Zeitschrift: Helvetica Physica Acta
Band: 58 (1985)
Heft: 6

Artikel: Superpositions of physical states : a metric viewpoint
Autor: Cantoni, Vittorio
DOI: https://doi.org/10.5169/seals-115628

Helvetica Physica Acta, Vol. 58 (1985) 956–968 0018–0238/85/060956–13$1.50 + 0.20/0

Superpositions of physical states: a metric viewpoint

By Vittorio Cantoni

Dipartimento di Matematica, Università di Milano, via Saldini 50, 20133 Milano, Italy.

(16. IV. 1985)

Abstract. Superpositions of states are defined in terms of a natural metric possessed by the state space of any physical system. The new concept does not presuppose a linear structure connected with the state space, and proves to be useful in a characterization of the complex separable projective Hilbert space.

1. Introduction

In setting about promoting to the rank of fundamental principle the occurrence, in quantum physics, of that special relation of a state to others ever since referred to as *superposition*, Dirac (1930) gave (among other clarifications) the following qualitative description of the new concept ([1] p. 13):

> "When a state is formed by the superposition of two states, it will have properties that are in some vague way intermediate between those of the original states and that approach more or less closely to those of either of them according to the greater or less 'weight' attached to the state in the superposition process. The new state is completely defined by the two original states when their relative weights in the superposition process are known, together with a certain phase difference, the general meaning of weights and phases being provided in the general case by the mathematical theory."

In later pages of his book, Dirac indicated the existence of superpositions as a suggestion for the introduction of a linear space connected with the physical states ([1], p. 15):

> "The superposition process is a kind of additive process and implies that states can in some way be added to give new states. The states must therefore be connected with mathematical quantities of a kind which can be added together to give quantities of the same kind."

Next came the development of the full mathematical scheme, which provided a precise, quantitative specification of the admittedly 'vague' original notion.

Thus, in Dirac's classic approach to quantum physics the qualitative aspects of superpositions suggest the plausibility of an underlying linear structure, which, in turn (together with an inner product) appears indispensable for establishing precisely in what sense, and to what extent, a state is a superposition of others.

Owing to its central role in the conceptual framework, as well as in the implementation of quantum theories, considerable attention has been devoted to the notion of superposition in the various axiomatic approaches to Quantum Mechanics and its modifications and generalizations [3 to 13]. While all definitions of superposition so far proposed seem to take into due account the *qualitative* a priori requirement that for any superposition σ of a set of states $\{\alpha_i\}$ and for every observable the values occurring on σ should also occur on some state α_i of the set, *quantitative* relations involving the statistical distributions of the observed values usually appear only *after* the introduction of an underlying linear structure, postulated or derived in some way from the axioms. The only exception known to the author is the approach due to Deĥiyannis [8], where such quantitative relations are indeed assumed beforehand, though designed to reconstruct the projective Hilbert space model for the set of pure states, so that by assuming them (together with a suitable set of other axioms) the underlying linear strucure is also, in fact, implicitly assumed.

We propose a definition of the concept of superposition of states that makes sense in any model of a physical system, provided that the model involves a set $\mathscr{S}$ representing the states, a set $\mathcal{O}$ representing the observables, a function $p(A, \alpha, E)$ interpretable as the probability that a measurement of the observable A on the state α will give a result in the Borel set E of the real numbers $\mathbb{R}$, and therefore satisfies the very natural and unrestrictive axioms of Mackey [14]:

I. $p(A, \alpha, E) \in [0, 1]$, $p(A, \alpha, \mathbb{R}) = 1$, $p(A, \alpha, \varnothing) = 0$, $p(A, \alpha, \bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} p(A, \alpha, E_i)$, where $\varnothing$ denotes the empty set, $[0, 1]$ the closed unit interval, and the last equations must hold for any countable family $\{E_i\}$ of pairwise disjoint elements of $\mathscr{B}(\mathbb{R})$ (the Borel sets of $\mathbb{R}$);

II. $p(A, \alpha, E) = p(A', \alpha, E)$ for all α and all E implies $A = A'$, $p(A, \alpha, E) = p(A, \alpha', E)$ for all A and E implies $\alpha = \alpha'$.

Our definition is based on a metric structure which can always be derived in $\mathscr{S}$ from the probability function p. It is *quantitative*, because it rests on equalities involving the distance function. When applied to classical systems it characterizes the mixtures of pure states, and produces the relative weights. When applied to the pure states of a quantum mechanical system, it agrees with the usual definition and produces the relative weights and phases. In full generality, it neither presupposes nor implies the existence of a linear space related to the states; however, a purely metric characterization of the projective complex separable Hilbert space is possible on special assumptions on the subsets of the state space that are "closed under superposition" in a well-defined sense. This, in principle, provides a criterion by which to establish, by direct inspection of the statistical distributions of its observables, whether any given physical system may admit a quantum-mechanical description.

In Section 2 we summarize some preliminary concepts from Distance Geometry to be used later on, following Menger, Birkhoff, Blumenthal, Busemann and Wang [15 to 19]. In Section 3 the metric on which our considerations rely is derived from the *generalized transition probability* function T, introduced in ref. [20] (see also [21 to 24]). In Section 4 the proposed definition of *superposition* is given, and some of its features are briefly analyzed in the most general context. Sections 5 and 6 examine the concept within the frameworks of classical systems and of Quantum Mechanics, where it is found to possess the properties one should expect. Section 7 notes that a purely metric characterization of finite-dimensional complex projective spaces is directly provided by a remarkable theorem of Wang [19], on a suitable additional assumption concerning the superpositions of two states. By modifying some of Wang's hypotheses a similar characterization is obtained for the infinite-dimensional separable complex projective Hilbert space.

2. Basic concepts and results from Distance Geometry

Let M be a metric space, and let us denote by xy the distance between two of its points x and y (so that $xy \geqslant 0$, $xy = 0$ if and only if $x = y$, $xy = yx$ and $xz \leqslant xy + yz$).

The point y is said to lie *between* x and z if and only if x, y and z are distinct and the relation $xz = xy + yz$ is satisfied (so that for the three points, in a suitable order, the triangle inequality holds, in fact, as an equality). Such an occurrence is denoted by the symbol xyz. We shall also introduce the symbol $[xyz]$ to denote a situation in which one of the distinct points x, y and z (no matter which) lies between the other two. If $[xyz]$ is verified, we shall say that the three points are *aligned.*

A subset of M is said to be *convex* (in Menger's sense) if it contains, together with any two distinct points x and z, a point y between them (so that xyz, and therefore $[xyz]$, are true).

A *segment* is a subset of M which is isometric with a segment of the real line. It can be proved that if a subset of M is complete (in the usual sense of convergence of Cauchy sequences) and convex, with any two distinct points it contains a segment which joins them (Blumenthal [17] p. 41).

A metric space is said to be *two-point homogeneous* if, given any two pairs (x_1, y_1) and (x_2, y_2) of its points such that $x_1y_1 = x_2y_2$, there is an isometry of the space carrying x_1 into x_2 and y_1 into y_2. Wang [19] has proved that if a metric space is convex, two-point homogeneous and compact, it can only be spherical, elliptic, complex elliptic, quaternion elliptic or the Cayley elliptic plane. (See Busemann [18] for an alternative proof and a detailed description of the spaces involved).

We shall also use the following additional concepts: a subset U of the metric space M will be called *closed under alignment* (briefly, *a-closed*) if and only if, together with any two distinct points x and y it contains every point of M aligned with them (so that $x, y \in U$ and $[xyz]$ imply $z \in U$). The subsets consisting of a single point will also be regarded as a-closed.

Let Σ be any nonempty subset of M. Since the whole space M is a-closed, there exist a-closed subsets of M containing Σ, and it is clear that their intersection is a-closed and contains Σ. Thus to every subset Σ of M there is associated a smallest a-closed subset containing it, which will be called the *envelope* of Σ, and denoted by $\tilde{\Sigma}$. Notice that $\tilde{\Sigma}$ is in general distinct from the "convex extension", or "convex hull", of Σ (Blumenthal [17] p. 51).

3. The metric

Consider any system $\{\mathscr{S}, \mathcal{O}, p\}$ consisting of a state space $\mathscr{S}$, a set $\mathcal{O}$ of observables, a probability function p, and satisfying the axioms I and II of Section

1. In [20] we have equipped $\mathscr{S}$ with a two-point function $T(\alpha, \beta)$, $(\alpha, \beta \in \mathscr{S})$, defined by

$$T(\alpha, \beta) = \inf_{A \in \mathcal{O}} T_A(\alpha, \beta), \tag{1}$$

where

$$T_A(\alpha, \beta) \equiv \left| \int_{\mathbb{R}} d\sqrt{\alpha_A \beta_A} \right|^2. \tag{2}$$

In (2) α_A is the probability measure on $\mathbb{R}$ defined by $\int_E d\alpha_A \equiv p(A, \alpha, E)$, β_A is defined analogously and $\sqrt{\alpha_A \beta_A}$ denotes the measure on $\mathbb{R}$ defined by

$$\int_E d\sqrt{\alpha_A \beta_A} \equiv \int_E \sqrt{\frac{d\alpha_A}{d\sigma} \frac{d\beta_A}{d\sigma}}\, d\sigma, \tag{3}$$

where σ is any finite measure on $\mathbb{R}$ with respect to which α_A and β_A are absolutely continuous, and $d\alpha_A/d\sigma$, $d\beta_A/d\sigma$ are Radon–Nikodym derivatives. The definition of T and the axioms imply:

$$0 \leqslant T(\alpha, \beta) \leqslant 1, \tag{4}$$

$$T(\alpha, \beta) = T(\beta, \alpha), \tag{5}$$

$$T(\alpha, \beta) = 1 \quad \text{if and only if} \quad \alpha = \beta. \tag{6}$$

We now define the *distance* $\alpha\beta$ between two states α and β of $\mathscr{S}$ by setting

$$\alpha\beta = 2 \operatorname{Arc\,cos} \sqrt{T(\alpha, \beta)}. \tag{7}$$

The two-point function $\alpha\beta$ is non-negative on account of (4), symmetric on account of (6) and (4) one has $\alpha\beta = 0$ if and only if $\alpha = \beta$. (4) also implies that $\alpha\beta$ is bounded by π:

$$0 \leqslant \alpha\beta \leqslant \pi. \tag{8}$$

We now show that the distance defined by (7) is indeed a metric by proving the triangle inequality

$$\alpha\gamma \leqslant \alpha\beta + \beta\gamma \qquad (\alpha, \beta, \gamma \in \mathscr{S}) \tag{9}$$

Proof of the triangle inequality

Given any observable $A \in \mathcal{O}$, define the two-point function $(\alpha\beta)_A$ by setting

$$(\alpha\beta)_A = 2 \operatorname{Arc\,cos} \sqrt{T_A(\alpha, \beta)} \tag{10}$$

(see (2)). Since $\sqrt{T_A(\alpha, \beta)}$ is the (positive definite) scalar product of the functions $\sqrt{d\alpha_A/\alpha\sigma}$ and $\sqrt{d\beta_A/d\sigma}$ in the real Hilbert space associated with the measure σ of $\mathbb{R}$ (see (3)), for any triple of states α, β, γ its restriction to the (at most three-dimensional) subspace generated by $\sqrt{d\alpha_A/d\alpha}$, $\sqrt{d\beta_A/d\sigma}$ and $\sqrt{d\gamma_A/d\sigma}$ is positive definite, and this implies

$$\begin{vmatrix} \sqrt{T_A(\alpha,\alpha)} & \sqrt{T_A(\alpha,\beta)} & \sqrt{T_A(\alpha,\gamma)} \\ \sqrt{T_A(\beta,\alpha)} & \sqrt{T_A(\beta,\beta)} & \sqrt{T_A(\beta,\gamma)} \\ \sqrt{T_A(\gamma,\alpha)} & \sqrt{T_A(\gamma,\beta)} & \sqrt{T_A(\gamma,\gamma)} \end{vmatrix}$$

$$\equiv \begin{vmatrix} 1 & \cos\frac{(\alpha\beta)_A}{2} & \cos\frac{(\alpha\gamma)_A}{2} \\ \cos\frac{(\beta\alpha)_A}{2} & 1 & \cos\frac{(\beta\gamma)_A}{2} \\ \cos\frac{(\gamma\alpha)_A}{2} & \cos\frac{(\gamma\beta)_A}{2} & 1 \end{vmatrix} \geqslant 0.$$

(The diagonal elements are equal to 1 on account of (2), (3) and of the fact that α_A, β_A and γ_A are probability measures).

By developing the determinant and factorizing it conveniently the last inequality can be rewritten as follows

$$\left[\cos\frac{(\alpha\beta)_A}{2}-\cos\frac{(\beta\gamma)_A+(\gamma\alpha)_A}{2}\right]\left[\cos\frac{(\beta\gamma)_A-(\gamma\alpha)_A}{2}-\cos\frac{(\alpha\beta)_A}{2}\right]\geqslant 0. \tag{11}$$

Since the two-point function $(\alpha\beta)_A$ is always non-negative and not greater than π, the arguments of the cosines in the first bracket fall in the interval $[0, \pi]$, where the function cos is decreasing, so that

$$\frac{(\alpha\beta)_A}{2}\leqslant\frac{(\beta\gamma)_A+(\gamma\alpha)_A}{2}$$

provided that the second bracket is non-negative. Indeed the second bracket cannot be negative, for this would imply

$$\frac{(\alpha\beta)_A}{2}<\left|\frac{(\beta\gamma)_A-(\gamma\alpha)_A}{2}\right|,$$

and a fortiori

$$\frac{(\alpha\beta)_A}{2}<\frac{(\beta\gamma)_A+(\gamma\alpha)_A}{2},$$

while on account of (11) the first bracket would have to be non-positive, implying

$$\frac{(\alpha\beta)_A}{2}\geqslant\frac{(\beta\gamma)_A+(\gamma\alpha)_A}{2},$$

a contradiction.

Thus for every observable A the function $(\alpha\beta)_A$ satisfies the triangle inequality

$$(\alpha\beta)_A\leqslant(\alpha\gamma)_A+(\gamma\beta)_A. \tag{12}$$

(So that $(\alpha\beta)_A$ is a pseudo-metric.) Consequently

$$\sup_{A\in\mathcal{O}}(\alpha\beta)_A\leqslant\sup_{A\in\mathcal{O}}(\beta\gamma)_A+\sup_{A\in\mathcal{O}}(\gamma\beta)_A,$$

and since

$$\alpha\beta = 2 \text{ Arc cos } \sqrt{T(\alpha, \beta)} = 2 \text{ Arc cos } \inf_{A \in \mathcal{O}} \sqrt{T_A(\alpha, \beta)}$$
$$= \sup_{A \in \mathcal{O}} 2 \text{ Arc cos } \sqrt{T_A(\alpha, \beta)} = \sup_{A \in \mathcal{O}} (\alpha\beta)_A,$$

the triangle inequality is established for our distance.

4. Superpositions

Since the state space $\mathcal{S}$, with the metric defined in the previous section, is a metric space, we can apply to it the terminology and notations introduced in Section 2.

We shall say that a state γ is a *strong superposition* (or, briefly, a *superposition*) of two states α and β if and only if γ lies between α and β.

We shall also say that a state γ is a *weak superposition* of two states α and β if and only if it belongs to their envelope $\{\widetilde{\alpha, \beta}\}$. Thus every (strong) superposition is also a weak superposition, but the converse is not necessarily true.

More generally, γ will be called a weak superposition of the states of a given subset Σ of $\mathcal{S}$ if and only if it belongs to the a-closure $\tilde{\Sigma}$ of Σ. The subsets of $\mathcal{S}$ which are a-closed will also be qualified as *closed under superposition.*

We start by illustrating the relation between these notions on a simple example, which will turn out to be particularly relevant in relation with Quantum Mechanics.

Assume that $\mathcal{S}$ is isometric with the 2-sphere of curvature 1, with the usual metric given by the shortest arclength between two points.

Suppose, first, that α and β are distinct and not diametral (i.e. $0 < \alpha\beta < \pi$). The strong superpositions of α and β are the points of the smallest arc of great circle with endpoints α and β. Their weak superpositions are the whole sphere, since the point α^* diametral to α (i.e. such that $\alpha\alpha^* = \pi$) satisfies the relation $[\alpha\beta\alpha^*]$, so that $\alpha^* \in \{\widetilde{\alpha, \beta}\}$, and together with α and α^* the envelope of α and β must obviously contain every point of $\mathcal{S}$.

Suppose next that α and β are already diametral ($\alpha\beta = \pi$). The set of their strong superpositions itself, together with α and β, is now the whole sphere.

For $\alpha\beta < \pi$ a superposition γ of α and β is uniquely determined by its distances from the component states α and β. For $\alpha\beta = \pi$ the superpositions with given distances from the components constitute a circle, and an additional parameter (phase) is necessary for their identification.

Let us come back to the general case. The structure of the set obtained by forming the superpositions of two given states with given distances from the components, depends of course on the structure of the whole system, and may also depend, as the above example shows, on the choice of the component states.

Let us now examine the betweenness relation $\alpha\gamma\beta$ from a more physical point of view. As already noted in Section 3, one has

$$\alpha\beta = \sup_{A \in \mathcal{O}} (\alpha, \beta)_A, \tag{13}$$

where the function $(\alpha\beta)_A$ defined by (10) and (2) expresses to what extent the states α and β can be 'resolved' by performing on them measurements of the observable A only (complete resolution being attained if and only if $(\alpha\beta)_A = \pi$, in which case the intersection of the ranges[1]) of A on the states α and β has measure zero with respect to the measure class of σ). Thus the distance $\alpha\beta$ represents the best resolution which can be reached, or approached as closely as one wishes, by a convenient choice (dependent on α and β) of the observable A. In order to simplify our phrasing we shall express ourselves as if the 'sup' in (13) was actually attained for some observable $A_{\alpha\beta}$ (but this is by no means essential, and everything could easily be recast in rigorous form with slight adaptations and by interpreting the equalities 'up to arbitraily small ε').

Since for every $A \in \mathcal{O}$ the inequality (12) holds, in particular one has

$$(\alpha\beta)_{A_{\alpha\beta}} \leqslant (\alpha\gamma)_{A_{\alpha\beta}} + (\gamma\beta)_{A_{\alpha\beta}}. \tag{14}$$

On the other hand, by (13), one has

$$(\alpha\gamma)_{A_{\alpha\beta}} \leqslant \alpha\gamma, \quad (\gamma\beta)_{A_{\alpha\beta}} \leqslant \gamma\beta, \tag{15}$$

so that

$$(\alpha\gamma)_{A_{\alpha\beta}} + (\gamma\beta)_{A_{\alpha\beta}} \leqslant \alpha\gamma + \gamma\beta. \tag{16}$$

If γ is a superposition of α and β, (14) and (16) imply

$$\alpha\beta \equiv (\alpha\beta)_{A_{\alpha\beta}} = (\alpha\gamma)_{A_{\alpha\beta}} + (\gamma\beta)_{A_{\alpha\beta}} = \gamma\alpha + \gamma\beta,$$

which is compatible with (15) only if $(\alpha\gamma)_{A_{\alpha\beta}} = \alpha\gamma$ and $(\gamma\beta)_{A_{\alpha\beta}} = \gamma\beta$.

Thus whenever γ is a superposition of α and β, an observable $A_{\alpha\beta}$ providing the best resolution for the pair of states $\{\alpha, \beta\}$ also provides the best resolution for each of the two pairs $\{\alpha, \gamma\}$ and $\{\gamma, \beta\}$. Conversely, it is obvious that if there exists an observable A^* providing the best resolution for the three pairs, and if $(\alpha\beta)_{A^*} = (\alpha\gamma)_{A^*} + (\gamma\beta)_{A^*}$, then γ is a superposition of α and β. The statistical distributions of any of the observables with the properties of A^* (or $A_{\alpha\beta}$) are sufficient for an estimate of the distance between α and β, and of their distances from any of their superpositions. Notice, however, that the $A_{\alpha\beta}$'s need not provide the best resolution for the pairs $\{\alpha, \gamma\}$ and $\{\gamma, \beta\}$ if γ is not between α and β, (even when it is aligned with them).

5. Superpositions and mixtures in classical systems

As usual a state γ will be called a *mixture* if there exist two distinct states α and β and two positive numbers t_α and t_β such that $t_\alpha + t_\beta = 1$ and

$$p(A, \gamma, E) = t_\alpha p(A, \alpha, E) + t_\beta p(A, \beta, E). \tag{17}$$

for all A's in $\mathcal{O}$ and all E's in $\mathscr{B}(\mathbb{R})$. A state is *pure* if it is not a mixture.

Let $\{\mathcal{S}, \mathcal{O}, p\}$ be the model of a classical system. By *classical* is meant that: (a) with every observable $A \in \mathcal{O}$ there is associated a real-valued function $a(\alpha)$ defined on the subset $S_p \subset S$ of all pure states, such that $p(A, \alpha, a(\alpha)) = 1$ (and therefore $p(A, \alpha, E) = 0$ whenever $a(\alpha) \neq E$); (b) given any pair of observables A

---

[1]) The *range* of an observable A on a state α is defined as the smallest of the closed subsets C of $\mathbb{R}$ such that $p(A, \alpha, C) = 1$.

and B and any real-valued function $f(x, y)$ of two real variables defined on the cartesian product of the ranges of $a(\alpha)$ and $b(\alpha)$ (the functions representing A and B on $\mathscr{S}_p$), there exists in $\mathcal{O}$ an observable, denoted $f(A, B)$, which is represented by $f(a(\alpha), b(\alpha))$ on $\mathscr{S}_p$, with the further property that its possible values on any given (not necessarily pure) state γ are contained in the set $\{f(x_\gamma, y_\gamma)\}$, where x_γ and y_γ run through the possible values of A and B on γ.

It is easy to see that assumption (a) and the definition of T_A imply, for any state γ and any pure state α:

$$T_A(\alpha, \gamma) = p(A, \gamma, a(\alpha)) \qquad (\gamma \in \mathscr{S}, \alpha \in \mathscr{S}_p). \tag{18}$$

In particular, if γ is itself a pure state distinct from α, since by xiom II (Section 1) there must exist an observable A such that $a(\gamma) \neq a(\alpha)$, (18) implies $T(\alpha, \gamma) = 0$.

Let now α and β be pure states, and γ one of their mixtures (so that $T(\alpha, \beta) = 0$, and (17) holds for some relative weights t_α and t_β). For every observable A such that $a(\alpha) \neq a(\beta)$ equations (17) and (18) imply $T_A(\alpha, \gamma) = t_\alpha$, $T_A(\gamma, \beta) = t_\beta$. Since, on the other hand, for observables A such that $a(\alpha) = a(\beta)$ one has $T_A(\alpha, \gamma) = T_A(\gamma, \beta) = 1$, by taking the inf's as required by Definition (1), and from Definition (7), one gets

$$T(\alpha, \gamma) = \cos^2 \frac{\alpha\gamma}{2} = t_\alpha; \qquad T(\gamma, \beta) = \cos^2 \frac{\gamma\beta}{2} = t_\beta \tag{19}$$

while, as we already know, $T(\alpha, \beta) = 0$. Consequently, $\alpha\beta = \pi$, and $\cos^2 \frac{\alpha\gamma}{2} + \cos^2 \frac{\gamma\beta}{2} = 1$, so that $\alpha\gamma + \gamma\beta = \pi$. Thus γ is a superposition of α and β, and its relative weights as a mixture are related by (19) to the distances $\alpha\gamma$ and $\gamma\beta$.

Conversely, let us show that if α and β are distinct pure states and γ is one of their superpositions, then γ is a mixture of α and β with weights given by (19).

In fact, the remarks at the end of Section 4 now imply that for any observable A such that $a(\alpha) \neq a(\beta)$ one has $(\alpha\gamma)_A = \alpha\gamma$ and $(\gamma\beta)_A = \gamma\beta$, which is equivalent to $T_A(\alpha, \gamma) = T(\alpha, \gamma)$ and $T_A(\gamma, \beta) = T(\gamma, \beta)$. Since, by (18), $T_A(\alpha, \gamma) = p(A, \gamma, a(\alpha))$ and $T_A(\gamma, \beta) = p(A, \gamma, a(\beta))$, one gets:

$$\begin{aligned} p(A, \gamma, a(\alpha)) &= T(\alpha, \gamma) \\ p(A, \gamma, a(\beta)) &= T(\gamma, \beta), \end{aligned} \tag{21}$$

which can be rewritten in the form

$$\begin{aligned} p(A, \gamma, a(\alpha)) &= t_\alpha p(A, \alpha, a(\alpha)) + t_\beta p(A, \beta, a(\alpha)) \\ p(A, \gamma, a(\beta)) &= t_\alpha p(A, \alpha, a(\beta)) + t_\beta p(A, \beta, a(\beta)) \end{aligned} \tag{22}$$

by setting $t_\alpha = T(\alpha, \gamma)$, $t_\beta = T(\gamma, \beta)$ and remembering that $p(A, \alpha, a(\alpha)) = p(A, \beta, a(\beta)) = 1$ and $p(A, \alpha, a(\beta)) = p(A, \beta, a(\alpha)) = 0$. Since we are assuming that $\alpha\gamma + \gamma\beta = \alpha\beta$, and since $\alpha\beta = \pi$ because $T(\alpha, \beta) = 0$, we have

$$\begin{aligned} T(\alpha, \gamma) &= \cos^2 \frac{\alpha\gamma}{2} = \cos^2\left(\frac{\pi}{2} + \frac{\gamma\beta}{2}\right) = \sin^2 \frac{\gamma\beta}{2} \\ &= 1 - \cos^2 \frac{\gamma\beta}{2} = 1 - T(\gamma, \beta). \end{aligned}$$

This shows that $t_\alpha + t_\beta = 1$, i.e. by (21)

$$p(A, \gamma, a(\alpha)) + p(A, \gamma, a(\beta)) = 1, \tag{23}$$

From this it follows, by the additivity of the probability measure $p(A, \gamma, \cdot)$, that $p(A, \gamma, E) = 0$ whenever $a(\alpha) \notin E$ and $a(\beta) \notin E$. This remark and eqs. (22) show that (17) is verified except, perhaps, if $a(\alpha) = a(\beta)$.

In order to show that (17) is verified in all cases we note, first, that equation (23) means that whenever the values of an observable on α and on β are distinct, the only values that the same observable can take on γ are the ones it takes on α and β. Now suppose that $a(\alpha)$ and $a(\beta)$ have the same value a, and consider an observable B represented on the pure states by a function b such that $b(\alpha) \neq b(\beta)$. By our assumption (b) on classical systems, for any nonzero real number x there is an observable $A + xB$ represented on $\mathcal{S}_p$ by the function $a + xb$, which also takes distinct values on α and β. Thus, as just remarked, while B can take on γ the values $b(\alpha)$ and $b(\beta)$ only, $A + xB$ can take on γ the values $a + xb(\alpha)$ and $a + xb(\beta)$ only, and by the last part of assumption (b) on classical systems the only possible values of A on γ are, at this stage, a, $a + xb(\alpha) - xb(\beta)$ and $a + xb(\beta) - xb(\alpha)$. By repeating the argument with a different value of x we conclude that A can only take the value a on γ, and from this the validity of (17) in full generality follows immediately.

Notice that in the proof of (17) for an observable A such that $a(\alpha) \neq a(\beta)$ the classical character of the system has only been used to assure that A has definite values on α and on β. Thus, if the system is non-classical, a superposition γ of two states α and β still behaves like a mixture if the set of observables is restricted to the ones which take definite, distinct values on the component studies. However (17) is no longer true, in general, for observables whose ranges on α and β are not concentrated at a point.

Notice also that the superpositions of two pure states of a classical system form a segment of length π with endpoints at the component states. This is no longer true in general when the system is not classical, since two superpositions with the same distances from a given pair of component states may then give rise to distinct distributions for some observable which does not take definite values on the component states, and therefore be distinct.

6. Pure superpositions in quantum mechanics

Let $\{\mathcal{S}, \mathcal{O}, p\}$ now denote a quantum-mechanical system, with the state space $\mathcal{S}$ *restricted to the pure states*. Thus $\mathcal{S}$ is the projective Hilbert space $\tilde{\mathcal{H}}$ associated with the complex, separable, infinite-dimensional Hilbert space $\mathcal{H}$; the elements of $\mathcal{O}$ are represented by self-adjoint operators in $\mathcal{H}$, and the probability function is given by $p(A, \alpha, E) = \langle \boldsymbol{\alpha}, P_E^A \boldsymbol{\alpha} \rangle$, where P_E^A is the projection operator associated with the Borel set E in the resolution of the self-adjoint operator representing A, $\boldsymbol{\alpha}$ is any unit representative of the pure state α in $\mathcal{H}$ and $\langle\ ,\ \rangle$ denotes the inner product. It can be proved that in this case one has [20]

$$T(\alpha, \beta) = |\langle \boldsymbol{\alpha}, \boldsymbol{\beta} \rangle|^2. \tag{24}$$

We shall show that *whenever a state γ is a superposition of two distinct states α*

and β (in the metric sense defined in Section 4), *it is also a linear superposition of* α *and* β (in the ordinary sense of Quantum Mechanics).

In fact, on account of (7) and (24), the relation $\alpha\beta = \alpha\gamma + \gamma\beta$ can now be written in the form $\operatorname{Arc}\cos|\langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle| = \operatorname{Arc}\cos|\langle\boldsymbol{\alpha}, \boldsymbol{\gamma}\rangle| + \operatorname{Arc}\cos|\langle\boldsymbol{\gamma}, \boldsymbol{\beta}\rangle|$, i.e., by taking the cosines:

$$|\langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle| = |\langle\boldsymbol{\alpha}, \boldsymbol{\gamma}\rangle|\,|\langle\boldsymbol{\gamma}, \boldsymbol{\beta}\rangle| - \sqrt{(1-|\langle\boldsymbol{\alpha}, \boldsymbol{\gamma}\rangle|^2(1-|\langle\boldsymbol{\gamma}, \boldsymbol{\beta}\rangle|^2)},$$

which implies (by solving with respect to the square root and taking the squares):

$$1-|\langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle|^2 - |\langle\boldsymbol{\beta}, \boldsymbol{\gamma}\rangle|^2 - |\langle\boldsymbol{\gamma}, \boldsymbol{\alpha}\rangle|^2 + 2\,|\langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle\langle\boldsymbol{\beta}, \boldsymbol{\gamma}\rangle\langle\boldsymbol{\gamma}, \boldsymbol{\alpha}\rangle| = 0. \tag{25}$$

On the other hand, for any triple of unit vectors $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ in $\mathscr{H}$, the positive-definiteness of the hermitian form $\langle\ ,\ \rangle$ implies the relation

$$\begin{vmatrix} 1 & \langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle & \langle\boldsymbol{\alpha}, \boldsymbol{\gamma}\rangle \\ \langle\boldsymbol{\beta}, \boldsymbol{\alpha}\rangle & 1 & \langle\boldsymbol{\beta}, \boldsymbol{\gamma}\rangle \\ \langle\boldsymbol{\gamma}, \boldsymbol{\alpha}\rangle & \langle\boldsymbol{\gamma}, \boldsymbol{\beta}\rangle & 1 \end{vmatrix} \geqslant 0,$$

i.e.

$$1-|\langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle|^2 - |\langle\boldsymbol{\beta}, \boldsymbol{\gamma}\rangle|^2 - |\langle\boldsymbol{\gamma}, \boldsymbol{\alpha}\rangle|^2 + 2\operatorname{Re}(\langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle\langle\boldsymbol{\beta}, \boldsymbol{\gamma}\rangle\langle\boldsymbol{\gamma}, \boldsymbol{\alpha}\rangle) \geqslant 0. \tag{26}$$

Since $\operatorname{Re}(\langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle\langle\boldsymbol{\beta}, \boldsymbol{\gamma}\rangle\langle\boldsymbol{\gamma}, \boldsymbol{\alpha}\rangle) \leqslant |\langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle\langle\boldsymbol{\beta}, \boldsymbol{\gamma}\rangle\langle\boldsymbol{\gamma}, \boldsymbol{\alpha}\rangle|$, (25) can be true (and γ can be a "metric" superposition of α and β) only if $\operatorname{Re}(\langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle\langle\boldsymbol{\beta}, \boldsymbol{\gamma}\rangle\langle\boldsymbol{\gamma}, \boldsymbol{\alpha}\rangle) = |\langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle\langle\boldsymbol{\beta}, \boldsymbol{\gamma}\rangle\langle\boldsymbol{\gamma}, \boldsymbol{\alpha}\rangle|$ (so that $\langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle\langle\boldsymbol{\beta}, \boldsymbol{\gamma}\rangle\langle\boldsymbol{\gamma}, \boldsymbol{\alpha}\rangle$ is real and non-negative). If this occurs, the left-hand sides of (25) and (26) coincide, so that (26) holds with the equality sign, and this expresses the linear dependence of the three vectors $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$.

On the other hand, from (25) one gets $|\langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle| = |\langle\boldsymbol{\alpha}, \boldsymbol{\gamma}\rangle|\,|\langle\boldsymbol{\beta}, \boldsymbol{\gamma}\rangle| \pm \sqrt{(1-|\langle\boldsymbol{\alpha}, \boldsymbol{\gamma}\rangle|^2)(1-|\langle\boldsymbol{\beta}, \boldsymbol{\gamma}\rangle|^2)}$, and by taking the cosines one sees that this relation is necessary and sufficient for α, β and γ to be aligned.

Notice that the value of the product $\langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle\langle\boldsymbol{\beta}, \boldsymbol{\gamma}\rangle\langle\boldsymbol{\gamma}, \boldsymbol{\alpha}\rangle$ (in particular, its non-negative character whenever it occurs) is independent of the choice of the unit representatives of the three states involved.

Whenever $\alpha\gamma\beta$ is true, so that $\langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle\langle\boldsymbol{\beta}, \boldsymbol{\gamma}\rangle\langle\boldsymbol{\gamma}, \boldsymbol{\alpha}\rangle > 0$, the representatives $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ can be chosen such that $\langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle > 0$ and, at the same time,

$$\boldsymbol{\gamma} = a\boldsymbol{\alpha} + b\boldsymbol{\beta} \tag{27}$$

holds with *positive* coefficients a and b. The coefficients satisfy the relation

$$a^2 + b^2 + 2ab\langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle = 1, \tag{28}$$

which expresses that $\boldsymbol{\gamma}$ is a unit vector. From (25), (24) and (7) we get

$$\cos\frac{\alpha\gamma}{2} = a + b\langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle$$

$$\cos\frac{\gamma\beta}{2} = b + a\langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle,$$

which yield

$$a = \frac{\cos\dfrac{\alpha\gamma}{2} - \cos\dfrac{\gamma\beta}{2}\cos\dfrac{\alpha\beta}{2}}{1-\cos^2\dfrac{\alpha\beta}{2}}, \qquad b = \frac{\cos\dfrac{\gamma\beta}{2} - \cos\dfrac{\alpha\gamma}{2}\cos\dfrac{\alpha\beta}{2}}{1-\cos^2\dfrac{\alpha\beta}{2}}.$$

If, keeping the representatives $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ fixed, the coefficients a and b in (27) are allowed to take *complex* values compatible with the normalization $\langle\boldsymbol{\gamma}, \boldsymbol{\gamma}\rangle = 1$, γ describes a 2-sphere through α and β (the complex projective space associated with the two-dimensional hermitian space generated by $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$).[1] If $\alpha\beta < \pi$, (28) shows that the strong superpositions of α and β belong to the shortest geodesic arc with endpoints α and β, while their weak superpositions constitute the whole sphere. If $\alpha\beta = \pi$ the strong superpositions, together with α and β themselves, constitute the whole sphere. (Compare with the example of Section 4.)

Thus, in the quantum-mechanical model, there is a correspondence between the envelopes of pairs of states of $\mathcal{S}$ and the linear subspaces of $\mathcal{H}$ generated by their representatives. It is easy to see that the linear subspaces of $\mathcal{H}$ of any finite dimension also correspond to subsets of $\mathcal{S}$ which are closed under superposition (in the sense of Section 4).

7. A metric characterization of the complex separable projective Hilbert space

Among the spaces listed in Wang's theorem (Section 2) the finite-dimensional complex elliptic spaces (i.e. the finite-dimensional complex projective spaces CP^n, $n = 1, 2, \ldots$) are the only ones with the property that the envelope of any pair of distinct points is isometric to a 2-sphere. Therefore, as an immediate corollary of the theorem, the complex projective space of dimension n can be characterized as

[1] In this case the number $\rho \equiv \langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle\langle\boldsymbol{\beta}, \boldsymbol{\gamma}\rangle\langle\boldsymbol{\gamma}, \boldsymbol{\alpha}\rangle$ is complex. Writing $\rho = |\rho|\, e^{i\theta}$, whenever $\rho \neq 0$ the "relative phase" θ of the triple α, β, γ is determined, up to the sign, by the equation

$$\cos\theta = \frac{|\langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle|^2 + |\langle\boldsymbol{\beta}, \boldsymbol{\gamma}\rangle|^2 + |\langle\boldsymbol{\gamma}, \boldsymbol{\alpha}\rangle|^2}{2\,|\langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle\langle\boldsymbol{\beta}, \boldsymbol{\gamma}\rangle\langle\boldsymbol{\gamma}, \boldsymbol{\alpha}\rangle|}$$

which can be written in purely metric terms:

$$\cos\theta = \frac{\cos^2\dfrac{\alpha\beta}{2} + \cos^2\dfrac{\beta\gamma}{2} + \cos^2\dfrac{\gamma\alpha}{2}}{2\cos\dfrac{\alpha\beta}{2}\cos\dfrac{\beta\gamma}{2}\cos\dfrac{\gamma\alpha}{2}}.$$

The condition for the right-hand side to belong to the interval $[-1, 1]$ is equivalent to the relation

$$\begin{vmatrix} 1 & \cos\alpha\beta & \cos\alpha\gamma \\ \cos\beta\alpha & 1 & \cos\beta\gamma \\ \cos\gamma\alpha & \cos\gamma\beta & 1 \end{vmatrix} \geqslant 0,$$

which is a necessary and sufficient condition in order that the triple α, β, γ be isometric to a triple of points of the 2-sphere of radius 1 (see [27]).

a metric space with the following properties:

(a) compactness;
(b) two-point homogeneity;
(c) the envelope of any pair of distinct points is isometric to a 2-sphere;
(d) there exists a set of n points (but not one of $n-1$ points) having the whole space as its envelope.

(Convexity is not explicitly mentioned, but is implied by (c)).

We shall now see that, again as a consequence of Wang's theorem, the infinite-dimensional complex separable projective Hilbert space $\tilde{\mathcal{H}}$ admits a characterization of the same kind. Namely, it is the only metric space M with the following properties:

(α) There exists a countable set of points $\Sigma \subset M$ having the whole space as its envelope, and no finite subset of Σ has this property;
(β) the envelope of any finite subset of Σ is compact and two-point homogeneous;
(γ) the envelope of any pair of distinct points of M is isometric to a 2-sphere.

Remarks

If the space is convex and condition (β) is assumed to hold for any finite subset of M, the envelope of any pair of distinct points must be one of the spaces listed in Wang's theorem and (γ) can be replaced by the weaker requirement that its dimension be 2 for some pair of distinct points.

Let us also note that on account of an extension of Wang's theorem due to Tits [25], in (β) "compact" could be replaced by "finitely-compact" ([17] Section 2) provided that it is also assumed that the space has finite diameter (a condition which is automatically satisfied, as we know, by the state space of any physical system).

Proof of the characterization of $\tilde{\mathcal{H}}$

Let M be a metric space with the properties (α), (β) and (γ). We must show that M is isometric with the projective space $\tilde{\mathcal{H}}$ associated with the complex separable Hilbert space $\mathcal{H}$.

Let $e_1, e_2, \ldots$ be the elements of the countable set Σ with the properties (α) and (β). For every integer k, denote by $\tilde{\Sigma}_k$ the envelope of the first k elements of Σ. By Wang's theorem and assumptions (β) and (γ), $\tilde{\Sigma}_k$ is isometric with a finite-dimensional complex projective space, so that we can find an isometric map $\tilde{\varphi}_k$ from $\tilde{\Sigma}_k$ to $\tilde{\mathcal{H}}_k$, the complex projective space associated with some finite-dimensional linear subspace $\mathcal{H}_k$ of $\mathcal{H}$, and each $\tilde{\varphi}_k$ can be chosen such that, for $h<k$, $\tilde{\varphi}_h$ is the restriction of $\tilde{\varphi}_k$ to $\tilde{\Sigma}_h$.

For an arbitrary point $x \in M$, let x_k be a point of the compact set $\tilde{\Sigma}_k$ such that the distance xx_k is equal to the distance from x to Σ_k (i.e. $xx_k = \min_{y\in\tilde{\Sigma}_k} xy$). The numerical sequence $\{xx_k\}$ is obviously non-increasing, and one has $\lim xx_k = 0$: in fact, if this were not the case, the distance from x to $\tilde{\Sigma}_k$ would be greater than some positive number ρ for every k, so that the distance from x to $\bigcup_{k=1}^{\infty} \tilde{\Sigma}_k$

would also be greater than ρ, and $\bigcup_{k=0}^{\infty} \tilde{\Sigma}_k$ would be an a-closed set containing Σ and distinct from M, in contradiction with (α).

Together with the sequence $\{x_k\}$ which converges to x in M, the sequence $\{\tilde{\varphi}_k(x_k)\}$ is a Cauchy sequence in $\tilde{\mathscr{H}}$ (on account of the isometric character of the $\tilde{\varphi}_k$'s), and by the completeness of Hilbert space some sequence $\{\varphi_k(x_k)\}$ of unit representatives of the $\tilde{\varphi}_k(x_k)$'s in $\mathscr{H}$ converges to a unit vector $\varphi(x)$ representing a well-determined element $\tilde{\varphi}(x) \in \tilde{\mathscr{H}}$. The map $x \to \tilde{\varphi}(x)$ is an isometry because, if y is any other point of M, one has $xy = \lim x_k y_k = \lim \tilde{\varphi}_k(x_k)\tilde{\varphi}_k(y_k) = \lim \tilde{\varphi}(x_k)\tilde{\varphi}(y_k) = \tilde{\varphi}(x)\tilde{\varphi}(y)$.

Denote by $\mathscr{H}_M$ the smallest closed subspace of $\mathscr{H}$ containing all the representatives of the points of the subset $\tilde{\varphi}(M)$ of $\tilde{\mathscr{H}}$ described by $\tilde{\varphi}(x)$ as x runs through M, and by $\tilde{\mathscr{H}}_M$ the associated projective space: every element of $\tilde{\mathscr{H}}_M$ is the limit of the sequence of its nearest points in the subsets $\tilde{\mathscr{H}}_k$, and therefore it is the image of some point of M. Thus $\tilde{\mathscr{H}}_M$ and M are isometric, and our proposition is proved.

We conclude by observing that while simple homogeneity is a very natural assumption for the set of pure states of any reversible physical system (see [24]), the two-point homogenity condition expresses a high degree of symmetry of the state space. It represents a true restriction which, together with the dimensionality of the whole space and of the superpositions of pairs of states, essentially characterizes the quantum-mechanical model.

REFERENCES

[1] P. A. M. DIRAC, *The Principles of Quantum Mechanics* (Clarendon Press 1958; first edition 1930).
[2] V. S. VARADARAJAN, *Geometry of Quantum Theory* vol 1, (Van Norstrand Co. 1968).
[3] J. M. JAUCH, *Foundations of Quantum Mechanics*, (Addison Wesley 1968).
[4] J. E. ROBERTS and G. ROEPSTORFF, Commun. Math. Phys. *11*, 321 (1969).
[5] S. P. GUDDER, J. Math. Phys. *11*, 1037 (1970).
[6] E. CHEN, J. Math. Phys. *14*, 1462 (1973).
[7] C. PIRON, *Foundations of Quantum Physics* (Benjamin 1976).
[8] P. DELIYANNIS, J. Math. Phys. *17*, 248 (1976).
[9] S. PULMANOVÀ, Commun. Math. Phys. *49*, 47 (1976).
[10] V. CANTONI, Commun. Math. Phys. *50*, 241 (1976).
[11] A. ZECCA, Int. J. Theoret. Phys. *19*, 629 (1980).
[12] A. ZECCA, Int. J. Theoret. Phys. *20*, 191 (1981).
[13] E. G. BELTRAMETTI and G. CASSINELLI, *The Logic of Quantum Mechanics* (Addison Wesley 1981).
[14] G. W. MACKEY, *Mathematical Foundations of Quantum Mechanics* (Benjamin 1963).
[15] K. MENGER, Mathematische Annalen *100*, 75 (1928).
[16] G. BIRKHOFF, Trans. Amer. Math. Soc. *55*, 465 (1944).
[17] L. M. BLUMENTHAL, *Theory and Applications of Distance Geometry* (Clarendon Press 1953).
[18] H. BUSEMANN, *The geometry of Geodesics*, (Academic Press 1955).
[19] H. C. WANG, Ann. Math. *55*, 177 (1952).
[20] V. CANTONI, Commun. Math. Phys. *44*, 125 (1975).
[21] S. P. GUDDER, Commun. Math. Phys. *63*, 265 (1978).
[22] S. P. GUDDER, *Stochastic Methods in Quantum Mechanics*, (North Holland 1979).
[23] N. HADJISAVVAS, Commun. Math. Phys. *87*, 153 (1982).
[24] V. CANTONI, Commun. Math. Phys. *87*, 153 (1982).
[25] J. TITS, Mém. Acad. Roy. Belg. XX1X fasc. 3 (1955).
[26] L. M. BLUMENTHAL and G. A. GARRETT, Am. J. of Math. *55*, 619 (1933).
[27] L. M. BLUMENTHAL, Am. J. of Math. *57*, 51 (1935).