EPR anomalies, “photon” amplitudes and Malus’ Law: an intuitive classical approach

Caroline H Thompson[*]

ABSTRACT

Observation shows that, to a very close approximation, real polarisers and photomultipliers reproduce Malus’ Law even at the “single photon” level, giving a sinusoidal variation of “photon” count with polarisation angle. At first sight, this appears to rule out various realist models of polarisers and detectors, and casts doubt on their ability to explain “paired-photon” EPR experiments. The observed violations of Bell inequalities appear to require narrower peaks and broader troughs than are produced by a perfect Malus’ Law. It is shown here in principle how amplitude restrictions on the signal pairs can contribute towards the reconciliation of the apparent conflicts.

This paper follows on from earlier work on “The Chaotic Ball” papers (Thompson, 1995a; Thompson, 1995c; Thompson, 1995b), which between them illustrate the possibility of complete realist explanations of many EPR experiments. The way is opened up for new understanding of the nature of the “quantum ensemble”, the mechanism behind Malus’ Law, and what it is that photomultipliers actually detect. It is asserted that the *only* satisfactory way to model complete EPR experiments is to treat light as *purely wave* in nature.

---

[*] Department of Computer Science, UW Aberystwyth, Dyfed, SY23 3DB, UK.
e-mail: cat@aber.ac.uk

INTRODUCTION

Several plausible realist (local) models exist (Marshall et al., 1983; Marshall, 1988; Thompson, 1995a) that can explain in principle the observed violations of Bell inequalities in Einstein-Podolsky-Rosen (EPR) experiments. One reason why they have not gained full acceptance, however, may be that they are not compatible with *all* the observations. The particular problems that inspired the present work are:

1. They predict non-sinusoidal "response functions", describing the relationships between probabilities of detection and polarisation angle, whereas the data that most nearly relates to this – that obtained during calibration runs intended to establish the characteristics of the polarisers – shows almost exact sine curves (Aspect in his thesis reports 1.5% discrepancies (Aspect, 1983, page 300)). The realist models predict narrower peaks and broader troughs than the sine curve.
2. They tend to predict sizeable variations in the total number of coincidences[1], whereas real experiments show that the total is almost constant (again with perhaps a 1.5% discrepancy).

In realist terms, therefore, it is evident that the models need some elaboration, taking into account at least one more variable. Marshall (Marshall and Santos, 1989) has produced a model that takes account of signal ellipticity. A simpler modification – and one that might be more amenable to experimental verification – might be to allow for amplitude correlations (assuming a classical wave model in which individual signals do *possess* amplitudes). This is the approach presented here. The simplest possible assumptions are made in order to demonstrate the principle involved. It emerges that the assumption of positive correlations does nothing to solve the problem, but a restriction of the population of emitted pairs to exclude the possibility of two high-amplitude signals occurring simultaneously can resolve *both* the above problems. Such a restriction would arise naturally if theory is correct in putting an upper limit on the total energy of the pair whilst real asymmetries at the source make the two signals unequal.

The approach used is similar to that of my ealier papers (Thompson, 1995a; Thompson, 1995c; Thompson, 1995b), in which a "chaotic ball" is used to illustrate realist EPR explanations. I shall work in the context of a two-channel experiment such as Aspect, Grangier and Roger's (Aspect et al., 1982; Aspect, 1983), in which pairs of signals are produced in opposite directions and each passed through a two-channel polariser with outputs labelled '+' and '−'. Photomultipliers and discriminators (or "photon dectors" (PD)) detect the signals and relay them to a coincidence monitor that records the counts (N_{++}, N_{--}, N_{+-} and N_{-+}) of '++', '−−', '+−' and '−+' events.

DEMONSTRATION OF AMPLITUDE EFFECTS

Let us look first at an idealised population of signal pairs that is all "low-amplitude", and results in low detection rates by PDs. Ignoring all other sources of variation, let us assume that we detect all those signals whose polarisation direction is within $\pi/8$ of a polariser axis and no others. The resulting "response function", the relationship between detection probability (p) and signal polarisation angle (λ), is shown in Fig. 1.

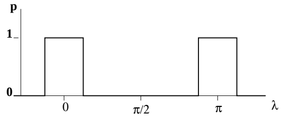

Fig. 1. Response function for low amplitude population

Assuming no complications (timing variations, asymmetries etc.), so that we have "factorability" (Thompson, 1995c) and can use the symmetries of the system to simplify the calculations, it is easy to calculate the resulting coincidence pattern. We have only to evaluate the integral

$$P_c(\phi) = \frac{1}{\pi}\int_0^{\pi} p_A(\lambda)\, p_B(|\lambda - \phi|)\, d\lambda,$$

where ϕ is the angle between detector settings and p_A and p_B are the response functions for the two signals. Alter-

---

[1] An exception is the asymmetrical case with a single "missing band", covered in "The Chaotic Ball".

natively (and with certain advantages, such as conceptual simplicity and easy extension to more general cases), we can get the same result by computer simulation.

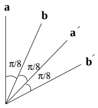

Fig. 2. Angles used in Bell test

We calculate the Bell statistic defined by

$$S_C = C(a,b) - C(a,b') + C(a',b) + C(a',b'), \qquad (1)$$

where a, a', b, b' are detector settings and C is the "correlation function" defined using the "correct" formula (Selleri, 1988, page 19),

$$C = \frac{N_{++} + N_{--} - N_{+-} - N_{-+}}{T}, \qquad (2)$$

where T is the number of pairs emitted, or

$$C = P_{++} + P_{--} - P_{+-} - P{-}{+}, \qquad (3)$$

putting $N_{++}/T = P_{++}$ etc..

Taking a, a', b, b' such that differences between relevant pairs are $\pi/8$ or $3\pi/8$ (see Fig. 2, a common choice), we find a mean value of $S_C = 1.00$, which does not, of course, violate the Bell inequality,

$$-2 \leq S_C \leq 2, \qquad (4)$$

as we have not modelled anything "non-local". (Note that symmetry considerations mean that, for example, $N_{++} = N_{--}$ and $N_{+-}(\pi/8) = N_{++}(3\pi/8)$, so that we need only in practice deal with $N_{++}(\phi)$. Also it will be found that all four terms in formula (1) are equal, so that S_C simplifies to $8(P_{++}(pi/8) - P_{++}(3\pi/8))$. The full formula should be used in real experiments, of course, as symmetry *cannot* be assumed (Thompson, 1995b)).

But, as stated emphatically my "The Chaotic Ball" paper (Thompson, 1995a), it has become customary in EPR experiments to use the erroneous "normalised" formula[2],

$$E = \frac{N_{++} + N_{--} - N_{+-} - N_{-+}}{N_{++} + N_{--} + N_{+-} + N_{-+}}, \qquad (5)$$

instead of C, and this results in replacing $S_C = 1.0$ by $S_E = 4.0$ (symmetry reduces the calculation to just $4(P_{++}(pi/8) - P_{++}(3\pi/8))/(P_{++}(pi/8) + P_{++}(3\pi/8))$, giving a gross violation of (4) *if* it applied! Admittedly, the formula would not be applied in this situation by any qualified researcher, as the accepted conditions for its validity are not met (for example, it will be found that the denominator varies markedly with ϕ, and this has been known since 1970 (Pearle, 1970) to invalidate the test). But the reader is asked to be patient. I am not assuming that the experimenter ever knowingly conducts a Bell test in such circumstances. He deals only with complete *ensembles* and these "low-low" cases we are considering are only part of the population. The data in front of him, showing coincidence rates from the whole population, shows no more than, say, 2% discrepancies from constancy.

We now repeat the exercise for a population of signal pairs that is entirely high amplitude, and assume we detect all signals except those within $\pi/8$ of the orthogonal polariser axis (see Fig. 3), and repeat it again for a population with all signals on one side low amplitude and all on the other high. Our definitions of "high" and "low" are guided by informed intuition, in that we know that the real experiment has "singles rates" just slightly less than half the rates with no polariser, but previous work has shown that violations of (4) arise easily from large "missing bands". These are typified here by our low amplitude response function, which corresponds to singles rates of 1/4. We attempt to balance this by using a high amplitude function that correponds to *double* bands, and has a singles rate of 3/4.

---

[2] The formula would be valid if the quantum theory assumption that all "photons" are identical, so that a single "quantum efficiency factor" for a PD applies to each, were correct. But in a classical wave theory this is not reasonable: the probability of detection depends on the signal amplitude and it exceeds the bounds of credibility to assume that changing the "quantum efficiency" by changing a voltage on the PD will affect all signals in the same proportion.

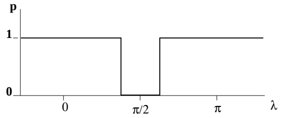

Fig. 3. Response function for high amplitude population

The results are given in the table below, in which we give individual coincidence rates as well as the Bell Statistic values. These individual values are vitally necessary for our purposes – a point that experimenters would do well to note, as we all too often see only S_E published and this on its own is not informative.

We are now in a position to see what would happen in various mixed populations, selecting whatever proportions we please for low-low, high-high and low-high (equalling high-low by symmetry) pairs. Chosing just low-low and high-high gives a rough simulation of positive amplitude correlation. The choices of most interest turn out to be (I) the "no-correlation" case, with proportions obtained by multiplying the assumed proportions for each side separately, and (II) the case that is the same except that the high-high possibility is excluded.

Thus we can calculate the values of S_C and S_E for a population that has $ll : lh : hh$ (using notation hh = number of high-high pairs etc.) in proportions $\alpha : \beta : \gamma$ by using weights α, β and γ to calculate new numerators and denominators for C and E. Alternatively, we can simulate mixed populations directly.

	Pure populations			Mixed populations	
	ll	lh	hh	49:42:9 Independent (I)	49:42:0 No high-high pairs (II)
$P_{++}(0)$	0.250	0.250	0.750	0.295	0.250
$P_{++}(\pi/8)$	0.125	0.250	0.625	0.223	0.183
$P_{++}(\pi/4)$	0	0.250	0.500	0.150	0.115
$P_{++}(3\pi/8)$	0	0.125	0.500	0.098	0.058
$P_{++}(\pi/2)$	0	0	0.500	0.045	0
S_C	1.00	1.00	1.00	1.00	1.00
S_E	**4.00**	**1.33**	**0.44**	**1.56**	**2.08**

The values chosen for the proportions were variations on the themes of mixes of low to high in the ratio 8:2 or 7:3. For example, a 7:3 mix with the assumption of *independence* results in proportions given by the body of the table:

	low	high	total
high	0.21	0.09	0.3
low	0.49	0.21	0.7
total	0.7	0.3	1.0

It was found that the 8:2 mix gave infringements of the S_E version of (4) even when pairs were assumed independent (giving the full 64:32:4 weightings), though omitting the high-high category did increase the extent of the infringement. The 7:3 mix gave results that might be more relevant to the real experimental situation. As shown in the table and graphically in Figs. 4 and 5, the omission of the high-high group was critical: there is greater "visibility" to the coincidence pattern, and greater violation of the inequality, without it but none with. Thus here we have demonstrated a situation in which the "singles" data – in practice, the polariser calibration data – would not suggest that violation was possible, yet the existence of a violating subpopulation determines the behaviour of the whole.

The singles data, in the idealised form of the response curve, is shown for the mixed cases in Fig. 6. There are just slight differences between the two cases modelled. It is not these small differences, however, that cause the differences in behaviour. It was checked that *independent* populations having response curves *identical* to the "no high-highs" one produced a Bell statistic, S_E, of only 1.88.

Another result that may be of interest to some is the case of a slight *positive* correlation, modelled by omitting the high-low and low-high groups. A 49:0:9 mix gave S_E = 1.78, again confirming that, to obtain a violation of the Bell inequality, we need to reduce the smoothing effect of the high-high group. It is *peaked* response curves, or ones with substantial zero regions, that produce violations.

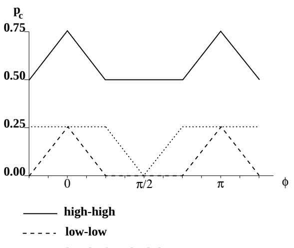

Fig. 4. Expected coincidences per emitted pair: pure populations

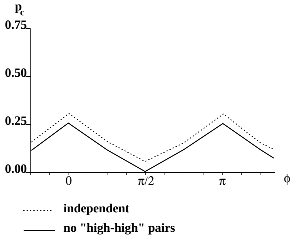

Fig. 5. Expected coincidences: mixed amplitude populations

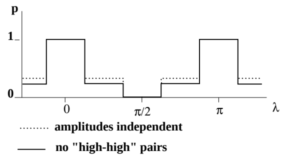

Fig. 6. Response functions for mixed amplitude populations

If Fig. 5 were presented in "normalised" form, i.e. the form corresponding to (5), it would show even more dramatically how our very cude assumptions have resulted in predictions that are surprisingly close to actual experimental results such as those of Aspect. I have chosen instead to present graphs (Figs. 7 and 8) of the total numbers of coincidences, the denominator of (5). These illustrate in striking manner how the near constancy of the total is a very dangerous guide to constancy within subpopulations.

Note that the total value plotted is not a probability but an expected number of coincidences per emitted pair, and our present model would allow this to go as high as 4. This is because there is no built-in assumption that the "photon" cannot be split, producing both + and − detections simultaneously. It would be possible in a more realistic model, allowing for low detection efficiencies, to reduce the assumed double detection rates to one compatible with experimental evidence such as Grangier, Roger and Aspect's of 1986 (Grangier et al., 1986). The reader should note that there is no evidence that double detections *never* occur, only that they are somewhat less frequent than a naïve classical theory might predict (Marshall and Santos, 1987).

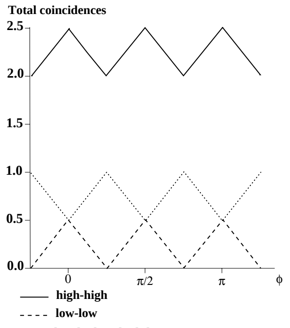

Fig. 7. Total number of coincidences per emitted pair for pure populations

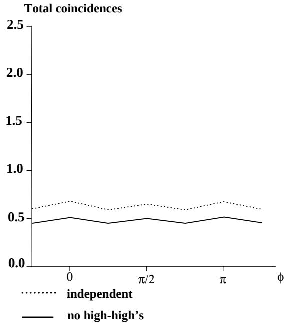

Fig. 8. Total number of coincidences per emitted pair for mixed populations

From an academic point of view, our findings on the effects of hidden subpopulations relate to work by Garuccio et al. (Garuccio et al., 1990) on "factorability": the full model, with its "amplitude dimension", is technically factorable, and this is what matters to ensure the validity of a *true* Bell's test (witness the non-violation by the S_C values). It is tempting to think that we should be able to predict coincidences from the "singles" data obtained when calibrating polarisers. But this data corresponds approximately (i.e. ignoring uncertainties in λ) to that of the mixed populations of Fig.6, and this averaged level is not factorable[3]. There is no way of deriving correct estimates of coincidences directly from here unless there are no amplitude correlations (we have independence). It is quite clear, though, that it is the use of an invalid formula, and not anything directly due to non-factorability,

---

[3] Contrast the genuinely non-factorable situation that can arise if we have timing biases (Thompson, 1995c).

that is in the present example causing inequality violations.

DISCUSSION

Could it be that what we actually obtain in, say, a paired-photon experiment using an atomic cascade source, is a population of pairs that are not only correlated in polarisation angle but also (negatively, in the manner illustrated above) in amplitude? Such an assumption certainly makes a contribution towards the solution of the two problems we set out to solve.

There is also an unexpected bonus to be gained from the idea that the observed adherence to Malus' Law by individual "photons" is partly due to being built up from whole distribution of different amplitudes. Aspect mentions (Aspect, 1983, pages 265-7) that his source of signals was not "rotationally invariant", and that this led to slight variations in the estimated values of one of the transmission coefficients as the first polariser was rotated, when using two polarisers to calibrate the second one. Now, what he would actually have observed would have been simply a slight change in *shape* of the probability-polarisation angle curve. In the current context, this should be interpreted simply as the direct effect of different *amplitudes* reaching the second polariser, the variations being, in turn, attributable to failure of rotational invariance, as he correctly states.

Amplitude variations are thus a distinct possibility, but they are unlikely to be the whole story. The ideas in this paper have, for instance, produced singles probabilities of less than 1/2, which conflicts with experimental evidence, and it may well be necessary to allow for enhancement, timing variations or both to get a value nearer the observed one. Also, of course, random variations, caused by interactions with the zero point field at both polarisers and PDs, will cause the response curves for fixed amplitudes to be rounded rather than plain rectangular.

The reader may well wonder why, if real PDs behave in anything like the assumed manner, this does not sometimes produce polariser calibration curves that are non-sinusoidal? I believe that part of the answer will be found to be in the nature of the amplitude distributions that occur naturally and details of the detection process (both very open subjects and too large to enter into here). The other part of the answer, though, lies in experimental practice: if the curve is *not* sinusoidal, we realise that our instrument is "too sensitive/insensitive" for our signal, and adjust either signal or instrument accordingly! This action is easily justified, as the official versions of both classical and quantum theories predict that detectors will reflect Malus' Law. So the custom is to force the apparent adherence to the rule, but the experimenter cannot do this for individual signals, and can only guess when parts of the quantum ensemble might be asserting their rights to individuality.

There is nothing new in the idea of assuming that non-conformity to an expected rule means a fault in the experiment. It would not be reasonable in most circumstances to interpret it as evidence of a different rule coming into play[4]. But EPR experiments furnish some rather extreme cases of "accepted theory bias". Alain Aspect, for example, openly states in his thesis his conviction that earlier experiments had already proved quantum mechanics right (he saw his role as checking that violations of Bell inequalities could not be caused by exchange of signals unless faster than light). He therefore felt little compunction in accepting Malus' Law, which his data seemed to support, adjusting parameters (thresholds, lens sizes etc.) to minimise "dissymetrie" (variability of the sum of + and − counts with polarisation angle) and keep coincidence counts as constant as possible. He also went along with accepted opinion to the effect that experimental imperfections could only decrease violations of Bell inequalities, not increase them[5], so that there was no real limit to the extent to which he felt justified in chosing the "best" parameters. Furthermore, he felt free to use the quality of the reproduction of quantum mechanics predictions as a measure of the correctness of the experimental conditions (achieving a "pure quantum state")[6]

At the time of Aspect's experiments (around 1980), there were no classical theories available that could have ex-

---

[4] " If all members of a community responded to each anomaly as a source of crisis or embraced each new theory advanced by a colleague, science would cease." (Kuhn, 1970, page 186).

[5] "De plus, toute imperfection expérimentale à tendance a faire disparaître le conflit entre prédictions quantiques et inégalités de Bell" (Aspect, 1983, page 3).

[6] "...la vérification de l'accord entre les mesures et les prévisions quantiques est un moyen privilégié pour s'assurer que les divers processus son parfaitement contrôlés" (Aspect, 1983, page 116).

plained satisfactorily all of his results (or, if there were, Aspect was evidently unaware of them). He was therefore justified to some extent in making adjustments so as to match quantum theory predictions. The situation is different now. It is possible, in hindsight, to see how various discrepancies that were to Aspect just curious anomalies all add up to a coherent body of facts that indicate that the accepted model is *not* quite correct: a classical theory has the potential to do better. More experimentation (or perhaps just access to laboratory notes for existing experiments) is needed to enable the details of the theory to be filled in. The classical model of paired-photon EPR experiment will not comprise such concise formulae as the conventionally-accepted quantum mechanical description – this would not be compatible with the complexity and intrinsically composite nature of the actual setup. It will, however, not only give better predictions but will have immeasurable advantages in terms of eliminating magic from science.

REFERENCES

Aspect, A., Grangier, P., and Roger, G., 1982, Experimental realization of Einstein-Podolsky-Rosen-Bohm gedankenexperiment: A new violation of Bell's inequalities: Physical Review Letters, **49**, 91.

Aspect, A., 1983, Trois tests expérimentaux des inégalités de Bell par mesure de corrélation de polarisation de photons: Ph.D. thesis, Université de Paris-Sud, Centre D'Orsay.

Garuccio, A., Lepore, V. L., and Selleri, F., 1990, Probabilities for correlated systems: Foundations of Physics, **20**, 1173–1189.

Grangier, P., Roger, G., and Aspect, A., 1986, Experimental evidence for a photon anticorrelation effect on a beam splitter: a new light on single-photon interferences: Europhysics Letters, **1**, 173–179.

Kuhn, T. S., 1970, The structure of scientific revolution: University of Chicago.

Marshall, T. W., and Santos, E., 1987, Comment on "experimental evidence for a photon anticorrelation effect on a beam splitter: a new light on single-photon interferences": Europhysics Letters, **3**, 293–6.

Marshall, T. W., and Santos, E., 1989, Stochastic optics: A local realistic analysis of optical tests of Bell inequalities: Physical Review A, **39**, 6271–6283.

Marshall, T. W., Santos, E., and Selleri, F., 1983, Local realism has not been refuted by atomic-cascade experiments: Physics Letters A, **98**, 5–9.

Marshall, T. W., 1988, Stochastic electrodynamics and the Einstein-Podolsky-Rosen argument in: *Quantum Mechanics versus Local Realism, The Einstein-Podolsky-Rosen Paradox* (F. Selleri ed.), Plenum Press, pages 413–432.

Pearle, P., 1970, Hidden-variable example based upon data rejection: Physical Review D, **2**, 1418–25.

Selleri, F., 1988, Quantum mechanics versus local realism. The Einstein-Podolsky-Rosen paradox: Plenum Press.

Thompson, C. H., 1995a, The chaotic ball: An intuitive analogy for EPR experiments: submitted for publication in Foundations of Physics Letters.

——— 1995b, A hidden variable explanation of Aspect's asymmetry anomalies: unpublished.

——— 1995c, A non-factorable local realist model for EPR experiments with timing variations: unpublished.