

The “Chaotic Ball” model, local realism and the Bell test loopholes

Caroline H Thompson* and Horst Holstein

Department of Computer Science, University of Wales, Aberystwyth, SY23 3DB, U.K.

(Submitted to *American Journal of Physics*, 17 October 2002; revised 30 March 2003)

It has long been known that the “detection loophole”, present when detector efficiencies are below a critical figure, could open the way for alternative “local realist” explanations for the violation of Bell tests. It has in recent years become common to assume the loophole can be ignored, regardless of which version of the Bell test is employed. We introduce a simple model to illustrate that this may not be justified. Two versions of the Bell test — the standard one of form $-2 \leq S \leq 2$ and the currently-popular “visibility” test — are at grave risk of bias. Statements implying that experimental evidence “refutes local realism” or shows that the quantum world really is “weird” should be reviewed. The detection loophole is on its own unlikely to account for more than one or two test violations, but when taken in conjunction with other loopholes (briefly discussed) it is seen that the experiments refute only a narrow class of “local hidden variable” models, applicable to idealised situations, not to the real world. The full class of local realist models may not only provide straightforward explanations for the publicised Bell-test violations but also account for some lesser-known “anomalies”.

Introduction

Pearle (1970¹) showed that failure to detect some particles during Bell test experiments can allow local realist (hidden variable) explanations to reproduce almost exactly the quantum-mechanical (QM) predictions. This fact, which has become known as the “detection”, “efficiency” or “fair sampling” loophole, has been re-discovered many times², for Pearle’s paper did not attract the wide publicity it deserved. It is cited in the review article by Clauser and Shimony (1978³) but few others. There seems, however, to have been general acceptance in the 1970’s that “fair sampling” could not legitimately be assumed. Experimenters such as Freedman and Clauser in 1972⁴ and Fry and Thompson in 1976⁵ recognised that neither Bell’s original test⁶ nor the variant introduced in 1969 by Clauser, Horne, Shimony and Holt⁷ was appropriate to real experimental conditions. They used instead even later versions, described by Clauser and Horne in 1974⁸ (CH74), that were not dependent upon near-perfect detection efficiency. “Normalisation” in these modified tests was by comparison with coincidence counts obtained with the polarisers removed, not, as in the modern interpretation of Clauser et al.’s 1969 test (hereafter referred to as the CHSH69 test⁹), with the total when both were present.

Clauser et al.’s paper of 1969⁷ had been inspired by Bell’s original work. Clauser and Horne’s of 1974⁸ took account of Bell’s later ideas¹⁰, making substantial improvements¹¹ and suggesting a test that did not demand high detector efficiency (instead, the

assumption of “no enhancement” was made¹²). In addition, detailed footnotes covered a wide range of other potential “loopholes” in real optical experiments. It is the 1969 paper, though, that is most frequently quoted and, since about 1980, the CHSH69 test⁹ and the almost-equivalent “visibility” test¹³ have come into favour. The assumption of “fair sampling” is accepted either with no comment at all or described by terms such as “plausible”, the hidden variable theories associated with its failure being dismissed as “bizarre” or requiring “conspiracies” between the detectors¹⁴.

Some experimenters, for example N. and B. Gisin¹⁵, have taken due note of the importance of the detection loophole, but others seem unaware of its implications. The CHSH69 and visibility tests can only safely be used if the detection rates are very high — a condition that has rarely if ever been met in practice. The trapped ion experiment of Rowe *et al.*¹⁶, acclaimed in some popular accounts¹⁷ as having closed the final loophole, did indeed have very high efficiency, but it failed another crucial test: the ions could not legitimately be described as “separated” (they were in the same trap, controlled by the same laser), so Bell’s inequality was not applicable.

The present paper, investigating conditions in which sampling is *not* “fair”, is based on the “Chaotic Ball” model devised by C. H. Thompson and first published in 1996¹⁸. It has now been improved by the addition of an analytical formula derived by one of us (HH), enabling readers to explore its possibilities for themselves. The logic is essentially the same as Pearle’s, but the geometry is very much

* Email: c.h.thompson@pgen.net; Web site: <http://users.aber.ac.uk/cat>

simpler¹⁹, aiming to illustrate principles rather than to reproduce the exact quantum-mechanical prediction.

The model is most appropriate for the thought experiments of Bohm²⁰, involving spin-1/2 particles, in which the direction of spin is taken to be a vector and used as the hidden variable. No claim is made, however, that it corresponds to the actual physics of Bohm's experiment, which has, it must be emphasised, never been performed and might not in practice ever be feasible.

The majority of actual Bell test experiments have involved light, whose particle nature is debatable²¹. Some, for example Aspect's well-known experiments of 1981-2²², involved the polarisation of light; other more recent ones, for example Tapster *et al.*, 1994²³, involved its phase and momentum. The basic local realist model that covers them all — or, at least, those for which emitted pairs can be identified on detection without ambiguity — predicts the probability of a coincidence to be

$$P(a,b) = \int_{\Lambda} d\lambda \mathbf{r}(\lambda) p_a(\lambda) p_b(\lambda), \quad (1)$$

where Λ is the space spanned by the "hidden variable" λ , a and b are the detector settings and p_a and p_b functions of λ giving the probabilities of detection at the two detectors. Some readers may prefer to think directly in terms of this formula, following Marshall, Selleri and Santos²⁴ or other articles by C. H. Thompson^{25,26}. The points raised in the current paper, though, are quite general, especially regarding the existence of less well known loopholes and the fact that not all Bell tests are identical or involve the same assumptions.

The Chaotic Ball, covering just one of the loopholes, corresponds to a deterministic case of expression (1), with p_a and p_b taking only values 0 or 1. The geometrical difference between spin (for which "opposite" means differing by 180°) and polarisation (for which it means differing by 90°) is of no significance so far as the principle illustrated is concerned. Even the apparently fundamental difference between a true particle (which can go only to one or other detector) and light (which can, despite the claims of quantum opticians²⁷, go to both at once) does not seriously affect the logic.

The "Chaotic Ball"

Let us consider Bohm's thought experiment, commonly taken as the standard example of the entanglement conundrum that Einstein, Podolsky and Rosen discussed in their seminal 1935 paper²⁸. A molecule is assumed to split into two atoms, A and B , of opposite spin, that separate in opposite directions. They are sent to pairs of "Stern-Gerlach" magnets, whose orientations can be chosen by the experimenter, and counts taken of the various

"coincidences" of spin "up" and spin "down". The obvious "realist" assumption is that each atom leaves the source with its own well-defined spin (a vector pointing in any direction), and it is the fact that the spins are opposite that accounts for the observed coincidence pattern. (The realist notion of spin cannot be the same as the quantum theory one, since in quantum theory "up" and "down" are concepts defined with respect to the magnet orientations, which can be varied. Under quantum mechanics, the particles exist in a superposition of up and down states until measured.)

Bell's original inequality was designed to apply to the estimated "quantum correlation"²⁹ between the particles. He proved that the realist assumption, based on the premise that the detection events for a given pair of particles are independent, leads to statistical limits on this correlation that are exceeded by the QM prediction. He did not, however, specify how it was to be estimated in cases where not all the particles were detected.

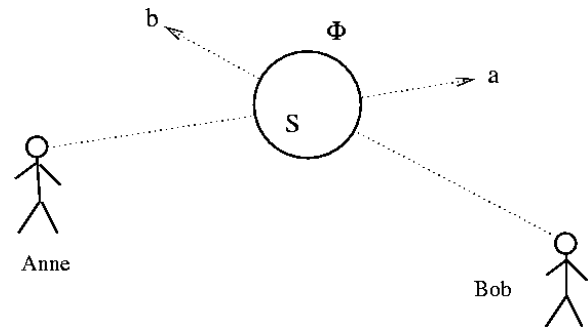


Fig. 1: Anne, Bob and the Chaotic Ball.

The letter S is visible while N , opposite to it, is out of sight. a and b are directions in which the assistants are viewing the ball; Φ the angle between them.

When detection is perfect there is no problem, but when it is not, the "detection loophole" creeps in. What assumptions can we reasonably make? Under quantum theory, the most natural one is that all emitted particles have an equal chance of non-detection (the sample detected is "fair", not varying with the settings of the detectors). The realist picture, however, is different.

Let us replace the detectors by two assistants, Anne (A) and Bob (B), the source of particles by a large ball on which are marked, at opposite points on the surface, an N and an S (fig. 1). The assistants look at the ball, which turns randomly about its centre (the term "chaotic", though bearing little relation to the modern use of the term, is retained for historical reasons). They record, at agreed times, whether they see an N or an S . When sufficient records have been made they get together and compile a list of the coincidences — the numbers of occurrences of NN ,

SS, NS and SN, where the first letter is Anne's and the second Bob's observation.

The astute reader will notice that, if the vector from S to N corresponds to the "spin" of the atom, the model covers the case in which the spins on the A and B sides are *identical*, not opposite. Anne and Bob are looking at identical copies of the ball, which can conveniently be represented as a single one. This simplification aids visualisation whilst having no significant effect on the logic. The difference mathematically is just a matter of change of sign, with no effect on numerical values. In point of fact, the assumption of identical spins makes the model better suited to some of the actual optical experiments. Aspect's, for example, involved plane-polarised "photons" (not, incidentally, circularly polarised, as frequently reported) with parallel, not orthogonal, polarisation directions.

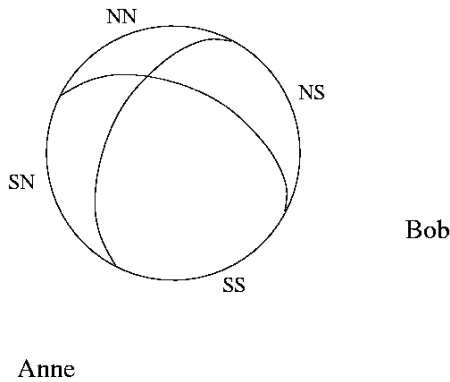


Fig. 2: The registered coincidences: Chaotic Ball with perfect detectors.

The first letter of each pair denotes what Anne records, the second Bob, when the S is in the region indicated.

With this simplification, geometry dictates that if the ball takes up all possible orientations with equal frequency (there is "rotational invariance") then the relative frequencies of the four different coincidence types will correspond to four areas on the surface of an abstract fixed sphere as shown in fig. 2.

Anne's observations correspond to two hemispheres, Bob's to a different pair, the dividing circles being determined by the positions of the assistants. We conduct a series of experiments, each with fixed lines of sight ("detector settings") a and b . It can readily be verified that the model will reproduce the standard "deterministic local realist" prediction, with linear relationship between the number of coincidences and Φ , the angle between the settings³⁰. This is shown in fig. 3, which also shows the quantum mechanical prediction, a sine curve.

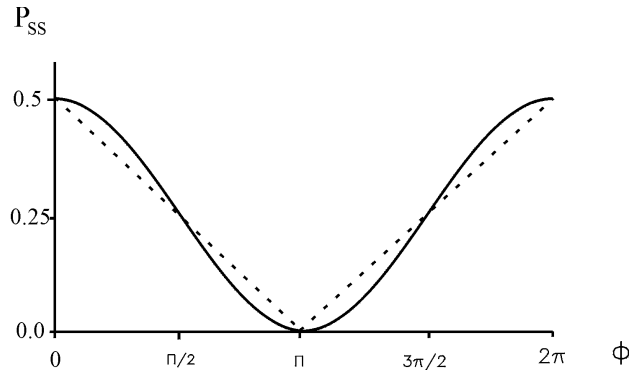


Fig. 3: Predicted coincidence curves.

Dotted lines give the local realist prediction for the probability that both Anne and Bob see an S , if there are no missing bands; the curve is the QM prediction, $\frac{1}{2} \cos^2(\Phi/2)$.

What happens, though, if the assistants do not both make a record at every agreed time? If the only reason they miss a record is that they are very easily distracted, this poses little problem. So long as the probability of non-detection can be taken to be random, the expected pattern of coincidences will remain unaltered. What if the reason for the missing record varies with the orientation of the a ball, though — with the "hidden variable", λ , the vector from S to N ?

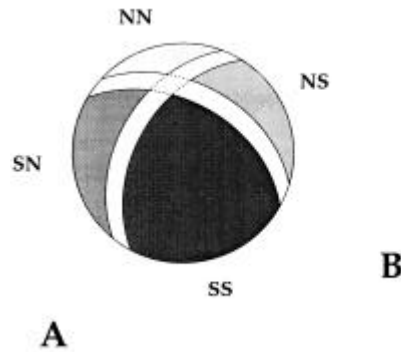


Fig. 4: Chaotic Ball with missing bands.

There is no coincidence unless both assistants make a record, so some data is thrown away.

Suppose the ball is so large that the assistants cannot see the whole of the hemisphere nearest to them. The picture changes to that shown in fig. 4, in which the shaded areas represent the regions in which, when occupied by the S , coincidences will be recorded as indicated. The ratios between the areas, which are what matter in Bell tests, change — indeed, some areas may disappear altogether. If the bands are *very* large, there will be certain positions of the assistants for which the estimated quantum correlation (E , equation (1) below) is not even defined, since there are *no* coincidences.

New decisions are required. Whereas before it was clear that if we wanted to normalise our coincidence rates we would divide by the total number of observations, which would correspond to the area of the whole surface, there is now a temptation to divide instead by the total shaded area. The former is correct if we want the proportion of coincidences to emitted pairs, but it is, regrettably, the latter that has been chosen in actual Bell test experiments. It is easily shown that the model will now inevitably, for a range of parameter choices, infringe the relevant Bell test if our estimates of “quantum correlation” are the usual ones, namely,

$$E(a,b) = \frac{NN + SS - NS - SN}{NN + SS + NS + SN}, \quad (2)$$

where the terms NN etc. stand for counts of coincidences in a self-evident manner.

The Bell test in question is the CHSH69 test referred to above. It takes the form $-2 \leq S \leq 2$, where the test statistic is

$$S = E(a,b) - E(a,b') + E(a',b) + E(a',b'). \quad (3)$$

The parameters a , a' , b and b' are the detector settings: to evaluate the four terms four separate sub-experiments are needed. The settings chosen for the Bell test are those that produce the greatest difference between the QM and standard local realist predictions, namely $a = 0$, $a' = \pi/2$, $b = \pi/4$ and $b' = 3\pi/4$. Since we are assuming rotational invariance, the value of E does not depend on the individual values of the parameters but on their difference, $\phi = b - a$, which is $\pi/4$ for three of the terms and $3\pi/4$ for the fourth. We can therefore immediately read off the required values from a graph such as that of fig. 5, where the curve is calculated from the geometry of fig. 4 (see next section).

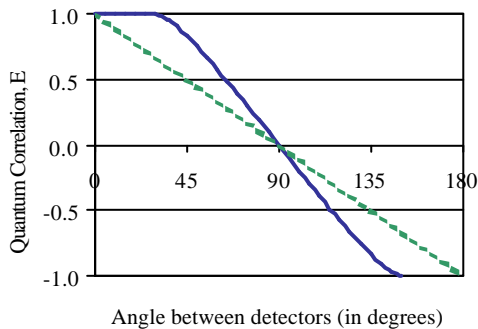


Fig. 5: Predicted quantum correlation E versus angle.

The curve corresponds to (moderate-sized) missing bands, the dotted line to none. See equation (4) of the text for the formula for the central section of the curve.

When there are no missing bands it is clear that the numerical value of each term is 0.5 and that they are all positive. Thus with no missing bands the model shows that we have exact equality, with S actually equalling 2.

If we do have missing bands, however, although the four terms are still all equal and all positive, each will have increased! *The Bell test will be infringed.*

An “imperfection” has *increased* the correlation, in contradiction to the opinion, voiced among others by Bell himself, that imperfections are unlikely ever to do this. It is not hard to imagine real situations, especially in the optical experiments, in which something like “missing bands” will occur, biasing this version of Bell’s test in favour of quantum mechanics. Note that the “visibility” test used in recent experiments such as Tittel’s long-distance Bell tests³¹ is equally unsatisfactory, biased from this same cause. The realist upper limit on the standard test statistic when there is imperfect detection is 4, not 2, well above the quantum-mechanical one of $2\sqrt{2} \approx 2.8$. The visibility can be as high as 1, not limited to the maximum of 0.5 that follows from the commonly-accepted assumptions.

Detailed Predictions for the basic model

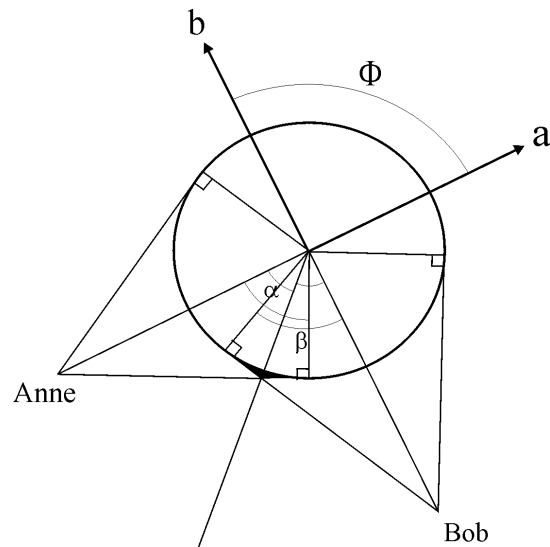


Fig. 6: Definition of angles used in equation (4).

The main formula for the proportion P_{SS} of “like” coincidences such as SS with respect to the number of emitted pairs N comes from the area of overlap of two circles on the surface of a sphere (see figs. 4 and 6). It can be shown to be

$$P_{SS}(\mathbf{a}, \mathbf{b}) = \frac{1}{p} \left(\cos^{-1} \left(\frac{\sin \mathbf{a}}{\sin \mathbf{b}} \right) - \cos^{-1} \left(\frac{\tan \mathbf{a}}{\tan \mathbf{b}} \right) \cos \mathbf{b} \right), \quad (4)$$

where $\alpha = \varphi/2$ and β is the half-angle defining the proportion of the surface for which each assistant makes a definite reading (zero corresponds to none; $\pi/2$ to the whole surface). P_{SS} achieves a maximum of $\frac{1}{2}(1 - \cos \beta)$ when $\alpha = 0$, which is less than the QM prediction of 0.5 unless β is $\pi/2$. When $\alpha \geq \beta$, it is zero (see fig. 7).

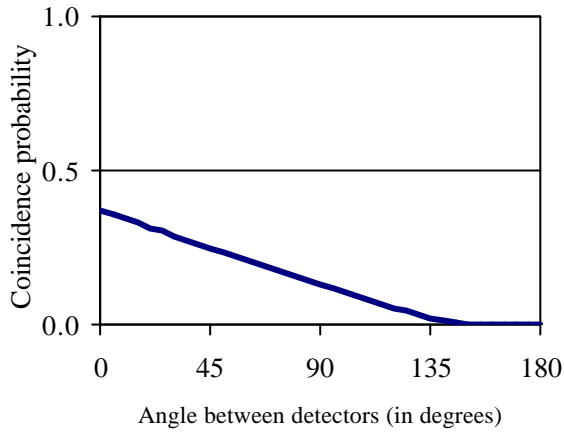


Fig. 7: Predicted coincidence rate P_{SS} for β set at 75° ($5\pi/12$).

The presence of the zero region has the interesting consequence that if β is too small (less than $\pi/4$) the derived quantum correlation is undefined for a region in the neighbourhood of $\alpha = \pi/4$ or $\varphi = \pi/2$. So far as actual experiments go, the interest is purely academic, though, since there are always background “dark counts” and other “accidentals” that ensure that the observed counts are never zero. In the actual experiments the whole curve would be smoother, as hard divisions between regions scoring 1 and those scoring zero would be unlikely to occur.

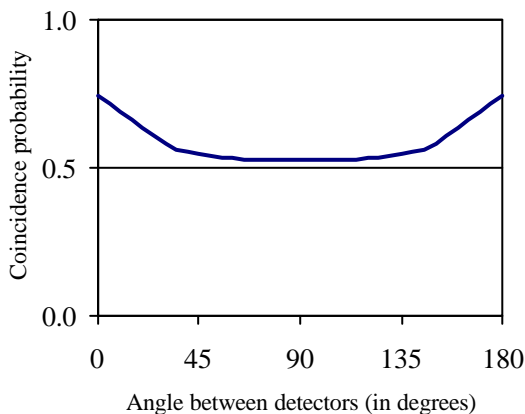


Fig. 8: Total coincidence rate, T_{obs}/N .

To obtain the prediction for “unlike” coincidences such as NS , we replace α by $\pi/2 - \alpha$. We can now

find the predicted value (fig. 8) for the total observed coincidence rate, $T_{obs}/N = (NN + SS - NS - SN)/N$.

The fact that this is not constant was recognised by Pearle in his 1970 paper, and can be used as a test for the validity of the QM model. It is, however, not a conclusive one, as can be seen by consideration of the situation in which one detector is perfect and the other has missing bands. The estimated quantum correlation could violate Bell’s limit despite the fact that T_{obs} was constant. An important fact (also noted by Marshall *et al.*²⁴) is that the values predicted for the “Bell test angles” of $\pi/4$ and $3\pi/4$ are equal. It is to be hoped that these are not the only ones that experimenters check!

We can derive the estimate of the ordinary “normalised” quantum correlation in which division is by T_{obs} , with results as shown in fig. 5, but it is of interest to look also at the “unnormalised” one, corresponding to division by N :

$$P_{NN} + P_{SS} - P_{NS} - P_{SN},$$

plotted in fig. 9.

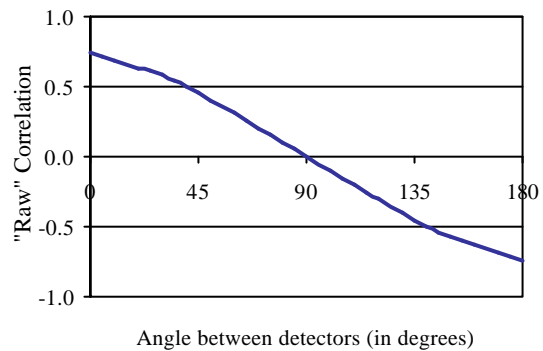


Fig. 9: Unnormalised “quantum correlation”.

The match with the QM prediction is considerably less impressive, the curve not reaching the maximum of 1 and not having the feature of a zero slope for parallel detectors ($\alpha = 0$). Whilst (for the chosen example, with β set at 75°) the model gives the CHSH69 test statistic of $S = 3.331 > 2$, the unnormalised estimate will never exceed 2. The values at the “Bell test angles” will always all be numerically less than 0.5.

Discussion

We have shown by means of this counter-example that the CHSH69 test as usually implemented (using the sum of observed coincidences in place of number of emitted pairs) rests on an assumption — that of fair sampling — that cannot be made lightly. Not only this, but the model demonstrates that local

hidden variable theories (or what Clauser and Horne⁸ prefer to call Objective Local Theories, since the hidden variables are not fully deterministic as originally defined) that violate Bell inequalities are not all “weird”. They do not, as often stated (see for example Laloë’s recent review article, ref 14) require “conspiracies” between the detectors. Furthermore, the model shows the fallacy of a belief, held by Bell and frequently quoted, that “imperfections” are unlikely to increase the significance of his tests (page 109 of his book, “Speakable and Unspeakable³²”).

Why, though, was all of the above not clear from Pearle’s 1970 paper (or, indeed, from the original intention of the CHSH69 test — see endnote 9)? Why did the modern interpretation of the CHSH69 test ever come to be used? Bell was well aware that in order to get a strictly valid test it was necessary to know the number of pairs you started with. He would have liked to see “event-ready detectors”, counting emissions as they occurred. If this number were to be used as denominator in the estimates E , the test would be valid and the “local realist” model represented by the ball would *not* violate the inequality.

It is, however, not hard to see why experimenters have taken the path they have. Neither Pearle’s nor later rediscoveries have succeeded in conveying at an intuitive level just *how* the detection loophole produces bias, and the term “fair sampling” does sound so reasonable. The use of event-ready detectors is not practical, since we are dealing with “quantum” events that are destroyed when observed. Also, in the real experiments we use light, and we have no perfect detectors for “single photons”. Since the very existence of the photon can be questioned, no universal agreement is to be expected on estimates of the number of pairs emitted. An estimate based on quantum-mechanical assumptions is not acceptable when it is the validity of QM that is being tested.

Is there any way of rescuing the situation? Can we not test the fairness of the sample? Though there are some cases in which sampling is clearly *not* fair³³, in practice, we cannot prove the converse. There is a better version of the Bell test that could be used, however. Clauser and Horne’s 1974 paper covers an alternative inequality (CH74) that does not have this particular weakness and which was (as mentioned earlier) the one used in practice in all experiments prior to about 1980, including all but the central one of Aspect’s 1981-2 set. Effectively, instead of using the total number of observed coincidences, they used the number of coincidences obtained when the detectors did not try and discriminate — when Anne and Bob were able to see the whole surface of the ball. In experiments using polarised light, they used the coincidence counts from sub-experiments in which the polarisers were removed. Another difference is that they looked only for N ’s, say, not

for N ’s and S ’s. The assumption that is part of the proof of the suspect CHSH69 or visibility tests — that we have fair sampling, with the sum of the probabilities of N and S results independent of the settings a and b — does not enter into the question.

In practice there are important parameters — in particular, beam intensity and characteristics of the photodetectors — that are at the experimenter’s discretion. Realist models, whether following the ball analogy or working directly from the basic theoretical formula, suggest that the coincidence curve will be strongly influenced by the choices made. If we are genuinely trying to find the best model for the physical situation, is it not necessary to include all these relevant parameters? They play a natural part in the realist approach but in the quantum-mechanical one are ignored: in quantum mechanics, detectors are characterised by a single “quantum efficiency”.

It is generally assumed that classical and quantum theory agree on an important consequence of this characterisation: adherence of the probability of detection to Malus’ Law, but, whether because the polarisers are imperfect or the detectors not exactly “square law”, this adherence will not be exact (see, for example, refs. 25 and 26). The actual relationship, and hence the observed coincidence curve, may not be quite sinusoidal. The majority of Bell test experiments have concentrated on just the few points needed to estimate the test statistic, not looking at enough data points to check the shape of the complete coincidence curve, let alone the constancy of the total coincidence count. Any deviation from the sine curve should be regarded as an indicator in favour of local realism.

Other loopholes

The detection loophole is, at least among professionals, well known, but the fact that it affects some versions of Bell’s test and not others is perhaps less well understood. Different loopholes apply to different versions, for each version comes with its attendant assumptions. Some come very much under the heading of “experimental detail” and have, as such, little interest to the theoretician. If we wish to decide on the value to be placed on a Bell test, however, such details cannot be ignored.

1. Subtraction of “accidentals”: Adjustment of the data by subtraction of “accidentals”, though standard practice in many applications, can bias Bell tests in favour of quantum theory. After a period in which this fact has been ignored by some experimenters, it is now once again accepted³⁴. The reader should be aware, though, that it invalidates many published results²⁶.

2. **Failure of rotational invariance:** The general form of a Bell test does not assume rotational invariance, but a number of experiments have been analysed using a simplified formula that depends upon it. It is possible that there has not always been adequate testing to justify this. Even where, as is usually the case, the actual test applied is general, if the hidden variables are not rotationally invariant, *i.e.* if some values are favoured more than others, this can result in misleading descriptions of the results. Graphs may be presented, for example, of coincidence rate against ϕ , the difference between the settings a and b , but if a more comprehensive set of experiments had been done it might have become clear that the rate depended on a and b separately³⁵. Cases in point may be Weihs' experiment, presented as having closed the "locality" loophole³⁶, and Kwiat's demonstration of entanglement using an "ultrabright photon source"³⁷.
3. **Synchronisation problems:** There is reason to think that in a few experiments bias could be caused when the coincidence window is shorter than some of the light pulses involved²⁵. These include one of historical importance — that of Freedman and Clauser, in 1972³⁸ — which used a test not sullied by either of the above possibilities.
4. **"Enhancement":** Tests such as that used by Freedman and Clauser (essentially the CH74 test) are subject to the assumption that there is "no enhancement", *i.e.* that there is no hidden variable value for which the presence of a polariser *increases* the probability of detection. This assumption is considered suspect by some authors, notably Marshall and Santos, but in practice, in the few instances in which the CH74 inequality has been used, the test has been invalidated by other more evident loopholes such as the subtraction of accidentals.
5. **Asymmetry:** Whilst not necessarily invalidating Bell tests, the presence of asymmetry (for instance, the different frequencies of the light on the two sides of Aspect's experiments) increases the options for local realist models³⁹.

A loophole that is notably absent from the above list is the so-called "locality" or "light-cone" one, whereby some unspecified mechanism is taken as conveying additional information between the two detectors so as to increase their correlation above the classical limit. In the view of many realists, this has never been a serious contender. John Bell supported Aspect's investigation of it (see page 109 of ref 32) and had some active involvement with the work, being on the examining board for Aspect's PhD. Weihs improved upon the test in his experiment of 1998³⁶, but nobody has ever put forward plausible ideas for the mechanism. Its properties would have to be quite extraordinary, as it is required to explain "entanglement" in a great variety of geometrical

setups, including over a distance of several kilometers in the Geneva experiments of 1997-8^{31, 34}.

There may well be yet more loopholes. For instance, in many experiments the electronics is such that simultaneous '+' and '-' counts from both outputs of a polariser can never occur, only one or the other being recorded. Under QM, they will not occur anyway, but under a wave theory the suppression of these counts will cause even the basic realist prediction (expression (1) above) to yield "unfair sampling". The effect is negligible, however, if the detection efficiencies are low, since the 3- or 4-fold coincidences involved (two on one side, one or more on the other) then hardly ever happen.

Explaining "anomalies"

Alain Aspect's set of three experiments is deservedly given pride of place in any listing of Bell test trials. His PhD thesis⁴⁰, which gives very much more detail of the experiments than could be included in Physical Review Letters, shows evidence of careful reasoning and meticulous attention to detail. That detail includes, however, mention of some anomalies that perhaps deserve more attention. None reached the level of statistical significance, but Aspect evidently recognised that they were potentially important.

One anomaly in particular the reader will recognise as giving cause for concern: the total number of coincidences in the experiment in which the CHSH69 test was used (the first 1982 one, using two-channel polarisers) varied slightly with detector setting, so the sample may not have been strictly "fair". Instead of increasing replication to check if the variations were consistent, though, Aspect derived (pp 124-7 of his thesis) a generalised Bell test designed to compensate for them. When applied to the 1982 experiment, this modified test seemed to show that the effect associated with the anomalous total could not have caused significant bias. The new test is mentioned briefly in the relevant paper, but the promised publication on the subject does not appear to have materialised. Its validity can be challenged and, especially since Aspect's methods have been used as a precedent by others, the matter deserves wider publicity.

Another minor anomaly in the two-channel experiment was a discrepancy between the counts of '+ -' and '- +' coincidences. A slight adaptation of the Chaotic Ball model to allow for two known facts of the experiment — asymmetry of A and B detectors related to the differences in wavelength, and slight bias in the polarisers — shows straightforwardly how this might arise.

Conclusion

The “Chaotic Ball” models a hypothetical Bell test experiment in a manner that encourages the use of intuition and realism. It illustrates the fact, well known to those working in the field, that if not all “particles” are detected there is risk of bias in the tests used, which are no longer able to discriminate between the “nonseparable” quantum-mechanical model and local realism. Since the Bell tests themselves fail to discriminate, we suggest that more comprehensive experiments, covering more data points and a wider range of conditions, are desirable. The full predictions of the model, not just the Bell test values, should be checked. If interpretation along the lines suggested in this paper proves fruitful, the possibility that quantum entanglement of separated particles may not happen may need to be taken seriously.

The ball model is a special case of a general class of local realist models, many of which are covered by equation (1) above — essentially the same formula that Bell assumed, reflecting the local realist assumption that *the observed correlations originate from shared properties acquired at the source, and the detection events are independent*. The formula presupposes the standard case, in which the experiment consists of a series of readily identifiable events, a single particle pair being associated with each. If this is not the case, for example where there are “accidentals” or where there is uncertainty in the time of detection, further generalisation is needed.

Note that the phrase in italics above is all we need in order to specify a local realist description. This simplicity is in stark contrast to the oft-stated objection (raised, e.g. by Laloë, ref 14) that local realist models are complicated, “weird” and “ad hoc”. Another frequent objection is that local realism cannot match quantum theory when it comes to accurate quantitative predictions. True, it cannot easily match *exactly* the quantum-mechanical coincidence formulae (the ball model, illustrating principles only, does not even attempt to), but what is required is surely a match with experimental results, not with the quantum theory predictions. We suggest that the quantum-mechanical predictions are not necessarily correct, and, indeed, that the apparently accurate matches observed are slightly illusory, being improved by several factors: (a) adaptations are made (as with the local realist model) to the formulae to allow for experimental conditions (b) various experimental parameters such as beam intensity and detector efficiency are at the experimenter’s discretion, and (c) it is usual to check only the minimum data points needed for a Bell test, leaving untested the remainder of the model, covering the shape of the coincidence curve.

Our underlying assumption when we look at the optical case (where we use not the deterministic ball model but the general “stochastic” model of equation (1)) is that the energy of each individual light pulse (“photon”) is split at the polariser — something no photon can do. The intensity of the emerging pulse then influences statistically the probability of the detector firing. Despite the success of the photon model of light in many applications of quantum optics, this success is at the price of recognised conceptual difficulties. The possibility that wave models that allow for the idiosyncrasies of the apparatus used — deviations from Malus’ Law for example — may be able to account with no such difficulties for *all* optical phenomena is the subject of ongoing research by one of us (CHT).

The spin-offs from experiments related to attempted applications of quantum entanglement — improved technology, for example, in the areas of optical and nano-scale communications and computing — amply justify continued research in this area. The technology stand in its own right. The true theory, however, is still very much an open question:

“Any theory will account for *some* facts; but only the true explanation will satisfy *all* the conditions of the problem ...” (William Crookes, 1875⁴¹)

Acknowledgements

Thanks are due to Franck Laloë for encouragement to air again the Chaotic Ball model. The work would not have been completed without the moral support of David Falla of the Physics Department of University of Wales, Aberystwyth, and of the many who have expressed appreciation of C H Thompson’s web site or contributions to Internet discussions. Her use of the computer and library facilities at Aberystwyth is by kind courtesy of the Department of Computer Science, of which she is an associate member.

¹ Pearle, P, “Hidden-Variable Example Based upon Data Rejection”, Phys. Rev. D **2**, 1418-25 (1970).

² The importance of non-detections is noted in, for example:

A. Fine, “Some local models for correlation experiments”, Synthese **50**, 279-294 (1982); T. Maudlin, “Quantum Non-Locality and Relativity”, (Cambridge, Mass., 1994); D. M. Greenberger, “Two-particle versus three-particle {EPR} experiments”, Annals of the New York Academy of Sciences **755**, 585-599 (1995); Gilbert, B and S Sulcs, “The measurement problem resolved and local realism preserved via a collapse-free photon detection model”, Found. Phys. **26**, 1401 (1996).

- ³ J. F. Clauser and A. Shimony, “Bell’s theorem: experimental tests and implications”, Rep. Prog. Phys. **41**, 1881 (1978).
- ⁴ S. J. Freedman and J. F. Clauser, “Experimental test of local hidden-variable theories”, Phys. Rev. Lett. **28**, 938 (1972).
- ⁵ Edward D. Fry, and Randall C. Thompson, “Experimental test of local hidden-variable theories”, Phys. Rev. Lett. **37**, 465-468 (1976).
- ⁶ Bell, John S, “On the Einstein-Podolsky-Rosen paradox”, Physics **1**, 195 (1964), reproduced as Ch. 2, pp 14-21, of ref 32.
- ⁷ J. F. Clauser, M. A. Horne, A. Shimony and R. A. Holt, “Proposed experiment to test local hidden-variable theories”, Phys. Rev. Lett. **23**, 880-884 (1969).
- ⁸ J. F. Clauser and M. A. Horne, “Experimental consequences of objective local theories”, Phys. Rev. D **10**, 526-35 (1974).
- ⁹ The term “CHSH69 test” will be used in the current paper with its commonly accepted meaning, in which the denominator in the estimate for each “correlation” term is the total number of observed coincidences. Careful reading of the 1969 paper reveals, however, that the authors had no intention of applying the inequality in this way. They describe how to use it in an experimentally realisable manner, which means effectively using the CH74 test. At no point do they mention the possibility of dividing by the total number of observed coincidences.
- ¹⁰ J. S. Bell, in *Foundations of Quantum Mechanics, Proceedings of the International School of Physics “Enrico Fermi”, Course XLIX*, edited by B. d’Espagnat (Academic, New York, 1971), p. 171 and Appendix B. Pages 171-81 are reproduced as Ch. 4, pp 29-39, of ref 32.
- ¹¹ Bell kept consistently to a model and notation that were adapted to a spin-1/2 experiment, in which the “outcome” on side A of the experiment was denoted by A. In the 1964 paper this could take only values +1 and -1. By 1971 zero outcomes were also admitted. However, this notation — used in all popular accounts of Bell’s inequalities — was both confusing, with the symbol *P* used for quantum correlation instead of for a probability, and unsuitable for use in optical applications. Clauser and Horne in 1974 broke away from this tradition. They extended Bell’s 1971 idea to establish a new concept of hidden variable, one that did not determine the outcome but only its probability. They considered in the first instance a setup using plane polarised light and polarisers that had only one output, and derived in a very straightforward manner an inequality restricting the probability of coincidence instead of the quantum correlation. The latter, though this was not explicitly stated, is not in fact well defined for optical work, in that it does not allow for “double detections” — the simultaneous occurrence of ‘+’ and ‘-’ values at the two outputs of a (two-channel) polariser.
- ¹² “No enhancement” means that there is no hidden variable value for which the presence of a polariser increases the probability of detection.
- ¹³ A test used in several recent experiments depends on the fact that in certain conditions the CH74 test reduces to $v \leq 1/\sqrt{2} \approx 0.71$, where *v* is the “visibility”, (max – min)/(max + min), of the coincidence curve. Its validity depends, among other things, on the assumption that this curve is truly sinusoidal.
- ¹⁴ Franck Laloë, “Do we really understand quantum mechanics? Strange correlations, paradoxes and theorems”, Am. J. Phys., **69**(6), 655-701, (June 2001).
- ¹⁵ Gisin, N and B Gisin, “A local hidden variable model of quantum correlation exploiting the detection loophole”, Physics Letters A **260**, 323-327 (1999), <http://arxiv.org/abs/quant-ph/9905018>.
- ¹⁶ M. Rowe et al., “Experimental violation of a Bell’s inequality with efficient detection” Nature **409**, 791 (2001).
- ¹⁷ “Quantum Loophole Shut”, Phys. World, p.3 (March 2001).
- ¹⁸ C. H. Thompson, “The Chaotic Ball: An Intuitive Analogy for EPR Experiments”, Found. Phys. Lett. **9**, 357 (1996), <http://arxiv.org/abs/quant-ph/9611037>.
- ¹⁹ For an excellent diagram illustrating Pearle’s geometry, see R. Risco-Delgado, “The Variable Detection Approach: a Wave Particle Model”, Found. Phys. Lett. **6**, 399-428 (1993).
- ²⁰ D. Bohm, *Quantum Mechanics* (Prentice-Hall, 1951).
- ²¹ W. E. Lamb Jr., “Antiphoton”, App. Phys. B **60**, 77-84 (1995).
- ²² A. Aspect, *et al.*, Phys. Rev. Lett. **47**, 460 (1981); **49**, 91 (1982) and **49**, 1804 (1982).
- ²³ P. R. Tapster, J. G. Rarity and P. C. M. Owens, “Violation of Bell’s inequality over 4 km of optical fibre”, Phys. Rev. Lett. **73**, 1923 (1994).
- ²⁴ T. W. Marshall, E. Santos and F. Selleri: “Local Realism has not been Refuted by Atomic-Cascade Experiments”, Phys. Lett. A **98**, 5-9 (1983).
- ²⁵ C. H. Thompson, “Timing, ‘accidentals’ and other artifacts in EPR Experiments” (1997), <http://arxiv.org/abs/quant-ph/9711044>.
- ²⁶ C. H. Thompson, “Subtraction of ‘accidentals’ and the validity of Bell tests” (1999), <http://arXiv.org/abs/quant-ph/9903066>.
- ²⁷ See for example C. H. Holbrow, A. Galvez and M. E. Parks, “Photon quantum mechanics and beam splitters”, Am. J. Phys. **70** (3), 260-265 (2002).

Experiments are described that enable students to confirm for themselves that light behaves like a particle at beam splitters. A different conclusion might, however, be drawn if they were to explore the possibilities of using different sources or different detectors. A more sensitive detector (not one of the modern “avalanche” photomultipliers which are, it seems, designed to ignore weak signals come what may, but a “cascade” type or earlier) would show that some light went each way. This is illustrated in papers such as P. Grangier, G. Roger and A. Aspect, “Experimental evidence for a photon anticorrelation effect on a beam splitter: a new light on single-photon interferences”, *Europhys. Lett.* **1**, 173-179(1986). Grangier *et al.* interpret their results as illustrating particle-like behaviour, but an alternative, wave, description of the experiment has been published: T. W. Marshall and E. Santos, *Europhys. Lett.* **3**, 293-6 (1987).

²⁸ A. Einstein, B. Podolsky, and N. Rosen, “Can Quantum-Mechanical Description of Physical Reality be Considered Complete?”, *Phys. Rev.* **47**, 77 (1935).

²⁹ The definition that Bell gave (page 15 of ref 32) for quantum correlation was the “expectation” value of the product of the “outcomes” on the two sides, where the “outcome” is defined to be +1 or -1 according to which of two possible cases is observed. It is to be assumed that he was using the word “expectation” in its usual statistical sense and that an unbiased estimate would be used.

³⁰ The prediction of a linear relationship for the “perfect” case is most easily verified by drawing diagrams of the ball as seen from above. The dividing circles are then straight lines through the centre and the areas required are proportional to the angles between them.

³¹ W. Tittel *et al.*, “Experimental demonstration of quantum-correlations over more than 10 kilometers”, *Phys. Rev. A*, **57**, 3229 (1997), <http://arxiv.org/abs/quant-ph/9707042>.

³² J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics*, (Cambridge University Press 1987).

³³ There are well-understood physical reasons why some recent experiments cannot be expected to have fair sampling. A clear example is the teleportation experiment of Jennewein *et al.* (T.Jennewein, G. Weihs, J-W Pan and A.Zeilinger, “Experimental nonlocality proof of quantum teleportation and entanglement swapping”, <http://arXiv.org/abs/quant-ph/0201134> (2002)). They found more coincidences when the detectors were at 0° or 90° than at ±45°, or, in other words, T_{obs} was not constant. As the authors themselves point out, at ±45° detection depends on interference effects, and hence on relative phase of vertical and horizontal components, whilst at 0° or 90°, when the axes of the analysing polarisers are parallel to those of the crystal within which the light originates, phase is irrelevant. See C. H.

Thompson’s paper, ref 35, for some related discussion.

³⁴ W. Tittel *et al.*, “Long-distance Bell-type tests using energy-time entangled photons”, <http://arxiv.org/abs/quant-ph/9809025> (1998).

³⁵ C. H. Thompson, “Rotational invariance, phase relationships and the quantum entanglement illusion”, <http://xxx.lanl.gov/abs/quant-ph/9912082> (1999).

³⁶ G. Weihs, *et al.*, “Violation of Bell’s inequality under strict Einstein locality conditions”, *Phys. Rev. Lett.* **81**, 5039 (1998) and <http://arXiv.org/abs/quant-ph/9910080>, and private correspondence.

³⁷ P.G. Kwiat, E. Waks, A. G. White, I. Appelbaum and P. H. Eberhard, “Ultrabright source of polarization-entangled photons”, *Physical Review A* **60** (2), R773-R776 (1999), <http://arXiv.org/abs/quant-ph/9810003>.

³⁸ S. J. Freedman and J. F. Clauser, *Phys. Rev. Lett.* **28**, 938 (1972).

³⁹ S. Caser, “Objective local theories and the symmetry between analysers”, *Phys. Lett. A* **102**, 152-8 (1984).

⁴⁰ A. Aspect, *Trois tests expérimentaux des inégalités de Bell par mesure de corrélation de polarisation de photons*, PhD thesis No. 2674, Université de Paris -Sud, Centre D’Orsay, (1983).

⁴¹ W. Crookes, “The Mechanical Action of Light”, *Quarterly Journal of Science* **VI**, 337-352 (July 1875).