would hint at a new exposition of  $\Psi$ -function in order to allow of (and treat of) transferring results between logic and quantum information.

At that time specifically, the logical equivalent of ‘negative probability’ could be investigated. It will be distributed along the boundary between syntax and semantics and thus, describing the transition between them.

Logic understood as a formal system is depicted as “syntax” and represented by a finite set of rules for constructing formally correct sentences. However, that set is absolutely indifferent to ‘name’ that is “semantics hidden behind symbol”.

In 20<sup>th</sup> century, a usage of terms like “the logic of A” has been propagated. “A” has been understood as one or another subject domain, and consequently as semantics. Thus the idea that the syntax depends on semantics has been involved implicitly, and vice versa: in which case syntax has been thought of as “logic adequate for the subject area at issue”.

Now, we try to formalize how that idea acquired scientific popularity and then we will show how the  $\Psi$ -function is isomorphic to such a formalizing. Even though, it is a new, syntactic-semantic exposition that is implied by an external, but similar approach.

A general sketch of the idea could be the following:

The complex coefficient is transformed and viewed as the Gödel number of the description; that is: as the terms of the description, which are also “coded” in terms of the corresponding member of the basis of Hilbert space!

---

<sup>3</sup> “The Wigner distribution, as considered in signal analysis, was the first example of joint time-frequency distribution ...” (Cohen 1995: 136).

A metaphor (which suggests also the useful idea of the relativity of ‘thing’ and ‘world’) can serve us as the basis for clearing: We divide our world into “orthogonal” things, i.e. without any common intersection between any two of them, and then, describe the investigated object using the set of its metaphors, that is by each one of the orthogonal things in the world.

As a humorous example, let’s say we want to make an inventory of ‘rabbit’ and the row of things chosen by guesswork, which turns out to be suitable, is: ‘bear’, ‘house’, ‘cockroach’, ‘concept’, ‘electron’, etc. The coefficient of ‘bear’ will be probably the biggest, as the scale of “resemblance” may be what we use, with ‘rabbit’ compared to ‘rabbit’ being the highest.

Respectively, the Gödel coding of the partial description of ‘rabbit’ in the “logic of bear” will give the biggest value of the coefficient, while “likening it to”, “comparing it to” ‘etc.’ more and more irrelevant things, will result in the coefficient expressing the complexity of the description tending to zero.

So the ‘rabbit’ is now represented by – the most convenient form would be – a simple sum of its description in any possible world: the rabbit as a bear, the rabbit as a house, the rabbit as a cockroach, the rabbit as the thing of “notion”, the rabbit as an electron, etc.

Applying such a humorous procedure, the idea of Hilbert space is represented as a mathematical equivalent of a semantic network, defined by its basis, and precise quantitative coefficients derived from the scale of the resemblance of the investigated object to anything in terms of the network.

The world as the set of things or of possible worlds may be discussed also by the mathematical concept of category.

Any one thing can be represented by any of the other things and by morphisms between the thing under discussion and the rest.

A subspace of Hilbert space will correspond to any category of this type. If the coding is one-to-one, then also that mapping is one-to-one. The fundamental nature of the notion of category reveals the fundamental nature of the notion of Hilbert space. The latter is a one-to-one coded image of the former. But more interestingly: The syntactic-semantic interpretation of  $\Psi$ -function hints that Hilbert space may be considered as a coded image of the generalization of ‘category’, loosely denoted above as “semantic network”. After which the focus is on meanings, and on the mapping morphisms between them.

In reverse, if we have restricted the categories to topoi, then logic in its usual sense can be defined on them as the axiomatic-deductive “kernel of all tautologies among the



set of all senses of meaning” (which are de facto pure semantic relations). At the same time logic as a “totality”, i.e. logic as a universal and omnipresence doctrine, is also the logic of a specific thing, implicitly taken for granted through its axioms. The deep philosophical essence of the notion of “topos”, itself serving as the basis of ‘formal’ or ‘mathematical logic’, consists in allowing “topology” to be defined and hence: “discontinuity” and “continuity”. With such a structure the generalization of Einstein’s principle of relativity for discrete transforms makes sense, I mean this in the sense of quantum mechanics.

Going immediately to the field of quantum information: – to apply the Skolemian relativity of ‘discrete’ and ‘continual’ implies that the distinction between ‘continuous’ and ‘discontinuous’ loses part of its sense, which leads us into stepping beyond logic, maybe to the “language” of being:

Let us divide an infinite set, e.g. that of integers, into two compact subsets, so that any element of the initial set belongs to just one of the two subsets. The method of diagonalization shows that such a dividing necessarily exists, after which both subsets can be put in one-to-one mapping with the initial set (Пенчев 2009: 306). Suppose we have already constructed a new, “actualist” representation of diagonalization, then we could build an arithmetic version of the so-called Skolem’s paradox, as follows:

A real number may be juxtaposed uniquely with any of the divided sets of the foregoing type, in such a way that when one of the sets is finite, the associated real number is a rational number, but when both are infinite, then it has an irrational one. Further, we can show that there exists such a one-to-one mapping between the real numbers and all divisions of that type. The set of all such divisions will be denoted by **A**.

Finally, it is evident that we can build another unique mapping between integers and the set **A**.

Since the composition of two one-one mappings is also a one-one mapping, therefore, using an “intermediate station”, we have now built a one-one mapping between the set of integers and that of real numbers and therefore the former and the last are equinumerous.

But, we have utilized an “actualist” version of diagonalization, which in its initial, “constructivist” version had been applied by Cantor to show that the cardinality of real numbers is different from that of natural or rational numbers; and since the enumerable cardinality had been alleged as “the least” cardinality of an infinite set, it therefore followed that cardinality of real numbers is one of “the greater”, though not necessarily obviously the greater cardinality (that it is clearly the greater cardinality is stated by the continuum hypothesis suggested yet again by Cantor) (Пенчев 2009: 330-331).

In the “language” of being, number is not different from word, and the physical “quantity” as information is the tool by which we measure the “unity” between words and numbers (but on the numerical, quantitative side).

The natural question, which we will just put, but without deciding is: What is its counterpart on the side of word. Is it metaphor?

Let us also mention the question of whether, or to what extent, descriptions using “terms taken for granted” can be considered as the logic of the ‘name’ reserved for the set of those terms (which name is precisely “the class of tautologies”; which will also turn out to be increased, since the variables are restricted to accepting values only in those terms, with respect to fulfilling the specified relations between those terms. Roughly speaking, ‘A is brown’ will be an additional tautology, if we have limited ourselves to the logic of all brown items.)

If we are able to build such a wide generalization or interpretation of the notion of Hilbert space and its element,  $\Psi$ -function, then we can generalize or interpret correspondingly phase space. And negative probability, considered after such a generalization would also be more widely generalized by means of generalizing the Wigner function. A milestone along this route is its generalization by Cohen.

Such a wide generalization may be denoted as a language, signal, or informational interpretation.

Essential to the scheme is Kripke’s conception (1975) that an exact logical notion of truth can be introduced by infinite syntax (below I will discuss whether “infinite” is necessary), which however remains “less” than actually infinite semantics, it seems (i.e., to be precise, that in Tarski’s “as if” semantic conception of truth, “infinity” is essential, moreover a constructive infinity, rather than semantics).

After that however, it means, using Kripke’s truth, semantic instability, which is more interesting than semantic stability (or “validity”, Lutkanov 2009: 119), that is, at least a logical tautology (law) would not be valid necessarily at the meta-level. That is to say, a new name will appear “ex nihilo” transiting from a finite to infinite syntax, or a name will disappear, which is the same in essence.

Feferman’s notion of “reflexive closure” (Feferman 1991: 1-2) will help us to clear up which is the syntactic “kernel” shared by two possible worlds (descriptions, theories). The complement of the kernel to the set of all syntactic (analytic) statements in a given world (from the participation in forming the kernel) is semantic (= syntactic) unstable in relation to all worlds (due to the participation in forming the kernel).

We may also mention the hypothesis that there is no ordinal between  $\Gamma_0$  (the [Feferman – Schütte ordinal](#)) and  $\varepsilon_0$  (including the case of coincidence between  $\Gamma_0$  and  $\varepsilon_0$ ).

An example may be given by intuitionism if we interpret the intuitionist theory of infinite set as a meta-theory towards the intuitionist theory of finite set. The law of excluded middle ceases to be valid. A new name (a semantic item) of “infinite set” appears. The semantic (= syntactic) conception of truth is presented together with a properly and only syntactic conception of truth as to infinite sets of the following kind: “Till now we have not yet known whether it is true, but we will ever understand”. Is such a concept of truth really syntactic? Yes, since truth is intended in the sequence of finite experiments. Being any experiment finite, it means that if we accomplish its finite procedure (algorithm), we can acquire a definite answer whether the tested statement (hypothesis) is true or not.

Bridging our consideration to Kripke’s concept of truth, we need a new, namely “semantic-syntactic” interpretation of  $\Psi$ -function in quantum mechanics. That is not too difficult as the ontology of possible logical world à la Kripke transfers the many-worlds interpretation of quantum mechanics into a logical language. Kripke’s concept of truth accepts the boundary between syntax and semantics as movable and it is what makes it fruitful in my opinion.

Then we are going to characterize a given possible world by the proper confinement each to its place of syntax and semantics. Consequently, for any two possible worlds, there will exist a statement which is syntactic (“analytical”) in one of them, but semantic (“synthetic”) in the other one.

The  $\Psi$ -function itself interpreted semantically-syntactically describes one and the same thing but in different ways in any possible world, and represents a catalog of all possible descriptions or of all expectations about its behavior (Schrödinger 1935 (49): 823-824).

Reality (in the usual empirical or experimental sense of the word) is not the true catalog, but its “change” since (or as) the catalog is invariant in anyone of all possible worlds, including also those, in which the thing is described as untrue or inexistent, through being exactly in that boundary between syntax and semantics: a boundary being uniquely characterised towards just that world.

The semantic-syntactic interpretation of von Neumann’s theorem (1932) about the absence of hidden parameters in quantum mechanics corresponds to “standard” quantum logic, whose base he founded in the same book: There is nothing which is true in a finite number of worlds, particularly in a single world. Truth or untruth can be defined on finite sets of any world, but only on an infinite number of possible worlds. Such a standard interpreta-

tion is consistent also with Kripke's truth. In addition, the boundary between (a) thing and (b) (or its) world remains absolute.

A semantic-syntactic interpretation of Bell's revision ([1964](#), [1966](#)), or in other words, defining the limits of validity for the foregoing theorem would correspond rather to "holistic semantics" (Cattaneo, Chiara, Giuntini, Paoli 2009: 193). If the interaction of possible worlds is introduced consisting in the interchange of possibilities (probabilities of events) in definite rules, with respect to those between the sets of possible worlds, then the semantic-syntactic interpretation of von Neumann's theorem is not valid. The boundary of validity is outlined by the missing of interaction, i.e. they outline the validity only as to isolated possible worlds. Hence, if events are progressing in more than one possible world<sup>4</sup>, then "truth" can be defined for a finite number of or even for a single world, and the boundary between a thing and a world is **not** already absolute.

This sketched "very wide" generalization and interpretation signposts us to Cohen's generalization by using its philosophical meaning drawn from the discussion of how low and high frequencies may be managed by "division" amongst the kernels:

"Therefore what the analytic procedure does, at least for signals that result in the above forms, is to put the low frequency content in the amplitude and the high frequency content in the term  $e^{j\varphi(t)}$ " (Cohen 1995: 36).

Mathematically such management is accomplished by the "kernel function":

"The approach characterizes time-frequency distributions by an auxiliary function, the kernel function. The properties of a distribution are reflected by simple constraints on the kernel, and by examining the kernel one readily can ascertain the properties of distribution. This allows one to pick and choose those kernels that produce distributions with prescribed, desirable properties. This general class can be derived by the method of characteristic functions" (Cohen 1995: 136).

That kernel function may be interpreted as a filter allowing amplitude (or time) and frequency-phase component to be divided from one another. However it can also be interpreted as an external influence of another quantum object, i.e. representing entanglement or its degree in the deep sense of quantum information.

When the generalization of the Wigner function "was subsequently realized" (Cohen 1966: 782; 1995: 136), it was also realized that an infinite number can be readily generated from:

---

<sup>4</sup> It means: The change of probability or the progressing of an event in one of the worlds changes the probability of progressing an event in another possible world.

$$P(t, \omega) = \frac{1}{4\pi^2} \iiint e^{-j\theta t - j\tau\omega + j\theta u} \phi(\theta, \tau) \cdot s^*\left(u - \frac{1}{2}\tau\right) s\left(u + \frac{1}{2}\tau\right) du d\tau d\theta$$

where  $\phi(\theta, \tau)$  is an arbitrary function called the kernel<sup>5</sup> by Claasen and Mecklenbrauker<sup>6</sup>” ([Cohen 1989: 943](#)).

The meaning can be made crystal clear and is complemented by the next two schemata:

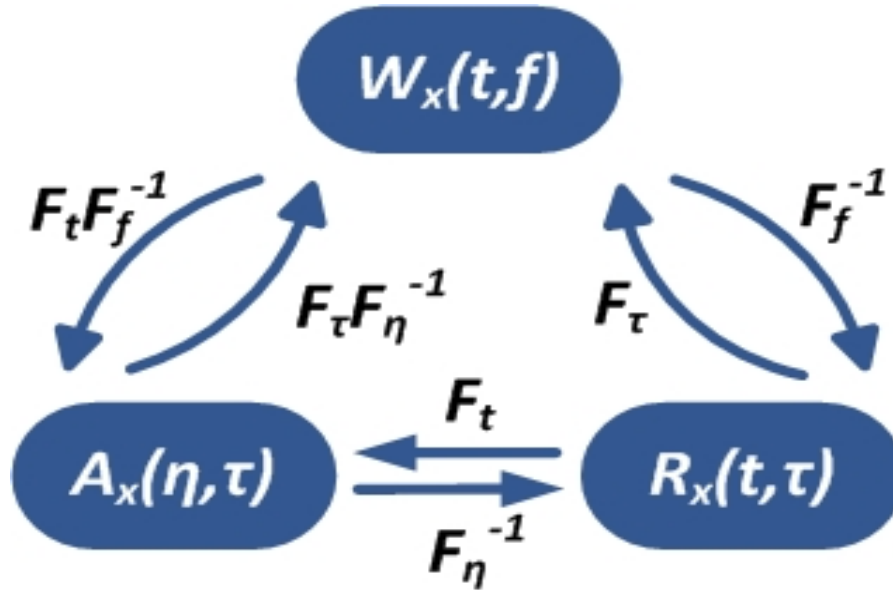


Fig. 3. The relationship between the Wigner distribution function, the auto-correlation function and the ambiguity function (from Wikipedia: [Cohen's class distribution function](#) )

The notations are the following:

$$W_x(t, f) = \int_{-\infty}^{\infty} R_x(\tau) e^{-j\pi f \tau} d\tau$$

$$R_x(t, \tau) = x(t + \tau/2) * x(t - \tau/2)$$

$$A_x(\eta, \tau) = \int_{-\infty}^{\infty} x(t + \tau/2) * x\left(t - \frac{\tau}{2}\right) e^{-j2\pi t \eta} dt$$

<sup>5</sup> “In general the kernel may depend explicitly on time and frequency and in addition may also be a functional of the signal” (Cohen 1989: 943, footnote 3).

<sup>6</sup> T. A. C. M. Claasen and W. F. G. Mecklenbrauker, "The Wigner distribution – a tool for time-frequency signal analysis; part III: relations with other time-frequency signal transformations," Philips J. Res., vol. 35, pp. 372-389, 1980.

The first one is Wigner function transformed into parameters of time and frequency from the parameters of coordinates and momentums. The second one is the autocorrelation function, and the third one the ambiguity function.

$t$  is time, and  $\tau$  the time of another instant (e.g. the function „decelerated” or „accelerated”).  $f$  is frequency in relation to the instant  $t$ , and  $\eta$  frequency in relation to the instant of  $\tau$ .  $*$  is the operation of convolution.

Obviously, the autocorrelation function expresses immediately interaction between different time instants, and the Wigner function and ambiguity function through the agency of frequency, do the same. Hence, it is clear that the introduction of negative probability for some small regions of phase space in the usual representation of the Wigner function reflects indirectly the interaction of different time instants, which is “unpacked” in consecutive order by transformation of the parameters to time-frequency, and ultimately, by the interactions connection with the autocorrelation function.

Of course, classical physics takes for granted that such a type of interaction is absent, for its availability would produce retro-causality. The additional kernel function of Cohen’s generalization allows for the degree of interaction between the instants to be regulated, particularly to be concentrated onto some of them. The Wigner function is the special case when the kernel function being 1 does not exert an effect – but it does not correspond, as we have seen, to the absence of negative probability!

The method and degree of separation between the time instants is regulated by the fundamental constant of light velocity in vacuum. From such a viewpoint Minkowski space represents the area of autocorrelation, i.e. of possible physical interaction. It is described using duality in two alternative ways: without correlations, i.e. by diffeomorphisms according to Einstein’s principle of relativity; or with correlations, i.e. in the standard way of quantum mechanics, according to which the Cohen function is reduced to the Wigner one since the kernel function is 1.

The next figure displays how kernel function can be used as a filter for the separation, with respect to increase, either of the interaction or the degree of distinction between the instants:

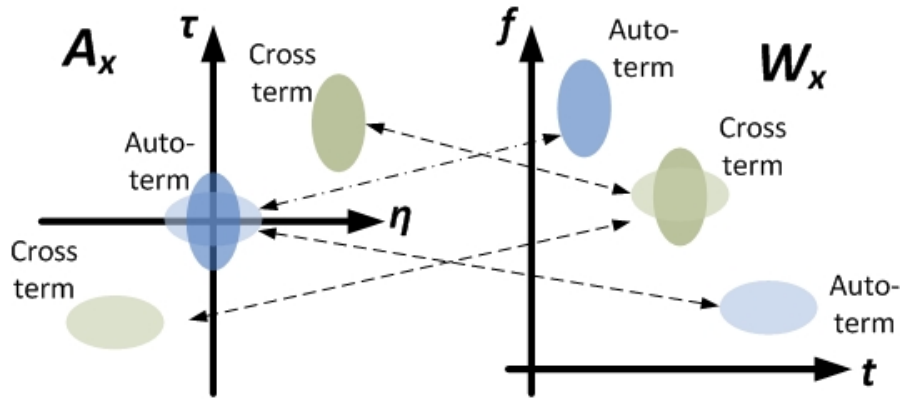


Fig. 4. What is the benefit of the additional kernel function? The figure shows the distribution of the auto-term and the cross-term of a multi-component signal in both the ambiguity and the Wigner distribution function. (Wikipedia: [Cohen's class distribution function](#) )

Furthermore it is obvious that if “straight” Wigner function can be generalized, then analogically the reverse transform of Weyl can be, and that was done by Cohen in a later paper:

Cohen’s generalization of the Weyl transform is the following:

$$“\mathcal{A}_a^\Phi(x, D) = \iint \hat{a}(\theta, \tau)\Phi(\theta, \tau)e^{i\theta x + i\tau D} d\theta d\tau,$$

where  $\Phi(\theta, \tau)$  is a two dimensional function called the kernel<sup>7</sup>. The kernel characterizes a specific transform and its properties” (Cohen 2008: 260).

It is well to keep in mind the manner of Cohen’s thought while discussing Groenewold’s statistical ideas (1946). Their essence was recapitulated by Groenewold himself as follows:

“Our problems are about:

- $\alpha$  the correspondence  $a \leftrightarrow \mathbf{a}$  between physical quantities  $a$  and quantum operators  $\mathbf{a}$  (quantization) and
- $\beta$  the possibility of understanding the statistical character of quantum mechanics by averaging over uniquely determined processes as in statistical classical mechanics (interpretation)” (Groenewold 1946: 405).

$\alpha$ , the correspondence  $a \leftrightarrow \mathbf{a}$  (quantization), in fact, generates two kinds of problems about the physical quantities  $a$ :

<sup>7</sup> Cohen 1966.

- **a** is not a continuous function (it is either a continuous one or a generalized one, i.e. a distribution);
- There exists quantities  $a_i$  whose product is not commutative.

The difficulties in  $\alpha$  (the quantization of physical quantities) reflect in the same fashion in  $\beta$  (statistical description). The negative probability of some states does appear, but they are easily interpreted physically by the regions of partial overlap between orthogonal probabilistic distributions.

The main ideas of Moyal’s statistical approach (1949) would be represented by a few significant textual citations from his works:

“Classical statistical mechanics is, however, only a special case in the general theory of dynamical statistical (stochastic) processes. In the general case, there is the possibility of ‘diffusion’ of the probability ‘fluid’, so that the transformation with time of the probability distribution need not be deterministic in the classical sense. In this paper, we shall attempt to interpret quantum mechanics as a form of such a general statistical dynamics” ([Moyal 1949: 99](#)).

A single system	Statistical description	Description by $\Psi$ -function
Determinism	Quantum objects	Probabilistic ( $P \geq 0$ ) distribution
Indeterminism	Probabilistic ( $P \geq 0, P < 0$ ) distribution	Quantum objects
Ensemble	Boltzmann non-standard	Gibbs standard

Fig. 5. Moyal’s statistical approach

According to Moyal, “*phase-space distributions are not unique for a given state, but depend on the variables one is going to measure.* In Heisenberg's words<sup>8</sup>, ‘the statistical predictions of quantum theory are thus significant only when combined with experiments

<sup>8</sup> “Heisenberg, W. *The physical principles of the quantum theory* (Cambridge, 1930), p.34.”



which are actually capable of observing the phenomena treated by the statistics” (Moyal 1949: 100);

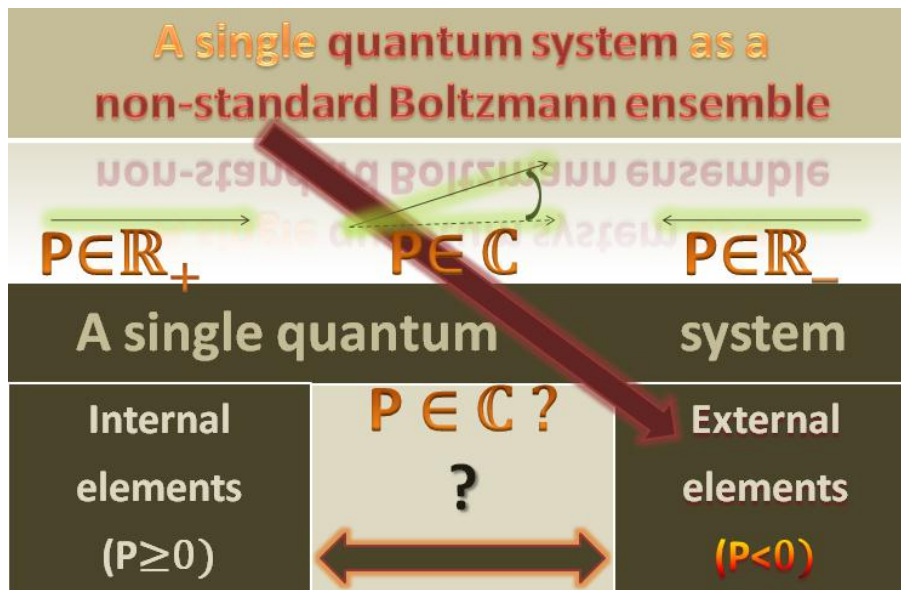


Fig. 6. Moyal’s statistical approach (1949)

## Equivalences between:

Statistical description	Description by $\Psi$ -functions
Non-standard Boltzmann ensemble	Gibbs ensemble
Parts (elements) of the system	Possible states (worlds) of the system
Essential parts (elements) of the system	Orthogonal possible states (separated worlds)
External “parts” (elements) of the system	Non-orthogonal possible states (interacting worlds)

Fig. 7. Statistical or  $\Psi$ -function description

We should emphasize the significance of ‘spin’, which is a characteristic physical quantity of any quantum object (in contrast to any object of classical physics), for making apparent and clear „probability dependence”: „symmetry (or antisymmetry) conditions *intro-*

duce a probability dependence between any two particles in B. E. (or F. D.) ensembles even in the absence of any energy interaction. ... It is this dependence which gives rise to the 'exchange energy' between the particles when they interact" (Moyal 1949: 116):

$$\mu(q_1, q_2) = \overline{q_1 q_2} - \bar{q}_1 \bar{q}_2 = \gamma \sum_{i,k} n_i n_k |Q_{nk}|^2$$

$$\mu(p_1, p_2) = \overline{p_1 p_2} - \bar{p}_1 \bar{p}_2 = \gamma \sum_{i,k} n_i n_k |P_{nk}|^2$$

$$Q_{nk} = \iint q f_{nk}(p, q) dp dq$$

$$P_{nk} = \iint p f_{nk}(p, q) dp dq$$

(Moyal 1949: 116, eq. 14.8)

The parameter  $\gamma$  accepts the following values: in a Maxwell – Boltzmann ensemble (the classical case)  $\gamma=0$ ; in a Bose – Einstein ensemble:  $\gamma=1$ ; in a Fermi – Dirac:  $\gamma= -1$ .  $n_i, n_k$  are average frequencies of the number of articles  $a_i$ , with respect to  $a_k$ , occupying a given micro-state  $\alpha_i$ , with respect to  $\alpha_k$  (Moyal 1949: 114).

Speaking of the three basic abstract mathematical spaces of our physics, namely Hilbert, phase, and Minkowski (or pseudo-Riemannian) space, they form a eureka “triangle” prompting the notion that one of its side is missing:

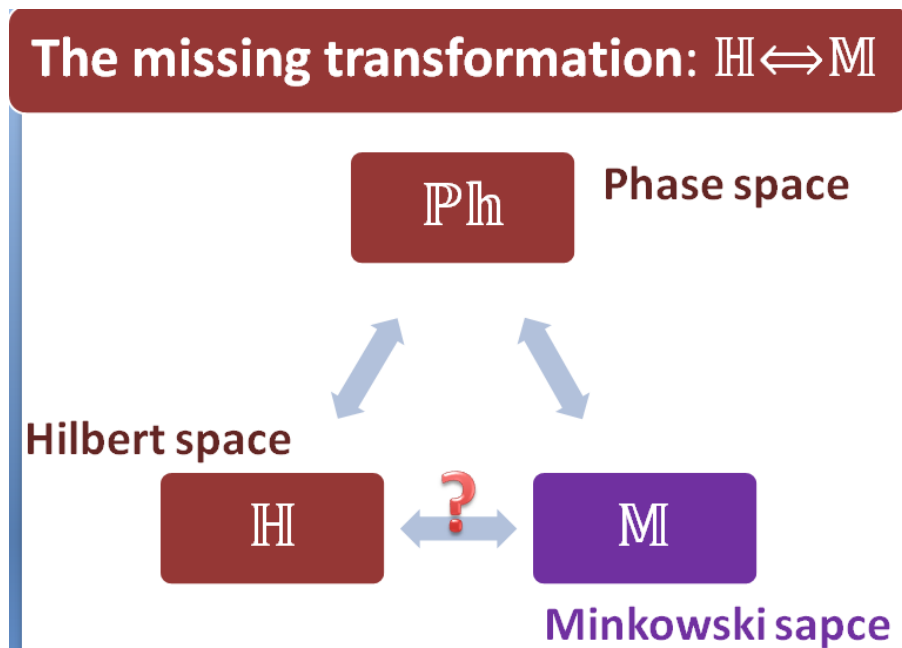


Fig. 8. Abstract mathematical spaces and transformations between them

**III. Negative probability in quantum correlations:** In the light of quantum information and quantum correlations the studied by it, the battle for or against “hidden parameters” in quantum mechanics can be interpreted and reformulated as local “hidden parameters” (causality) against nonlocal ones (quantum correlations).

The initial notions should be sought for in: [the famous „paradox” \(or argument, in fact and essence\) of Einstein, Podolsky, and Rosen](#) and [Schrödinger’s “cat paradox paper”](#), in 1935. The quantum correlation deducted and discussed in those articles implies negative probability in the final analysis.

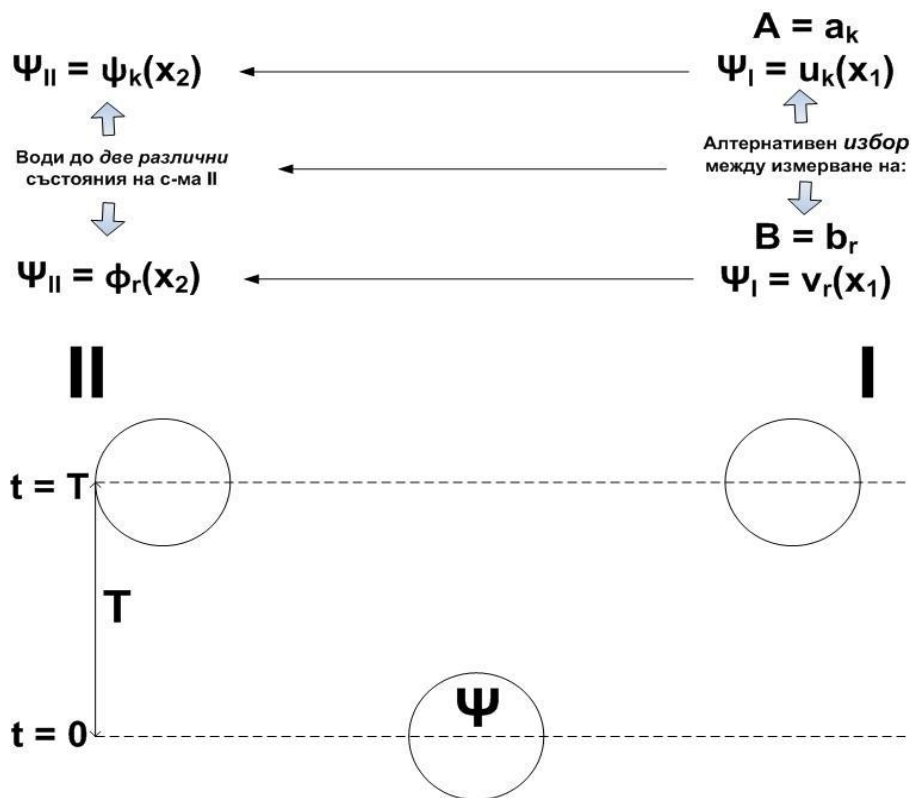


Fig. 9. [Gedanken experiment Einstein – Podolsky – Rosen \(1935\)](#)<sup>9</sup>

Negative probability appears “effectively”, i.e. by the restriction of the degrees of freedom (DOF) of any correlating quantum object by the others. The mechanism of such transformation is bellow discussed.

The quantum allegory of Schrödinger’s alive-and-dead cat helps us to understand the actual state by restricting DOF of any possible states, and consequently, working back-

<sup>9</sup> The second line of the figure translated in English: *Leads to two different states of the system II ← an alternative choice between the measuring of:* (see the arrows towards either bellow or above); in details, Penchev 2009: 55ff).

wards, the quantum correlations to be suggested between the states forming any quantum superposition; and hence, presents a possible form of the mechanism of the mystic “collapse” of wave packet during the real process of measurement.



Fig. 10. [Schrödinger's \(1935: \(48\) 812\)](#) poor “cat”

Von Neumann's theorem (1932) about the absence of hidden parameters in quantum mechanics underlies both quantum correlation and quantum superposition. Its conclusion is: “There are no ensembles which are free from dispersion. There are homogeneous ensembles...” (Neumann 1932: 170)<sup>10</sup>. Consequently, there are no homogeneous ensembles by contradiction, i.e. for example those of a single quantity being free of dispersion. Any quantity has dispersion, which is not due to any cause, to any hidden variable. The premises of the theorem explicated by von Neumann himself are six (A', B',  $\alpha'$ ,  $\beta'$ , I, II) as follows:

**A'**:  $\mathfrak{K} \geq 0 \Rightarrow \text{Erw}(\mathfrak{K}) \geq 0$  (Neumann 1932: 165); **B'**:  $\text{Erw}(a \cdot \mathfrak{K} + b \cdot \mathfrak{G} + \dots) = a \cdot \text{Erw}(\mathfrak{K}) + b \cdot \text{Erw}(\mathfrak{G}) + \dots$ , where  $a, b \in \mathbb{R}$  (Neumann 1932: 165);  **$\alpha'$** :  $\mathfrak{K}$  is a dispersion free quantity  $\stackrel{\text{def}}{=} \text{Erw}(\mathfrak{R}_1) = 1$ ;  $[\text{Erw}(\mathfrak{K}, \varphi) = (\mathfrak{R} \varphi, \varphi)]$  (Neumann 1932: 166);  **$\beta'$** :  $\mathfrak{K}$  is a homogenous one  $\stackrel{\text{def}}{=} \{ a, b \in \mathbb{R}, a+b=1, \text{Erw}(\mathfrak{K}) = a \text{Erw}'(\mathfrak{K}) + b \text{Erw} \} \Rightarrow \{ a=0 \Leftrightarrow b=0 \}$  (Neumann 1932: 166); **I.**  $\{ \mathfrak{K} \mapsto \mathfrak{R} \} \Rightarrow \{ f(\mathfrak{K}) \mapsto f(\mathfrak{R}) \}$  (Neumann 1932: 167); **II.**  $\{ \mathfrak{K} \mapsto \mathfrak{R}, \mathfrak{G} \mapsto \mathfrak{S}, \dots \} \Rightarrow \{ \mathfrak{K} + \mathfrak{G} + \dots \mapsto \mathfrak{R} + \mathfrak{S} + \dots \}$  (Neumann 1932: 167).

We should emphasize the correspondence of “one-to-one” between a mathematical entity as hypermaximal Hermitian operator and a physical entity as quantity. “There corresponds to each physical quantity of a quantum mechanical system, a unique hypermaximal

<sup>10</sup> Here and bellow the translation of the textual citations from von Neumann's book from German to English is according to: J. von Neumann. 1955. *Mathematical Foundations of Quantum Mechanics*. Princeton: University Press.

Hermitian operator, as we know ... and it is convenient to assume that this correspondence is one-to-one – that is, that actually each hypermaximal operator corresponds to a physical quantity.” (Neumann 1932: 167)

What is the connection between von Neumann’s theorem and negative probabilities? By introducing negative probability, expectation is not an additive one in general, the premises of the theorem are not fulfilled, and the deduction is not valid.

<b>Parts (elements) of the system</b>	<b>Possible states (worlds) of the system</b>
<b>Essential parts (elements) of the system</b>	<b>Orthogonal possible states (separated worlds)</b>
<b>External “parts” (elements) of the system</b>	<b>Non-orthogonal possible states (interacting worlds)</b>

Fig. 11. Von Neumann’s theorem and negative probability

Here are a few equivalent expressions for the boundary of the validity of von Neumann’s theorem:

1. Non-negative probability
2. Orthogonal possible states
3. Separated “worlds”
4. An isolated quantum system
5. The additivity of expectation

[Bell’s criticism \(1966\)](#) about von Neumann’s theorem partly rediscovered [Grete Hermann’s objections \(1935\)](#) and is very important for revealing the connection between causality, quantum correlation, and negative probability: “The demonstrations of von Neumann and others, that quantum mechanics does not permit a hidden variable interpretation, are re-considered. It is shown that their essential axioms are unreasonable. It is urged that in further examination of this problem an interesting axiom would be that mutually distant systems are independent of one another” ([Bell 1966: 447](#)). “His essential assumption is: *Any real linear combination of any two Hermitian operators represents an observable, and the same linear combination of expectation values is the expectation value of the combination.* This is true for quantum mechanical states; it is required by von Neumann of the hypothetical dispersion free states also” ([Bell 1966: 448-449](#)).

The idea of [Bell's inequalities \(1964\)](#) can be expressed by negative probability. Since von Neumann's theorem is valid only about nonnegative probability (expectation additivity), and quantum mechanics permits negative probability, the idea is that the domain of the theorem validity is to be described by an inequality of the expectation of two quantities (the spin of two particles) according to the EPR conditions:

$$1 + P(\vec{b}, \vec{c}) \geq |P(\vec{a}, \vec{b}) - P(\vec{a}, \vec{c})| \text{ (Bell 1964: 198).}$$

The last inequality can be rewritten by means of an alleged hidden parameter  $\lambda$ , by which the connection with von Neumann's theorem about the absence of any hidden parameters such as  $\lambda$  becomes clearer:

$$1 + P[b(\lambda), c(\lambda)] \geq |P[a(\lambda), b(\lambda)] - P[a(\lambda), c(\lambda)]|$$

In the light of von Neumann's theorem, [Bell's paper \(1966\)](#) reveals that an eventual violation of the inequality above would require generalizing the notion of hidden parameter introducing nonlocal one. The main steps are:

1. Von Neumann's theorem as well as the theories of hidden parameters interprets them as local ones implicitly.
2. Bell's inequalities discuss the distinction between local and nonlocal parameter because quantum mechanics allows nonlocal variables.

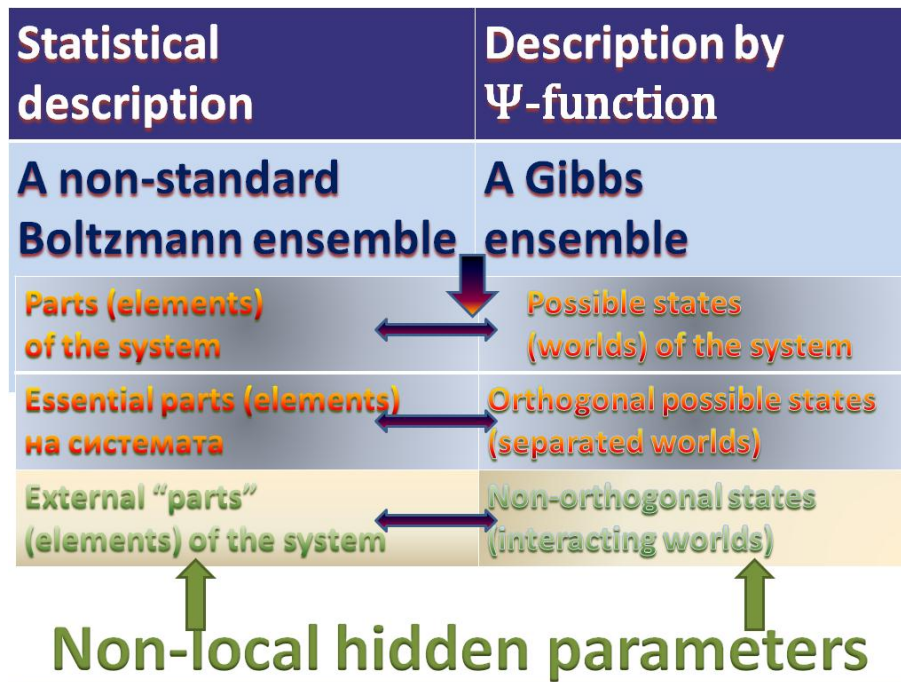


Fig. 12. Non-local hidden parameter



The notion of non-local hidden parameter suggests the notion of the externality of a system considering it as a nonstandard, namely ambient, or “environmental” part of the system. Not any external neighborhood, but only a small one (of the order of a few  $\hbar$ ), at that correspondingly, only in phase space. In the corresponding small neighborhood in Minkowski space, Lorentz invariance is not valid. The following three figures show that:

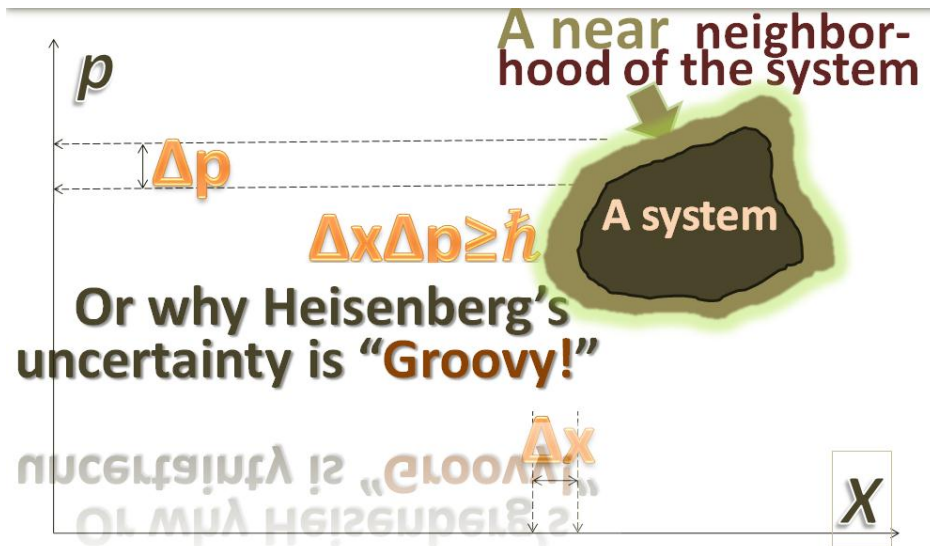


Fig. 13. Groove uncertainty!

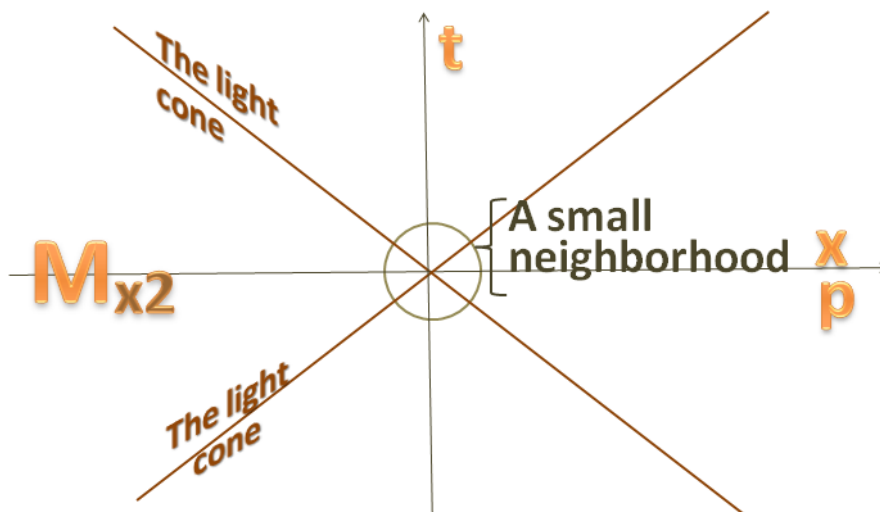


Fig. 14. A small neighborhood in Minkowski space

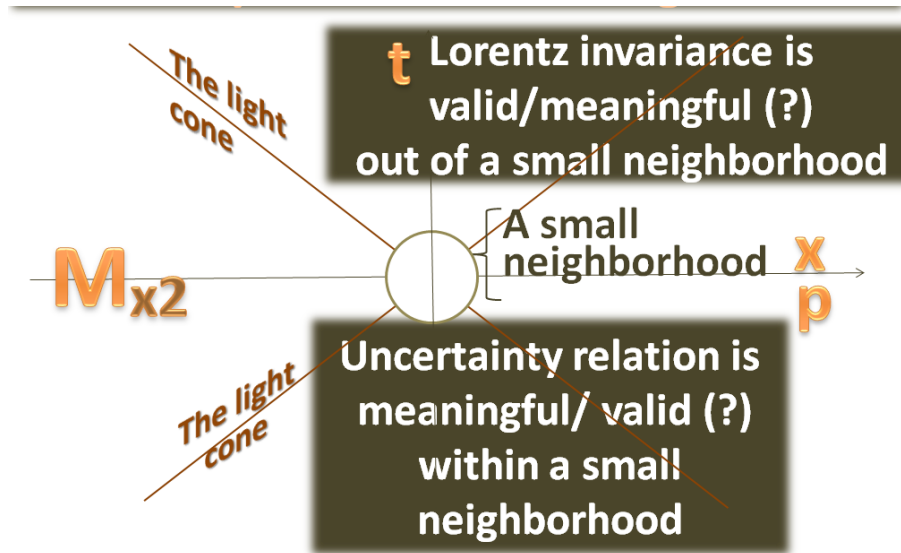


Fig. 15. Lorentz invariance and uncertainty relation in a small neighborhood in Minkowski space

The topic may be illustrated by the notion of an absolutely immovable body: Heisenberg’s uncertainty excludes any absolutely immovable body as well as any exactly constant phase volume. A body is outlined rather by an undetermined “aura” or “halo” than by a sharp outline. The aura is outlined within phase space and its magnitude is comparable with the Planck constant. It consists of the states of negative probability, which “push out” the states of any other bodies beyond the outline region:

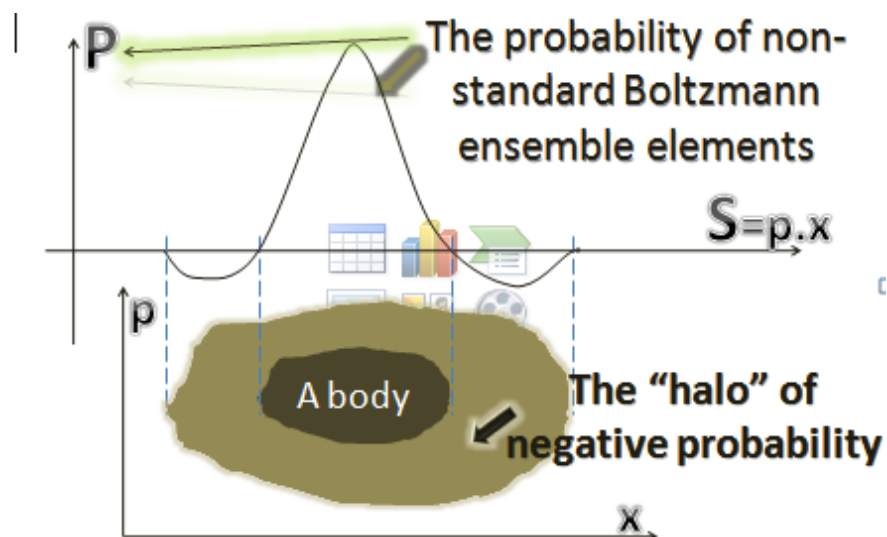


Fig. 16. The “halo” of negative probability in phase space



The next table shows the appearance of negative probability when quantum mechanics is reformulated from the standard language of  $\Psi$ -functions into that of a standard statistical description containing, however, nonstandard, or “external to” the whole, „parts” of it:

<b>A Gibbs ensemble</b>	<b>A non-standard Boltzmann one</b>
<b>An ensemble of the states of the system as a whole</b>	<b>An ensemble of parts (<math>P&gt;0, P&lt;0</math>)</b>
<b>Description by <math>\Psi</math>-functions</b>	<b>Statistical description</b>
<b>No states of <math>P&lt;0</math></b>	<b>Parts of <math>P&lt;0</math></b>
<b>Simultaneously immeasurable quantities</b>	<b>No such quantities</b>

Fig. 17. Again about the comparison of a Gibbs and of a non-standard Boltzmann ensemble

Such a kind of explanation can be applied also to true Bell’s inequalities:

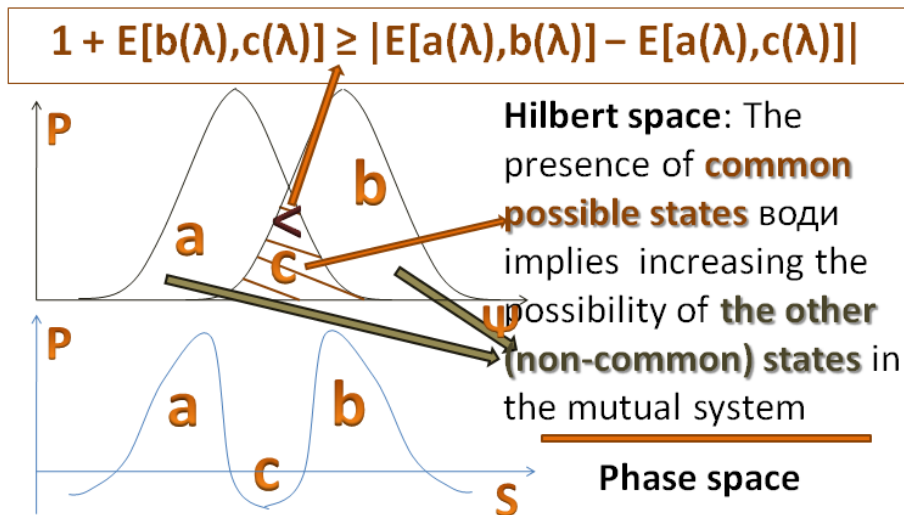


Fig. 18. A mechanism of violating Bell’s inequalities ([Bell 1964](#))

Negative probability cannot help but violate Bell’s inequalities:

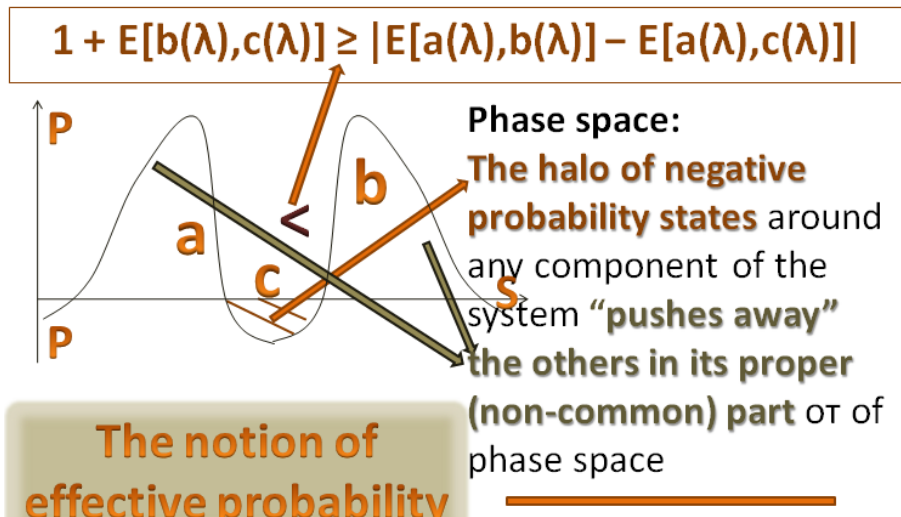


Fig. 19. How does negative probability violate Bell’s inequalities?

The notion of effective probability aids us in bridging physical ‘interaction’ and mathematical ‘probability’. It is effective probability that exerts impact from a given probability onto another. Effective probability is a probabilistic “force” by means of which both the probabilities interact (or more than 2). Since [Dirac \(1942: 8\)](#), negative probability has been thought of just as an effective probability rather than a “real” probability of whatever: Negative probability has been likened to money since it assists “the balance”, while being non-existent at that! Following after [Mermin \(1998\)](#), we could at least question the role of negative probabilities or of any description of probability as a new, nonstandard, but maybe omnipresent, universal and all-embracing substance. This substance includes matter and energy, which are ordinary and generally accepted substances in physics.

Both the aspects symbolized by the EPR argument and Schrödinger’s cat, respectively of two or more “probabilistically” interacting systems and of the superposition of the states of a single and isolated system suggest a Skolemian type of relativity between isolation and interaction as well as an extension of von Neumann’s theorem. The corollary, after generalizing from an isolated to two or more interacting systems is the immediate interaction of probabilities:

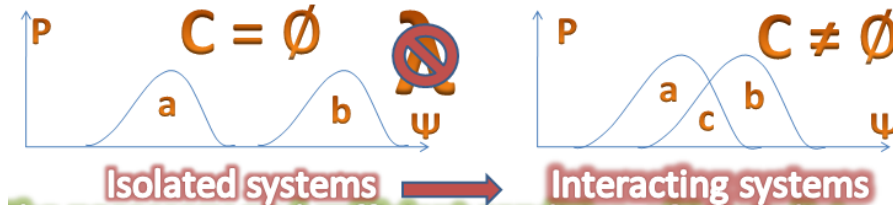


Fig. 20. The notion of effective probability (immediate probabilistic interaction)

The necessary and sufficient condition of immediate probabilistic interaction is: shared common possible states of nonzero probability. The same equivalence refers to the superposition of all the states of an isolated system. The notion of “effective probability” or the more particular, but correlate one of “negative probability”, are the basis of the perceptible commonality, that “sameness” of the Skolemian relativity:

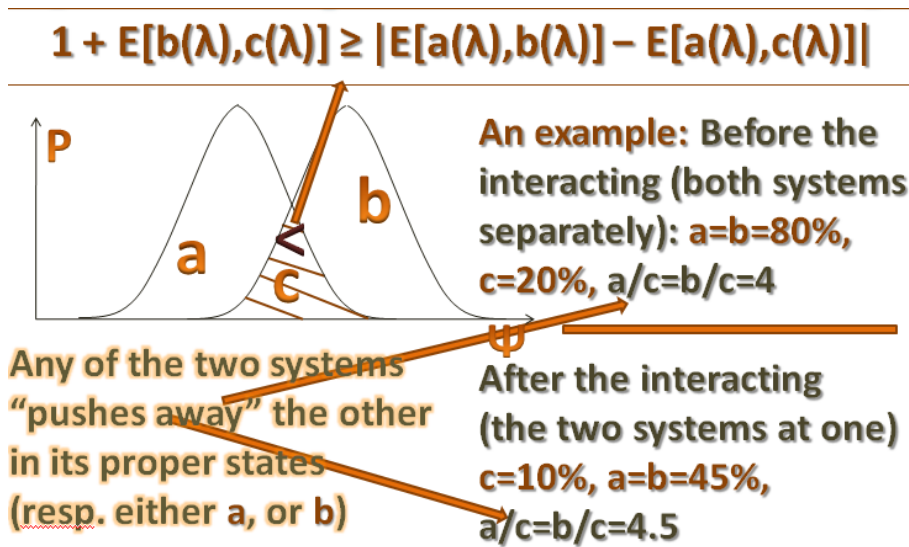


Fig. 21. How do probabilities interact?

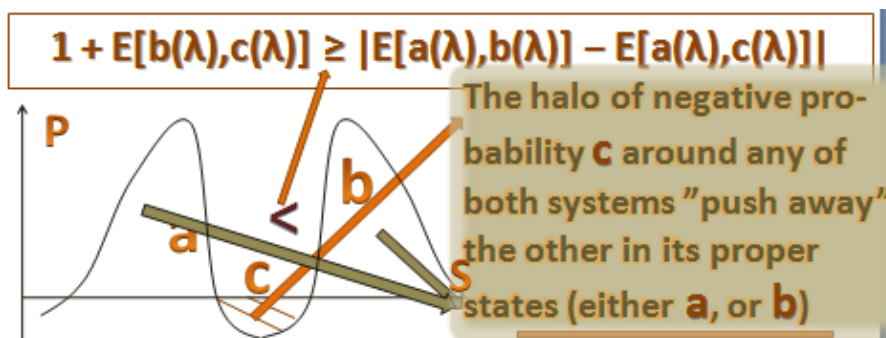


Fig. 22. The notion of effective probability (immediate probabilistic interaction)

However the halo of negative probability “pushes away”: – the system itself; any of the 2; and the states of positive probability, too. They become relatively more probable. Consequently, a Skolemian or an Einsteinian relativity, by means of negative probability, reveals the intimate mechanism of any physical interaction manifesting itself. In the final analysis by restricting DOF of any system participating in the interaction in question any physical interaction manifests itself.

We have two kinds of description, which are equivalent:

STATISTICAL ONE	STANDARD ONE
<b>Different parts of the system</b>	<b>Different states of the system as a whole</b>
<b>The parts of negative probability</b>	<b>Any state is one of non-negative probability</b>
<b>Relatively higher positive probability also within a separated system</b>	<b>Relatively higher positive probability of the proper states only within the common system</b>

Fig. 23. A comparison of the statistical and standard formalism by means of “effective probability”

Negative probability is only one of two ways to represent physical reality. Which way corresponds more to the Boltzmann statistical consideration than to the Gibbs one? For each, there exist two kinds of ontological projection:

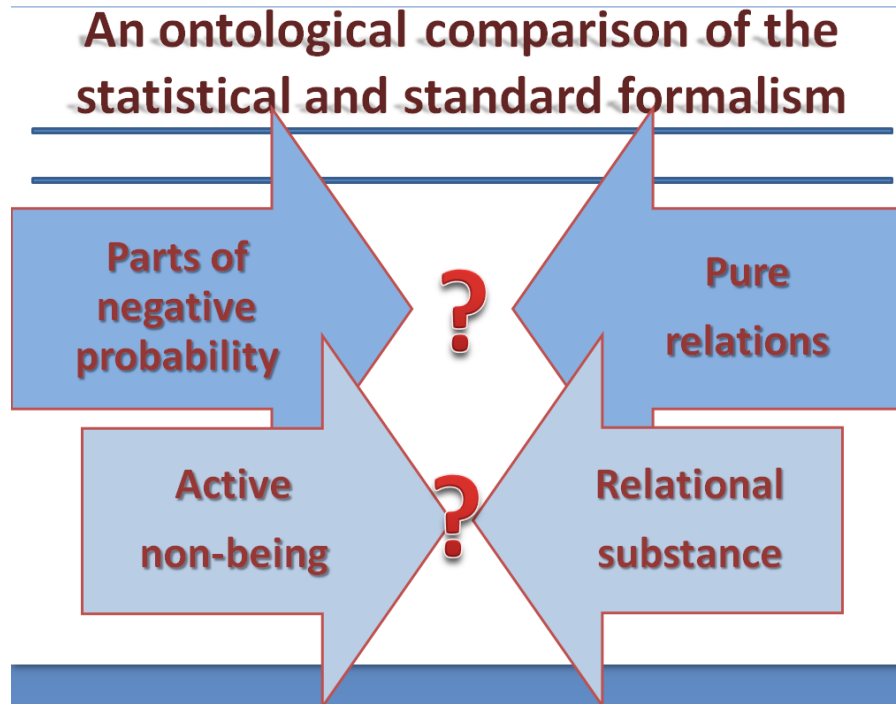


Fig. 24. The statistical vs. standard formalism

The statistical formalism enables the simultaneously immeasurable quantities of the standard formalism to be calculated together. [Kochen-Specker's theorem](#) shows that homogenous quantities have dispersion, even in quantities simultaneously measurable in a mathematical sense, and correspondingly, that there cannot be “hidden parameters”.

Also, [Kochen – Specker's theorem](#) can be discussed as a generalization of von Neumann's theorem in the following way:

- Von Neumann's theorem (1932: 157-173) covers isolated systems and simultaneously immeasurable quantities.
- [Bell's theorem](#) (or inequalities, 1964) clarifies the effect and interpretation of the absence of hidden parameters in interacting systems.
- [Kochen – Specker's theorem](#) deals with isolated systems and simultaneously measurable quantities as follows: the dispersion of homogenous quantities is conditioned by quantum “leaps”.

The two authors formalized the notion of simultaneous measurability: [Kochen and Specker \(1967: 63 ...\)](#) interpreted simultaneous measurability as the availability of a common measure which required the measure of the set of points of discontinuity (quantum leaps) to be zero, i.e. simultaneously measurable quantities should not be continuous, but *almost* con-

tinuous. They proved on such a condition that homogenous quantities have necessarily a dispersion, i.e. there are no hidden parameters.

Next, there isn't any homomorphism of the algebra of statements about commuting quantum quantities into Boolean algebra.

An immediate corollary follows: there is no mapping of a qubit even of two simultaneously measurable quantities, into a bit.

There is a propositional formula which is a classical tautology, but which is "not true" after substituting quantum propositions.

[Kochen – Specker's theorem](#) is connected to the Skolemian relativity of the discrete and continuous: In § 5 of their paper a model of hidden parameter in  $\mathbb{H}^2$  of the particle of spin  $\frac{1}{2}$  is constructed; however no such model exists according von Neumann's theorem ([Kochen, Specker 1967: 74-75](#)). That model of hidden parameter is isomorphic, in fact, to the mapping of a qubit into a bit.

The immediate corollary comes to bear: There is no mapping of a qubit of two simultaneously measurable quantities into a bit. Thus an explanation would be that the notion of simultaneous measurability introduces implicitly Skolemian relativity.

In fact, wave-particle dualism in quantum mechanics is a form of the Skolemian relativity of the discrete (in quantum mechanics) and continuous (in classical physics). That's why a Skolemian type of relativity appears also between the availability and absence of hidden parameters. A qubit can and cannot (à la Skolem) be represented as a bit ([Kochen, Specker 1967: 70, esp. "Remark"](#)).

#### IV. Negative probability for wave-particle dualism

We should return to [Einstein's paper \(1905I\)](#) about mass and energy:

"Gibt ein Körper die Energie  $L$  in Form von Strahlung ab, so verkleinert sich seine Masse um  $L/V^2$ . Hierbei ist es offenbar unwesentlich, daß die dem Körper entzogene Energie gerade in Energie der Strahlung übergeht, so daß wir zu der allgemeineren Folgerung geführt werden: Die Masse eines Körpers ist ein Maß für dessen Energieinhalt; ändert sich die Energie um  $L$ , so ändert sich die Masse in demselben Sinne um  $L/9 \cdot 10^{20}$ , wenn die Energie in Erg und die Masse in Grammen gemessen wird" ([Einstein 1905I: 641](#)). ( $V=3 \cdot 10^{10}$  [cm/s] is the speed of electromagnetic radiation in vacuum.)

Next, we are going to juxtapose the quote above with a no less famous paper of the same year ([Einstein: 1905Ü](#)) about quanta and energy, by means of which I search for the connection between mass and quanta:

In contemporary designations its content may be abstracted by the formula:  $E = \hbar \nu$ . In original designations and Einstein's words: "monochromatische Strahlung ... wie ein diskontinuierliches Medium verhält, welches aus Energiequanten von der Größe  $R\beta\nu/N$  besteht" ([Einstein 1905Ü: 143](#)); "das erzeugte Licht aus Energiequanten von der Größe  $(R/N)\beta\nu$  bestehe, wobei  $\nu$  die betreffende Frequenz bedeutet" ([Einstein 1905Ü: 144](#)) „R die absolute Gaskonstante, N die Anzahl der „wirklichen Moleküle" in einem Grammäquivalent ... bedeutet" ([Einstein 1905Ü: 134](#)), and " $\beta = 4.866 \cdot 10^{-11}$ " ([Einstein 1905Ü: 136](#)). „Monochromatische Strahlung ... verhält sich ..., wie wenn sie aus voneinander unabhängigen Energiequanten von der Größe  $R\beta\nu/N$  bestünde" ([Einstein 1905Ü: 143](#)).

Light is discrete (corpuscular) in its interaction with matter, but it is continuous (wave) as a medium „by itself" which is propagated in space. Quantum mechanics transferred initially that property formulated about electromagnetic radiation to all the quanta, and in fact, to all the physical objects. Afterwards the separation of different relations has been abandoned replacing it by the  $\Psi$ -function description. The hypothesis of hidden parameters from such a viewpoint, in fact, conserves the original and already clearly formulated opinion of Einstein (1905) that the two contradicting aspects, continuity and discreteness, are to be separated, and thus *to be separated into different relations*.

Consequently, energy à la Einstein is already determined in two incompatible ways: as a continual quantity of mass and as a discrete number of quanta. Mechanical energy in classical physics is the sum of kinetic and potential energy:  $E = E_k(p) + E_p(x)$ . Transferred into quantum mechanics, it is the sum of momentum and location functions, which are simultaneously non-measurable. That sum is measured by a third way, by frequency (Neumann 1932: 256, Anm. 164). Von Neumann (1932: 164) had already pointed out that non-commutability refers only to the multiplication, not to the addition of operators and had given the example (just above) of energy as a sum of simultaneously immeasurable quantities. A possible solution is to define the presence of discreteness (quanta) in quantum mechanics to be ascribed to the sum or availability of non-commuting (simultaneously immeasurable) quantities.

Before passing to Einstein's principle of general relativity, let us attempt an ordering of the "mess" of notions:



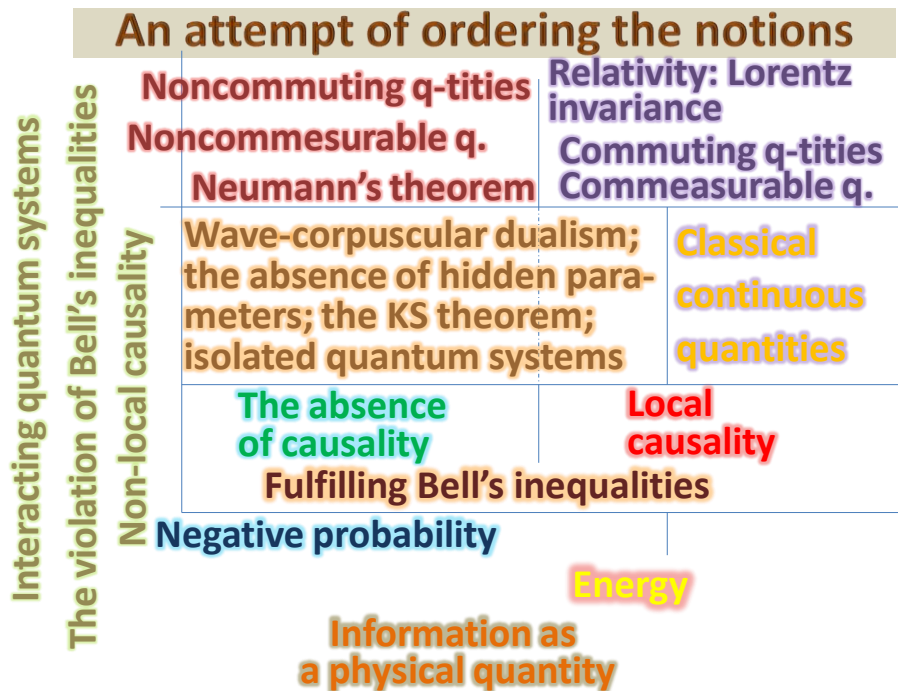


Fig. 25. An attempt at ordering the notions

[Einstein \(1918\)](#) formulated the first two principles of general relativity as follows:

- „a) *Relativitätsprinzip*: Die Naturgesetze sind nur Aussagen über zeiträumliche Koinzidenzen; sie finden deshalb ihren einzig natürlichen Ausdruck in allgemein kovarianten Gleichungen.
- b) *Äquivalenzprinzip*: Tragheit und Schwere sind wesensgleich” ([Einstein 1918: 241](#)).

Einstein reformulated the principle of relativity in mathematical language as the invariance of physical laws towards diffeomorfisms. In this way discrete transformations were excluded. However they are the essential subject of quantum mechanics. Correspondingly, the uniting of quantum mechanics and relativity may be researched in terms of a generalization of invariance to a wider class of morphisms, which should include discrete ones.

The wave-particle dualism considered as a definite type of generalization of the principle of relativity introduces negative probability. The following hypothesis can be advanced: that any relevant generalization of the principle of relativity which includes discrete morphisms should introduce negative probability explicitly or implicitly.

[The principles of general relativity \(1918\)](#):

The original variant ([Einstein 1918: 243](#)) of the basic equation is:  $G_{\mu\nu} = -\kappa \left( T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T \right)$ .

The repaired variant ([Einstein 1918: 243](#)) with the cosmological constant  $\lambda$  is:  $G_{\mu\nu} - \lambda g_{\mu\nu} = -\kappa \left( T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T \right)$ .



Here follows a possible generalization of the principle of relativity and negative probability. Its three sequential levels are as follows:

1.  $\lambda = 0$  ( $\lambda g_{\mu\nu} = 0$ )
2.  $\lambda g_{\mu\nu} = \lambda(g_{\mu\nu})$
3.  $\lambda g_{\mu\nu} = f(g_{\mu\nu})$

There exists a link between any possible generalization of relativity and a rejection (or a generalization) of so-called (by Einstein) Mach's principle:

„c) *Machsches Prinzip*: Das G-Feld ist *restlos* durch die Massen der Körper bestimmt. Da Masse und Energie nach den Ergebnissen der speziellen Relativitätstheorie das Gleiche sind und die Energie formal durch den symmetrischen Energietensor ( $T_{\mu\nu}$ ), beschrieben wird, so besagt dies, daß das durch den Energietensor der Materie bedingt und bestimmt sei” ([Einstein 1918: 241-242](#))

It seems that the colligation of general relativity (introducing negative probabilities) implies restricting the validity or generalization of Mach's principle: The negative probability restriction of the degrees of freedom of a part of a system can be equivalently equated with energy and hence, with mass. For the formula  $E=\hbar\nu$  the frequency  $\nu$  should interpret as the density of informational exchange between the parts of the system (i.e. bits per time).

Non-commutability, wave-particle dualism (and by it, the rejecting Mach's principle material or physical information), and “hidden parameters” can be juxtaposed straightforwardly:

- Non-commutability is a sufficient, but not necessary condition of discreteness (quanta) and hence, of wave-particle dualism.
- Wave-particle dualism is a necessary and sufficient condition of the absence of local hidden parameters as well as of non-local ones.
- An introduction of non-local hidden parameters à la Bell implies the violation of wave-particle dualism: either only “waves”, or only “particles”.
- Mathematical non-commutability is interpreted as simultaneous non-commeasurability of corresponding quantities.
- The commensurability of physical quantities is interpreted as mathematical commensurability, i.e. as the availability of general measure.
- Being interpreted, the availability of common measure covers the cases both of continuous quantities and of their discrete leaps.

Returning now to the fig. 25 and our attempt at ordering the notions, we can already emphasize Kochen-Specker's theorem as a generalization of von Neumann's theorem:

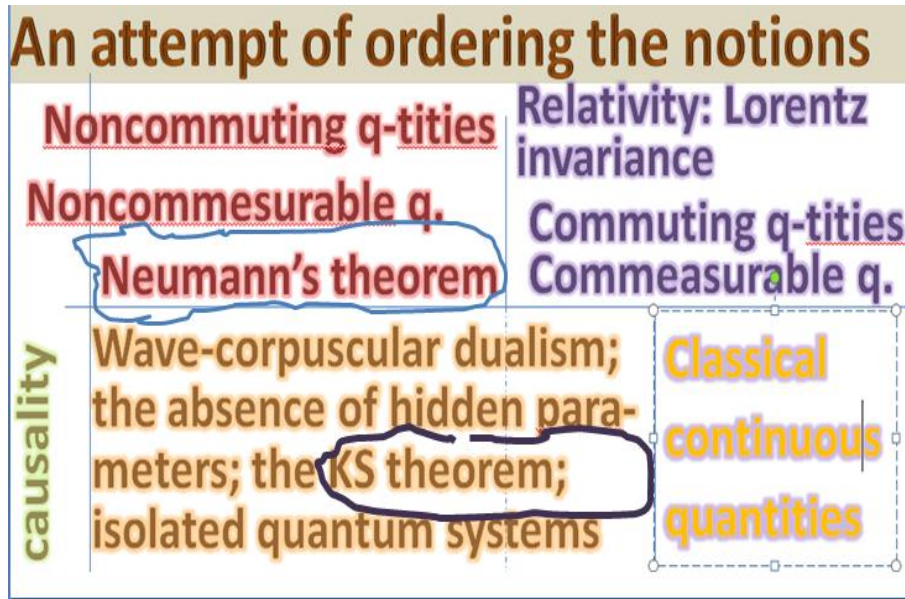


Fig. 26. Von Neumann's and Kochen – Specker's theorem

A Skolemian relativity of “hidden parameters” is clearly seen: If the KS theorem is equivalent to wave-particle dualism, then the very absence of hidden parameters in quantum mechanics should consider as equivalent with a Skolemian relativity of the continual and discrete. However their counterexample concerning also their proper construction shows that even the true absence of hidden parameters is relative. Consequently, even the true relativity of the continual and discrete is also relative, which represents “super-Skolemian” relativity.

We may introduce the term of “contramotion” borrowed from A. and B. Strugatsky's novel “Monday begins on Saturday” (1965) to illustrate the opposite directions of the discrete and the continual of wave-particle dualism:

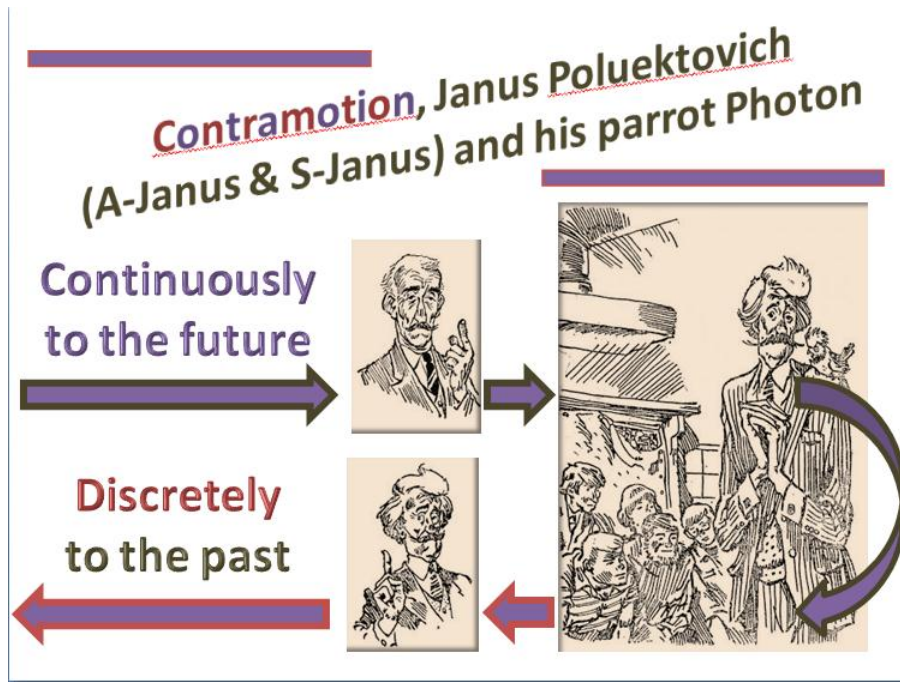


Fig. 27. Janus Poluektovich and his parrot named Photon

The novel represents a two-faced character, unambiguously called Janus. Each face however is then separated into isolated bodies which are never together in the same place and time. Each body is a different man, respectively Janus-Administrator (A-Janus) and Janus-Scientist (S-Janus). In a distant future Janus will manage (A-Janus's point of view) or has managed (S-Janus's) to reverse himself towards the arrow of time. On his shoulder, his parrot Photon will stay (or has stayed) at that instant, that's why it will transform (or has transformed) in a "contramotion", too.

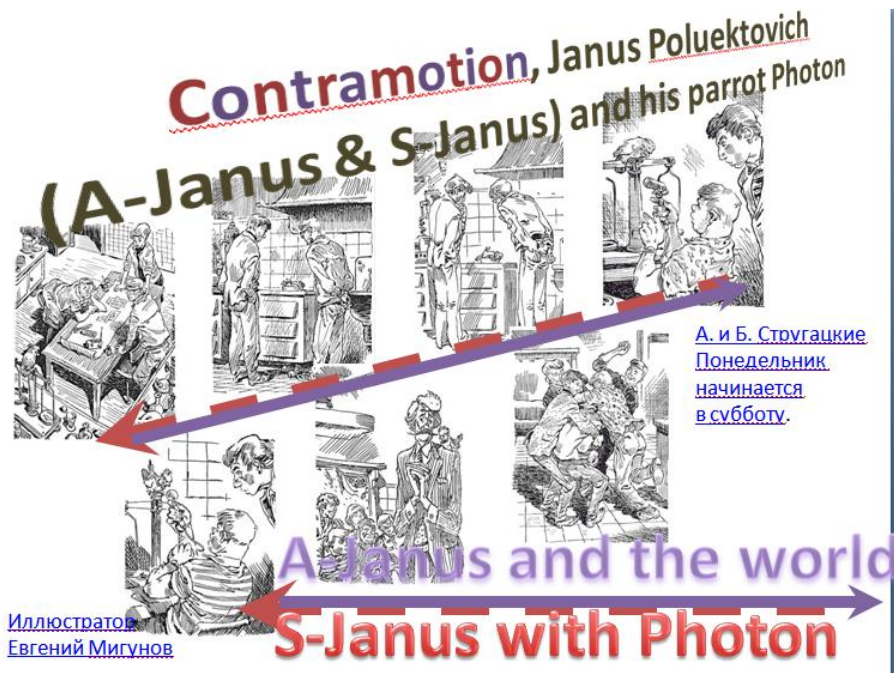


Fig. 28. Contramotion: Janus Poluektovich and his parrot named Photon against the arrow of time

So A-Janus with all the rest is moved “correctly” in time arrow, but S-Janus and Photon being contramotioners are moved “inversely” towards the arrow of time. However there is an over ruling circumstance: Contramotion is discontinuous. Exactly at midnight S-Janus takes his parrot with him, stays alone in a deep rest, and ... both jump into the day before instead of passing in the next day as all normal people, incl. A-Janus.

Consequently, the personage of Janus with his two faces embodied in two persons may be interpreted as an allegory of wave-corpuscular dualism pointing out the possibility of the two aspects being disjunctively separated forwards or backs in time:

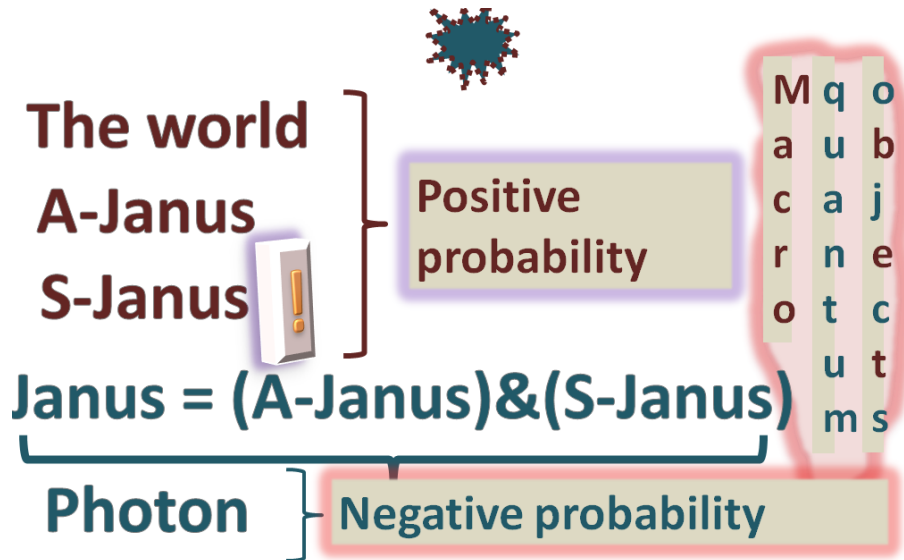


Fig. 29. Wave-corpuseular dualism, contramotion and negative probability

Discontinuous contramotion implies negative probability since the two measures, correspondingly of the discrete and of the continuous motion, are directed oppositely.

This way, we can create and suggest a new, quantum, fable, “Parrot and Cat”, combining Strugatsky’s parrot and Schrödinger’s cat. The condition of its precept or moral is: Contramotion is the sufficient, but not necessary condition of coherent superposition. Our allegory would be: Schrödinger’s cat has eaten Photon, the parrot – contramotioner (rather black humor).

[P. Dicae’s conception \(1942\)](#) helps us to put the question about the ontological status of negative probability, or figuratively said, whether negative probability might be “eaten”, and in such a way, “transferred” being rather a property of a material thing (like the parrot named “Photon”) than a relation between two or more things:

“Negative energies and probabilities should not be considered as nonsense. They are well-defined concepts mathematically, like a negative sum of money, since the equations which express the important properties of energies and probabilities can still be used when they are negative. Thus negative energies and probabilities should be considered simply as things which do not appear in experimental results” ([Dirac 1942: 8](#)) As to bosons “there is the added difficulty that states of negative energy occur with a negative probability” ([Dirac 1942: 1](#)). Obviously, Dirac’s position is categorically in favor of only the relative nature of negative probability, or in his words, only “like a negative sum of money” in its balance.

I would mention [Feynman’s article \(1991\)](#), which is often cited. He did not go beyond Dirac’s approach of introducing negative probabilities only conventionally, in the course

of calculations, in an analogy of “negative money”. He gave many examples from classical and quantum physics. Negative probability in them meant that the happening of an event decreased the realizing of another: i.e. negative probability could have only relative character.

Much more interesting is Pauli’s (2000: 71-72 ) consideration about negative probability: On the subject of [the theory of Gupta-Bleurer](#) Pauli discussed the formalism of “negative probabilities”, which they had introduced. The norm and the expectation had been generalized correspondingly as follows:

$$\sum_n \Psi_n^* \Psi_n \rightarrow \sum_n \Psi_n^* \eta \Psi_n;$$

$$\langle A \rangle = \frac{\sum_{nm} \Psi_n^* A_{nm} \Psi_n}{\sum_n \Psi_n^* \Psi_n} \rightarrow \frac{\sum_{nm} \Psi_n^* \eta A_{nm} \Psi_n}{\sum_n \Psi_n^* \eta \Psi_n}$$

Here  $\eta$  is a Hermitian operator which introduces an operator measure in Hilbert space. In both cases, the norm remains constant in time if the Hamiltonian of the system is a Hermitian operator. If  $\eta$  is not a Hermitian operator, the norm still conserves the constant. Our interpretation: The operator  $\eta$  can be thought as a bivalent tensor which transforms a Hilbert space into another:

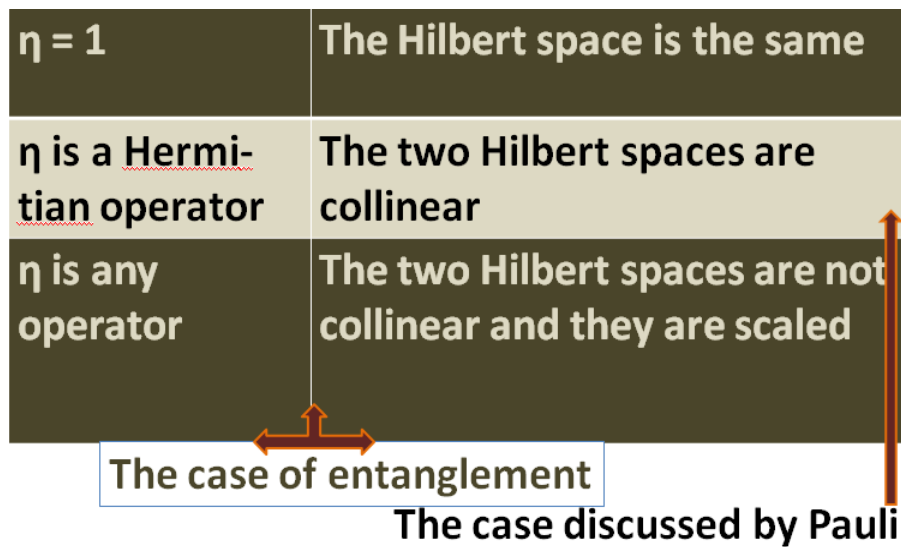


Fig. 30. The operator measure  $\eta$  interpreted physically

The approach of [Riesz’s representation theorem](#) can be used for very important conclusions about the operator measure  $\eta$  and its physical interpretation. This theorem establishes an important connection between a Hilbert space and its (continuous) dual space: If the



underlying field is that of the complex numbers as is in quantum mechanics, the two are isometrically anti-isomorphic.

All the physical quantities can be interpreted as such operator measures  $\eta$ , restricting them into the set of Hermitian operators. All the visible world of many different and divided things is that set, i.e. a clearly restricted class of the transformation of the whole into and within itself. A natural question is about the interpretation rather of the full than restricted class of such transformations. The answer is: their states, i.e. all the  $\Psi$ -functions, the true whole, are represented by Hilbert space: Moreover, it is that set which transformations form in and of itself, or mathematically speaking, are isomorphic to itself. That it is the case may be demonstrated by the following mapping:

In general, one-to-one correspondence is valid between any operator in Hilbert space and a point in it:  $\mathbf{c}_j \mathbf{e}^{i\varphi_j} = \mathbf{e}^{x_j + iy_j} \mathbf{e}^{i\varphi_j} \leftrightarrow \mathbf{x}_j + i(\mathbf{y}_j + \varphi_j)$ . A tensor of any finite valence  $\mathbf{k}+1$  can be represented as an operator in Hilbert space and consequently, as a point in it:  $(\prod_{\mathbf{k}} \alpha_{\mathbf{k}} \mathbf{c}_{j\mathbf{k}}) \mathbf{e}^{i \sum_{\mathbf{l}} \beta_{\mathbf{l}} \varphi_{j\mathbf{l}}}$ .

The physical homogeneity of the imaginary member, namely  $i(\mathbf{y}_j + \varphi_j)$ , or correspondingly,  $e^{i(\varphi_j + y_j)}$ , needs to be interpreted, and that returns us to Einstein's two "incompatible" views both of discrete quanta and of differentiable, consequently continuous, mechanical motions (morphisms). The energy-Janus of two faces, combines the two: – potential energy depending only on coordinates, and in that way suggesting a discrete concept of bodies, and also kinetic energy depending solely on speeds, and in this way suggesting a continuous concept of motions of those bodies.

We should emphasize two innovative or even revolutionary points:

1. The wave-particle dualism of quantum mechanics already existed implicitly in classical mechanics as both aspects of a body at rest or in motion: that is its potential or kinetic energy. It is the Planck constant and Heisenberg uncertainty that force the two aspects to transform between themselves (it being impossible to give separated distinctions), and then to be added as in classical mechanics and physics. The case is such that it points to a hidden common essence behind the macro visible, i.e. a body and its motion.

It is as if quantum mechanics for the Planck constant is concentrated on the boundary between them, which is demonstrated to be rather an area than a contour, within which particulate body and continuous wave motion are the same, namely a single quantum

object. Negative probability describes the immediate interaction of probabilities, which makes their relation that which is necessary information to be a physical entity.

2. The above, pinpointed, physical homogeneity of the imaginary member for the physical quantity of energy hints that not only mass and energy, but also time share a common essence, and consequently their mutual conversion, in the same way as that mutual conversion of the formerly two, might or even should be admitted.

In the spirit of a Skolemian relativity (Skolem 1970: 138; Пенчев 2009: 307-325), any entangled subsystem of a system can be represented equivalently both as an independent, isolated, indivisible system and as an arbitrary operator (an operator measure  $\eta$ ) transforming a point of the Hilbert space of the one system into that of another system.

The conclusion made above implies particularly (entanglement = 0) that any measured value of a quantity in the system ('an apparatus – a quantum object') is 'the objective value' of that quantity of the quantum object. There are no hidden parameters which dictate deterministically among its random values in this case. The Hermitian character of any physical quantity generalizes the requirement that the value is in an exactly given point of time. It also generalizes the requirements about discrete functions.

$\Psi$ -function represents the condition that a quantum quantity can obtain non-zero values only on areas whose common measure is zero. The absence of hidden parameter is due to the zero measure of any area with non-zero values. The same case is represented by Dirac's  $\delta$ -function. The zero measure of any area with non-zero values is a mathematical way to represent uncertainty relations. Any  $\Psi$ -function can be interpreted as the operator measure  $\eta$  of a quantum object entangled with its environment.

## V. A few mathematical questions

Let us consider:

- Bartlett's approach for introducing negative probability by means of the characteristic function of random quantities;
- [Gleason's theorem](#) about the existence of measure in Hilbert space;
- [Kochen – Specker's theorem](#) again, but now in a statistical interpretation.

Bartlett's properly mathematical approach to negative probability is the following:

"Since a negative probability implies automatically a complementary probability greater than unity, we shall reconsider  $p_r A_r \equiv p_1 A_1 + p_2 A_2 + \dots + p_m A_m$  with all restrictions on the values of the individual  $p_r$  removed, provided that the sum remains finite equal to the conventional sum of unity. For those familiar with the correspondence between



probability theory and the theory of measure, it is noted that the parallel extension in this more general form of probability theory corresponds to the use of an additive set function which is always the algebraic difference of two positive functions” (Bartlett 1944: 71-72).

Further, the interval  $[0, 1]$  of probabilities conserves a unique meaning:

“Thus probabilities in the original range 0 to 1, as we might reasonably expect, still retain their special significance. It is only these probabilities which we can immediately relate with actual frequencies; it is only these probabilities, for example, for which the theoretical frequency ratio  $r/n$  tends to  $p$  with probability one, as  $n$  tends to infinity” (Bartlett 1944: 72).

Consequently, the rest, or nonstandard, probabilities cannot be defined within the usual understanding of whole and part: Their introduction implies necessarily “external” parts (examined above) of a whole like a quantum system. Speaking purely mathematically and returning back to Bartlett’s proper approach, we are going to introduce such kinds of generalized probabilities indirectly, namely as corresponding to an appropriate generalization of random quantity in such way that a random quantity for any characteristic function is required:

“Random variables are correspondingly generalized to include extraordinary random variables; these have been defined in general, however, only through their characteristic functions” (Bartlett 1944: 73).

As Dirac, Bartlett also suggested negative probability only “in the balance” of probabilities. Implicitly, empirical and physical reality can be only the usual type of whole consisting only of normal (internal) parts, with respect to which, probabilities within  $[0, 1]$  could be actually observed in experiments:

“Negative probabilities must always be combined with positive ones to give an ordinary probability before a physical interpretation is admissible. This suggests that where negative probabilities have appeared spontaneously in quantum theory it is due to the mathematical segregation of systems or states which physically only exist in combination” (Bartlett 1944: 73).

However why not make the  $\Psi$ -function be the characteristic function of a random quantity? Bartlett’s approach leads to the idea of considering the  $\Psi$ -function as the characteristic function of a physical quantity which is “random”, or more exactly, a function of the coordinates in configuration space. The utilization of the  $\Psi$ -function as the characteristic function instead of the probabilistic distribution of random quantity has the advantage of describing its behavior in general: including also a discrete change of probability (a quantum leap) when the probabilistic distribution itself in that point is represented by  $\delta$ -function. In

fact, the discrete change of probability is what is available in all the phenomena of entanglement: when the probabilistic distribution of a quantum object restricts immediately the degrees of freedom of another; as a result of that “informational” interaction, the probability of a given point endures a discrete leap in the general case. The differential value in that point of the probabilistic distribution is  $\infty$ , and the most significant condition is that it cannot have a finite norm, conventionally accepted as unity. It should be  $> 1$ ! If however the differential probability is  $> 1$  in the point of the discrete change of probability, then it implies the appearance and introduction of negative probability following Bartlett’s approach:  $p=p_1+ p_2$ . If  $p_1 > 1$ , keeping  $p=1$ , then  $p_2 < 0$ ! Consequently in the final analysis, the appearance of negative probability is due to the availability of discrete leaps of probability in some points. It is what strongly suggests that just the  $\psi$ -function (it is the characteristic function) be utilized instead of the probabilistic distribution itself of the random quantity.

The question remains “How does the  $\psi$ -function comply with the characteristic function?” Let us see:

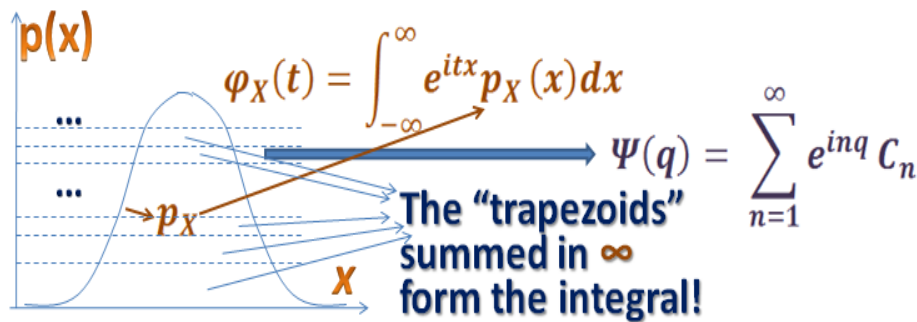


Fig. 31.  $\Psi$ -function as the characteristic function of a random quantity: The  $\Psi$ -function (on the right) is obtained by the probabilistic distribution (on the left) as the integral is substituted by an infinite sum of constants (“trapezoids”) necessarily  $\rightarrow 0$  for the factor  $1/n, n \rightarrow \infty$ .

Our newly acquired demonstration of the  $\Psi$ -function as characteristic function of random configuration-space coordinates leads us to a new perspective on [Gleason’s theorem](#) about measures in Hilbert space:

Till now we have discussed negative or complex probability as a relation of measures (any of which may be a nonnegative real number). Gleason’s theorem states that any measure of such an ordinary type in a Hilbert space of dimension more than two necessarily conserves the orthogonality of the dimensions. The theorem inclines us to a subverse, logically negated idea: to generalize the notion of measure to complex (particularly, negative), and to investigate violating the orthogonality of the dimensions in Hilbert space. In

other words, whether to be the “not complex” measure is adequate on any “curved” (i.e. having a non-orthogonal basis) Hilbert space of dimension more than 2?

What is the significance of the exception of [Gleason’s theorem](#), in the case of 2 dimensions? The idea is that the exception, 2 dimensions, guarantees the “backdoor”, through which a Skolemian type of relativity between “flat” and “curved” Hilbert space can pass, also the Skolemian relativity of a measure on “flat” and a measure on “curved” Hilbert space.

[Gleason’s theorem](#) itself states:

“Let  $\mu$  be a measure on the closed subspaces of a separable (real or complex) Hilbert space  $\mathcal{H}$  of dimension at least three. There exists a positive semi-definite self-adjoint operator  $\mathbf{T}$  of the trace class such that for all closed subspaces  $\mathbf{A}$  of  $\mathcal{H}$

$$\mu(\mathbf{A}) = \text{trace}(\mathbf{TP}_A),$$

where  $\mathbf{P}_A$  is the orthogonal projection of  $\mathcal{H}$  onto  $\mathbf{A}$ ” ([Gleason 1957: 892-893](#)).

[Bell’s interpretation \(1966\)](#) of [Gleason’s theorem \(1957\)](#) was the following: “... if the dimensionality of the state space is greater than two, the additivity requirement for expectation values of *commuting operators* cannot be met by dispersion free states” ([Bell 1966: 450](#)). In other words, [Bell \(1966\)](#) interpreted [Gleason’s theorem](#) according to the yet to come [theorem of Kochen and Specker \(1967\)](#).

“It was tacitly assumed that measurement of an observable must yield the same value independently of what other measurements may be made simultaneously” ([Bell 1966: 451](#)). Otherwise, [Bell’s objection \(1966\)](#) to Gleason’s theorem can cover the [theorem of Kochen and Specker \(1967\)](#).

Now we have a necessary justification to look in a new way at [Kochen – Specker’s theorem](#): “statistically” interpreting it. In fact, the theorem has a clearly expressed “anti-statistical” meaning: quantum mechanics uses probabilistic distributions, which are not statistics. According to the discussion so far, there should exist states of negative probability, and that’s why they cannot be “states” in a restricted, properly statistical sense. They represent immediate interactions of statistical states, i.e. the distributions of system states. The system is not a priori or independent of its states. The parts of a system are not identifiable with “the substance” of its states.

**Instead of conclusions** (as the necessary conclusions have mentioned in the abstract and the beginning of the paper), a few questions:

1. Whether negative probability is only a mathematical construction, or there exist physical objects of negative probability?
2. Whether negative probability and pure relation (such a one which cannot be reduced to predications) are equivalent, expressing the same case in different ways?
3. Whether negative probability does imply the physical existence of probability?

4. Can probabilities interact immediately (i.e. without any physical interaction of the things, phenomena, or events possessing those probabilities)?
5. Whether physical, existing information is equivalent to the interaction of probabilities?

## LITERATURE:

**Bartlett, M.** 1945. Negative Probability. – *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 41, No 1 (June 1945), 71-73 (cited or mentioned on the following pages above: XXXIX-XLI).

**Bell, J.** 1964. On the Einstein – Podolsky – Rosen paradox. – *Physics* (New York), 1, 195-200. Also in: Bell, J. *Speakable and unspeakable in quantum mechanics: collected papers in quantum mechanics*. Cambridge: University Press, 1987, 14-21.

([http://www.physics.princeton.edu/~mcdonald/examples/QM/bell\\_physics\\_1\\_195\\_64.pdf](http://www.physics.princeton.edu/~mcdonald/examples/QM/bell_physics_1_195_64.pdf) cited or mentioned on the following pages above: I, XI, XXI, XXIV, XXV, XXVIII, XXXII).

**Bell, J.** 1966. On the Problem of Hidden Variables in Quantum Mechanics. – *Reviews of Modern Physics*. Vol. 38, No 3 (July), 447-452. Also in: Bell, J. *Speakable and unspeakable in quantum mechanics: collected papers in quantum mechanics*. Cambridge: University Press, 1987, 1-13. ([http://www.physics.princeton.edu/~mcdonald/examples/QM/bell\\_rmp\\_38\\_447\\_66.pdf](http://www.physics.princeton.edu/~mcdonald/examples/QM/bell_rmp_38_447_66.pdf)); cited or mentioned on the following pages above: XI, XX, XXI, XLII.)

**Cattaneo, G., M. Chiara, R. Giuntini, F. Paoli.** 2009. Quantum Logic and Nonclassical Logics. – In: *Handbook of Quantum Logic and Quantum Structures. Quantum Logic* (eds. K. Engesser, D. Gabbay, D. Lenmann). Amsterdam, ets.: Elsevier, 127-226 (cited or mentioned on the following pages above: XI).

**Cohen, L.** 1966. Generalized phase-space distribution functions. – *Journal of Mathematical Physics*. Vol. 7, No 5, 781-786 (cited or mentioned on the following pages above: VI, IX, XI-XII, XIII, XIV).

**Cohen, L.** 1989. Time-Frequency Distributions – A Review. – *Proceedings of the IEEE*. Vol. 77, No. 7, 941–981, 1989. (<http://www.geo.unizh.ch/oldrsl/research/SARLab/GMTILiterature/Ver09/PDF/Coh89.pdf>); cited or mentioned on the following pages above: VI, IX, XI-XII, XIII, XIV).

**Cohen, L.** 1995. *Time-Frequency Analysis*. New York: Prentice-Hall (cited or mentioned on the following pages above: VI, VII, IX, XI-XII, XIII, XIV).

**Cohen, L.** 2008. The Weyl transformation and its generalization. – *Rendiconti del Seminario Matematico. Università Politecnico di Torino*, Vol. 66, No 4, 259-270. – <http://seminariomatematico.dm.unito.it/rendiconti/66-4/259.pdf> (cited or mentioned on the following pages above: XIV).

**Dirac, P.** 1942. Bakerian Lecture. The Physical Interpretation of Quantum Mechanics. – *Proceedings of the Royal Society of London, A*, vol. 180 (980; March 18 1942), 1-40. (<http://rspa.royalsocietypublishing.org/content/180/980/1.full.pdf+html>) (cited or mentioned on the following pages above: XXV, XXXVI, XL.)

**Einstein, A.** 1905. Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig? – *Annalen der Physik*. Bd. 18, No 13, 639–641. – [http://www.physik.uni-augsburg.de/annalen/history/einstein-papers/1905\\_18\\_639-641.pdf](http://www.physik.uni-augsburg.de/annalen/history/einstein-papers/1905_18_639-641.pdf) (cited or mentioned on the following pages above: XXIX, XXX, XXXVIII).

**Einstein, A.** 1905. Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. – *Annalen der Physik*. Bd. 17, No 6, 132–148. ([http://www.physik.uni-augsburg.de/annalen/history/einstein-papers/1905\\_17\\_132-148.pdf](http://www.physik.uni-augsburg.de/annalen/history/einstein-papers/1905_17_132-148.pdf) (cited or mentioned on the following pages above: XXIX-XXX, XXXVIII.)

**Einstein, A.** 1918. Prinzipielle zur allgemeinen Relativitätstheorie. – *Annalen der Physik*. Bd. 55, № 4, 241-244. ([http://www.physik.uni-augsburg.de/annalen/history/einstein-papers/1918\\_55\\_241-244.pdf](http://www.physik.uni-augsburg.de/annalen/history/einstein-papers/1918_55_241-244.pdf) ; cited or mentioned on the following pages above: VIII, XIII, XXVII, XXX, XXXI-XXXII.)

**Einstein, A., B. Podolsky and N. Rosen.** 1935. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? – *Physical Review*, 1935, 47, 777-780. ([http://www.phys.uu.nl/~stiefelh/epr\\_latex.pdf](http://www.phys.uu.nl/~stiefelh/epr_latex.pdf) .) In Russian: А. Эйнштейн, Б. Подолски, Н. Розен. 1936. Можно ли считать, что квантово-механическое описание физической реальности является полным? – *Успехи физических наук*. Т. XVI, № 4, 440-446 – [http://ufn.ru/ufn36/ufn36\\_4/Russian/r364\\_b.pdf](http://ufn.ru/ufn36/ufn36_4/Russian/r364_b.pdf) . (Cited or mentioned on the following pages above: XVIII, XXI, XXV).

**Feferman, S.** 1991. Reflecting and Incompleteness. – *The Journal of Symbolic Logic*. Vol. 56, No 1 (March 1991), 1-49 (cited or mentioned on the following pages above: X).

**Feynman, R.** 1991. Negative Probability. – In: *Quantum Implications: Essays in Honour of David Bohm* (eds. B.Hiley, D. Bohm, D. Peat). London – New York: Routledge, 235-248 ([books.google.com](http://books.google.com)). (Cited or mentioned on the following pages above: XXXVI-XXXVII).

**Gleason, A.** 1957. Measures on the Closed Subspaces of a Hilbert Space. – *Journal of Mathematics and Mechanics*. Vol. 6, No 6, 885-893. (<http://www.iujm.indiana.edu/IUMJ/FTDLOAD/1957/6/56050/pdf> ; cited or mentioned on the following pages above: XXXIX, XLI, XLII.)

**Groenewold, H.** 1946. On the Principles of Elementary Quantum Mechanics. – *Physica*. Vol. XII, No 7 (Oct 1946), 405-460; cited or mentioned on the following pages above: XIV).

**Hermann, G.** 1935. The circularity in von Neumann's proof. (Translation by Michiel Seevinck of "Der Zirkel in NEUMANNs Beweis", section 7 from the essay by Grete Hermann, *Die Naturphilosophischen Grundlagen der Quantenmechanik*. Abhandlungen der fries'schen Schule, 6, 1935 – <http://www.phys.uu.nl/igg/seevinck/trans.pdf> (cited or mentioned on the following pages above: XX).

**Kochen, S., E. Specker.** 1967. The problem of hidden variables in quantum mechanics. – *Physical Review A*. Vol. 17, № 1, 59-87. ([http://www.hep.princeton.edu/~mcdonald/examples/QM/kochen\\_iujm\\_17\\_59\\_68.pdf](http://www.hep.princeton.edu/~mcdonald/examples/QM/kochen_iujm_17_59_68.pdf) (cited or mentioned on the following pages above: XXVIII, XXIX, XXXIII, XXXIX, XLII).

**Kripke, S.** 1975. Outline of a Theory of Truth. – *The Journal of Philosophy*, Vol. 72, No. 19, Seventy-Second Annual Meeting American Philosophical Association, Eastern Division. (Nov. 6, 1975), 690-716. ([http://philo.ruc.edu.cn/logic/reading/Kripke\\_%20Theory%20of%20Truth.pdf](http://philo.ruc.edu.cn/logic/reading/Kripke_%20Theory%20of%20Truth.pdf) ; cited or mentioned on the following pages above: IX-X).

**Lutskanov, R.** 2009. What is the definition of ‘logical constant’? – In: M. Pelis. (ed.) *Logica Year-book’08*. College Publishers: London, 111-121 (cited or mentioned on the following pages above: X).

**Mermin, D.** 1998. What Is Quantum Mechanics Trying to Tell Us? – *American Journal of Physics*. Vol. 66, No 9, 753-767 (<http://arxiv.org/abs/quant-ph/9801057v2> ; cited or mentioned on the following pages above: XXV).

**Moyal, J.** 1949. Quantum mechanics as a statistical theory. – *Proceedings of the Cambridge Philosophical Society*. Vol. 45, No 1, 99-124. ( [http://epress.anu.edu.au/maverick/mobile\\_devices/apc.html](http://epress.anu.edu.au/maverick/mobile_devices/apc.html); cited or mentioned on the following pages above: XV-XVII).

**Neumann, J. von.** 1932. *Mathematische Grundlagen der Quantenmechanik*. Berlin: Verlag von Julius Springer (<http://www.calameo.com/read/000187019bb347837a6bf> ). In English: J. von Neumann. 1955. *Mathematical Foundations of Quantum Mechanics*. Princeton: University Press. In Russian: Й. фон Нейман. 1964. *Математические основы квантовой механики*. Москва: „Наука”. (Cited or mentioned on the following pages above : XI, XIX-XX, XXI, XXV, XXVIII, XXIX, XXX, XXXIII).

**Pauli, W.** 2000. *Pauli lectures on physics. Vol. 6. Selected topics in field quantization*. (ed. C. Enz.) New York: Courier Dover Publications, 71-73 (<books.google.com/> ; cited or mentioned on the following pages above: XXXVII).

**Schrödinger, E.** 1935. Die gegenwärtige situation in der Quantenmechanik. – *Die Naturwissenschaften*, Bd. 48, 807-812; Bd. 49, 823-828, Bd. 50, 844-849. (In English: <http://www.tu-harburg.de/rzt/rzt/it/QM/cat.html>; превод на руски: Шредингер, Э. 1971. Современное положение в квантовой механике. – В: Э. Шредингер. *Новые пути в физике*. Москва: „Наука”, 1971, 66-106. The text is cited or mentioned on the following pages above: X, XVIII-XIX, XXV, XXXVI).

**Skolem, T.** 1970. Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre. – In: *Selected works in logic* (ed. E. Fenstad), Oslo etc: Univforlaget, 137-152 (cited or mentioned on the following pages above: VIII, XXV-XXVII, XXIX, XXXIII, XXXIX, XLII).

**Ville, J.** 1948. Théorie et Applications de la Notion de Signal Analytique. – *Cables et Transmission*, 2A (janvier), 61-74 ([http://qss.stanford.edu/~godfrey/noise\\_and\\_information/Ville\\_Sig\\_Analytique\\_Cables\\_et\\_Trans\\_1948.pdf](http://qss.stanford.edu/~godfrey/noise_and_information/Ville_Sig_Analytique_Cables_et_Trans_1948.pdf); English translation by I. Selin, “Theory and Applications of the Notion of Complex Signal” – <http://www.rand.org/pubs/translations/2008/T92.pdf> ; cited or mentioned on the following pages above: IV-VI.)

**Weyl, H.** 1927. Quantenmechanik und Gruppentheorie. – *Zeitschrift für Physik*. Bd. 46, No 1-2 (Nov 1927), 1-46 (cited or mentioned on the following pages above: IV, XIV).

**Wigner, E.** 1932. On the Quantum Correction For Thermodynamic Equilibrium. – *Physical Review*. Vol. 40, No 5 (June 1932), 749-759 (cited or mentioned on the following pages above: III-IV, VI, VII, IX, XII-XIV).

**Пенчев, В.** 2009. Философия на квантовата информация. Айнщайн и Гьодел. С.: ИФИ-БАН (cited or mentioned on the following pages above: VIII-IX, XXXIX).