האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM

QUANTIC SUPERPOSITIONS AND THE GEOMETRY OF COMPLEX HILBERT SPACES

by

DANIEL LEHMANN

Discussion Paper # 447 February 2007

מרכז לחקר הרציונליות

CENTER FOR THE STUDY OF RATIONALITY

Feldman Building, Givat-Ram, 91904 Jerusalem, Israel
PHONE: [972]-2-6584135 FAX: [972]-2-6513681
E-MAIL: ratio@math.huji.ac.il
URL: http://www.ratio.huji.ac.il/

Quantic Superpositions and the Geometry of Complex Hilbert spaces *

Daniel Lehmann
Selim and Rachel Benin School of Computer Science and Engineering
and Center for the Study of Rationality
Hebrew University,
Jerusalem 91904, Israel

February 2007

Abstract

The concept of a superposition is a revolutionary novelty introduced by Quantum Mechanics. If a system may be in any one of two pure states x and y, we must consider that it may also be in any one of many *superpositions* of x and y. This paper proposes an in-depth analysis of superpositions. It claims that superpositions must be considered when one cannot distinguish between possible paths, i.e., histories, leading to the current state of the system. In such a case the resulting state is some compound of the states that result from each of the possible paths. It claims that states can be compounded, i.e., superposed in such a way only if they are not orthogonal. Since different classical states are orthogonal, the claim implies no non-trivial superpositions can be observed in classical systems. It studies the parameters that define such compounds and finds two: a proportion defining the mix of the different states entering the compound and a phase difference describing the interference between the

*This work was partially supported by the Jean and Helene Alfassa fund for research in Artificial Intelligence, by the Israel Science Foundation grant 183/03 on "Quantum and other cumulative logics" and by EPSRC Visiting Fellowship GR/T 24562 on "Quantum Logic"

different paths. Both quantities are geometrical in nature: relating one-dimensional subspaces in complex Hilbert spaces. It proposes a formal definition of superpositions in geometrical terms. It studies the properties of superpositions. Keywords: Superpositions in Quantum Mechanics, Geometry of Hilbert Spaces, Quantum measurements, Measurement algebras, Quantum Logic. PACS: 02.10.-v.

1 Introduction and Previous Work

During the elaboration of [1] John von Neumann wrote to Garret Birkhoff: "I would like to make a confession which may seem immoral: I do not believe absolutely in Hilbert space any more. After all Hilbert-space (as far as quantum-mechanical things are concerned) was obtained by generalizing Euclidean space, footing on the principle of "conserving the validity of all formal rules". This is very clear, if you consider the axiomatic-geometric definition of Hilbert-space, where one simply takes Weyl's axioms for a unitary-Euclidean space, drops the condition on the existence of a finite linear basis, and replaces it by a minimum of topological assumptions (completeness + separability). Thus Hilbert-space is the straightforward generalization of Euclidean space, if one considers the *vectors* as the essential notions. Now we begin to believe that it is not the *vectors* which matter but the *lattice of all linear (closed) subspaces*. Because:

1. The vectors ought to represent the physical *states*, but they do it redundantly, up to a complex factor only.
2. And besides the *states* are merely a derived notion, the primitive (phenomenologically given) notion being the *qualities*, which correspond to the *linear closed subspaces*" (see [4], p. 59, letter dated Nov. 13, Wednesday, 1935).

The goal of this work is to pursue von Neumann's program of describing Quantum Logic in terms of closed subspaces and without vectors one step further. This work presents two original features:

- it takes a logical approach to Quantum Physics, where states and propositions take the main roles, and

- while it assumes the formalism of Hilbert spaces that fits Quantum Physics it tries the utmost to use only notions, such as states, propositions, projections, orthogonality and so on, that have a meaning, albeit mostly trivial, in Classical Physics. Special care will be taken to ensure that the quantic principles proposed hold classically.

2 Quantum Logic

One may say that Logic is the study of the relation between states of the world and propositions used to talk about those states. Quantum logic must therefore be the study of the relation between quantum states and quantum propositions. The accepted view is that both quantum states and quantum propositions should be represented by closed subspaces of a Hilbert space. Quantum states are one-dimensional subspaces. Quantum logic is therefore the study of the relation between one-dimensional subspaces and arbitrary closed subspaces. One obvious topic for Quantum logic is therefore the study of the properties of projections in Hilbert spaces: a one-dimensional subspace projects onto a one-dimensional or zero-dimensional subspace of any closed subspace. Projections are also central to Quantum physics since they correspond to the change brought about by the measurement of a physical property. In a previous paper [3], a first study of some of the properties of such projections has been presented: it deals only with qualitative properties. The present paper inaugurates the quantitative study of the projective geometry of complex Hilbert spaces.

The purpose of the exercise is to shed light on the notion of measurement in Quantum Physics by developing a geometry of Hilbert spaces whose entities are physically meaningful: states of physical systems and measurements on physical systems. Our goal can be understood in considering the history of geometry. Euclidean plane geometry was the starting point. Its elements are points and lines. Mathematical developments (due to Descartes in particular) enabled a treatment of geometry in the vector space $\mathcal{R}^n$. A new definition of geometry, abstracting from the vector space structure and returning to the basic notions of points and lines, enabled the development of non-Euclidean geometries. For Hilbert spaces, historically the algebraic presentation came first. The purpose of this paper is to extract from the algebraic presentation a leaner presentation similar in spirit to Euclid's geometry. Our basic entities are one-dimensional subspaces and, more generally, closed subspaces *and not vectors.*

In an obvious way, two elements (vectors) of a Hilbert space define a number, their inner product. We are looking for numbers that characterize subspaces, not vectors. This paper proposes to associate a real number with any pair of one-dimensional subspaces x, y: $p(x, y)$ and, by extension, to any pair of a one-dimensional subspace and a closed subspace α: $p(x, \alpha)$. This number is always in the interval $[0, 1]$ and behaves in many ways like *the probability that the proposition α is found true when it is tested on state x*, in line with the probabilistic interpretation of Quantum Physics. It satisfies further properties that are more difficult to interpret and characterize the linear dependence structure and the definition of projections.

Another numerical quantity, an angle, θ, is defined by any triple of one-dimensional subspaces. It represents the phase difference created by two different measurements and will be interpreted as the source of the interference occurring between alternative paths a particle could take. This paper is devoted to the study of those aspects of the geometry of Hilbert spaces related to the numbers p and θ. The study of those M-algebras (see [3]) that admit quantities satisfying the properties of p and θ is left for further study.

3 Background and Notations

We assume a Hilbert space $\mathcal{H}$ on the field $\mathcal{C}$ of complex numbers is given. The complex conjugate of a complex number c is $\bar{c}$. For any complex number c, $|\, c \,|$ represents its modulus, which is a nonnegative real number. For any complex number c different from 0, $\arg(c)$ represents its complex argument: $c = |\, c \,| \; e^{i \arg(c)}$. Elements of $\mathcal{H}$ will typically be: $\vec{u}, \vec{v} \ldots$. The zero vector is denoted by $\vec{0}$. The inner product of $\vec{u}$ and $\vec{v}$ is $\langle \vec{u}, \vec{v} \rangle$. The inner product is linear in its first argument and conjugate-linear in its second argument. Two vectors $\vec{u}$ and $\vec{v}$ are perpendicular, written $\vec{u} \perp \vec{v}$, iff $\langle \vec{u}, \vec{v} \rangle = 0$. The norm of $\vec{u}$ is $\parallel \vec{u} \parallel$. A unit-vector is a vector of norm 1.

The set of all closed subspaces of $\mathcal{H}$ will be denote by M. The elements of M should be thought of representing propositions, or, results of physical measurements. Greek letters from the beginning of the alphabet will be used to denote elements of M. The reader may think of a typical element of M, α as meaning *the spin in the z-direction is nonnegative.* Note that propositions represent measurements with a specified result or a set of possible results: such as measuring the value 1/2 for the spin in the z-direction or measuring a nonnegative value for this spin. To every $\alpha \in M$ one may associate its orthogonal complement, which will be denoted $\neg\alpha$. The proposition $\neg\alpha$ is interpreted as the measurement that measures the quantity measured by α but provides a value that is not in the set specified by α. If α claims that the spin in the z-direction is nonnegative, $\neg\alpha$ measures the spin along the same direction but finds it negative. Two specific propositions are worth mentioning: falsehood, $\bot$ is the null subspace $\{\vec{0}\}$ and truth, $\top$ is the whole space $\mathcal{H}$. Any closed subspace α of $\mathcal{H}$ defines of projection of $\mathcal{H}$ onto α. For any $\vec{u} \in \mathcal{H}$ its projection on α will be denoted $\alpha(\vec{u})$. The relation between physical measurements and projections will be explained after we discuss states.

Among the closed subspaces of $\mathcal{H}$ particular attention will be paid to one-dimensional subspaces. The set of one-dimensional subspaces of $\mathcal{H}$ is denoted X and the elements of X are typically letters from the end of the alphabet: x, y and so on. As mentioned just above: $X \subseteq M$. Elements of X will be called *states*. A one-dimensional subspace x represents a possible (pure) state of the physical system. Think of the state in which the spin in the z-direction is $1/2$, for example. Our assumption that states are propositions. The fact that $X \subseteq M$ reflects the situation in which every pure state has an associated measurement that characterizes it: one may measure the spin in the z-direction and one of the possible values is $1/2$. The proposition "the spin in the z-direction is nonnegative" is not a state.

Since a proposition $\alpha \in M$ is a closed subspace of $\mathcal{H}$, for any $x \in X$, either $x \subseteq \alpha$ or α contains no vector of x except the zero vector. Any proposition is the union of the states it includes and any proposition can be seen as the set of all the states it includes. We shall indeed prefer the notation $x \in \alpha$ to $x \subseteq \alpha$.

Note that if $\vec{v} \in x \in X$ and $\vec{u} \perp \vec{v}$ then $\vec{u} \perp \vec{w}$ for every $\vec{w} \in x$. We denote such a situation by $\vec{u} \perp x$. If every $\vec{u} \in \alpha$ is orthogonal to x we say that $x \perp \alpha$. If every $x \in X$, $x \in \alpha$ is orthogonal to β we say that $\alpha \perp \beta$. The image of any $x \in X$ by any (projection) $\alpha \in M$ is either a one-dimensional subspace $y \in X$ or the zero-dimensional subspace. This second possibility occurs exactly when x is orthogonal to α. We shall denote by $\alpha(x)$ the one-dimensional or zero-dimensional subspace that is the projection of x onto α. Note that $\alpha(x) = x$ iff $x \in \alpha$. We write $\alpha(x) = 0$ to denote the case $\alpha(x)$ is zero-dimensional, i.e., the case $x \perp \alpha$. The projection of the zero-dimensional subspace on any α is the zero-dimensional subspace and we shall extend the action of α by setting $\alpha(0) = 0$.

In Quantum Physics measurements may change the state of the system. The state obtained when measuring α in state x is precisely $\alpha(x)$, the projection of x on the subspace α. If x is orthogonal to α, then the measurement α is impossible in state x: this happens precisely when the quantity measured by α has, in x, a well-defined value that is not in the set specified by α. Equivalently, this happens precisely when x is in the subspace $\neg\alpha$, or $(\neg\alpha)(x) = x$.

4 Classical Physics

The notions described in Section 3 have been given a meaning grounded in the Hilbert space formalism of Quantum Mechanics. This seems to preclude their application to Classical Mechanics, since, classically, states are not rays in a Hilbert space. Nevertheless, the common wisdom is that Quantum Mechanics should apply everywhere and that Classical Mechanics should be a limiting case of Quantum Mechanics. Indeed, both Classical Mechanics and Quantum Mechanics can be studied in structures that abstract from the definitions of Section 3, preserving the properties of states and measurements. A full treatment is left for future work, but the following remark explains the main feature of classical systems.

Classically, measurements do not change the state of a system, therefore if a state x is not orthogonal to a proposition α, we have $\alpha(x) = x$, expressing the fact that either x possesses the property α or it possesses its negation $\neg\alpha$. We have:

Principle of Classical Physics Any two different states are orthogonal.

5 The Reciprocity Principle

Before proceeding to an analysis of the notion of a superposition, a first principle will be described. It states that if the measurement $\neg x$ acting on states y and z produces the same state, then x, y and z must sit in the same plane, and therefore the measurement $\neg y$ must produce the same state when acting on x and on z.

Reciprocity Principle Let $x, y, z \in X$, be pairwise different.

$$\text{Then } (\neg x)(y) = (\neg x)(z) \;\Rightarrow\; (\neg y)(z) = (\neg y)(x).$$

The Reciprocity Principle suggests the following definition.

Definition 1 *We shall say that states x, y and z are* coplanar*, written* $coplanar(x, y, z)$ *iff either two out of the three are equal, or they are pairwise different and* $(\neg x)(y) = (\neg x)(z)$.

The Reciprocity Principle says that *coplanarity* is a property of the set $\{x, y, z\}$, i.e., for any permutation x', y', z' of x, y, z $coplanar(x', y', z')$ is equivalent to $coplanar(x, y, z)$.

The Reciprocity Principle is experimentally testable: if the *no* answer to a test x gives the same state when performed on y and on z, the *no* answer on a test y will give the same answer on z and x.

In Hilbert space, indeed, if y and z have the same projection on the subspace orthogonal to x, call it x', then all four one-dimensional subspaces: x, x', y and z are in the same two-dimensional subspace, call it α, and therefore the projections of z and x on the subspace orthogonal to y are both the one-dimensional subspace of α orthogonal to y.

In Classical Physics, the Reciprocity Principle holds trivially, since its assumptions are never satisfied. Indeed if $x \neq y$, we have $(\neg x)(y) = y$, and similarly $(\neg x)(z) = z$ and therefore the assumption implies $y = z$.

6 Superpositions: Conceptual Analysis

The concept of a superposition is a revolutionary novelty introduced by Quantum Mechanics. If a system may be in any one of two pure states x and y, we must consider that it may also be in any one of many *superpositions* of x and y. This paper is devoted to an in-depth analysis of superpositions.

A remark that has resulted in a vast literature is the following: the revolutionary character of quantic superpositions is the consequence of the fact no such superpositions have to be considered, or may be seen in classical systems. In Schrödinger's colorful thought experiment: the cat is either dead or alive, but nobody has evidence of a superposition of a dead and a live cat. This seems to contradict the principle exposed in Section 4, of the universality of Quantum Mechanics. If everything in the universe is quantic and any two quantic states can be superposed, then any two classical states, such as a live and a dead cat, can be superposed. Many explanations have been proposed and this is not the place for a survey. Most explanations accept the existence of superpositions of classical states and explain why such superpositions are not seen. The analysis of the superposition concept to be developed below proposes a radically different explanation. It is not the case that, in Quantum Mechanics, any two states can be superposed: no superposition of orthogonal states can ever be considered. Since different classical states are orthogonal, the only superpositions of classical states that can ever occur are trivial: superpositions of a state with itself. Trivial superpositions are indeed observed and unproblematic.

First, we wish to reflect on the nature of superpositions and their origin: what are they and how do they come into consideration, without trying to describe formally such superpositions. Then, we shall propose a formalization.

The reader should notice that the linear combination of vectors of a Hilbert space provides a formal operation, not a conceptual analysis, and also that, since vectors do not represent states, the linear combination of vectors cannot offer a proper formalization for the superpositions of states. Even though we announced above that orthogonal states cannot be superposed, it is clear that orthogonal unit vectors can be combined linearly to form unit vectors. This should convince the reader that we shall not formalize superposition as a straightforward linear combination.

6.1 Nature and Origin

Superpositions must be considered to describe systems about which all we know is that they are the result of one of a number of different possible paths (or histories), i.e., if we have no way of knowing which history indeed took place. In such a case, we must consider that the system is in some state that is a superposition, i.e., a *compound* of the states that are the produced by each of the possible paths. The term *compound* is used here where, chemically-speaking, the term *mixture* may be more appropriate because this last term is used in Quantum Mechanics with a different meaning.

If one knows which path has been taken, or one could discover which path has been taken, then one must consider that the system is in the state that results from the path taken, and one must use probability theory to describe one's ignorance about the state of the system. If one does not know and cannot know which path has been taken, then one must consider that the system is in some specific superposition of the states resulting from the different possible paths. This is a general principle: if one cannot know which path has been taken, then those paths *interfere* and therefore the system cannot be described using only probability theory, but must be described by a state that is a compound, i.e., a superposition of the states resulting from the different interfering paths. This general principle holds also in Classical Physics, as will be seen in Section 6.3. The way in which the different paths may interfere, i.e., the parameters that characterize the different possible superpositions will be described in Section 6.2.

The paradigmatical example of such a situation is a the two-slits experiment in which a particle travels through one of two slits and one does not know which.

6.2 Parameters

To leave things simple we shall consider only the superpositions of two states, without loss of generality as long as we consider only a finite number of possible paths. Generalizing to path integrals is beyond the scope of this paper. Suppose therefore that we must deal with a system that may result from two different paths. If path p_1 was taken, the system is in state y; if path p_2 was taken, the system is in state z. If one cannot know which path was taken, one must consider that the system is in a state that is some superposition of the two states y and z. Many such superpositions are possible and the purpose of this section is to describe the experimental parameters that influence the superposition to be used. In Section 6.3, the question of whether we can know which path was taken will be given and unequivocal answer.

It is natural to consider that the superposition of y and z obtained as the result of the interference between the two paths p_1 and p_2 is characterized by two parameters. The proper value to be chosen for each of those parameters is a function of the experimental setup and of the path p_1 and p_2.

The first parameter, that will be denoted by r, describes the composition of the *compound* that is a superposition. It describes, in a sense, the respective proportions (ratios) of y and z present in the superposition. The parameter r is therefore a real number: $0 < r < 1$ that describes the *weight* of y relative to z in the superposition. We do not allow the values 0 or 1 since they do not describe a true superposition.

In the two-slits experiment, where y represents the state resulting from the electron moving through the upper slit and z the state resulting from the electron moving through the lower slit, the parameter r will depend on the respective widths of the two slits and the respective distance of those slits to the origin. For example, if the upper slit is twice as large than the lower one and both slits are at equal distance from the origin we shall put $r = 2/3$ expressing the fact that we judge the possibility of y twice as high as that of z. In the spin experiment, r will depend on the respective a priori probabilities we give to each of the measurements.

The second parameter characterizing the interference between the two paths, and therefore the superposition, is an angle φ, $0 \leq \varphi < 2\,\pi$. This angle represents the phase difference incurred in path p_2 with respect with p_1. When we add or subtract angles, this has to be understood modulo $2\,\pi$. This phase difference is determined by the experimental set-up: in a two-slits experiment by the relative distance of the two slits to the origin and the wavelength of the electron.

It seems to be a principle of Physics that any superposition of two states y and z can be described with the help of those two parameters only. The only superpositions we shall ever consider are therefore of the form $super(y, z, r, \varphi)$ for states $y, z \in X$ and real numbers $0 < r < 1$ and $0 \leq \varphi < 2\,\pi$. The telling notation $r\,y \;+\; (1 - r)\,e^{i\varphi}\,z$ will be used in place of the more austere $super(y, z, r, \varphi)$, but the reader is warned that $+$ does not mean addition, " " does not mean multiplication and some of the properties one would expect from our notation do *not* hold. In particular the composition of superpositions does not possess the properties suggested by the notation.

6.3 Conditions

Section 6.2 indicated that superpositions of states y and z should be considered only if there is no way to know which one of the paths p_1 or p_2 leading to y and z respectively has been traveled. It is time to reflect on this condition.

If the states y and z are orthogonal: $y \perp z$, then there is a way to find out for sure which of the two paths has been traveled: perform on the resulting state a measurement testing whether the state is y or not: a test y, $(\neg y)$. If path p_1 has been traveled, the result will be a *yes* for sure since the state is y. If path p_2 has been traveled, the result, for sure, will be a *no* since the state is z, orthogonal to y. Similarly, we could have tested for z or for any proposition satisfied by one of the states y or z and orthogonal to the other one. We see that no superposition of orthogonal states can ever be defined. This is is stark contrast with the linear combination of vectors in a Hilbert space.

Further reflection shows that if the states y and z are not orthogonal, one can never find out for sure which of the paths p_1 or p_2 has been traveled. Indeed the only situation in which one could find out would be to test for some proposition α satisfied, for sure, by one of the two states y or z and not satisfied, for sure, by the other state. In other terms, a closed subspace α containing one of y or z and orthogonal to the other one. But this implies $y \perp z$. We see that:

Principle of Superposition The superposition $r\, y \;+\; (1 - r)\, e^{i\varphi}\, z$

is defined if and only if $y \not\perp z$.

A list of desirable properties of superpositions will be presented now. In Section 9 a definition of superpositions in the formalism of Hilbert spaces will be provided, that satisfies all those desirable properties. It is a thesis of this paper that the epistemological status of the properties below should be considered a ground more secure than the Hilbert space apparatus. We shall now deal, first, with trivial superpositions and, then, with a first principle concerning generic superpositions.

6.4 Trivial Superpositions

Let us consider, first, the superpositions of a state y with itself: $r\, y \;+\; (1 - r)\, e^{i\varphi}\, y$. By the Principle of Classical Physics, these are the only superpositions possible in classical physics.

Evidence from both classical and quantum physics shows that such superpositions are trivial:

Principle of Triviality $\forall y \in X, \forall r \in]0, 1[, \forall \varphi \in [0, 2\,\pi[, r\, y + (1 - r)\, e^{i\varphi}\, y = y$.

Having disposed of the cases $y \perp z$ and $y = z$, let us study the generic case of superpositions.

6.5 Principle of Coplanarity

Our first principle is that a superposition is coplanar with its components. Assume $y \not\perp z$.

Principle of Coplanarity $\forall r \in]0, 1[, \forall \varphi \in [0, 2\,\pi[,$

$$coplanar(r\,y\ +\ (1-r)\,e^{i\varphi}\,z,\ y,\ z).$$

This principle can be justified in the following way. The superposition $x = r\,y\ +\ (1-r)\,e^{i\varphi}\,z$ results from our inability to know which of p_1, resulting in y or p_2, resulting in z has been traveled. Measuring $\neg y$ on x shows that the path p_1 has not been traveled and therefore p_2 has been traveled and the current state $(\neg y)(x)$ is in fact $(\neg y)(z)$.

The following is some kind of a converse of the principle of coplanarity.

Principle of Uniqueness If $y \neq z, y \not\perp z, x \not\perp y, x \not\perp z$ and $coplanar(x, y, z)$,

$$\text{then } \exists r \in]0,1[, \varphi \in [0, 2\,\pi[,\ \text{such that } x = r\,y\ +\ (1-r)\,e^{i\varphi}\,z.$$

Moreover the r and the φ above are unique.

In the sequel, it will be shown that the geometry of Hilbert spaces allows for the definition of two quantities $\rho(x, y, z)$ and $\theta(x, y, z)$ for states x, y and z that satisfy assumptions slightly more liberal than those of the Principle of Superposition, such that $x = \rho(x, y, z)\,y\ +\ (1 - \rho(x, y, z))\,e^{i\theta(x,y,z)}\,z$.

In Section 7, the quantity $\rho(x, y, z)$ will be defined. In Section 8, the quantity $\theta(x, y, z)$ will be defined. In Section 9, a definition of superpositions in the formalism of Hilbert spaces will be proposed, and it will be shown that this definition satisfies all the properties deemed desirable. In Section 10, additional properties of superpositions will be studied.

7 Definition of $\rho(x, y, z)$

Given any three states $x, y, z \in X$, we shall define a real number $\rho(x, y, z)$. This definition will take place in three stages. First, we shall define a geometrical property of two states.

7.1 The quantity $a(x, y)$

We shall now define the first geometric quantity we wish to consider. When considering the geometry of Hilbert spaces it is useful to begin by reflecting on the geometry of Euclidean spaces, about which we know much more and have a much better intuition. Consider two lines, i.e, one-dimensional linear (not affine) subspaces, in $\mathcal{R}^n$. The only invariant characterizing their relation is their angle. Two lines define a plane and four angles. Those four angles are two pairs of equal angles. Therefore only two quantities are defined by two lines. Moreover those two angles add up to π, therefore there is essentially only one quantity defined. One can take as the fundamental quantity either the acute or the obtuse angle. Let us consider the acute angle as the quantity of interest. Two lines in Euclidean space define an angle φ in the interval $[0, \pi/2]$. Equivalently, they define a real number in the interval $[0, 1]$, the value of $\cos(\varphi)$.

The same quantity may be defined in Hilbert spaces.

Consider two states $x, y \in X$. We are trying to associate a numerical quantity to this pair of states. The most natural thing to consider is the inner product of two vectors contained in x and y respectively. It is very natural to choose two unit-vectors $\vec{u} \in x$ and $\vec{v} \in y$ and consider the inner product $\langle \vec{u}, \vec{v} \rangle$. This will not do since the quantity depends on the choice of the unit-vectors $\vec{u}$ and $\vec{v}$ and we are looking for a quantity that depends only on x and y. The inner product depends on the choice of the unit-vectors, but its modulus does not. Consider therefore the quantity

$$a(x, y) \stackrel{\text{def}}{=} |\langle \vec{u}, \vec{v} \rangle|$$

for arbitrary unit-vectors $\vec{u}$ and $\vec{v}$ of x and y respectively. Any unit-vector $\vec{u}'$ of x has the form: $\vec{u}' = e^{i\theta}\vec{u}$ and any $\vec{v}'$ of y has the form: $\vec{v}' = e^{i\varphi}\vec{v}$. Therefore $\langle \vec{u}', \vec{v}' \rangle = e^{i(\theta-\varphi)}\langle \vec{u}, \vec{v} \rangle$, and $|\langle \vec{u}', \vec{v}' \rangle| = |\langle \vec{u}, \vec{v} \rangle|$.

The following is easily proved.

Lemma 1 *For any $x, y \in X$:*

1. *$a(x, y)$ is a real number of the interval $[0, 1]$,*
2. *$a(x, y) = 1$ iff $x = y$,*
3. *$a(x, y) = 0$ iff $x \perp y$,*
4. *$a(y, x) = a(x, y)$.*

7.2 Similarity: the quantity p

It turns out that the square of the quantity $a(x, y)$, akin to the $\cos^2$ of an angle has even more remarkable properties.

Definition 2 *Given any states $x, y \in X$, we shall define their similarity $p(x, y)$ by*

$$p(x, y) = a^2(x, y).$$

The quantity p will be called *similarity* because it measures how similar, i.e., close, are its arguments x and y. Its physical interpretation is straightforward: $p(x, y)$ is the probability that, when, on state x, one tests whether y is the case, one gets a positive answer. With probability $1 - p(x, y)$ one gets the the answer that y is not the case. This physical interpretation is the reason $p = a^2$ and not a has been chosen as the quantity of reference. Note that p can be directly obtained experimentally. Below, we shall extend the definition of p to measure the *similarity* between any state $x \in X$ and any proposition $\alpha \in M$, i.e., the degree to which state x satisfies proposition α.

A straightforward result on Hilbert spaces will be recalled now.

Lemma 2 *Let $\vec{u}, \vec{v} \in \mathcal{H}$. Assume $\vec{v}$ is a unit-vector and $\vec{v} \in x \in X$. Then the projection $x(\vec{u})$ of $\vec{u}$ on x is $\langle \vec{u}, \vec{v} \rangle \, \vec{v}$.*

Proof: $\vec{u} - \langle \vec{u}, \vec{v} \rangle \, \vec{v}$ is indeed orthogonal to $\vec{v}$ and therefore to x. ∎

First properties of p are described in the following.

Lemma 3 *For any $x, y \in X$:*

1. *$p(x, y)$ is a real number in the interval $[0, 1]$,*
2. *$p(x, y) = 1$ iff $x = y$,*
3. *$p(x, y) = 0$ iff $x \perp y$,*
4. *$p(y, x) = p(x, y)$,*
5. *for any unit-vector $\vec{u} \in x$, $p(x, y) = \langle \vec{u}, y(\vec{u}) \rangle$ where $y(\vec{u})$ is the projection of $\vec{u}$ on y,*
6. *for any unit-vector $\vec{u} \in x$, $p(x, y) = \parallel y(\vec{u}) \parallel^2$.*

Proof: For 5, note that for any unit-vector $\vec{v}$ of y, we have, by Lemma 2, $y(\vec{u}) = \langle \vec{u}, \vec{v} \rangle \, \vec{v}$, and therefore $\langle \vec{u}, y(\vec{u}) \rangle = \overline{\langle \vec{u}, \vec{v} \rangle} \, \langle \vec{u}, \vec{v} \rangle = | \langle \vec{u}, \vec{v} \rangle |^2$. Note that this implies that the inner product $\langle \vec{u}, y(\vec{u}) \rangle$ is a real number. For 6, note that projections are Hermitian and idempotent, and therefore $\langle y(\vec{u}), y(\vec{u}) \rangle = \langle \vec{u}, y(y(\vec{u})) \rangle = \langle \vec{u}, y(\vec{u}) \rangle$. ∎

The next result is central. It shows that, for states of any given proposition α, the projection on α is determined by the p-structure.

Theorem 1 *For any proposition $\alpha \in M$ and any states $x, y \in X$, if $x \not\perp \alpha$ and $y \in \alpha$ then $p(x, y) = p(x, \alpha(x)) \, p(\alpha(x), y)$.*

Proof: Let $\vec{u}$ be a unit-vector of x. Since $y \in \alpha$, the projection of any vector on y can be obtained by projecting the vector first on α and then projecting the result on y. In particular, $y(\vec{u}) = y(\alpha(\vec{u}))$. Therefore

$$p(x, y) = \parallel y(\vec{u}) \parallel^2 = \parallel y(\alpha(\vec{u})) \parallel^2 / \parallel \alpha(\vec{u})) \parallel^2 / : \times \ \parallel \alpha(\vec{u}) \parallel^2$$

Let $\vec{v} = \alpha(\vec{u}) / \parallel \alpha(\vec{u} \parallel$. Notice that $\vec{v}$ is a unit-vector of $\alpha(x)$ and therefore

$$p(x, y) = \parallel \vec{v} \parallel^2 \ \times \ \parallel \alpha(\vec{u}) \parallel^2 = p(\alpha(x), y) \ \times \ p(x, \alpha(x))$$

since $\alpha(\vec{u})$ is the projection of $\vec{u}$ on $\alpha(x)$, and by Lemma 3. ∎

The following shows that the action of projections can be read of the similarity p.

Corollary 1 *For any proposition $\alpha \in M$ and any state $x \in X$, if $x \not\perp \alpha$ then $\alpha(x)$ is the unique state y of α on which the value of $p(x, y)$ is maximal.*

In short, there is a unique state of α that is most similar to x, this is x's projection on α.
Proof: By Theorem 1, since $p(\alpha(x), y) \leq 1$ by Lemma 3, $p(x, y) \leq p(x, \alpha(x))$ for any $y \in \alpha$.

For uniqueness, suppose $y \in \alpha$ and $p(x, y) = p(x, \alpha(x))$. By Theorem 1, $p(x, \alpha(x)) = p(x, \alpha(x)) \, p(\alpha(x), y)$. Since x is not orthogonal to α, $p(x, \alpha(x)) > 0$ and therefore $p(\alpha(x), y) = 1$ and $\alpha(x) = y$. ∎

It is now only natural to extend the definition of p to an arbitrary proposition as second argument. For any $x \in X$ and $\alpha \in M$, we define $p(x, \alpha)$ in the following way:

- $p(x, \alpha) = 0$ if $x \perp \alpha$, and
- $p(x, \alpha) = p(x, \alpha(x))$ otherwise.

The following is known, in Physics, as Born's rule. The quantity $p(x, \alpha)$ is the probability of measuring the property α in state x.

Lemma 4 *For any state $x \in X$ and any proposition $\alpha \in M$, if $\vec{u} \neq \vec{0} \in x$, $p(x, \alpha) = \parallel \alpha(\vec{u}) \parallel^2 / \parallel \vec{u} \parallel^2$.*

The proof is obvious. The following is an obvious consequence of Corollary 1.

Corollary 2 *For any state x and any proposition α, $x \in \alpha$ iff $\alpha(x) = x$ iff $p(x, \alpha) = 1$.*

The next two sections prove additional properties of the quantity p. Section 7.3 shows that, for any given x and different α's, $p(x, \alpha)$ behaves very much as a probability on the propositions. Exactly so, for propositions that commute as projections. Section 7.4 proves an intriguing inequality that provides a numerical strengthening of the **Interference** property of [3]. Section 7.3 and 7.4 are mainly technical and can be skipped without conceptual harm.

7.3 Similarity as probability

The following results will show that, for any fixed $x \in X$, the quantities $p(x, \alpha)$ for different measurements α play the role of a probability on the propositions. For any two propositions $\alpha, \beta \in M$ we shall define, as traditional since [1], their conjunction $\alpha \wedge \beta$ as their intersection $\alpha \cap \beta$ (note the intersection of closed subspaces is a closed subspace) and their disjunction $\alpha \vee \beta$ as the topological closure of their linear sum: $cl(\alpha + \beta)$. Note that these notations are inconsistent with those of [3] where conjunction and disjunction were defined only for *commuting* propositions. We shall demonstrate a particular interest in *commuting* propositions. For the sake of obtaining a straightforward definition of commutation, we shall extend our notation for projections.

Definition 3 *Let $\alpha, \beta \in M$ be two propositions. We shall say that α and β* commute *iff for any $x \in X$ $\alpha(\beta(x)) = \beta(\alpha(x))$.*

Lemma 5 *Any two propositions $\alpha, \beta \in M$ commute iff there are three pairwise orthogonal propositions $\gamma_i, i = 1, \ldots, 3$ such that $\alpha = \gamma_1 \vee \gamma_2$ and $\beta = \gamma_1 \vee \gamma_3$.*

Note that one of the propositions γ_i may be the falsehood $\perp$.
Proof: The *if* claim is obvious. The *only if* claim follows from the fact that projections are Hermitian and that Hermitian operators commute iff they have a joint basis of eigenvectors. ∎

Corollary 3 *For any $\alpha, \beta \in X$, if $\alpha \subseteq \beta$ or $\alpha \perp \beta$, then α and β commute.*

Proof: In the first case, take $\gamma_1 = \alpha$, $\gamma_2 = \perp$ and $\gamma_3 = \neg\alpha \wedge \beta$. In the second case, take $\gamma_1 = \alpha$, $\gamma_2 = \perp$ and $\gamma_3 = \beta$. ∎

Corollary 4 *For any $\alpha, \beta \in X$, if α and β commute then $\neg\alpha$ and β commute.*

Proof: Let $\alpha = \gamma_1 \vee \gamma_2$ and $\beta = \gamma_1 \vee \gamma_3$. Then $\neg\alpha = \neg\gamma_1 \wedge \neg\gamma_2$. Since $\gamma_3 \subseteq \neg\alpha$, we have, by the orthomodular property, $\neg\alpha = \gamma_3 \vee \neg\gamma_1 \wedge \neg\gamma_2 \wedge \neg\gamma_3$. But $\beta = \gamma_3 \vee \gamma_2$ and $\gamma_2 \perp \neg\gamma_1 \wedge \neg\gamma_2 \wedge \neg\gamma_3$. ▮

First, we shall consider disjunctions of orthogonal propositions.

Lemma 6 *If $\alpha \perp \beta$ then, for any $x \in X$, $p(x, \alpha \vee \beta) = p(x, \alpha) + p(x, \beta)$.*

Proof: Consider any $\vec{u} \neq \vec{0} \in x$. Now $(\alpha \vee \beta)(\vec{u}) = \alpha(\vec{u}) + \beta(\vec{u})$ (see [2] Theorem 2, page 46). Therefore $\langle \vec{u}, (\alpha \vee \beta)(\vec{u}) \rangle = \langle \vec{u}, \alpha(\vec{u}) \rangle + \langle \vec{u}, \beta(\vec{u}) \rangle$. ▮

Corollary 5 *If α_i is a family of pairwise orthogonal measurements, then for any $x \in X$ we have $p(x, \bigvee_{i \in I} \alpha_i) = \sum_{i \in I} p(x, \alpha_i)$.*

Proof: By induction on the size of I, and associativity of disjunction. ▮

The following lemmas are fundamental characteristics of probabilities.

Lemma 7 *For any $\alpha \in M$ and any $x \in X$: $p(x, \alpha) + p(x, \neg\alpha) = 1$.*

Proof: By Lemma 6, $p(x, \alpha) + p(x, \neg\alpha) = p(x, \alpha \vee \neg\alpha)$. But $\alpha \vee \neg\alpha = \top$ and therefore $(\alpha \vee \neg\alpha)(x) = x$ and, by Corollary 2, $p(x, \alpha \vee \beta) = 1$. ▮

Lemma 8 *For any $\alpha \in M$ and any $x \in X$: $0 \leq p(x, \alpha) \leq 1$.*

Proof: By Lemmas 4 and 7. ▮

Lemma 9 *Let $\alpha, \beta \in M$ be any* commuting *measurements. For any $x \in X$ $p(x, \alpha \vee \beta) = p(x, \alpha) + p(x, \beta) - p(x, \alpha \wedge \beta)$.*

Proof: We know that $\alpha \vee \beta = (\alpha \wedge \beta) \vee (\alpha \wedge \neg\beta) \vee (\neg\alpha \wedge \beta)$. The three parts of the disjunction above are pairwise orthogonal, therefore Corollary 5 implies that $p(x, \alpha \vee \beta) = p(x, \alpha \wedge \beta) + p(x, \alpha \wedge \neg\beta) + p(x, \neg\alpha \wedge \beta)$. But, by Lemma 6: $p(x, \alpha \wedge \beta) + p(x, \alpha \wedge \neg\beta) = p(x, \alpha)$ and $p(x, \alpha \wedge \beta) + p(x, \neg\alpha \wedge \beta) = p(x, \beta)$. ▮

The lemmas above dealt mostly with the properties of disjunction. The next result concerns conjunction and parallels the consideration of conditional probabilities.

Lemma 10 *Let $\alpha, \beta \in M$ be any* commuting *measurements. For any $x \in X$: $p(x, \alpha \wedge \beta) = p(x, \alpha)\, p(\alpha(x), \beta)$.*

Proof: Since $\alpha \wedge \beta = \alpha \circ \beta$, by the definition of p, taking any $\vec{u} \neq \vec{0} \in x$:

$$p(x, \alpha \wedge \beta) = \frac{\| (\alpha \circ \beta)(\vec{u}) \|^2}{\| \vec{u} \|^2} = \frac{\| (\alpha \circ \beta)(\vec{u}) \|^2}{\| \alpha(\vec{u}) \|^2} \frac{\| \alpha(\vec{u}) \|^2}{\| \vec{u} \|^2} = p(\alpha(x), \beta)\, p(x, \alpha).$$

∎

Corollary 6 *Let $\alpha, \beta \in M$ be any measurements such that $\alpha \leq \beta$. Then for any $x \in X$, $p(x, \alpha) \leq p(x, \beta)$.*

Proof: If $\alpha \leq \beta$, the two measurements commute and $\alpha = \beta \wedge \alpha$. By Lemma 10, then $p(x, \alpha) = p(x, \beta)\, p(\beta(x), \alpha) \leq p(x, \beta)$ by Lemma 8. ∎

Corollary 7 *Let $\alpha, \beta \in M$ be any* commuting *measurements. Then for any $x \in X$, $p(x, \beta) = p(x, \alpha)\, p(\alpha(x), \beta) \; + \; p(x, \neg\alpha)\, p((\neg\alpha)(x), \beta)$.*

Proof: Since α and β commute, by Theorem 1 of [3], $\beta = (\alpha \wedge \beta) \vee (\neg\alpha \wedge \beta)$. By Lemma 6 we have: $p(x, \beta) = p(x, \alpha \wedge \beta) \; + \; p(x, \neg\alpha \wedge \beta)$. We conclude, by Lemma 10, that $p(x, \beta) = p(x, \alpha)\, p(\alpha(x), \beta) \; + \; p(x, \neg\alpha)\, p((\neg\alpha)(x), \beta)$. ∎

In Corollary 7 one cannot omit the requirement that α and β commute. The consideration of a two-dimensional Euclidean space where α is the x-axis and x makes an angle θ with the x-axis is sufficient. If β is x, then $p(x, \beta) = 1$ whereas $p(x, \alpha) = \cos^2(\theta) = p(\alpha(x), \beta)$ and $p(x, \neg\alpha) = \sin^2(\theta) = p((\neg\alpha)(x), \beta)$. Also taking β orthogonal to x gives $p(x, \beta) = 0$ and $p(x, \alpha) = \cos^2(\theta) = p((\neg\alpha)(x), \beta)$ and $p(x, \neg\alpha) = \sin^2(\theta) = p(\alpha(x), \beta)$. Nevertheless the result holds in the following case.

Lemma 11 *For any $x \in X$ and any $\alpha, \beta \in M$ such that $\alpha(x) \in \beta$ and $(\neg\alpha)(x) \in \beta$, one has*

$$p(x, \beta) = p(x, \alpha)\, p(\alpha(x), \beta) \; + \; p(x, \neg\alpha)\, p((\neg\alpha)(x), \beta) = 1.$$

Proof: By assumption both $\alpha(x)$ and $(\neg\alpha)(x)$ are subspaces of β. Given any $\vec{u} \in x$, both $\alpha(\vec{u})$ and $(\neg\alpha)(\vec{u})$ are in β. But β is a subspace and therefore $\alpha(\vec{u}) + (\neg\alpha)(\vec{u}) = \vec{u} \in \beta$. ∎

Lemma 12 *For any $x \in X$ and any $\alpha, \beta \in M$ such that $(\alpha \circ \beta)(x) = (\beta \circ \alpha)(x)$, we have $p(x, \beta) = p(x, \alpha)\, p(\alpha(x), \beta) \; + \; p(x, \neg\alpha)\, p((\neg\alpha)(x), \beta)$.*

Proof: Assume that $(\alpha \circ \beta)(x) = (\beta \circ \alpha)(x)$. By Lemma 4, $(\neg\alpha \circ \beta)(x) = (\beta \circ \neg\alpha)(x)$. Take any $\vec{u} \neq \vec{0} \in x$. Then,

$$p(x, \beta) \; = \parallel \beta(\vec{u}) \parallel^2 \; / \; \parallel \vec{u} \parallel^2 = \parallel \alpha(\beta(\vec{u}))) + (\neg\alpha)(\beta(\vec{u})) \parallel^2 \; / \; \parallel \vec{u} \parallel^2 =$$

$$\parallel \alpha(\beta(\vec{u}))) \parallel^2 \; / \; \parallel \vec{u} \parallel^2 + \parallel (\neg\alpha)(\beta(\vec{u})) \parallel^2 \; / \; \parallel \vec{u} \parallel^2 =$$

$$\parallel \beta(\alpha(\vec{u}))) \parallel^2 \; / \; \parallel \vec{u} \parallel^2 + \parallel (\beta)((\neg\alpha)(\vec{u})) \parallel^2 \; / \; \parallel \vec{u} \parallel^2 =$$

$$\frac{\parallel \beta(\alpha(\vec{u})) \parallel^2}{\parallel \alpha(x) \parallel^2} \, \frac{\parallel \alpha(x) \parallel^2}{\parallel \vec{u} \parallel^2} + \frac{\parallel (\beta)((\neg\alpha)(\vec{u})) \parallel^2}{\parallel (\neg\alpha)(x) \parallel^2} \, \frac{\parallel (\neg\alpha)(x) \parallel^2}{\parallel \vec{u} \parallel^2} \quad =$$

$$p(\alpha(x), \beta)\, p(x, \alpha) \; + \; p((\neg\alpha)(x), \beta)\, p(x, \neg\alpha).$$

∎

7.4 An inequality

The next result strengthens the Interference property of [3] by presenting a quantitative version of the principle.

Theorem 2 *For any $\alpha, \beta \in M$ and any $x \in X$ such that $\alpha(x) = x$,*

$$p(x, \beta)\, (1 - p(\beta(x), \alpha))^2 \; \leq \; p(\beta(x), \alpha)\, (1 - p(\alpha(\beta(x)), \beta))$$

Note that, by Theorem 1, $p(x, \beta) \leq p(\beta(x), \alpha)$ but $(1 - p(\beta(x), \alpha)) \geq (1 - p(\beta(x), \alpha))$. The fact that the quantity $1 - p(\beta(x), \alpha)$ appears squared seems inevitable. An examination of $\mathcal{R}^3$ shows that it may be the case that $p(x, \beta)\, (1 - p(\beta(x), \alpha)) \; > p(\beta(x), \alpha)(1 - p(\alpha(\beta(x)), \beta))$.

Proof: Assume $\vec{t} \neq \vec{0} \in x$. Let $\vec{u} = \beta(\vec{t})$, $\vec{v} = \alpha(\vec{u})$ and $\vec{w} = \beta(\vec{v})$.

In a first step we want to show that:

$$\parallel \vec{u} - \vec{v} \parallel^2 = \langle \vec{t}, \vec{v} - \vec{w} \rangle.$$

Indeed: $\parallel \vec{u} - \vec{v} \parallel^2 = \; \langle \vec{u} - \vec{v}, \vec{u} - \vec{v} \rangle \; = \; \langle \vec{u}, \vec{u} - \vec{v} \rangle - \; \langle \vec{v}, \vec{u} - \vec{v} \rangle$. But the last term is null since $\vec{u} - \vec{v}$ is orthogonal to α in general and in particular to $\vec{v}$. We have:

$$\parallel \vec{u} - \vec{v} \parallel^2 = \langle \vec{u}, \vec{u} - \vec{v} \rangle.$$

But $\vec{t} - \vec{u}$ is, similarly, orthogonal to $\vec{u}$ and $\langle \vec{u}\,, \vec{u} \rangle = \langle \vec{t}\,, \vec{u} \rangle$. Since $\vec{u} - \vec{v}$ is orthogonal to $\vec{t}$, $\langle \vec{t}\,, \vec{u} \rangle = \langle \vec{t}\,, \vec{v} \rangle$. We have:

$$\| \vec{u} - \vec{v} \|^2 = \langle \vec{t}\,, \vec{v} \rangle - \langle \vec{u}\,, \vec{v} \rangle.$$

Again, $\vec{v} - \vec{w}$ is orthogonal to $\vec{u}$ and therefore: $\langle \vec{u}\,, \vec{v} \rangle = \langle \vec{u}\,, \vec{w} \rangle$ and $\vec{t} - \vec{u}$ is orthogonal to $\vec{w}$ and we have: $\langle \vec{u}\,, \vec{w} \rangle = \langle \vec{t}\,, \vec{w} \rangle$. Therefore:

$$\| \vec{u} - \vec{v} \|^2 = \langle \vec{t}\,, \vec{v} \rangle - \langle \vec{t}\,, \vec{w} \rangle = \langle \vec{t}\,, \vec{v} - \vec{w} \rangle.$$

By Cauchy-Schwarz therefore we have:

$$\| \vec{u} - \vec{v} \|^2 \leq \| \vec{t} \| \, \| \vec{v} - \vec{w} \| \, .$$

and:

$$\| \vec{u} - \vec{v} \|^4 \leq \| \vec{t} \|^2 \, \| \vec{v} - \vec{w} \|^2 \, .$$

But: $\| \vec{u} \|^2 = \| \vec{v} \|^2 + \| \vec{u} - \vec{v} \|^2$, and $\| \vec{v} \|^2 = \| \vec{w} \|^2 + \| \vec{v} - \vec{w} \|^2$. Therefore we have:

$$(\| \vec{u} \|^2 - \| \vec{v} \|^2)^2 \ \leq \| \vec{t} \|^2 \ (\| \vec{v} \|^2 - \| \vec{w} \|^2).$$

and

$$\frac{\| \vec{u} \|^2}{\| \vec{t} \|^2} \, (1 - \frac{\| \vec{v} \|^2}{\| \vec{u} \|^2})^2 \ \leq \ \frac{\| \vec{v} \|^2 - \| \vec{w} \|^2}{\| \vec{u} \|^2},$$

$$p(x, \beta) \, (1 - p(\beta(x), \alpha))^2 \ \leq \ \frac{\| \vec{v} \|^2}{\| \vec{u} \|^2} \, (1 - \frac{\| \vec{w} \|^2}{\| \vec{v} \|^2}).$$

We conclude that:

$$p(x, \beta) \, (1 - p(\beta(x), \alpha))^2 \ \leq \ p(\beta(x), \alpha) \, (1 - p(\alpha(\beta(x)), \beta)).$$

∎

Theorem 2 is a quantitative strengthening of the **Interference** property of projections in Hilbert spaces that plays a central role in the definition of an M-algebra [3]. Indeed, assuming that $x \in \alpha$, if $\alpha(\beta(x)) \in \beta$, then, by Corollary 2, $p(\alpha(\beta(x)), \beta) = 1$ and by Theorem 2, either $p(x, \beta) = 0$ or $p(\beta(x), \alpha) = 1$. In both cases we have $p(\beta(x), \alpha) = 1$ and, by Corollary 2, $\beta(x) \in \alpha$.

7.5 The quantity $\rho(x, y, z)$

The stage is now set for the definition of the quantity $\rho(x, y, z)$ announced in the principle of superposition. Very naturally $\rho(x, y, z)$ measures the relative similarity of x to y and z respectively.

Definition 4 *For any $x, y, z \in X$ define*

$$\rho(x, y, z) = p(x, y)/(p(x, y) + p(x, z)).$$

The following is obvious by Lemma 3.

Lemma 13 *If $x \not\perp y$ and $x \not\perp z$, then $0 < \rho(x, y, z) < 1$.*

Lemma 14 *For any $x, y, z \in X$ $\rho(x, z, y) = 1 - \rho(x, y, z)$.*

In relation with the principle of superposition, we shall consider quantities $\rho(x, y, z)$ only for states such that $coplanar(x, y, z)$ and $y \not\perp z$, but the definition of ρ above can be used for any triple of states. It is only the definition of the quantity $\theta(x, y, z)$ below that requires that y and z be non-orthogonal.

8 Phases for Triangles: the quantity $\theta(x, y, z)$

We may now proceed to the definition of a second geometric quantity relating three states: $\theta(x, y, z)$. This quantity does not seem to have been studied previously.

In section 7.1 a quantity was attached to any pair of states. This quantity was the modulus of some inner product. It seems natural that the argument of a similar inner product represents another important geometrical quantity. But, clearly some thinking must be done to define, out of such an argument, a quantity that does not depend on the vectors chosen, but only on states. A new quantity, $\theta(x, y, z)$, an angle in the interval $[0, 2\pi]$ will be attached to triples of states. This quantity can be defined only if no two of the three states x, y and z are orthogonal. Under such a condition, if $y \neq z$ and $coplanar(x, y, z)$, the quantity $\theta(x, y, z)$ should be understood in the following way. The state x is a superposition of y and z in which the phase difference is $\theta(x, y, z)$, i.e., $x = ry + (1 - r)e^{i\varphi}z$ for some $r \in]0, 1[$, with $\varphi = \theta(x, y, z)$.

Definition 5 *Let $x, y, z \in X$ be such that $x \not\perp y$, $y \not\perp z$ and $z \not\perp x$. We shall define $\theta(x, y, z)$ in the following way. Choose arbitrary unit-vectors $\vec{u}$, $\vec{v}$ and $\vec{w}$ in x, y and z respectively and let:*

$$\theta(x, y, z) \; = \; \arg(\langle \vec{u}, \vec{v} \rangle) \; + \; \arg(\langle \vec{v}, \vec{w} \rangle) \; + \; \arg(\langle \vec{w}, \vec{u} \rangle).$$

Note that each of those three inner products is different from zero, by assumption, and therefore the three complex arguments are well-defined.

We need to justify the definition by showing that the quantity $\theta(x, y, z)$ depends only on x, y and z and does not depend on the vectors $\vec{u}$, $\vec{v}$ and $\vec{w}$. For example, the definition is independent of the vector $\vec{u}$ chosen in x since any unit-vector $\vec{s}$ of x has the form $\vec{s} = e^{i\varphi}\vec{u}$ for some $\varphi \in [0, 2\pi]$. Had we used $\vec{s}$ instead of $\vec{u}$ we would have obtained:

$$\arg(\langle e^{i\varphi}\vec{u}, \vec{v} \rangle) \; + \; \arg(\langle \vec{v}, \vec{w} \rangle) \; + \; \arg(\langle \vec{w}, e^{i\varphi}\vec{u} \rangle) \; =$$

$$\arg(e^{i\varphi}\langle \vec{u}, \vec{v} \rangle) \; + \; \arg(\langle \vec{v}, \vec{w} \rangle) \; + \; \arg(e^{-i\varphi}\langle \vec{w}, \vec{u} \rangle) \; =$$

$$\varphi + \arg(\langle \vec{u}, \vec{v} \rangle) \; + \; \arg(\langle \vec{v}, \vec{w} \rangle) \; - \; \varphi \; + \; \arg(\langle \vec{w}, \vec{u} \rangle).$$

A similar line shows that the choice of none of $\vec{v}$ or $\vec{w}$ influences $\theta(x, y, z)$.

We shall now prove some properties of θ. First, $\theta(x, y, z)$ is invariant under a circular permutation of the arguments and antisymmetric under transpositions.

Lemma 15 *For any generic states x, y and z, we have: $\theta(y, z, x) = \theta(x, y, z)$, $\theta(x, z, y) = -\theta(x, y, z)$ and $\theta(x, y, w) = \theta(x, y, z) + \theta(x, z, w) + \theta(z, y, w)$.*

Proof: Obvious. ▮

The behavior of θ under (planar) orthogonal complements is also antisymmetric.

Lemma 16 *Assume $x, y, z \in X$ are states no two of them are equal and no two of them are orthogonal and such that coplanar(x, y, z). Let $x' = (\neg x)(y) = (\neg x)(z)$, $y' = (\neg y)(z) = (\neg y)(x)$ and $z' = (\neg z)(x) = (\neg z)(y)$. Then $\theta(x', y', z') = -\theta(x, y, z)$.*

Proof: Choose an arbitrary unit-vector $\vec{u}$ in x. Let $\vec{v}$ be the unit-vector of y such that $\langle \vec{u}, \vec{v} \rangle > 0$. Let $\vec{u}'$ be the unit-vector of x' such that $\langle \vec{v}, \vec{u}' \rangle > 0$. Let us have $\vec{v} = r_1\vec{u} + r_2\vec{u}'$ for positive real numbers $r_i, i = 1, 2$. The vector

$r_2\vec{u} - r_1\vec{u}'$ is a unit-vector in y'. Let $\vec{v}' = r_2\vec{u} - r_1\vec{u}'$. Let $\vec{w}$ be the unit-vector of z such that $\langle \vec{u}, \vec{w} \rangle > 0$. Let $\vec{w} = r_3\vec{u} + r_4 e^{i\varphi}\vec{u}'$ for positive r_i's $i = 3, 4$ and some angle φ. Let $\vec{w}' = r_4 e^{-i\varphi}\vec{u} - r_3\vec{u}'$, a unit vector of z'.

We see that:

$$\theta(x, y, z) \ = \ \arg(\langle \vec{u}, \vec{v} \rangle) \ + \ \arg(\langle \vec{v}, \vec{w} \rangle) \ + \ \arg(\langle \vec{w}, \vec{u} \rangle) \ = \ 0 + \arg(\langle \vec{v}, \vec{w} \rangle) + 0.$$

and

$$\theta(x', y', z') = \arg(\langle \vec{u}', \vec{v}' \rangle) + \arg(\langle \vec{v}', \vec{w}' \rangle) + \arg(\langle \vec{w}', \vec{u}' \rangle) = \pi + \arg(\langle \vec{v}', \vec{w}' \rangle) + \pi.$$

We are left to show that $\arg(\langle \vec{v}', \vec{w}' \rangle) = -\arg(\langle \vec{v}, \vec{w} \rangle)$. In fact, we shall show that $\langle \vec{v}', \vec{w}' \rangle = \langle \vec{w}, \vec{v} \rangle$. Indeed, $\langle \vec{v}', \vec{w}' \rangle = r_2 r_4 e^{i\varphi} + r_1 r_3$ and $\langle \vec{w}, \vec{v} \rangle = r_1 r_3 + r_2 r_4 e^{i\varphi}$. ∎

9 Formal Definition of Superpositions

We shall now present the definition of the superposition $r\, y \ + \ (1 - r)\, e^{i\varphi}\, z$. The definition is not straightforward.

Definition 6 *We shall now define the state $r\, y \ + \ (1 - r)\, e^{i\varphi}\, z$, for any $r \in]0, 1[$, any $\varphi \in [0, 2\,\pi[$, any $y, z \in X$ such that $y \not\perp z$.*

If $y = z$, let $r\, y \ + \ (1 - r)\, e^{i\varphi}\, z = y$.

Suppose now that $y \neq z$. Choose some arbitrary unit-vector $\vec{v}$ in y. Since $y \not\perp z$, there is a unique unit-vector $\vec{w}$ of z such that $\langle \vec{v}, \vec{w} \rangle \ > \ 0$. Let z' be the one-dimensional subspace orthogonal to z in the two-dimensional subspace generated by y and z: $z' \ = \ (\neg z)(y)$. Since $z \neq y$, $z' \not\perp y$. Let, then, $\vec{w}'$ be the unique unit-vector of z' such that $\langle \vec{v}, \vec{w}' \rangle \ > \ 0$. Let y' be the one-dimensional subspace orthogonal to y in the two-dimensional subspace generated by y and z: $y' \ = \ (\neg y)(z)$. Since $y \neq z$, we have $y' \not\perp z$. Let, then, $\vec{v}'$ be the unique unit-vector of y' such that $\langle \vec{w}, \vec{v}' \rangle \ > \ 0$. In summary, the following inner products are positive real numbers: $\langle \vec{v}, \vec{w} \rangle$, $\langle \vec{v}, \vec{w}' \rangle$ and $\langle \vec{v}', \vec{w} \rangle$. Define, now:

$$\vec{u} \ = \ \sqrt{1 - r}\, \vec{v}' \ + \ \sqrt{r}\, e^{i\varphi}\, \vec{w}'. \tag{1}$$

Note that $\vec{u} \neq \vec{0}$ since $\vec{v}'$ and $\vec{w}'$ are linearly independent and one at least (in fact both) of r or $1 - r$ is different from zero. We may now define $r\, y \ + \ (1 - r)\, e^{i\varphi}\, z$ to be the one-dimensional subspace generated by $\vec{u}$.

Note that this definition would not square well with a convention that $1\,y + 0\,e^{i\varphi}\,z = y$, since $\vec{u}$ approaches $\vec{w}'$, i.e., a state orthogonal to z and not the state y for r approaching 1. This is the reason we do not hold such a convention.

Lemma 17 *Under the assumptions of Definition 6, one has* $\langle \vec{v}', \vec{w}' \rangle = -\sqrt{p(y,z)}$.

Proof: We know that $\vec{v} = r_1 \vec{w} + r_2 \vec{w}'$ for some real positive numbers r_1, r_2. The unit vector $\vec{v}'$ is orthogonal to $\vec{v}$ in the two-dimensional subspace spanned by $\vec{w}$ and $\vec{w}'$. Since $langle \vec{v}', \vec{w} \rangle > 0$, we have $\vec{v}' = r_3 \vec{w} + r_4 e^{i\varphi} \vec{w}'$ for positive numbers r_3, r_4 and some angle φ. We have $r_1 r_3 + r_2 r_4 e^{i\varphi} = 0$. Therefore $r_3 = r_2$, $r_4 = r_1$ and $\varphi = \pi$. We see that $\langle \vec{v}', \vec{w}' \rangle = -r_4 = -r_1 = -\langle \vec{v}, \vec{w} \rangle = -\sqrt{p(y,z)}$. ∎

10 Properties of Superpositions

The following summarizes properties of superpositions that were discussed above.

Lemma 18 *For any* $y, z \in X$ *such that* $y \not\perp z$ *and* $y \neq z$ *and for any* $r \in]0,1[$, $\varphi \in [0, 2\pi[$, *if* $x = r\,y \; + \; (1-r)\,e^{i\varphi}\,z$ *then*

1. *coplanar*(x,y,z) *(Principle of Coplanarity),*
2. $x \not\perp y$ *and* $x \not\perp z$,
3. $\rho(x,y,z) = r$, *and*
4. $\theta(x,y,z) = \varphi$.

Proof: Let $\vec{y}'$ and $\vec{z}'$ be the one-dimensional spaces containing $\vec{v}'$ and $\vec{w}'$ respectively. From Equation 1 one sees that $coplanar(x, y', z')$. Since $coplanar(y', y, z)$ and $coplanar(z', y, z)$, we have $coplanar(x, y, z)$.

For (2), note that $\vec{u}$ cannot be co-linear with $\vec{v}'$ since $r > 0$. But $coplanar(x,y,z)$ and $coplanar(y', y, z)$ then implies that $x \not\perp y$. Similarly $x \not\perp z$ since $r < 1$.

Let us evaluate $p(x,y)$. Let $a = \parallel \vec{u} \parallel^2$. We have, by Lemma 3: $p(x,y) = \parallel y(\vec{u}) \parallel^2 \; / \, a$. But $y(\vec{u}) = y(\sqrt{1-r}\,\vec{v'}) + y(\sqrt{r}\,e^{i\varphi}\,\vec{w'}) = \sqrt{r}\,e^{i\varphi}\,y(\vec{w'})$. Therefore $p(x,y) = r \parallel y(\vec{w'}) \parallel^2 \; / \, a = r\,p(z', y) \,/\, a$. Similarly: $p(x,z) = (1-r)\,p(y', z) \,/\, a$. But $p(z', y) = p(y, z') = 1 - p(y,z)$ by Lemma 7 since $z' = (\neg z)(y)$. Similarly: $p(y', z) = 1 - p(z,y)$. We see that $p(x,y) = r\,(1 - p(y,z)) \,/\, a$ and

$p(x,y)+p(x,z) = (r\,(1-p(y,z)) \;+\; (1-r)\,(1-p(y,z)))\,/\,a = (1-p(y,z))\,/\,a$. Therefore $p(x,y)/(p(x,y)+p(x,z)) = r$. We conclude that $\rho(x,y,z) = r$.

To compute $\theta(x,y,z)$, consider the vectors $\vec{u}$, $\vec{v}$ and $\vec{w}$ of x, y and z respectively. Since, by construction, $\arg(\langle \vec{v}, \vec{w} \rangle) = 0$, we have: $\theta(x,y,z) = \arg(\langle \vec{u}, \vec{v} \rangle) + \arg(\langle \vec{w}, \vec{u} \rangle)$. But $\langle \vec{u}, \vec{v} \rangle = \langle \sqrt{1-r}\, e^{i\varphi}\, \vec{w'}, \vec{v} \rangle = \sqrt{1-r}\, e^{i\varphi} \langle \vec{w'}, \vec{v} \rangle$. Since this last inner product is strictly positive, as is $\sqrt{1-r}$, $\arg(\langle \vec{u}, \vec{v} \rangle) = \varphi$. We have: $\langle \vec{w}, \vec{u} \rangle = \langle \vec{w}, \sqrt{r}\, \vec{v'} \rangle > 0$. Therefore $\arg(\langle \vec{w}, \vec{u} \rangle) = 0$. ▮

Lemma 19 *Let $x, y, z \in X$ such that $y \neq z$, $x \not\perp y$, $y \not\perp z$, $z \not\perp x$ and such that $coplanar(x,y,z)$, then $x = \rho(x,y,z)\, y \;+\; (1-\rho(x,y,z))\, e^{i\theta(x,y,z)}\, z$.*

Proof: Choose $\vec{v}$, $\vec{v'}$, $\vec{w}$ and $\vec{w'}$ such as in Definition 6: unit vectors, $\vec{v} \in y$, $\vec{w} \in z$, $y' = (\neg y)(z)$, $z' = (\neg z)(y)$. Take $\vec{t}$ to be any non-zero vector of x. Since $coplanar(x, y', z')$ and $x \neq y'$, $x \neq z'$, there are real numbers $r_i \in]0,1[$, $\psi_i \in [0, 2\,\pi[$, for $i = 1,2$ such that $\vec{t} = r_1\, e^{i\psi_1}\, \vec{v'} \;+\; r_2\, e^{i\psi_2}\, \vec{w'}$. The vector $\vec{s} = 1/\sqrt{r_1^2 + r_2^2}\, e^{-i\psi_1}\, \vec{t}$ is a non-zero vector of x. But, if we let $r = r_2^2/\sqrt{r_1^2 + r_2^2}$, we have: $\vec{s} = \sqrt{1-r}\, \vec{v'} \;+\; \sqrt{r}\, e^{i(\psi_2 - \psi_1)}\, \vec{w'}$ and therefore $x = r\, y \;+\; (1-r)\, e^{i(\psi_2-\psi_1)}\, z$. By Lemma 18, now, $r = p(x,y) + (p(x,y) + p(x,z))$ and $\psi_2 - \psi_1 = \theta(x,y,z)$.
▮

The following shows that any state is a non-trivial superposition of itself and any other, non-orthogonal, state.

Corollary 8 *Let $y, z \in X$ such that $y \not\perp z$. Then $y = ry + (1-r)z$ for $r = 1/(p(y,z)+1)$.*

Proof: If $y = z$, by the definition of trivial superpositions. If $y \neq z$, by Lemma 19 since $\rho(y,y,z) = 1/(1+p(y,z))$ and $\theta(y,y,z) = 0$. ▮

We shall now evaluate $p(r\, y \;+\; (1-r)\, e^{i\varphi}\, z, x)$. First, we need to evaluate the norm of the vector $\vec{u}$ of equation (1).

$$\parallel \vec{u} \parallel^2 = \langle \sqrt{1-r}\, \vec{v'} \;+\; \sqrt{r}\, e^{i\varphi}\, \vec{w'}, \sqrt{1-r}\, \vec{v'} \;+\; \sqrt{r}\, e^{i\varphi}\, \vec{w'} \rangle =$$

$$(1-r) + \sqrt{r(1-r)} e^{-i\varphi} \langle \vec{v'}, \vec{w'} \rangle + \sqrt{r(1-r)} e^{i\varphi} \langle \vec{w'}, \vec{v'} \rangle + r$$

But, by construction there are $r_1, r_2 > 0$ such that $\vec{v} = r_1 \vec{w} + r_2 \vec{w'}$. Since $\langle \vec{v'}, \vec{w} \rangle > 0$, we have $\vec{v'} = r_2 \vec{w} - r_1 \vec{w'}$. Therefore $\langle \vec{v'}, \vec{w'} \rangle = -r_1 = -\sqrt{p(y,z)}$. We conclude that:

$$\parallel \vec{u} \parallel^2 = 1 - 2\cos\varphi \sqrt{r(1-r)p(y,z)}.$$

Notation: $\omega(r,y,z,\varphi) = 1/\sqrt{1 - 2\cos\varphi\sqrt{r(1-r)p(y,z)}}$.

Lemma 20 *For any* $x, y, z \in X$, $y \not\perp z$, *for any* $r \in]0,1[$, $\varphi \in [0, 2\pi[$, *we have:*

$$p(r\, y \; + \; (1-r)\, e^{i\varphi}\, z, x) =$$
$$\omega(r, y, z, \varphi)^2 \mid (\sqrt{(1-r)(1-p(x,y))} + \sqrt{r(1-p(x,z))}e^{i(\varphi - \theta(x,y',z'))} \mid^2 .$$

Proof: We shall write ω for $\omega(r, y, z, \varphi)$. Let $\vec{u} = \sqrt{1-r}\vec{v}' + \sqrt{r}e^{i\varphi}\vec{w}'$. The vector $\omega\,\vec{u}$ is a unit-vector of $r\,y \; + \; (1-r)\,e^{i\varphi}\,z$. Let $\vec{t}$ be the unique unit vector of x such that $\langle \vec{v}', \vec{t} \rangle > 0$. We have: $\langle \omega\vec{u}, \vec{t} \rangle = \omega(\sqrt{1-r}\langle \vec{v}', \vec{t} \rangle + \sqrt{r}e^{i\varphi}\langle \vec{w}', \vec{t} \rangle)$. But $\langle \vec{v}', \vec{t} \rangle = \sqrt{p(y', x)}$. Also

$$\theta(x, y', z') = \arg(\langle \vec{t}, \vec{v}' \rangle) + \arg(\langle \vec{v}', \vec{w}' \rangle) + \arg(\langle \vec{w}', \vec{t} \rangle)$$

and $\langle \vec{v}', \vec{w}' \rangle = -\sqrt{p(y,z)}$ by Lemma 17. Therefore

$$\theta(x, y', z') = \pi + \arg(\langle \vec{w}', \vec{t} \rangle) = -\arg(\langle \vec{w}', \vec{t} \rangle).$$

We conclude that:

$$p(r\, y \; + \; (1-r)\, e^{i\varphi}\, z, x) =$$
$$\mid \omega(\sqrt{(1-r)(1-p(x,y))} + \sqrt{r(1-p(x,z))}e^{i(\varphi - \theta(x,y',z'))}) \mid^2 .$$

∎

The expression found could be more elegant if we knew how to express $\theta(x, y', z')$ in terms of $\theta(x, y, z)$. Unfortunately simple considerations show that $\theta(x, y', z')$ is not determined by $\theta(x, y, z)$. In view of Lemma 16 it is enough to show that $\theta(x', y, z)$ is not determined by $\theta(x, y, z)$. Consider $\vec{v}$ and $\vec{v}'$ orthogonal unit vectors. Then $\vec{u} = 1/\sqrt{2}(\vec{v} + \vec{v}')$ is a unit vector. The vector $\vec{u}' = 1/\sqrt{2}(\vec{v} - \vec{v}')$ is a unit vector orthogonal to $\vec{u}$. Let $\vec{w} = r_1\vec{v} + r_2\vec{v}'$ be a unit vector with $r_1, r_2 > 0$. If we take for x, y and z the one-dimensional subspaces spanned by $\vec{u}$, $\vec{v}$ and $\vec{w}$ respectively, we have $\theta(x, y, z) = 0$. We also have, using $\vec{u}'$, $\theta(x', y, z) = \arg(1/\sqrt{2}(r_1 - r_2))$. Take, first, $r_1 > r_2$, we then have $\theta(x, y, z) = 0 = \theta(x', y, z)$. But if we take $r_1 < r_2$, we have $\theta(x, y, z) = 0$ and $\theta(x', y, z) = \pi$. We see that $\theta(x', y, z)$ is not determined by $\theta(x, y, z)$.

11 Superp-structures and their mappings

It is a thesis of this paper that the structure of superpositions is the fundamental structure of Hilbert spaces that is meaningful for Quantum Physics.

To support this thesis one should, now, analyze the fundamental constructions used in Quantum Physics, such as tensor products and quotients as universal, i.e., categorical constructions in the category of superposition preserving mappings. Such an analysis has not been performed yet. One preliminary step must be the proper definition of the category of superposition structures and their superposition preserving mappings. This paper does not provide for a proper definition of such a category, whose objects must include both structures defined by Hilbert spaces, studied here, and classical structures in which any two distinct states are orthogonal, and all structures in-between.

We shall, therefore, consider only superposition structures defined by some Hilbert space. A more general definition abstracting from Hilbert spaces and based on the properties of the quantities ρ and θ is left for future work.

Definition 7 *For any Hilbert space $\mathcal{H}$ on the complex field, the set of its one-dimensional subspaces X together with the operation of superposition associating an element $r\,y \,+\, (1-r)\,e^{i\varphi}\,z$ of X with any quadruple $y, z \in X$, $r \in]0,1[$ and any $\varphi \in [0, 2\,\pi[$ such that $y \not\perp z$ is called a* superp-structure. *We shall, in the sequel, and without harm, denote by X the superp-structure above, forgetting to give an explicit name to the superposition operation. A* superp-morphism *f between two superp-structures X and Y is a function $f : X \longrightarrow Y$ that preserves superpositions, i.e., such that for any $y, z \in X$, such that $y \not\perp z$ and for any $r \in]0,1[$ and any $\varphi \in [0, 2\,\pi[$ the superposition, in Y, $r\,f(y) \,+\, (1-r)\,e^{i\varphi}\,f(z)$ is defined, i.e., $f(y) \not\perp f(z)$ and is equal to $f(r\,y \,+\, (1-r)\,e^{i\varphi}\,z)$.*

This notion of a superp-morphism is original. Note that if $f : X \rightarrow Y$ preserves superpositions and $x \not\perp y$ then $f(x) \not\perp f(y)$ since the superpositions $r\,f(x) \,+\, (1-r)\,e^{i\varphi}\,f(y)$ must be defined.

We shall now present some preliminary results concerning superp-morphisms. Our first two results are properties of all superp-morphisms. First, a superp-morphism preserves ρ and θ for suitable triples.

Lemma 21 *Let $f : X_1 \rightarrow X_2$ is a superp-morphism. For any $y, z \in X_1$ such that $y \not\perp z$ and $f(y) \neq f(z)$, for any $r \in]0,1[$ and any $\varphi \in [0, 2\,\pi[$ we have $\rho(f(r\,y \,+\, (1-r)\,e^{i\varphi}\,z), f(y), f(z)) = r$ and $\theta(f(r\,y \,+\, (1-r)\,e^{i\varphi}\,z), f(y), f(z)) = \varphi$.*

Note that $f(y) \neq f(z)$ cannot be weakened to $y \neq z$.
Proof: Since f preserves superpositions: $f(r\,y \,+\, (1-r)\,e^{i\varphi}\,z) = r\,f(y) \,+\, (1-r)\,e^{i\varphi}\,f(z)$ and the claims obtain from Lemma 18. ▮

As a consequence, if f is a superp-morphism and x, y are not orthogonal, then either they have the same image under f or f preserves their p.

Lemma 22 *If $f : X_1 \rightarrow X_2$ is a superp-morphism, then, for any $x, y \in X_1$ such that $x \not\perp y$ and $f(x) \neq f(y)$ we have $p(f(x), f(y)) = p(x, y)$.*

Proof: Assume f preserves superpositions, $x, y \in X_1$ and $x \not\perp y$. By Corollary 8, $x = rx + (1-r)y$ for $r = 1/(p(x, y) + 1)$. By Lemma 21, now, $\rho(f(x), f(x), f(y)) = 1/(p(x, y) + 1)$. But, by Definition,

$$\rho(f(x), f(x), f(y)) = p(f(x), f(x))/(p(f(x), f(x)) + p(f(x), f(y)) = 1/(1 + p(f(x), f(y))).$$

Therefore $p(f(x), f(y)) = p(x, y)$. ▮
The author does not know whether the assumption $x \not\perp y$ is necessary in Lemma 22. Lemma 25 will show that, when superp-morphism f is obtained from a linear mapping of Hilbert spaces, then the condition is not necessary.

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces and let X_1 and X_2 be the superp-structures corresponding to $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively. A natural way to obtain a superp-morphism $f : X_1 \rightarrow X_2$ is to start from a linear map $m : \mathcal{H}_1 \rightarrow \mathcal{H}_2$. Such a map m associates, with every one-dimensional subspace of X_1, i.e., every member of X_1, a subspace of X_2 that is either one-dimensional or zero-dimensional. Any injective, i.e., left-invertible, linear map m provides an application $f_m : X_1 \rightarrow X_2$ defined by: $f_m(x)$ is the subspace generated by $m(\vec{u})$, for any non-null $\vec{u} \in x$.

Definition 8 *A superp-morphism obtained from an injective linear mapping between Hilbert spaces in the way described just above will be called* regular.

Note that not all superp-morphisms are regular. Let $\mathcal{H}_1$ be any Hilbert space of dimension two at least and $\mathcal{H}_2$ be a one-dimensional Hilbert space. The superp-structure X_2 has one element only. The unique mapping $f : X_1 \rightarrow X_2$ is obviously a superp-morphism. But no linear map $m : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ is injective and therefore no linear map defines f.

The regular superp-morphisms will now be fully characterized.

Lemma 23 *Let H_1, H_2 be Hilbert spaces. If $f : H_1 \rightarrow H_2$ is a linear function, the following three conditions are equivalent. A linear function f satisfying them is called a* near-isometry.

1. *f is injective and for any $\vec{u}, \vec{v} \in H_1$, if $\vec{u} \perp \vec{v}$, then $f(\vec{u}) \perp f(\vec{v})$),*

2. *there is some real number $r > 0$ such that, for any $\vec{u} \in H_1$ one has: $\| f(\vec{u}) \| = r \| \vec{u} \|$,*

3. *there is some real number $r > 0$ such that, for any $\vec{u}, \vec{v} \in H_1$ one has: $\langle f(\vec{u}), f(\vec{v}) \rangle = r \langle \vec{u}, \vec{v} \rangle$.*

Proof: Assume 1), we shall prove 2. Let $\vec{u}_i$ be an orthonormal basis for $\mathcal{H}_1$. Since f is injective $\| f(\vec{u}_i) \| > 0$ and since f preserves orthogonality, the family $f(\vec{u}_i) / \| f(\vec{u}_i) \|$ is an orthonormal basis for $f(\mathcal{H}_1)$. Let $r = \| f(\vec{u}_1) \|$. It is enough to prove that for any i, $\| f(\vec{u}_i) \| = r$. Take any complex numbers c_1, c_2 different from zero. We have: $c_1 \vec{u}_1 + c_2 \vec{u}_i \perp \bar{c_2} \vec{u} - \bar{c_1} \vec{u}_i$. Since f is linear and preserves orthogonality, we have: $c_1 f(\vec{u}_1) + c_2 f(\vec{u}_i) \perp \bar{c_2} f(\vec{u}) - \bar{c_1} f(\vec{u}_i)$. Since $f(\vec{u}) \perp f(\vec{u}_i)$ we have: $c_1 c_2 (\| f(\vec{u}) \| - \| f(\vec{u}_i) \| = 0$ and $\| f(\vec{u}_i) \| = r$.

Assume now that 2) holds. The fact that 3) holds follows immediately from the fact that scalar products in Hilbert spaces can be evaluated from norms. The real part of $\langle \vec{u}, \vec{v} \rangle$ is equal to a fourth of $\| \vec{u} + \vec{v} \|^2 - \| \vec{u} - \vec{v} \|^2$ and its imaginary part is equal to a fourth of $i(\| \vec{u} + i\vec{v} \|^2 - \| \vec{u} - i\vec{v} \|^2)$. It follows that 2 implies for any $\vec{u}, \vec{v} \in H_1$ one has: $\langle f(\vec{u}), f(\vec{v}) \rangle = r^2 \langle \vec{u}, \vec{v} \rangle$.

Assume 3. The function f is injective since $f \; \| f(\vec{u}) \| = \sqrt{r} \| \vec{u} \|$. It obviously preserves orthogonality. ∎

We shall now show that any near-isometry provides a superp-morphism.

Lemma 24 *Let $m : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ be a near-isometry. Then f_m is a superp-morphism of X_1 into X_2.*

In particular any unitary mapping is a superp-morphism. But, note that even not all isometries are self-adjoint.

Proof: Assume $\langle m(\vec{u}, m(\vec{v}) \rangle = r \langle \vec{u}, \vec{v} \rangle$ for every $\vec{u}, \vec{v} \in \mathcal{H}_1$. Since a near-isometry is injective, the function f_m is well-defined.

Now, for any $x, y, z \in X_1$ $a(f_m(x), f_m(y)) = a(x, y)$, $p(f_m(x), f_m(y)) = p(x, y)$, $\rho(f_m(x), f_m(y), f_m(z)) = \rho(x, y, z)$, and $\theta(f_m(x), f_m(y), f_m(z)) = \theta(x, y, z)$. Assume now that $x \not\perp y$ and $z = rx + (1-r)e^{i\varphi} y$. We have $f_m(x) \not\perp f_m(y)$ and therefore the superposition $r f_m(x) + (1-r)e^{i\varphi} f_m(y)$ is defined. But, by the Principle of Uniqueness this last superposition is the unique element w of X_2 that is coplanar with $f_m(x)$ and $f_m(y)$ and such that $\rho(w, f_m(x), f_m(y)) = r$. and $\theta(w, f_m(x), f_m(y)) = \varphi$. But, since m is linear, we have $coplanar(f_m(z), f_m(x), f_m(y))$ and, by the above and Lemma 18 we have: $\rho(f(z), f(x), f(y)) = \rho(z, x, y) = r$

and $\theta(f(z), f(x), f(y)) = \theta(z, x, y) = \varphi$. We conclude that $f_m(rx + (1-r)e^{i\varphi}y) = rf_m(x) + (1-r)e^{i\varphi}f_m(y)$. ∎

We now move to prove that if f_m is any regular superp-morphism, then m is a near-isometry. The following lemma shows that Lemma 22 may be strengthened for regular superp-morphisms.

Lemma 25 *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces and let X_1 and X_2 be the superp-structures corresponding to $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively. Assume $m : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ is an injective linear mapping and that $f_m : X_1 \rightarrow X_2$ is a (regular) superp-morphism. For any $x, y \in X_1$ $p(f(x), f(y)) = p(x, y)$.*

Proof: If $x \not\perp y$, we conclude by Lemma 22. Assume, now, that $p(x, y) = 0$. We may assume $x \neq y$. Since m is injective $f(x) \neq f(y)$. By Lemma 7, for any $z \in X_1$, such that *coplanar*(z, x, y), we have $p(z, x) + p(z, y) = 1$. Now, if $z \neq x$ and $z \neq y$ we have $z \not\perp y$ and $z \not\perp x$ and we conclude by Lemma 22 that $p(f(z), f(x)) + p(f(z), f(y)) = 1$.

But, since f is regular, any state coplanar, in X_2, with $f(x)$ and $f(y)$ is the image, by f, of some z coplanar, in X_1, of x and y.

We conclude that, for any $w \in X_2$ such that *coplanar*$(w, f(x), f(y))$ we have $p(w, f(x)) + p(w, f(y)) = 1$. This implies that $f(x) \perp f(y)$. ∎

We can now fully characterize regular superp-morphisms.

Theorem 3 *Let $m : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ be any linear injective mapping. The function $f_m : X_1 \rightarrow X_2$ is a (regular) superp-morphism iff m is a near-isometry.*

Proof: The *if* part is Lemma 24. The *only if* part follows from Lemmas 25 and 23. ∎

12 Conclusion and future work

We have shown that the properties of superpositions are governed by two geometrical quantities ρ and θ defined for triples of one-dimensional subspaces in a Hilbert space, thus moving forward John von Neumann's program of focusing on subspaces and not on vectors.

The most pressing task is probably now to provide an abstract definition of superp-structures, generalizing those structures provided by Hilbert spaces.

A quantic system composed of two sub-systems is represented by the tensor product of the Hilbert spaces representing the two sub-systems. An alternative, better grounded, description could be that such a quantic system is represented by all superpositions of pairs consisting of a state of the first sub-system and a state of the second sub-system. Such a definition equips the tensor product with an algebraic structure that has not been studied so far.

13 Acknowledgements

I am most grateful to Kurt Engesser and Dov Gabbay for extremely fruitful discussions during the elaboration of this paper. I thank Dorit Aharonov and Jean-Marc Lévy-Leblond for their interest and help.

References

[1] Garret Birkhoff and John von Neumann. The logic of quantum mechanics. *Annals of Mathematics*, 37:823–843, 1936.

[2] Paul R. Halmos. *Introduction to Hilbert Space and the Theory of Spectral Multiplicity*. Chelsea Publishing Company, New York 1, NY, 1951.

[3] Daniel Lehmann, Kurt Engesser, and Dov M. Gabbay. Algebras of measurements: the logical structure of quantum mechanics. *International Journal of Theoretical Physics*, 45(4):698–723, April 2006. DOI 10.1007/s10773-006-9062-y.

[4] Miklós Rédei, editor. *John von Neumann: Selected Letters*, volume 27 of *History of Mathematics*. London Mathematical Society - American Mathematical Society, 2005.