

On Cantoni’s Generalized Transition Probability

N. Hadjisavvas

Laboratoire de Mecanique Quantique, Université de Reims, F-51062 Reims Cedex, France

Abstract. We obtain simple expressions of the “generalized transition probability” proposed by V. Cantoni, for both classical and quantum mechanics. We compare the result with the ordinary quantum mechanical transition probability.

I. Introduction

Using Mackey’s axiomatization of physical theories ([1], p. 63), V. Cantoni introduced a function $T(\alpha, \beta)$ defined on pairs of states α, β . When the physical theory in question is quantum mechanics and for the special case of pure states, Cantoni proved that $T(\alpha, \beta)$ equals the “transition probability” $|\langle \alpha, \beta \rangle|^2$. In that sense, he named the function T “generalized transition probability.”

Our main purpose in this work is to give simpler expressions of T for both classical and quantum theories since Cantoni’s definition is quite involved. We shall prove furthermore that in the quantum case, $T(\alpha, \beta)$ equals the quantum mechanical “transition probability” between the states α and β , each time that this concept has an unambiguous sense.

We now recall that in Mackey’s system, one considers a set \mathcal{O} of observables and a set \mathcal{S} of states. For each $A \in \mathcal{O}$ and $\alpha \in \mathcal{S}$, and for any Borel subset E of the real line \mathbb{R} , one denotes by $p(A, \alpha, E)$ the probability that a measurement of A performed on a system in the state α will yield a result lying in E . Accordingly, $\alpha_A(E) \equiv p(A, \alpha, E)$ is a probability measure on \mathbb{R} . Mackey then imposes some axioms involving \mathcal{O}, \mathcal{S} and the probability p . We shall make use only of the first three, which trivially hold in all known physical theories.

Cantoni’s definition now runs as follows: for any pair of states α, β and any observable A , define the expression $T_A(\alpha, \beta)$ by

$$T_A^{1/2}(\alpha, \beta) = \int_{\mathbb{R}} \left(\frac{d\alpha_A}{d\sigma} \frac{d\beta_A}{d\sigma} \right)^{1/2} d\sigma, \tag{1}$$

where σ is any measure with respect to which both α_A and β_A are absolutely continuous. It is easy to see that T_A is independent of σ . Finally define

$$T(\alpha, \beta) = \inf_{A \in \mathcal{O}} T_A(\alpha, \beta). \tag{2}$$

This is Cantoni’s “generalized transition probability” [2, 3, 4].

II. The Main Theorem

Let us recall that in Mackey's formalism, the spectrum of an observable A can be defined, roughly, as the set of all possible values of A (for a rigorous definition see [1], p. 69). Our main result, from which all others follow, is that in (2) we can take the infimum over the observables with purely discrete spectrum only. The proof is based on two simple lemmas from measure theory:

Lemma 1. *Let (X, \mathcal{B}, μ) be a measure space and $\mathcal{L}^2(X, \mu)$ the Hilbert space of all square integrable real-valued functions on X . Then for any positive element $f \in \mathcal{L}^2(X, \mu)$ and any $\varepsilon > 0$, there exists a partition $D_i, i = 1, 2, \dots, n$ of X such that, for any partition $B_j, j = 1, 2, \dots, m$ finer than D_i , one has*

$$\left\| f - \sum_{\mu(B_j) \neq 0} \sqrt{\int_{B_j} f^2 d\mu} \frac{\chi_{B_j}}{\sqrt{\mu(B_j)}} \right\| < \varepsilon,$$

where χ_{B_j} is the characteristic function of B_j .

Proof. Since the step functions are dense in $\mathcal{L}^2(X, \mu)$, for any $\varepsilon > 0$ and $f \in \mathcal{L}^2(X, \mu)$, there exists a partition $D_i, i = 1, 2, \dots, n$, and numbers a_i such that

$$\left\| f - \sum_i a_i \chi_{D_i} \right\| \equiv \left(\int_X \left| f - \sum_i a_i \chi_{D_i} \right|^2 d\mu \right)^{1/2} < \varepsilon. \quad (3)$$

We recall that, by the usual definition of step functions ([5], p. 231), $\mu(D_i) = \infty$ implies $a_i = 0$.

If in addition $f \geq 0$, then we can take $a_i \geq 0$. For any partition B_j finer than D_i , define the numbers b_j by $b_j = a_i$ iff $B_j \subseteq D_i$. Then relation (3) implies

$$\left\| f - \sum_j b_j \chi_{B_j} \right\| < \varepsilon. \quad (4)$$

Setting

$$f_1 = \sum_{\mu(B_j) \neq 0} \sqrt{\int_{B_j} f^2 d\mu} \frac{\chi_{B_j}}{\sqrt{\mu(B_j)}},$$

we find

$$\|f - f_1\| \leq \left\| f - \sum_j b_j \chi_{B_j} \right\| + \left\| \sum_j b_j \chi_{B_j} - f_1 \right\|, \quad (5)$$

and also

$$\begin{aligned} \left\| \sum_j b_j \chi_{B_j} - f_1 \right\|^2 &= \sum_{\mu(B_j) \neq 0, \infty} \left| b_j - \frac{\left(\int_{B_j} f^2 d\mu \right)^{1/2}}{\sqrt{\mu(B_j)}} \right|^2 \mu(B_j) \\ &\leq \sum_j \|f \cdot \chi_{B_j}\| - \|b_j \chi_{B_j}\|^2 \leq \sum_j \|f \cdot \chi_{B_j} - b_j \chi_{B_j}\|^2 \\ &= \|f - \sum_j b_j \chi_{B_j}\|^2. \end{aligned} \quad (6)$$

Combining (4), (5), (6) we deduce that $\|f - f_1\| \leq 2\varepsilon$. ■

Lemma 2. *If $f, g \in \mathcal{L}^2(X, \mu)$ and $f \geq 0, g \geq 0$, then $\forall \varepsilon > 0$ there exists a partition $B_j, j = 1, 2, \dots, m$ such that*

$$\left| \int_X fg d\mu - \sum_j \left(\left(\int_{B_j} f^2 d\mu \right) \left(\int_{B_j} g^2 d\mu \right) \right)^{1/2} \right| \leq \varepsilon.$$

Proof. By Lemma 1, for every $\varepsilon > 0$ we can find partitions B_j, B'_j , such that $\|f - f_1\| \leq \varepsilon, \|g - g_1\| \leq \varepsilon$, where f_1 was defined in the preceding proof and g_1 is defined analogously. In addition, we can suppose that the two partitions are identical. (If not, consider the partition $B_j \cap B'_j$). One has

$$\begin{aligned} \left| \int_X fg d\mu - \int_X f_1 g_1 d\mu \right| &\leq \int |f - f_1| g d\mu + \int f_1 |g - g_1| d\mu \\ &\leq \varepsilon (\|f\| + \|g\|). \end{aligned} \quad (7)$$

Obviously

$$\int_X f_1 g_1 d\mu = \sum_{\mu(B_j) \neq \infty} \left(\left(\int_{B_j} f^2 d\mu \right) \left(\int_{B_j} g^2 d\mu \right) \right)^{1/2}, \quad (8)$$

while Lemma 1 again shows that for this partition B_j , one has

$$\sum_{\mu(B_j) = \infty} \int_{B_j} f^2 d\mu < \varepsilon^2, \quad \sum_{\mu(B_j) = \infty} \int_{B_j} g^2 d\mu < \varepsilon^2,$$

which imply

$$\sum_{\mu(B_j) = \infty} \left(\left(\int_{B_j} f^2 d\mu \right) \cdot \left(\int_{B_j} g^2 d\mu \right) \right)^{1/2} < \varepsilon^2. \quad (9)$$

A combination of (7), (8) and (9) entails Lemma 2. ■

We now turn to the statement and proof of the main theorem. Let us denote by \mathcal{O}_f the set of all observables the spectrum of which consists of a finite number of points.

Theorem 1. *In any physical theory obeying axioms I, II, III of Mackey's axiomatics, $T(\alpha, \beta)$ is also given by the formula*

$$T(\alpha, \beta) = \inf_{A \in \mathcal{O}_f} T_A(\alpha, \beta). \quad (10)$$

Proof. For any pair of states α, β and for any observable A , let σ be a measure with respect to which both α_A and β_A are absolutely continuous. Since α_A and β_A are probability measures on \mathbb{R} , $\sqrt{\frac{d\alpha_A}{d\sigma}}$ and $\sqrt{\frac{d\beta_A}{d\sigma}}$ belong to $\mathcal{L}^2(\mathbb{R}, \sigma)$. Accordingly, Lemma 2 and relation (1) imply that for any $\varepsilon > 0$ there exists a partition $B_k, k = 1, 2, \dots, n$ of \mathbb{R} such that

$$\left| T_A^{1/2} - \sum_k (\alpha_A(B_k) \beta_A(B_k))^{1/2} \right| < \varepsilon. \quad (11)$$

Define now the function $h(x) = \sum_{k=1}^n k \chi_{B_k}(x)$ and the observable $A' = h(A)$ (see

Mackey's axiom III, [1], p. 63). It is easy to see that (a) the spectrum of A' is discrete, consisting of the points $k = 1, 2, \dots, n$, (b) one has $\alpha_{A'}(\{k\}) = \alpha_A(B_k)$ and $T_{A'}^{1/2}(\alpha, \beta) = \sum_k (\alpha_A(B_k)\beta_A(B_k))^{1/2}$. Substitution in (11) gives

$$|T_{A'}^{1/2} - T_A^{1/2}| < \varepsilon. \quad (12)$$

It is now trivial that (12) and (2) imply (10). ■

III. The Classical and the Quantal Case

Let us see what the forgoing theorem implies when the physical theory in question is classical statistical or quantum mechanics. In the first case, a physical system is described by its phase space Ω . The invariant Liouville measure μ is defined on the Borel subsets of Ω . States are probability densities on Ω , i.e. non-negative functions ρ such that $\int_{\Omega} \rho d\mu = 1$ (see [1], p. 48) while the observables are measurable real valued functions f on Ω (i.e. random variables). The probability measure $p(f, \rho, E) \equiv \rho_f(E)$ is given by

$$\rho_f(E) = \int_{f^{-1}(E)} \rho d\mu. \quad (13)$$

The value of $T(\rho_1, \rho_2)$ can be easily found by applying Theorem 1:

Proposition 1. *In the case of classical statistical mechanics the value of $T(\rho_1, \rho_2)$ is given by the formula*

$$T(\rho_1, \rho_2) = \left(\int_{\Omega} \sqrt{\rho_1 \rho_2} d\mu \right)^2.$$

Proof. Let the random variable f represent an observable whose spectrum consists of the points $\lambda_1, \lambda_2, \dots, \lambda_n$. This means simply that f takes on only the values λ_i , so that $f^{-1}(\lambda_i), i = 1, 2, \dots, n$ is a partition of Ω . Consequently, (1) and (13) yield

$$T_f^{1/2}(\rho_1, \rho_2) = \sum_i \left(\int_{f^{-1}(\lambda_i)} \rho_1 d\mu \cdot \int_{f^{-1}(\lambda_i)} \rho_2 d\mu \right)^{1/2}. \quad (14)$$

Since f can be chosen freely, so can the partition $f^{-1}(\lambda_i)$. Accordingly (14) and Theorem 1 entail

$$T^{1/2} = \inf_{\{B_j\}} \sum_j \left(\int_{B_j} \rho_1 d\mu \cdot \int_{B_j} \rho_2 d\mu \right)^{1/2}, \quad (15)$$

the infimum being taken over all partitions. On the other hand, one has by the Schwartz inequality

$$\int_{\Omega} \sqrt{\rho_1 \rho_2} d\mu \leq \sum_j \left(\int_{B_j} \rho_1 d\mu \cdot \int_{B_j} \rho_2 d\mu \right)^{1/2},$$

which together with (15) and Lemma 2 imply the proposition. ■

We come now to quantum theory. The states of a quantal system are represented by density operators W (i.e. positive operators of unit trace) in a Hilbert space

\mathcal{H} , while the observables are represented by self-adjoint operators A . The measure $p(A, W, E)$ is given by

$$p(A, W, E) = \text{Tr}(F_E^A W), \quad (16)$$

where F_E^A is the projection-valued measure which corresponds to the spectral decomposition of A . We now give for $T(\alpha, \beta)$ a much simpler expression than that given by its definition (1) + (2), holding in the quantum case.

Proposition 2. *In the case of Quantum Mechanics, Cantoni's "generalized transition probability" between two states represented by the density operators W_1, W_2 is given by*

$$T^{1/2}(W_1, W_2) = \inf_{\{\psi_i\}_{i \in N}} \sum_{i \in N} ((\psi_i, W_1 \psi_i)(\psi_i, W_2 \psi_i))^{1/2}, \quad (17)$$

the infimum being taken over all bases $\{\psi_i\}_{i \in N}$ of the Hilbert space \mathcal{H} .

Proof. Theorem 1 shows that for any $\varepsilon > 0$, there exists an operator A with purely discrete spectrum $\lambda_1, \lambda_2, \dots$ such that $0 < T_A^{1/2} - T^{1/2} < \varepsilon$, or equivalently

$$0 \leq \sum_k (\text{Tr}(W_1 F_{\{\lambda_k\}}^A) \text{Tr}(W_2 F_{\{\lambda_k\}}^A))^{1/2} - T^{1/2} < \varepsilon. \quad (18)$$

Let $\{x_{k_i}\}_{k_i \in I_k}$ be a basis of the subspace onto which $F_{\{\lambda_k\}}^A$ projects.

Then $\bigcup_k \{x_{k_i}\}_{k_i \in I_k}$ is a basis of \mathcal{H} and

$$\begin{aligned} \sum_k (\text{Tr}(W_1 F_{\{\lambda_k\}}^A) \text{Tr}(W_2 F_{\{\lambda_k\}}^A))^{1/2} &= \sum_k \left(\sum_{I_k} (x_{k_i}, W_1 x_{k_i}) \sum_{I_k} (x_{k_i}, W_2 x_{k_i}) \right)^{1/2} \\ &\geq \sum_k \sum_{I_k} ((x_{k_i}, W_1 x_{k_i})(x_{k_i}, W_2 x_{k_i}))^{1/2}. \end{aligned} \quad (19)$$

On the other hand, if B is any self-adjoint operator having a purely discrete nondegenerate spectrum and all x_{k_i} as eigenvectors, we deduce from (1) and (2)

$$\sum_k \sum_{I_k} ((x_{k_i}, W_1 x_{k_i})(x_{k_i}, W_2 x_{k_i}))^{1/2} = T_B^{1/2}(W_1, W_2) \geq T^{1/2}(W_1, W_2). \quad (20)$$

The proposition follows by combining (18), (19) and (20). ■

The result of the forgoing proposition is considerably simplified if one of the two states is pure, i.e. represented by a unit vector g or, equivalently, by the projector P_g on g .

Proposition 3. *Cantoni's "generalized transition probability" between a pure state P_g and a state W equals (g, Wg) .*

Proof. For any two states W_1, W_2 and any basis $\{f_n\}_{n \in N}$, one has

$$\begin{aligned} \left(\sum_n ((f_n, W_1 f_n)(f_n, W_2 f_n))^{1/2} \right)^2 &= \sum_{n,m} ((f_n, W_1 f_n)(f_m, W_1 f_m)(f_n, W_2 f_n)(f_m, W_2 f_m))^{1/2} \\ &\geq \sum_{n,m} |(f_n, W_1 f_m)(f_n, W_2 f_m)| \\ &\geq \left| \sum_{n,m} (W_1 f_n, f_m)(f_m, W_2 f_n) \right| = \text{Tr}(W_1 W_2). \end{aligned}$$

Combining this result with Proposition 2, we deduce that

$$\forall W_1, W_2 : T(W_1, W_2) \cong \text{Tr}(W_1 W_2). \quad (21)$$

For $W_1 = W, W_2 = P_g$, we obtain from (21)

$$T(W, P_g) \cong (g, Wg). \quad (22)$$

On the other hand, it can be easily verified that

$$(g, Wg) = T_{P_g}(W, P_g) \cong T(W, P_g) \quad (23)$$

which, together with (22), proves the proposition. ■

References

1. Mackey, G. : The Mathematical foundations of Quantum Mechanics. New York : W. A. Benjamin Inc., 1963
2. Cantoni, V. : Commun. Math. Phys. **44**, 125 (1975)
3. Cantoni, V. : Commun. Math. Phys. **56**, 189 (1977)
4. Gudder, S. : Commun. Math. Phys. **63**, 265 (1978)
5. Lang, S. : Real analysis Reading, Mass. : Addison-Wesley, 1973

Communicated by R. Haag

Received February 10, 1981 ; in revised form May 24, 1981