

Classical three-box 'paradox'

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2003 J. Phys. A: Math. Gen. 36 4891

(<http://iopscience.iop.org/0305-4470/36/17/315>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 171.66.16.96

The article was downloaded on 02/06/2010 at 11:38

Please note that [terms and conditions apply](#).

Classical three-box ‘paradox’

K A Kirkpatrick

New Mexico Highlands University, Las Vegas, NM 87701, USA

E-mail: kirkpatrick@physics.nmhu.edu

Received 10 February 2003

Published 16 April 2003

Online at stacks.iop.org/JPhysA/36/4891

Abstract

A simple classical probabilistic system (a simple card game) classically exemplifies Aharonov and Vaidman’s ‘three-box ‘paradox’’ (1991 *J. Phys. A: Math. Gen.* **24** 2315), implying that the three-box example is neither quantal nor a paradox and leaving one with less difficulty to busy the interpreters of quantum mechanics. An ambiguity in the usual expression of the retrodiction formula is shown to have misled Albert *et al* (1985 *Phys. Rev. Lett.* **54** 5) to a result not, in fact, ‘curious’; the discussion illustrates how to avoid this ambiguity.

PACS numbers: 03.65.Bz, 01.70.+w

1. Introduction

Aharonov and Vaidman (1991; henceforth ‘AV’) introduced what has come to be called the ‘three-box paradox,’ a postselected process in which each of two disjoint events occurs with certainty. They express this example in terms of a particle and three boxes; the process starts with the particle in a state smeared over all three boxes and is postselected to end with the particle in another, similarly smeared, state. These terminal states are chosen so that, if the first of the boxes is opened during the process, the particle is certain to be found there, while if the second box is opened, the particle is found there: a single particle is certain to be found in each of two boxes! Vaidman (1996) sums this up with

The elements of reality for pre- and postselected quantum systems have unusual and counterintuitive properties. But, may be this is not because of the illness of their definition, but due to bizarre feature of quantum systems which goes against the intuition built during thousands of years, when the results of quantum experiments were not known.

But this, I think, is much ado about nothing. I will present an example of this behaviour in a classical-probability setting. (The example is in terms of playing cards, rather than particles in boxes; the paradoxical result will be that the card drawn must at one and the same time be a diamond and a spade.) Because the setting of this system is so ordinary, we cannot be tempted to dismiss its behaviour as another ‘bizarre feature of quantum mechanics’—instead, we are

led to look for the misapprehension which has led us astray. We certainly will not accede to the proposed identification of ‘elements of reality’ with ‘certainty of occurrence’—the elements of reality of this classical system are quite visible and identifiable, not easily conflated with ghosts.

In section 2, I present AV’s three-box example, restate it in terms of the physically less obscure triple-slit atomic Young device and analyse it in classical-probability terms. In section 3, I present a classical system (a deck of cards) with probabilistic properties which allow the construction of a classical three-box system.

Finally, in section 4, I describe how an ambiguity in the presentation by Aharonov, Bergmann and Lebowitz (1964; henceforth ‘ABL’) of the retrodiction formula misled Albert *et al* (1985) to a result they thought ‘curious’ (which, in turn, misled AV to the three-box example).

2. The three-box system and its ‘paradox’

The ‘three-box paradox’ of AV consists of a single particle and three boxes; the value p_j of the observable P denotes the presence of the particle in the corresponding box j . The system is prepared in a certain state s at the time t_0 , and only those occurrences for which the system is detected, at time t_1 , to be in a particular state q are considered. At an intermediate time $t \in (t_0, t_1)$ we may look for the particle by opening a box. AV showed that, with

$$|s\rangle = (|p_1\rangle + |p_2\rangle + |p_3\rangle)/\sqrt{3} \quad \text{and} \quad |q\rangle = (|p_1\rangle + |p_2\rangle - |p_3\rangle)/\sqrt{3} \quad (1)$$

if we open box 1, we (always) find the particle in it; however, they showed, it is also true that if we instead open box 2, we (always) find the particle in it: the retrodictive observation of p_1 is certain, and the retrodictive observation of p_2 is certain—the three-box ‘paradox.’

2.1. Three-slit Young implementation of the three-box system

The reader may find the physical significance of the initial and final states of the three-box system rather obscure, expressed in terms of particles in boxes. These states, as well as the entire system, are much more easily understood expressed as a three-slit atomic diffraction system. Expressing the three-box system in this form emphasizes that its paradoxical behaviour is as ordinary (to the extent that any atomic Young system may be considered ‘ordinary’!) as a double-slit interference apparatus.

Three slits are equally spaced with a separation a . The top and bottom slits are labelled 1 and 2, and the middle slit, 3. A detector D is placed on-axis, at a distance L from slit 3 so that $\sqrt{L^2 + a^2} - L = \lambda/2$ (with λ the wavelength). A detector d is placed at slit 1 or slit 2 (for atoms, a micromaser; for photons, a quarter-wave plate, the photon source linearly polarized). This system’s initial and final states are described by (1), and it behaves exactly as the three-box example. If d is placed at slit 1, every detection at D is in coincidence with a detection at d (implying passage through slit 1); if d is placed at slit 2, every detection at D is in coincidence with a detection at d (implying passage through slit 2). This behaviour is easily understood: placing a detector at slit 1 creates the disjunction ‘either the particle passed through slit 1 or it passed through the double-slit apparatus comprising slits 2 and 3’; but passage through the double-slit destructively interferes at the detector D , so the second term of the disjunction must be false, forcing the first to be true. This apparatus is symmetric under exchange of slits 1 and 2—hence the ‘paradox.’

The three slits correspond to the three boxes; a detector at slit 1 (only) corresponds to the opening of box 1 (only). The ‘paradox’ of the three-box system reduces to the phenomenon (paradox?) of destructive interference in a two-slit Young apparatus, nothing more.

2.2. *Classical-probability derivation of the three-box retrodiction formula*

Let us now derive the expression for the retrodictive probability applicable to the three-box example. We carry out the derivation in classical-probability terms to avoid being misled by any quantum ‘bizarreness.’ (See appendix A for notational matters.)

The observation of the contents of a single box is a *partial* observation: the opening of box 1, for example, determines ‘ $p_1 \vee p_1^\sim$,’ where p_1 denotes ‘particle in box 1’ and p_1^\sim denotes ‘particle not in box 1.’ Note that p_1^\sim is not the same as $p_2 \vee p_3$; p_1^\sim does not signify the ignorance of ‘either p_2 or p_3 , but we do not know which,’ but rather signifies the lack of a fact of the matter regarding these two possibilities¹. The condition of such a partial observation is the *partial manifestation*

$$M_{p_j}^\sim \equiv p_j \vee p_j^\sim \tag{2}$$

of ‘in box j ’ and ‘not in box j ,’ exactly *one* is true.

Accounting for this manifestation, and carefully labelling each event’s position in the sequence with a bracketed superscript ordinal (see appendix A), the expression for the retrodictive probability of finding the particle when box j is opened is

$$\Pr_{s^{[0]}}(p_j^{[1]} \mid M_{p_j}^{\sim [1]} \wedge q_k^{[2]}) = \frac{\Pr_{s^{[0]}}(p_j^{[1]} \wedge q_k^{[2]} \mid M_{p_j}^{\sim [1]})}{\Pr_{s^{[0]}}(q_k^{[2]} \mid M_{p_j}^{\sim [1]})} \tag{3}$$

where we have used (A.1). (Note that, were the manifestation not expressed explicitly, the denominator would be ambiguous as to the identity or nature of event^[1]).

Because the condition $p_j \wedge M_{p_j}^\sim = p_j$, the numerator is

$$\Pr_{s^{[0]}}(p_j^{[1]} \wedge q_k^{[2]} \mid M_{p_j}^{\sim [1]}) = \Pr_{s^{[0]}}(q_k^{[2]} \mid p_j^{[1]})\Pr_{s^{[0]}}(p_j^{[1]}). \tag{4}$$

Further, the condition $M_{p_j}^\sim \wedge (p_j \vee p_j^\sim) = M_{p_j}^\sim$; thus the denominator may be written as

$$\begin{aligned} \Pr_{s^{[0]}}(q_k^{[2]} \mid M_{p_j}^{\sim [1]}) &= \frac{\Pr_{s^{[0]}}((p_j \vee p_j^\sim)^{[1]} \wedge q_k^{[2]} \mid M_{p_j}^{\sim [1]})}{\Pr_{s^{[0]}}((p_j \vee p_j^\sim)^{[1]})} \\ &= \Pr_{s^{[0]}}(p_j^{[1]} \wedge q_k^{[2]} \mid M_{p_j}^{\sim [1]}) + \Pr_{s^{[0]}}(p_j^{\sim [1]} \wedge q_k^{[2]} \mid M_{p_j}^{\sim [1]}) \\ &= \Pr_s(q_k \mid p_j) \Pr_s(p_j) + \Pr_s(q_k \mid p_j^\sim) \Pr_s(p_j^\sim) \end{aligned} \tag{5}$$

where we have used equation (4) as well as the disjointness and completeness of $\{p_j, p_j^\sim\}$. Thus the retrodictive probability equation, (3), may be written as

$$\Pr_{s^{[0]}}(p_j^{[1]} \mid M_{p_j}^{\sim [1]} \wedge q_k^{[2]}) = \frac{\Pr_s(q \mid p_j) \Pr_s(p_j)}{\Pr_s(q \mid p_j) \Pr_s(p_j) + \Pr_s(q \mid p_j^\sim) \Pr_s(p_j^\sim)}. \tag{6}$$

2.3. *The three-box paradox*

If the second term of the denominator of (6) were to vanish, the probability would be 1; if there is a choice of s and q such that the term were to vanish for more than one value of j (e.g. for both $M_{p_1}^\sim$ and $M_{p_2}^\sim$), we would obtain a three-box ‘paradox.’ In the succeeding section we will present a classical system which exhibits exactly this behaviour.

¹ Aristotle gave this example: if P is the proposition ‘Tomorrow there will be a sea battle,’ then $P \vee P^\sim$ is true, but neither P nor P^\sim has a truth value today—there is no fact of the matter regarding either.

The three-box result requires *partial* manifestation; if, instead, we make a complete observation $M_P = p_1 \vee p_2 \vee p_3$ (e.g., by looking into at least two boxes), then the manifestation is complete,

$$\Pr_{s^{[0]}}(p_j^{[1]} | M_P^{[1]} \wedge q_k^{[2]}) = \frac{\Pr_s(q_k | p_j) \Pr_s(p_j)}{\sum_{t=1}^3 \Pr_s(q_k | p_t) \Pr_s(p_j)}. \quad (7)$$

From this it is clear that $\Pr_{s^{[0]}}(p_j^{[1]} | M_P^{[1]} \wedge q^{[2]}) = 1$ is not possible for more than one j , no matter the choice of s and q .

2.4. The quantal three-box paradox

If the particle is first placed in box 1 and then box 2 is opened (of course it is not there), it should be that if box 3 were next opened, the particle would not be there, but if, instead, box 1 were next opened, the particle *would* be found. This is the requirement of *stability*, (B.1): the act of looking in box 2 must not move the particle out of box 1. Thus (B.4) applies to $M_{\tilde{p}_j}$; applying equations (B.2) and (B.4) to (6), we obtain the latter in its quantum mechanical form

$$\Pr_{s^{[0]}}(p_j^{[1]} | M_{\tilde{p}_j}^{[1]} \wedge q^{[2]}) = \frac{|\langle q | p_j \rangle|^2 |\langle p_j | s \rangle|^2}{|\langle q | p_j \rangle|^2 |\langle p_j | s \rangle|^2 + |\sum_{t \neq j} \langle q | p_t \rangle \langle p_t | s \rangle|^2}. \quad (8)$$

(This is a specialization of ABL's (2.4) and (2.5), which, on p 1413, they extended to incomplete measurements. AV treat it as a new result, their (5).)

If $|s\rangle$ and $|q\rangle$ are such that

$$\langle q | p_1 \rangle \langle p_1 | s \rangle = \langle q | p_2 \rangle \langle p_2 | s \rangle = -\langle q | p_3 \rangle \langle p_3 | s \rangle \quad (9)$$

(such as equation (1), for example), then the retrodictive observation of p_1 is certain, and the retrodictive observation of p_2 is certain—the three-box ‘paradox.’

2.5. The quantal retrodiction (ABL) formula

Using $\Pr_s(y | x) = \Pr_x(y) = |\langle y | x \rangle|^2$ to express (7) in quantum mechanical terms, we obtain the retrodiction formula for a single complete intermediate observation

$$\Pr_{s^{[0]}}(p_j^{[1]} | M_P^{[1]} \wedge q_k^{[2]}) = \frac{|\langle q_k | p_j \rangle|^2 |\langle p_j | s \rangle|^2}{\sum_t |\langle q_k | p_t \rangle|^2 |\langle p_t | s \rangle|^2}. \quad (10)$$

(This is the result ABL stated in their (2.4) and (2.5), specialized to a single intermediate complete observation.)

3. A classical system with pseudo-quantal properties

I present here a strictly non-quantal system which has many statistical properties similar to quantum mechanics, and which, in particular, allows the construction of a three-box ‘paradox.’ (I discuss in greater depth the use of such classical-probability systems in Kirkpatrick (2003).) Constructed using ordinary playing cards, this system is as distinct from quantum mechanics as is possible. Each card carries two marks, the ‘face’ and the ‘suit’ (traditionally with names such as King, Queen, Jack, 10, . . . and spades, hearts, diamonds, clubs, respectively); these marks will be treated as system variables, *Face* (with values K, Q, J) and *Suit* (with values S, D, H). I will refer to these variables generically as P and Q : $P, Q \in \{\text{Face, Suit}\}$, $P \neq Q$, with values $\{p_j\}$ and $\{q_k\}$, respectively.

3.1. The classical system

The system is a deck of playing cards, each card marked with a *Face* value and a *Suit* value, and a memory containing the name (*not* the value) of the variable of the preceding observation; the content of the memory is denoted \mathcal{M} . The deck is divided into two parts, which we call *Active* and *Passive*.

To prepare the system in the state $P = p_j$:

- (i) Place all cards with $P = p_j$ in *Active* and the remainder of the deck in *Passive*.
 - (ii) Set the memory to the variable name: $\mathcal{M} \leftarrow P$.
- (11a)

To observe the variable P :

If $\mathcal{M} = P$ (i.e. the observation of P is being repeated)

- (i) Select a card at random from *Active*.
 - (ii) Report p_j the value of the variable P of this card.
- else ($\mathcal{M} \neq P$ —the preceding observation was not of P)
- (11b)

- (i) Select a card at random from *Passive*.
- (ii) Report p_j , the value of the variable P of this card.
- (iii) Prepare the system in the state $P = p_j$ (as (11a)).

The variables each take on the same number of values, V . Duplicate cards are allowed under the restriction that each value of each variable appears N times in the deck (so their *a priori* probabilities are equal). For example, we might use the deck $\{(2)KS, KH, QS, QH\}$ (the ‘(2)’ indicates *two* KS cards): $V = 2$ and $N = 3$.

3.2. The incomplete observation

In order to create a classical three-box-type system, we must make a partial, or incomplete, observation. Fortunately, the partial observation $M_{p_j}^{\sim} = p_j \vee p_j^{\sim}$ is easily implemented in our example system: to prepare the system in the state p_j^{\sim} , we simply follow the instructions literally, placing every card with P -value p_j^{\sim} (i.e., every card which satisfies $P \neq p_j$) in *Active*, and all the others (all of which satisfy $P = p_j$) in *Passive*. The observation of P under the manifestation rule $M_{p_j}^{\sim}$ is accomplished exactly as before: we report ‘ p_j ’ if the card’s value of P is p_j , and ‘ p_j^{\sim} ’ if the card’s value of P is *not* p_j ; for the purpose of the test ‘ $\mathcal{M} = P$,’ a variable is considered the same variable whether partially or fully observed.

Note from (B.1) that $M_{p_j}^{\sim}$ is stable (as we required of the quantal manifestation). Formulae for this system’s probabilities are given in appendix C.

3.3. The classical three-box ‘paradoxical’ system

We can now create a classical ‘three-box paradox’ example: take the deck to be $\{(2)KH, QS, QD, JD, JS\}$ ($V = 3, N = 2$). We will prepare the system in the state $Face = Q$, and filter to (accept only those processes which end in) the final state $Face = K$. Let $Suit = S$ correspond to box 1, and $Suit = D$ correspond to box 2; thus ‘opening box 1 (only)’ corresponds to the manifestation $M_S^{\sim} = S \vee S^{\sim}$ (‘is the card a S or not’), and ‘opening box 1 (only)’ corresponds to the manifestation $M_D^{\sim} = D \vee D^{\sim}$ (‘is the card a D or not’). This leads to an exact analogue to the three-box example, but in terms of a deck of cards.

Let us express the deck in the notation [*Active*, |*Passive*]; then preparation of the state Q leaves the deck as [QS, QD | (2)KH, JD, JS]. The partial manifestation M_S^{\sim} then leads, with probability 1/4, to the occurrence of S, [QS, JS | (2)KH, QD, JD], and, with probability 3/4, to the occurrence of S^{\sim} , [(2)KH, QD, JD | QS, JS]. Clearly $\Pr_Q(K | S^{\sim}) = 0$, so, by (6),

$\Pr_{Q^{[0]}}(S^{[1]} | M_S^{[1]} \wedge K^{[2]}) = 1$, and the card is certain to be a S. A parallel analysis of the partial manifestation $M_D^{[1]}$ leads to $\Pr_{Q^{[0]}}(D^{[1]} | M_D^{[1]} \wedge K^{[2]}) = 1$, so the card is certain to be a D. Thus, in this postselected process, at the intermediate point the card is *both* S and D, each with certainty—the ‘three-box paradox.’

3.4. Classical interference shown by p_j^{\sim}

The ‘paradox’ arises because both $\Pr_Q(K | S^{\sim})$ and $\Pr_Q(K | D^{\sim})$ vanish. The vanishing of these terms is not trivial: for example, S^{\sim} would seem to include D, but $\Pr_Q(K | D^{\sim}) \neq 0$.

The manifestation $M_{S^{[1]}} = S \vee H \vee D$ leads to S (as before), with a probability of 1/4, to H, [(2)KH | QS, QD, JS, JD], with probability 1/2, and to D, [QD, JD | (2)KH, QS, JS], with probability 1/4. In this case ‘not S’ is ‘ $H \vee D$,’ the mixture $\{(H, 2/3), (D, 1/3)\}$; this is easily seen to be [(4)KH, QD, JD | (2)KH, (3)QS, (2)QD, (3)JS, (2)JD] (with probability 3/4). Clearly, the mixture corresponding to $H \vee D$ differs from the pure state S^{\sim} .

It seems to be generally assumed that in classical probability there could be no difference between p_1^{\sim} and $p_2 \vee p_3 \vee \dots$; on the other hand, such a difference is known to occur in quantum mechanics, where it is called ‘interference,’ and is generally discussed as a mystery specific to quantum mechanics. However, the example presented here demonstrates interference in a classical system: the rules (11a) and (11b) show us that, for every state s , $\Pr_s(S^{\sim}) = \Pr_s(H \vee D)$, suggesting that $S^{\sim} = H \vee D$; however (for the deck of this example) $\Pr_Q(K | S^{\sim}) = 0 \neq \Pr_Q(K | H \vee D)$, so $S^{\sim} \neq H \vee D$. Because quantum interference exhibits exactly this probability behaviour, by analogy we say that H and D interfere.

If two or three of the ‘boxes’ are examined (i.e. the complete observation $S \vee H \vee D$ is made), so the results are governed by the probabilities $\Pr_{Q^{[0]}}(p_j | M_P^{[1]} \wedge K^{[2]})$, then no p_j has a retrodicted probability of 1. The ‘three-box paradox’ is an interference effect and that interference is destroyed by the facts-of-the-matter established by examining any two of the boxes—in the original quantal three-box setting, in the three-slit interference form (whether atoms or classical waves) *and* in the classical card game.

4. Not-so-curious statistics

The three-box example grew out of the ‘curious’ example of Albert, Aharonov and D’Amato (1985) (henceforth ‘AAD’); in that work, a pre- and postselected system is introduced with a contextual behaviour which, the authors claim, contradicts the theorems of Gleason and Kochen and Specker.

This conclusion is unwarranted; it arises out of ambiguity regarding the nature, complete or partial, of the intermediate manifestation.

AAD’s system has an observable X with eigenstates $\{|x_j\rangle\}$ an observable A with an eigenstate $|a\rangle = (|x_1\rangle + |x_2\rangle)/\sqrt{2}$ and an observable B with an eigenstate $|b\rangle = (|x_2\rangle + |x_3\rangle)/\sqrt{2}$. The process is preselected at the time t_0 for $A = a$ and postselected at t_2 for $B = b$. According to the ABL retrodiction formula, an observation of X at $t_1 \in (t_0, t_2)$ yields $X = x_2$ with certainty; thus $\Pr(X = x_1) = \Pr(X = x_3) = 0$ at t_1 . (Hence, the authors claim, $A = a$, $B = b$ and $X = x_2$ are ‘simultaneously well defined’ throughout the interval (t_0, t_2) , though they are all pairwise incompatible.) A fourth variable, Q , is introduced, with eigenstates

$$|q_1\rangle = \alpha|x_1\rangle + \beta|x_3\rangle \quad |q_2\rangle = |x_2\rangle \quad |q_3\rangle = \beta^*|x_1\rangle - \alpha^*|x_3\rangle. \quad (12)$$

Now, according to AAD,

something very curious arises. Suppose that Q is observed within the interval (t_0, t_2) . It might be expected, since $X = x_2$ within that interval, and since $|q_2\rangle = |x_2\rangle$, that such an observation will find, with certainty, that $Q = q_2$. But that is not so. . .

Well, whether that is so or not depends entirely on how Q and X are to be observed. These are different variables, but they share an eigenstate ($|x_2\rangle = |q_2\rangle$); because $x_2 = q_2$, also $x_2 \sim q_2$. Thus one would expect that observing $M_{Q_2} \equiv q_2 \vee q_2 \sim$ would yield the same result for q_2 as observing $M_{X_2} \equiv x_2 \vee x_2 \sim$ would yield for x_2 . And this is exactly right: applying (8) to this situation,

$$\Pr_{a^{[0]}}(x_2^{[1]} | M_{X_2}^{[1]} \wedge b^{[2]}) = 1 = \Pr_{a^{[0]}}(q_2^{[1]} | M_{Q_2}^{[1]} \wedge b^{[2]}). \tag{13}$$

On the other hand, $q_1 \vee q_3$ is physically different from $x_1 \vee x_3$, so the result for q_2 when observing $M_Q \equiv q_1 \vee q_2 \vee q_3$ may well differ from the result for x_2 when observing $M_X \equiv x_1 \vee x_2 \vee x_3$. Again, this is the case; applying (10), we find

$$\Pr_{a^{[0]}}(x_2^{[1]} | M_X^{[1]} \wedge b^{[2]}) = 1 \quad \text{but} \quad \Pr_{a^{[0]}}(q_2^{[1]} | M_Q^{[1]} \wedge b^{[2]}) < 1. \tag{14}$$

AAD inappropriately used the complete observation M_Q rather than the partial observation M_{Q_2} ; this led them to (14) rather than to (13), and hence to an unwarranted sense of ‘curiousness’: AAD continue

But that is not so: Albeit $\langle a | x_3 \rangle = \langle b | x_1 \rangle = 0$, yet $\langle a | q_1 \rangle \neq 0$ and $\langle b | q_1 \rangle \neq 0$. Consequently, albeit $\Pr(x_1) = 0$ and $\Pr(x_3) = 0$ within that interval, $\Pr(q_1) \neq 0$ there.

An ambiguity in the ABL expression for the retrodiction probability led AAD to this confusion: what we have expressed as $\Pr_{s^{[0]}}(p_j^{[1]} | M_p^{[1]} \wedge q_k^{[2]})$, ABL denoted $p(p_j/s, q_k)$ (and AAD denoted merely $P(p_j)$), leaving out any mention of the details of the manifestation at the intermediate observation. But, as we have just seen, it is necessary to take into account the degree of completeness of the manifestation, *even of those values not under direct consideration*.

The obvious lesson: use a complete, unambiguous notation which cannot fail to call attention to this need.

5. The values of a pair of incompatible variables in the interval between their observations

ABL suggested that, if a system were prepared in the state $P = p_j$ at time t_1 and observed at time $t_2 > t_1$ in the state $Q = q_k$, then *both* P and Q are sharp at all times $t \in (t_1, t_2)$, so ‘we are led into assigning the state $[|q_k\rangle]$ to the period of time *preceding* the observation of $[Q]$ yielding the eigenvalue $[q_k]$.’ This is based on the fact that, according to (7), an observation of P at any such time t would yield, with certainty, the value p_j , but also an observation of Q at any such time t would yield, with certainty, the value q_k . (Of course, (7) allows for only *one* such observation within that time interval.) A considerable controversy has arisen (cf Vaidman (1999), Kastner (1999) and citations within each) regarding exactly what such ‘counterfactual’ sharpness might mean.

It is interesting to examine this claim within our card system. Prepare the system in the state K at time t_1 , and at time $t_2 > t_1$ observe *Suit* and select the occurrence if and only if *Suit* = H. Then, exactly in the sense of ABL, *both* K and H are sharp throughout the interval (t_1, t_2) . But we can see inside this system: if during (t_1, t_2) either no observation, or an

observation of *Face*, is made, then $Face = K$ throughout but *Suit* has no value until t_2 ; on the other hand, if at time $t \in (t_1, t_2)$ an observation of *Suit* is made, then during the interval (t_1, t) K is sharp and *Suit* has no value, and during the interval (t, t_2) H is sharp and *Face* has no value. That the observation of *Suit* has the definite result of H implies nothing about its prior-to-observation value; the final selection simply throws away all occurrences of D and S . There is no warrant for the claim that these variables have simultaneously sharp values.

Note also how inappropriate it is to apply the term ‘measurement’ to these procedures: clearly, no value of Q exists at time t_1 to be measured; even the term ‘observation’ is misleading. In fact, an interaction of a specific nature has occurred which has brought to the variable Q a value; it is this that I have called ‘manifestation’. It is interesting to read the Vaidman–Kastner discussion (as well as the much older and more extensive literature concerning quantum ‘reality’) with this example in mind.

6. Conclusion

The three-box example arose in a quantum setting, and was taken (somewhat uncritically) to be another example of the ‘bizarre’ nature of quantum mechanics. The restatement of the example as a three-slit atomic Young system shows it to be a straightforward example of quantal interference. But exemplified in a perfectly ordinary setting, a deck of cards without a quantum in sight, the three-box phenomenon becomes merely an interesting phenomenon of ordinary probability systems, exhibited by a quantum mechanical system *qua* probability system, in no way a quantal phenomenon—hence my characterization of Vaidman’s comment as ‘much ado about nothing’.

Quantum mechanics provides only the probability of events; it is generally agreed that there is no underlying mechanism (and certainly, if there is, we know nothing of it). Thus all we know of the quantum three box example is this: a system is prepared in a certain way; one or another of two partial observations of a system variable is made; a final observation is used to filter out a single outcome; under a certain choice of the preparation and the final filter, each of the intermediate observations has a sharp outcome. Quantum mechanics tells us nothing more. This card game satisfies the same probabilistic description; every probabilistic quality is common to both systems, and the quantitative differences in probabilities fail to provide a qualitative distinguishing feature.

Their significant difference between them is that the card system has a known underlying mechanism which can be analysed and understood. First, such analysis shows no *ad hoc* devices—the card system follows its own internal logic consistently to the three-box-like result. Second, analysis of the interior state of this classical system makes it clear that there is no ontic significance to these retroactively sharp-if-and-when-observed values—the ‘observation’ actually *brings the value into existence*. Thus, the mere fact that a value is statistically definite does not imply that the observable ‘has’ that value (in any reasonable sense of ‘to have’): *a sharp value is not necessarily a possessed value*. This ‘failure of realism’ in an obviously real system undercuts metaphysical concerns regarding such failure in quantum mechanics and supports the conclusion that the three-box example has no ontic significance regarding possessed values.

Appendix A. Notation

Throughout this paper, P and Q represent distinct system variables with possible values $\{p_j\}$ and $\{q_k\}$, respectively. The proposition that a variable has a certain value, e.g., ‘ $P = p_j$ ’, is abbreviated with the variable’s value, ‘ p_j ’. A general preparation of the system will be

denoted s ; we write probability expressions with the preparation as a subscript: $\Pr_s(p_j)$ is the probability of the proposition $P = p_j$ after the preparation in the state s .

The conditional probability (probability conditioned on an occurrent fact), defined by²

$$\Pr_s(b | a) = \begin{cases} \Pr_s(a \wedge b) / \Pr_s(a) & \Pr_s(a) > 0 \\ \text{undefined} & \text{otherwise} \end{cases} \tag{A.1}$$

is the probability of the truth of the proposition b given that the fact stated by the proposition a occurs.

The set of propositions $\{a_j\}$ is *disjoint* iff, whenever all $\{a_j\}$ take on values, $a_j \wedge a_{j'} = F$, $j \neq j'$. A disjoint set satisfies $\sum_t \Pr_s(a_t) = \Pr_s(\bigvee_t a_t)$ for all preparations s .

The set $\{a_j\}$ is *complete* iff, whenever all $\{a_j\}$ take on values, $\bigvee_t a_t = T$; hence, a disjoint, complete set satisfies $\sum_t \Pr_s(a_t) = 1$ for all preparations s .

An *event* is an occurrence at which at least one variable of the system takes on a value randomly; this is brought about by a physical interaction of the system with its exterior. *Which* variable takes on a value randomly depends on the details of the physical interaction, or *manifestation*; at each event, then, a particular variable is manifested. Manifestation will be indicated as a condition of the probability (as discussed in section 2.2).

The ordinal position of an event in a sequence of events will be denoted by a superscript in brackets: the event E followed by the event F is denoted $E^{[1]} \wedge F^{[2]}$. However, when the terms in the probability expressions are in the ‘natural’ order and no ambiguity arises, the sequence ordinals will be dropped; thus $\Pr_s(E)$ always means $\Pr_{s^{[0]}}(E^{[1]})$, and $\Pr_s(F | E)$ always means $\Pr_{s^{[0]}}(F^{[2]} | E^{[1]})$.

Appendix B. Stability

In quantum mechanics a proposition p is represented by the projector $\mathbf{P}[p]$. If the manifestation $M_{p_k}^{\sim}$ is stable³ in the sense that

$$\Pr_{p_j^{[0]}}(p_k^{\sim[1]} \wedge p_{j'}^{[2]} | M_{p_k}^{\sim[1]}) = \delta_{jj'}(1 - \delta_{jk}) \tag{B.1}$$

then, applying the Wigner ‘sandwich’ formula for the probability of successive non-disjunctive events,

$$\Pr_s(p^{[1]} \wedge q^{[2]}) = \text{Tr}\{\rho[s]\mathbf{P}[p]\mathbf{P}[q]\mathbf{P}[p]\} \tag{B.2}$$

we obtain

$$|\langle p_j | \mathbf{P}[p_k^{\sim}] | p_{j'} \rangle|^2 = \delta_{jj'}(1 - \delta_{jk}). \tag{B.3}$$

In order that $\mathbf{P}[p_k^{\sim}]$ be a projector, the diagonal elements of its matrix must be the positive square roots, hence

$$\mathbf{P}[p_k^{\sim}] = \sum_{t \neq k} \mathbf{P}[p_t] = \mathbf{1} - \mathbf{P}[p_k]. \tag{B.4}$$

Appendix C. Probabilities of the card system

The example of section 3 satisfies the following probability expressions:

$$\begin{aligned} \Pr_s(\cdot | p_j) &= \Pr_{p_j}(\cdot) & \text{if } \Pr_s(p_j) \neq 0 \\ \Pr_{q_j}(\cdot | p_k^{\sim}) &= \Pr_{p_k^{\sim}}(\cdot) & \text{if } \Pr_{q_j}(p_k^{\sim}) \neq 0 \\ \Pr_{p_j}(\cdot | p_k^{\sim}) &= \Pr_{p_j}(\cdot) & \text{if } j \neq k \end{aligned} \tag{C.1}$$

² Disjunction (‘or’) is indicated by \vee ; conjunction (‘and’), by \wedge ; negation (‘not’) by \sim .

³ Stability is implied by, but weaker than, Wigner’s ‘morality’ (cf Goldberger and Watson 1964).

$$\begin{aligned} \Pr_{p_j}(p_k) &= \delta_{jk} & \Pr_{p_j}(q_k) &= \frac{N - N(p_j \cdot q_k)}{N(V-1)} \\ \Pr_{p_k}(p_j) &= (1 - \delta_{jk}) \frac{1}{V-1} & \Pr_{p_j}(q_k) &= \frac{N(p_j \cdot q_k)}{N} \end{aligned} \quad (\text{C.2})$$

$$\Pr_s(p_j) = 1 - \Pr_s(p_j). \quad (\text{C.3})$$

References

- Aharonov Y, Bergmann P G and Lebowitz J L 1964 Time symmetry in the quantum process of measurement *Phys. Rev. B* **134** 1410–6
- Aharonov Y and Vaidman L 1991 Complete description of a quantum system at a given time *J. Phys. A: Math. Gen.* **24** 2315–28
- Albert D Z, Aharonov Y and D'Amato S 1985 Curious new statistical prediction of quantum mechanics *Phys. Rev. Lett.* **54** 5–7
- Goldberger M L and Watson K M 1964 Time correlations for quantum-mechanical systems *Phys. Rev.* **134** B919–28
- Kastner R E 1999 TSQT Elements of possibility *Stud. Hist. Phil. Mod. Phys.* **30** 399–402 (*Preprint quant-ph/9812024*)
- Kirkpatrick K A 2003 'Quantal' behavior in classical probability *Found. Phys. Lett.* **16** at press (*Preprint quant-ph/0106072 v6*)
- Vaidman L 1996 Weak-measurement elements of reality *Found. Phys.* **26** 895–906
- Vaidman L 1999 The meaning of elements of reality—reply to kastner *Found. Phys.* **29** 865–76 (*Preprint quant-ph/9903095*)