

TWO NON-KOLMOGOROVIAN GENERALIZATIONS OF REICHENBACH'S COMMON CAUSE DEFINITION*

Gábor SZABÓ

Department of Philosophy¹
Technical University of Budapest
H-1521 Budapest, Hungary
phone: +36 1 463-1181, fax: +36 1 463-1042
E-mail: gszabo@hps.elte.hu

Received: Oct. 14, 1996

Abstract

Given a probabilistic correlation between two events, this correlation might be explained in terms of a common cause. In [1] Reichenbach defines the notion of common cause and shows that the definition is consistent with the explicable correlation, i.e. if two events have a common cause then they do correlate. In this paper we generalize the notion of common cause to Hilbert lattices in two different ways according to the two different definitions of conditional probability in the quantum case, and show that Reichenbach's theorem does not hold in either case. There will be given counter-examples when a common cause 'causes' correlation, anticorrelation and independence, respectively.

Keywords: Reichenbach's common cause definition, conditional probability, Hilbert lattice.

1. Introduction

Investigating the connection between time direction and macrostatistics REICHENBACH develops a theory of probabilistic causation in [1]. He distinguishes the two time directions by two types of so called conjunctive fork both constructed of two correlating events A and B and of a third event C regarded as the common cause or of an event E regarded as the common effect of the correlation, respectively. The ACB and the AEB forks are open toward the future or the past, respectively, so by their help the time direction can be defined. In this paper we do not investigate the physical and philosophical problem arising in the evolution of the causal theory of time (see [2]), we rather turn our attention to one of the key concepts of the theory, namely the concept of common cause.

¹This paper is the outline of two forthcoming papers submitted to the *Int. Journ. of Theor. Phys.* under the title *Reichenbach's Common Cause Definition on Hilbert lattices* and to the *Found. of Phys.* under the title *Can Reichenbach's Common Cause Definition be Generalized to Non-commutative Event Structures?*

Reichenbach gives several examples of how a correlation between two events can be explained by means of a common cause. *'Suppose both lamps in a room go out suddenly. We regard it as improbable that by chance both bulbs burned out at the same time and look for a burned out fuse or some other interruption of the common power supply. The improbable coincidence is thus explained as the product of a common cause... Or suppose several actors in a stage play fall ill showing symptoms of food poisoning. We assume that the poisoned food stems from the same source – for instance, that it was contained in a common meal – and then look for an explanation of the coincidence in terms of a common cause.'*

Reichenbach defines common cause in the following way. Let A and B be two events which happen simultaneously more frequently than can be expected for chance coincidences, that is

$$p(AB) > p(A)p(B) . \quad (1)$$

In order to explain this correlation, let us assume that there exists a common cause C . We introduce the assumption that the fork ACB satisfies the following relations:

$$p(AB|C) = p(A|C)p(B|C) , \quad (1)$$

$$p(AB|\bar{C}) = p(A|\bar{C})p(B|\bar{C}) , \quad (3)$$

$$p(A|C) > p(A|\bar{C}) , \quad (4)$$

$$p(B|C) > p(B|\bar{C}) . \quad (5)$$

We denote by $p(\cdot|C)$ and $p(\cdot|\bar{C})$ probabilities conditioned on C and non- C , respectively. (2) and (3) express the so called 'screening off' property of the common cause: if the probabilities of the correlating events are conditioned on the common cause or on its complement then they become independent. Or, in other words, the common cause C is the connecting link which transforms independence into dependence. (4) and (5) establish asymmetry between C and \bar{C} : they express that C makes A and B more frequent than \bar{C} . In this sense the cause of the correlation is C and not \bar{C} .

Let us see two quantitative examples. Let there be ten balls in a box, nine painted white, made of wood and one painted black, made of plastic. The probability of pulling a ball made of artificial material $p(a)$, pulling a black ball $p(b)$, or pulling a ball which is black and plastic $p(ab)$ equals $1/10$. So there is a correlation between these events:

$$\frac{1}{10} = p(ab) > p(a)p(b) = \frac{1}{100} .$$

What can be regarded as a common cause in this case? The common cause c which screens off the correlation is the event of pulling the black, plastic ball. The conditional probabilities are the following:

$$p(ab|c) = 1, \quad p(a|c) = 1, \quad p(b|c) = 1,$$

$$p(ab|\bar{c}) = 0, \quad p(a|\bar{c}) = 0, \quad p(b|\bar{c}) = 0,$$

and so they satisfy (2)–(5).

Let the other example be a set of experiments. There are two coins on the right hand side of a table, two coins on the left hand side and a die in the middle. In every run first we throw the dice. If the result is an even number, we take the first coin on both sides and throw them up. If the result is an odd number, we take the second coin on both sides and throw these up. We register the results and repeat the experiment. The possibilities are the following:

Table 1

DIE			
Even		Odd	
Right	Left	Right	Left
T_R^1	T_L^1	T_R^2	T_L^2
T_R^1	H_L^1	T_R^2	H_L^2
H_R^1	T_L^1	H_R^2	T_L^2
H_R^1	H_L^1	H_R^2	H_L^2

Let $p(H_R^{1,2})$ and $p(T_R^{1,2})$ denote the relative frequency of getting head or tail, respectively, by the first or the second coin, respectively, on the right side. Let $p(H_L^{1,2})$ and $p(T_L^{1,2})$ denote the same situation on the left side. Let the two events in question be that getting head by the first coin on the different sides. The relative frequencies $p(H_R^1)$ and $p(H_L^1)$ equal $1/4$. The relative frequency of the joint event $p(H_R^1, H_L^1)$ equals $1/8$. So there is a correlation between these events:

$$\frac{1}{8} = p(H_R^1, H_L^1) > p(H_R^1)p(H_L^1) = \frac{1}{16}.$$

What is the common cause in this case? C is the event that we throw an even number by the dice. The probabilities conditioned on the common cause and its complement are the following:

$$p(H_R^1, H_L^1|C) = \frac{1}{4}, \quad p(H_R^1|C) = \frac{1}{2}, \quad p(H_L^1|C) = \frac{1}{2},$$

$$p(H_R^1, H_L^1 | \bar{C}) = 0, \quad p(H_R^1 | \bar{C}) = 0, \quad p(H_L^1 | \bar{C}) = 0,$$

which satisfy again (2)–(5).

It is worth seeing how the definition of common cause contains direct causation, i.e. when the cause of the correlation of the events A and B is not a third event C but A or B itself, respectively. Both A and B written in the place of C satisfy (2) and (3). Furthermore (4) is fulfilled for $C = A$, and (5) for $C = B$. So the requirements for A or B to be the common cause are reduced to the inequalities:

$$p(B|A) > p(B|\bar{A}),$$

$$p(A|B) > p(A|\bar{B}),$$

respectively.

From now on we take this definition of the common cause for granted and turn our attention from its physical motivation to its mathematical structure.

2. The Classical Case

Let (i) (Ω, F, p) be a Kolmogorovian probability measure space and let (ii) the conditional probability of E given F be defined as it is usual:

$$p(E|F) = \frac{p(E \cap F)}{p(F)}.$$

Let $A, B \in \Omega$ be two correlating events, i.e.

$$p(A \cap B) > p(A)p(B). \quad (6)$$

Reichenbach defines the common cause of the correlation as follows:

Definition An event C is said to be the *common cause* of the correlation between A and B if the events A , B and C satisfy the following relations:

$$p(A \cap B|C) = p(A|C)p(B|C), \quad (7)$$

$$p(A \cap B|\bar{C}) = p(A|\bar{C})p(B|\bar{C}), \quad (8)$$

$$p(A|C) > p(A|\bar{C}), \quad (9)$$

$$p(B|C) > p(B|\bar{C}). \quad (10)$$

Now we do not investigate the question under what conditions a common cause satisfying (7)–(10) exists. We rather turn our attention to the question whether the existence of a common cause really yields correlation. The answer is given by the following

Theorem (Reichenbach, 1956) *Let A, B and C be elements of a Kolmogorovian probability measure space and let them satisfy (7)–(10). Then A and B correlate, i.e. they satisfy (6).*

Proof In the proof we use the following three equations:

$$(\alpha) \quad p(A) = p(C)p(A|C) + p(\bar{C})p(A|\bar{C}),$$

$$(\beta) \quad p(B) = p(C)p(B|C) + p(\bar{C})p(B|\bar{C}),$$

$$(\gamma) \quad p(A \cap B) = p(C)p(A|C)p(B|C) + p(\bar{C})p(A|\bar{C})p(B|\bar{C}).$$

(α) and (β) are identities in a Kolmogorovian probability measure space, (γ) is true if (7)–(8) are true. From these relations we find by some simple computations that

$$p(A \cap B) - p(A)p(B) = p(C)p(\bar{C})[p(A|C) - p(A|\bar{C})][p(B|C) - p(B|\bar{C})].$$

Because of (9)–(10) and under the assumption $0 < p(C) < 1$, we get that $p(A \cap B) - p(A)p(B) > 0$, which was to be proven.

Finally, we list some relations following from (7)–(8) and (α)–(γ) which show how the common cause increases the probability of happening A, B and $A \cap B$:

$$\begin{array}{rcccl} p(A|C) & > & p(A) & > & p(A|\bar{C}), \\ p(B|C) & > & p(B) & > & p(B|\bar{C}), \\ p(A \cap B|C) & > & p(A \cap B) & > & p(A \cap B|\bar{C}). \end{array}$$

These equations together with the derivability of the correlation from the existence of the common cause show the power of the definition in the classical case. But let us go over to the quantum case!

3. First Generalization

Let (i) $P(H)$ be a Hilbert lattice and W be a pure state represented by the unit vector w . For the projections E and F in the lattice let (ii) the conditional probability of E given F in a state W be defined in the following way:

$$p_w(E|F) = \frac{p_w(E \wedge F)}{p_w(F)} = \frac{\text{Tr}(W(E \wedge F))}{\text{Tr}(WF)}.$$

(Now we disregard the logical and mathematical difficulties arising from this generalization of the Bayes rule.) Let $A, B \in P(H)$ and assume a

correlation between A and B in the state W , i.e.

$$p_w(A \wedge B) > p_w(A)p_w(B). \quad (11)$$

We define now the common cause of the correlation in the quantum case:

Definition An event C is said to be the *common cause* of the correlation between A and B if the events A , B and C satisfy the following relations:

$$p_w(A \wedge B|C) = p_w(A|C)p_w(B|C), \quad (12)$$

$$p_w(A \wedge B|C^\perp) = p_w(A|C^\perp)p_w(B|C^\perp), \quad (13)$$

$$p_w(A|C) > p_w(A|C^\perp), \quad (14)$$

$$p_w(B|C) > p_w(B|C^\perp). \quad (15)$$

Now we show that the analogue of Reichenbach's theorem does not hold in this case. So we claim the following

Theorem Let A, B and C be elements of a Hilbert lattice and let them satisfy (12)–(15). Then A and B can either correlate, i.e. $p_w(A \cap B) > p_w(A)p_w(B)$; or anticorrelate, i.e. $p_w(A \cap B) < p_w(A)p_w(B)$; or be independent, i.e. $p_w(A \cap B) = p_w(A)p_w(B)$.

Proof Let $P(H_3)$ be the projection lattice of the three dimensional real Hilbert space H_3 with the basis $\{x, y, z\}$ (See Fig. 1). Let $RanC$ be the plane xy , $RanC^\perp$ be the axis z , $RanA$ and $RanB$ be two planes intersecting each other in line x , both having an angle α with z . Let w be in the plane xz meeting with z at an angle β .

We claim that for all $\alpha, \beta \in (0, \frac{\pi}{2})$, (12)–(15) are satisfied. The conditional probabilities are the following:

$$p_w(A|C) = \frac{p_w(A \wedge C)}{p_w(C)} = \frac{Tr(W(A \wedge C))}{Tr(WC)} = \frac{\cos^2 \beta}{\cos^2 \beta} = 1,$$

$$p_w(B|C) = \frac{p_w(B \wedge C)}{p_w(C)} = \frac{Tr(W(B \wedge C))}{Tr(WC)} = \frac{\cos^2 \beta}{\cos^2 \beta} = 1,$$

$$p_w(A \wedge B|C) = \frac{p_w(A \wedge B \wedge C)}{p_w(C)} = \frac{Tr(W(A \wedge B \wedge C))}{Tr(WC)} = \frac{\cos^2 \beta}{\cos^2 \beta} = 1,$$

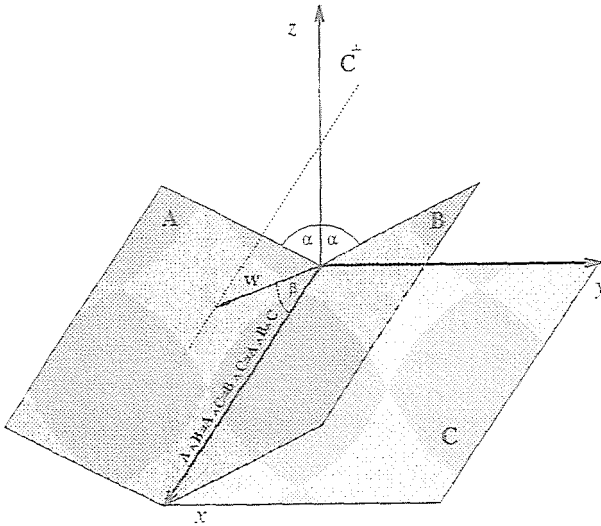


Fig. 1. The projections A, B and C in $P(H_3)$

since w is in the plane xz , so its projection onto the plane xy and the axis x are equal.

$$\begin{aligned}
 p_w(A|C^\perp) &= \frac{p_w(A \wedge C^\perp)}{p_w(C^\perp)} = \frac{\text{Tr}(W(A \wedge C^\perp))}{\text{Tr}(WC^\perp)} = 0, \\
 p_w(B|C^\perp) &= \frac{p_w(B \wedge C^\perp)}{p_w(C^\perp)} = \frac{\text{Tr}(W(B \wedge C^\perp))}{\text{Tr}(WC^\perp)} = 0, \\
 p_w(A \wedge B|C^\perp) &= \frac{p_w(A \wedge B \wedge C^\perp)}{p_w(C^\perp)} = \frac{\text{Tr}(W(A \wedge B \wedge C^\perp))}{\text{Tr}(WC^\perp)} = 0,
 \end{aligned}$$

since the intersection of A, B and $A \wedge B$ with C^\perp are 0-projections. By these numbers Eqs. (12)–(15) are satisfied:

$$\begin{aligned}
 1 &= p_w(A \wedge B|C) = p_w(A|C)p_w(B|C) = 1, \\
 0 &= p_w(A \wedge B|C^\perp) = p_w(A|C^\perp)p_w(B|C^\perp) = 0, \\
 1 &= p_w(A|C) > p_w(A|C^\perp) = 0, \\
 1 &= p_w(B|C) > p_w(B|C^\perp) = 0.
 \end{aligned}$$

So C can be regarded as the common cause of the correlation between A and B by the above definition.

Let us examine whether there exists a correlation between A and B indeed, i.e. whether (11) is satisfied. The two sides of Eq. (11) are the following:

$$p_w(A \wedge B) = \text{Tr}(W(A \wedge B)) = \cos^2 \beta ,$$

$$p_w(A)p_w(B) = \text{Tr}(WA)\text{Tr}(WB) = (\cos^2 \beta + \sin^2 \beta \cos^2 \alpha)^2 .$$

In Fig. 2 we represent the relation between the sides of (11) in the parameter space $(\alpha, \beta)_{0,0}^{\frac{\pi}{2}, \frac{\pi}{2}}$. We can see that the parameter space is divided into two regions by a curve reaching from the line $(0, \alpha)$ to the point $(\frac{\pi}{2}, \frac{\pi}{2})$ representing the places where $p_w(A \cap B) = p_w(A)p_w(B)$, i.e. where the events A and B are independent. The region 'under' the curve represents the places where $p_w(A \cap B) < p_w(A)p_w(B)$, i.e. where the events A and B anticorrelate. Finally, the region 'above' the curve represents the correlating places where $p_w(A \cap B) > p_w(A)p_w(B)$.

So we have found an example where for two events A and B a third event C can be chosen which can be regarded as the common cause, but A and B do not necessarily correlate; they can anticorrelate or be independent.

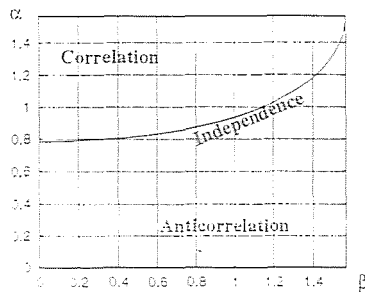


Fig. 2. The parameter space $(\alpha, \beta)_{0,0}^{\frac{\pi}{2}, \frac{\pi}{2}}$

In the next section we take another definition of the common cause on the Hilbert lattice using another definition of the conditional probability and examine the validity of the analogue of Reichenbach's theorem.

4. Second Generalization

Let (i) $P(H)$ be a Hilbert lattice and W be a pure state determined by the unit vector w . For the projections E and F in the lattice let (ii) the

conditional probability of E given F in a state W be defined in the following way:

$$p_w(E|F) = \frac{\text{Tr}(FWFE)}{\text{Tr}(FWF)}.$$

The motivation of this definition comes from the theory of measurement. If we carry out a measurement of an observable represented by the projection F in a pure state W then the state transforms as follows:

$$W \longmapsto \frac{FWF}{\text{Tr}(FWF)}.$$

It can easily be seen that the new state is pure again. Let us introduce the following notation for the new pure state: $W_F \equiv \frac{FWF}{\text{Tr}(FWF)}$. The $W \longmapsto W_F$ transformation can be regarded as the 'renormalized projection' of the state W onto the subspace $\text{Ran}F$. This rule is due to Lüders (see [3], [4]). Using the above notation now we are able to define the common cause in terms of this new conditional probability: Let $A, B \in P(H)$ and let there be a correlation between A and B in the state W , i.e.

$$p_w(A \wedge B) > p_w(A)p_w(B). \quad (16)$$

Definition An event C is said to be the *common cause* of the correlation between A and B if the events A , B and C satisfy the following relations:

$$\text{Tr}(W_C(A \wedge B)) = \text{Tr}(W_C A)\text{Tr}(W_C B), \quad (17)$$

$$\text{Tr}(W_{C^\perp}(A \wedge B)) = \text{Tr}(W_{C^\perp} A)\text{Tr}(W_{C^\perp} B), \quad (18)$$

$$\text{Tr}(W_C A) > \text{Tr}(W_{C^\perp} A), \quad (19)$$

$$\text{Tr}(W_C B) > \text{Tr}(W_{C^\perp} B). \quad (20)$$

Now we ask the question again whether A and B correlate, provided there exists a third event C such that conditions (17)–(20) hold. The answer is again negative.

Theorem Let A, B and C be elements of a Hilbert lattice and let them satisfy (17)–(20). Then A and B do not necessarily correlate.

Proof In the proof we give a rather technical counter-example which satisfies (17)–(20) but does not satisfy (16). Let us take the same three dimensional Hilbert lattice $P(H_3)$ as before with the basis $\{x, y, z\}$ (See Fig. 3). Since in Eqs. (17)–(20) C and C^\perp do not appear explicitly, in the first step

we do not determine these projections; instead we search for two unit perpendicular vectors w_C and $w_{C\perp}$ which satisfy (17)–(20), and at the end we return to the projections. Let $RanA$ and $RanB$ be two planes intersecting each other in x meeting with z at an angle α . By Fig. 2 there exists uniquely a vector v in the plane xz for which $p_v(A \wedge B) = p_v(A)p_v(B)$. Let this vector v be $w_{C\perp}$, so (18) is satisfied.

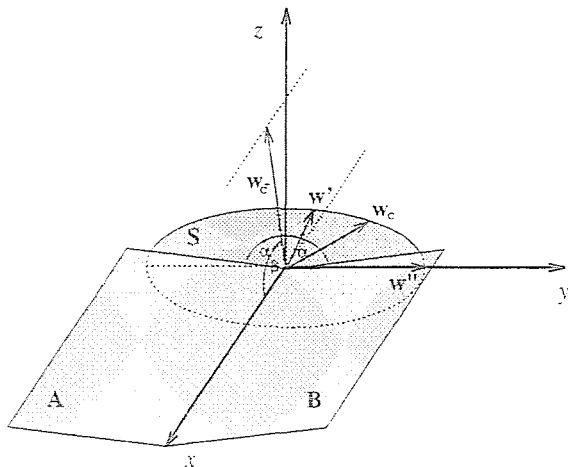


Fig. 3. The position of w_C and $w_{C\perp}$ in $P(H_3)$

Our task is now to find a vector w_C perpendicular to $w_{C\perp}$ so that (17) and (19)–(20) be satisfied. The last two inequalities can be satisfied as follows: Let α tend to $\frac{\pi}{2}$, i.e. let $RanA$ and $RanB$ tend to the plane xy . Then by Fig. 2 β also tends to $\frac{\pi}{2}$, i.e. $w_{C\perp}$ tends to the axis z . Let us denote the plane perpendicular to $w_{C\perp}$ by S . Now this plane is infinitesimally close to the plane xy and to $RanA$ and $RanB$. All from these follows that for arbitrary small ε_1 and ε_2 we can choose a δ so that for any α for which $|\frac{\pi}{2} - \alpha| < \delta$, $p_{w_{C\perp}}(A) = p_{w_{C\perp}}(B) < \varepsilon_1(\delta)$, and for every vector u in the plane S $p_u(A) > 1 - \varepsilon_2(\delta)$, $p_u(B) > 1 - \varepsilon_2(\delta)$. So (19)–(20) are satisfied for every u in S .

Now let us pick out the vector from the plane S which satisfies also (17). Instead of searching for a vector w_C satisfying $p_{w_C}(A \wedge B) = p_{w_C}(A)p_{w_C}(B)$, we pick out two other vectors w' and w'' for which inequalities hold with the opposite sign, i.e. $p_{w'}(A \wedge B) > p_{w'}(A)p_{w'}(B)$ and $p_{w''}(A \wedge B) < p_{w''}(A)p_{w''}(B)$. Let w' be the vector determined by the intersection of the planes xz and S . In Fig. 2 we can see that w' is in the correlating region, so for w' $p_{w'}(A \wedge B) > p_{w'}(A)p_{w'}(B)$. Let the other vector

w'' be determined by the intersection of the planes yz and S which is the axis y itself. For w'' $p_{w''}(A \wedge B) = 0$ since $w'' \perp A \wedge B$, but $p_{w''}(A) \neq 0$ and $p_{w''}(B) \neq 0$, so $p_{w''}(A \wedge B) < p_{w''}(A)p_{w''}(B)$. Now let us use the continuity of the $p_w(\cdot)$ -function on the plane S . If there is a vector w' for which $p_{w'}(A \wedge B) > p_{w'}(A)p_{w'}(B)$ and a vector w'' for which $p_{w''}(A \wedge B) < p_{w''}(A)p_{w''}(B)$, then there must be a vector between them in the plane S for which $p_w(A \wedge B) = p_w(A)p_w(B)$. Let this vector be w_C , so (17) is fulfilled.

So we have found two vectors w_C and w_{C^\perp} for which (17)–(20) are satisfied. What are the projections C and C^\perp , and what is the original w vector? Let C be the projection for which $RanC$ is the plane S , let C^\perp be the projection determined by w_{C^\perp} . Then w can be any of the vectors in the plane T spanned by w_C and w_{C^\perp} except for w' and w'' .

Now let us choose a possible w for which independence or anticorrelation happens. Let w be the vector determined by the intersection of the planes yz and T (see Fig. 4).

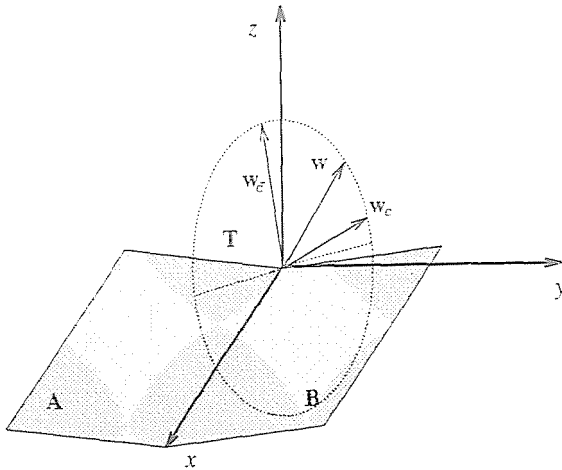


Fig. 4. The position of w in $P(H_3)$

For w $p_w(A \wedge B) = 0$, since w is in the plane yz . Now there are two possibilities: In the case that $p_w(B) = 0$ or $p_w(A) = 0$, then $p_w(A \wedge B) = p_w(A)p_w(B)$, i.e. A and B are independent; in the case that $p_w(B) \neq 0$ and $p_w(A) \neq 0$, then $p_w(A \wedge B) < p_w(A)p_w(B)$, i.e. A and B anticorrelate. So our counter-example satisfies (17)–(20) but not (16) and this was to be proven.

5. Conclusions

In our paper we have generalized the original notion of common cause given by Reichenbach in two different ways. The definitions differed from each other in (i) the type of the probability space and (ii) the definition of conditional probability. The possibilities – included the original one – were the following:

1. (i) (Ω, F, p) is a Kolmogorovian probability measure space,
 (ii) $p(E|F) = \frac{p(E \cap F)}{p(F)}$,
2. (i) $P(H)$ is a Hilbert lattice,
 (ii) $p_w(E|F) = \frac{p_w(E \wedge F)}{p_w(F)}$,
3. (i) $P(H)$ is a Hilbert lattice,
 (ii) $p_w(E|F) = \frac{\text{Tr}(FWFE)}{\text{Tr}(FWF)}$.

We have investigated the question whether the existence of a common cause for two events defined by (2)–(5) implies a correlation between the events. The answer in the first, classical case was affirmative, so it showed the deep consistency of the definition. The meaning of the notion of common cause in the other two, quantum cases is not so obvious since it is although true that correlation can sometimes be explained in terms of a common cause but also independence can sometimes be ‘explained’ by that. So in the quantum case not even the statement is false that we always find a common cause for a correlation (as it can be shown on an appropriate small Hilbert lattice) but also the opposite statement, namely that common cause always ‘leads to’ correlation.

References

1. REICHENBACH, H. (1956): *The Direction of Time*. University of California Press, Berkeley.
2. SALMON, W. C. (1980): Probabilistic Causality, *Pacific Philosophical Quarterly*, Vol. 61, Nos. 1–2., pp. 50–74.
3. LÜDERS, G. (1951): Über die Zustandsänderung durch den Messprozess, *Ann. d. Phys.*, Vol. 8, pp. 332–328.
4. BUB, J. (1979): Conditional Probabilities in non Boolean Possibility Structures, in: *The Logico Algebraic Approach of Quantum Mechanics Vol. II. Contemporary Consolidation*, C.A. Hooker ed., D. Reidel Publishing Co. Dordrecht, Holland.