PROBABILITIES INVOLVING DIRECTIONAL SIMILARITY

JOHN F. NEWELL,

*University of Pittsburgh**

◇

Abstract

This paper explores the mathematics for discovering, comparing, or computing the degrees of similarity or dissimilarity between the orientations of two objects, such as polarizing filters, by an analogy with the directional aspect of Euclidean vectors. The method presented here involves projecting the orthogonal components of a vector back onto it, thus forming constituent parts which reveal the influence of the components, and so allow for the formation of percentages and probabilities which reflect the degree of similarity (or difference) in the orientations of the components and the vector. A similar procedure can reveal the directional similarity (or difference) between two vectors, by treating one as an orthogonal component of the other. The resulting equations provide the mathematical basis for the Law of Malus and for violations of Bell Inequalities. In consequence, the anomalous aspect of such violations is shown to arise from expectations that result from applying the wrong sort of probabilities (i.e., probabilities which are not related to directional similarity) to the experiments (the key variable of which is the relative rotations of filtering devices).

Keywords: Euclidean vector; direction; degree of similarity; new method; EPR correlation; optics; Malus' law; polarization; entanglement; quantum mechanics; Bell's theorem; Bell's inequality; CHSH inequality; CH inequality

2010 Mathematics Subject Classification: Primary 51P05
Secondary 78M25; 81P40; 81V80

1. Introduction

In quantum mechanics, a problem arises when two verifiably independent polarizers are used to test photons with similar properties. It arises due to the following reasons:

(i) the results for unpolarized photons passing through a given polarizer show a random distribution of the possible outcomes,

(ii) the two polarizers show a correlation rate which varies with an equation whose sole variable is θ, the angle they would form if they were superimposed,

(iii) the equation, which emerged as a curious byproduct of the Schrödinger equation, cannot be accounted for by any known process for calculating probabilities.

This was first flagged as problematic by Einstein *et al.* [4], and then clarified as a problem by Bell [1].

This same equation (essentially, $cos^2\theta$) also arises in optics when a beam of unpolarized light is passed through two successive polarizers. Half of the light passes through the first polarizer, while the amount passing through the second depends on the angle of rotation, θ,

*Postal address: Department of Classics, University of Pittsburgh, 1501 Cathedral of Learning, 4200 Fifth Avenue, Pittsburgh, PA 15260, USA. Email address: jfn2@pitt.edu

◇*Based on an article of the same name first printed in Math. Scientist* **39,** 37–44 (2014)
and a correction printed in Math. Scientist **41** 74 (2016)
by The *Applied Probability Trust*
Received 3 May 2013.

between the two polarizers. Étienne-Louis Malus discovered this in the early 1800s when he noticed that the equation provided a close fit to his data.

These situations suggest the following conclusions

(i) there exists a set of hitherto unknown probabilities which are governed by relative rotational orientations,

(ii) a group of these probabilities is formed from a distinct set of equations which feature the term $cos^2\theta$,

(iii) they reveal what we may call ‘directional similarity’, that is, the degree to which two or more axial orientations share a commonality in their spatial positions.

Since these phenomena are restricted to light and polarizers, we might contend that what is needed is a more exact description of their physical interactions. These experiments, however, have merely called our attention to a commonplace occurrence which has, until now, escaped notice. A few examples can help to illustrate this fact.

1.1. Commonplace examples

The hands of a clock provide two such examples. One occurs when people report the time as, say, 1:30 or 1:25 when the clock actually reads, say, 1:28. We might provide some vague account for this based on proximity, but is there a mathematical basis, an actual equation, which can supply the probability for a given response?

The second example arises when we consider the second hand as it moves in relation to the minute hand: each moment, it is either closer to or further from the minute hand, so that every position of the second hand should have a degree of similarity (and difference) with the position of the minute hand. Yet, what is the equation which provides the value for the degree of similarity at any moment? What is its basis in geometry? And how does it relate to θ, the angle formed between the two hands?

A third everyday example is supplied by a common compass. Compasses often feature a diagram, called a compass rose, which indicates at least the four cardinal directions (labeled N, S, E, and W). In some cases, an eight-point rose is provided by halving the arcs defined by the four cardinal directions (inserting directions labeled NE, SE, SW, and NW). The same process is sometimes used to produce a sixteen-point rose (introducing directions labeled NNE, ENE, ESE, SSE, and so forth) and a thirty-two-point rose can be formed in the same way (leading to points labeled NbE, NEbN, NEbE, EbN, and so forth, which are read ‘north by east’, ‘northeast by north’, and so forth).

These labels indicate a hunch that any given direction is somehow composed of a combination of the two closest cardinal directions, and this suggests that there should be a mathematically precise recipe (two parts N, one part E, and so forth) for every position of the needle.

We may arrive at this recipe by taking the number of degrees in the angle, θ, arising between the position of the compass needle and one of the neighboring cardinal directions, and dividing it by the number of degrees in the angle, ξ, between the cardinal directions on either side of the needle (that is, 90°). Thus, for example, if θ, measured from N in the direction of E, is 17°, then we have an 18.9% deviation from N, and a 71.1% deviation from E, and we can say that the position is 71.1% N and 18.9% E.

This system, however, is not the one we are looking for, because the experiments in quantum mechanics and optics discussed earlier use an equation which features $cos^2\theta$, and, for most values of θ, such a system is not compatible with the one we have just described for the compass rose.

Such incompatibility, however, is no more fatal to our search than the existence of the harmonic mean is fatal to the existence of the geometric mean, the median, or the average. Having found one system, we may even be encouraged, since this may be taken to indicate that there should be more. Why nature might use one system rather than another would be an interesting question for physicists or philosophers to pursue, but our task is to find the mathematical meaning of the system that features $\cos^2\theta$.

Aside from forming part to whole ratios between θ and ξ, the known mathematical relationships which include indications of directional similarity include the relation of vectors to their components, and the relation between lines and their projections. We will, therefore, begin our investigation with a closer look at these.

2. Euclidean vectors and their components

The practice of resolving Euclidean vectors into orthogonal components has long played a useful and respected role in various fields of applied mathematics. The process itself is well known: take any Euclidean vector, $\mathbf{v}$, draw a line, l, through its head, R, or tail, P, which will form an acute angle θ; on the opposite end of the vector, form the perpendicular to l which meets it at Q. The orthogonal components, a and b, are the legs of the right-angled triangle thus formed (see Figure 1). In many cases, the directions of the orthogonal components are known in advance (as, for example, when we seek the vertical and horizontal components, or when we seek, say, the north and east components of a given directional heading). In such cases, a and b are found by first drawing l parallel to one of the desired components, and then completing the right-angled triangle.

Since the components form a right-angled triangle, Pythagoras' theorem is applied when we seek to determine the quantitative values of a, b, or $\mathbf{v}$. Thus, the magnitude of $\mathbf{v}$ (or v, which may be arrived at by dividing $\mathbf{v}$ by a parallel unit vector) is equal to $\sqrt{a^2+b^2}$. In the special case that arises when $\mathbf{v}$ is a unit vector, its magnitude, $v = c = 1$, is equal to $a^2 + b^2$.

Pythagoras' theorem is, of course, well known, and there are many variations of proofs for it. Of interest to us here will be the form, shown in Figure 1, which makes use of similar triangles. The elegance and cleverness of this demonstration lies in the simplicity of forming a single perpendicular from the right angle to the hypotenuse, cutting it into segments e and f. At this point, we see that $c^2 = ce + cf$, and that $a/c = e/a$ while $b/c = f/b$, so that $ce = a^2$ and $cf = b^2$; it then follows that $a^2 + b^2 = c^2$.

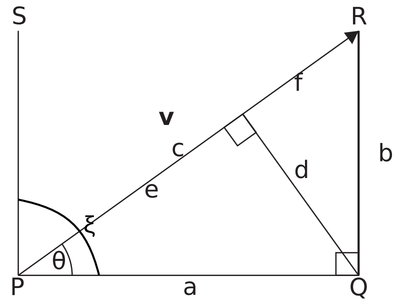

Figure 1: Component a projected onto $\mathbf{v}$ at e, and b projected onto $\mathbf{v}$ at f

If we set $c = 1$, and if the angle formed at the intersection, P, of c and a is θ, then $c^2 = c$, $a = \cos\theta$, $b = \sin\theta$, and $c = \sin^2\theta + \cos^2\theta$. Since $ce = a^2$, it follows that

$$e = \cos^2\theta\,, \tag{1}$$

and similarly (or by subtraction),

$$f = \sin^2\theta\,. \tag{2}$$

This yields, substituting e and f for a^2 and b^2 in Pythagoras' theorem,

$$e + f = c = 1, \tag{3}$$

which we can readily see by simply looking at Figure 1 or by knowing the relation of the parts to the whole when a single straight line is cut into segments by another straight line in Euclidean space.

3. Incorporating components into a vector

These well-known relations become important to us when we also consider the practice of projecting one line onto another in Euclidean space. This is done by forming a perpendicular from the end of the line we want to project, say a at Q, onto the line we want it projected upon, say c. Thus, in Figure 1, we may regard e as the projection of a onto c, or f as the projection of b onto c. Figure 1 then shows a and b projected onto, c and the result is no overlap or shortfall: the projection of a plus the projection of b equals c (see (3)). Consequently, we can form part to whole ratios, e/c and f/c, or, multiplying these ratios by 100, percentages, which tell us to what extent c is composed of the projection of a, and to what extent it is composed of the projection of b. These percentages can then serve to indicate the degree to which c is similar or compatible with a or b, because the process, in effect, *incorporates* the legs into the hypotenuse: the projection is the transformation of the leg into an ingredient.

Thus, to return to discussing vectors, the component isolates a directional attribute, and its projection onto the vector instantiates that attribute as an ingredient or part of the vector. So, for example, if a aligns with the horizontal, and b with the vertical, while $\mathbf{v}$ is a unit vector in some direction in the xy plane, then $\mathbf{v}$ is $(100e)\%$ horizontal and $(100f)\%$ vertical, and e and f can be readily calculated by knowing the value of θ and using (1) and (2).

We might easily see, then, how this system could apply to similar situations involving orthogonal components. Let us now expand this discussion to include degrees of similarity (or difference) between two vectors.

3.1. Comparing two vectors

When dealing with two vectors, we must first understand that we are only interested in their direction, not their magnitude. If, for example, we consider a truck that is driving south, it does not matter, in regard to its southward orientation, whether the truck is full or empty, or whether it is speeding or crawling along, or whether it is an eighteen wheeler or a children's toy. And, once your heading is south, you can not make your heading more southward. So, with vectors, quantities such as speed or mass are expressed by the magnitude of the vector, i.e., its length. Since our interest is in the direction of the vectors, we can let their lengths be any arbitrary values, but life (and the math) will be simpler if we set one vector, $\mathbf{v}$, as a unit vector, and set the other, $\mathbf{w}$, equal to $\cos\theta$ (which is the length, c, we will cut of on $\mathbf{w}$ (or its extension), if we form a perpendicular, b, to it at Q from the head, R, of $\mathbf{v}$), where θ is the angle formed if we join them tail to tail at P (Figure 2), Thus, a is the component of $\mathbf{v}$ in the direction of $\mathbf{w}$, and b is the orthogonal component of $\mathbf{v}$. So, by incorporating a into $\mathbf{v}$, we can find the degree of similarity between $\mathbf{v}$ and $\mathbf{w}$, and, by incorporating b into $\mathbf{v}$, we can find the degree of difference between them. This

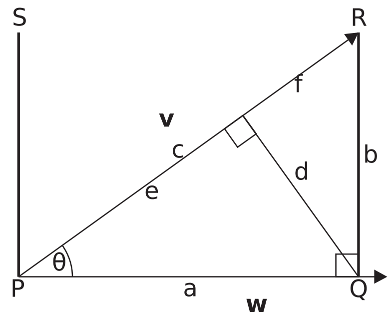

Figure 2: Degrees of similarity, e, and difference, f, in $\mathbf{v}$ with respect to $\mathbf{w}$.

is done, as above, by forming the altitude, d, which cuts $\mathbf{v}$ into two segments, e and f; (1)–(3) follow, with (1) giving the degree of similarity or compatibility and (2) giving the degree of dissimilarity or incompatibility.

We may now develop a more general understanding of directional similarity, since, to this point, our analysis has considered only situations involving unit vectors which form an acute angle, that is, situations in which the confining angle, ξ, is equal to 90°. Incidentally, we can discover ξ by forming the perpendicular PS (resulting in angle SPQ or ξ; see Figure 1), and seeing that e ranges from 1 to 0 as θ changes from 0° to 90° within ξ. For a more complete understanding, we must next account for what happens when ξ has any given value.

3.2. Cases involving nonorthogonal reference axes

Let us now regard the point P in Figure 1 as a hinge for SP and PQ, so that we may adjust ξ to any value in the plane of the page. No matter what the value of ξ is, the probabilities must still range from 0 to 1, satisfy (3), reach 0.5 when $\theta = \xi/2$, and show a smooth gradation of values as the value of θ is changed by moving c from, say, coinciding with PQ (that is, a) to coinciding with SP (which may be regarded as b since it is parallel). We should also expect that the equations for the probabilities should have the same basic form as they do when $\xi = 90°$, that is, the equation for the orthogonal case may show some variation from the equations for other values of ξ (e.g. a variable may have ‘vanished’ by becoming a 1 or 0), but its terms will not be completely unrelated to theirs.

Since (1) features the cosine function, let us take a more generalized look at this function. The standard cosine function, which graphs as a wave, may be written as

$$y = A\cos^{n} B(x + C) + D, \tag{4}$$

where A indicates the amplitude of the wave, C and D govern horizontal and vertical displacements, respectively, and $B = (m+p)\,\pi/\lambda$ where λ is the wavelength in radians, and m and p govern any horizontal compression or stretching of λ. In (1), for example, $A = B = m = 1$, $n = 2$, $x = \theta$ in radians, and $C = D = 0$. The quantity m is governed by n, such that (for whole number values of n) $m = 1$ if n is even, and $m = 2$ if n is odd (this occurs because even powers of the cosine function only produce positive values for the standard wave, so the wavelength, λ, recurs with π; with odd powers, the curve alternates into positive and negative segments, each extending horizontally to a length of π, and so it takes two such segments to constitute a single wavelength). In addition, ξ, the angle within which the probabilities are confined, governs λ, for it is essentially a continuous portion of λ which runs from 0 to $+1$ (or vice versa) without repeating any values (a full wavelength for a standard wave, in contrast, runs, say, from 0 to $+1$ and back again to 0 when $m = 1$, and, when $m = 2$, it extends not only this far but also extends down to -1 and then returns again to 0 before it is completed). This indicates that the relationship between λ and ξ is also a function of m (ξ being half of λ when $m = 1$, and a quarter of it when $m = 2$). In general, if we express ξ in radians, then $\lambda = 2^m\xi$ and the extended form of (1) is

$$e = 1\cos^2\left(\frac{1\pi}{(2^1)\xi}\theta + 0\right) + 0$$

This simplifies to

$$e = \cos^2\left(\frac{\pi}{2\xi}\right)\theta. \tag{5}$$

If $\xi = 90^\circ$ or, in radians, $\pi/2$, we get (1). If $\xi = 180^\circ$ (or π radians) we get

$$e = \cos^2\left(\frac{1}{2}\right)\theta\ . \tag{6}$$

For other values of ξ, we could now readily arrive at analogous equations by inserting the correct value for ξ (in radians) into (5).

We can get (2) and its variants for values of ξ simply by substituting f for e and sine for cosine in the foregoing discussion.

4. Application to Known Problems

EPR and Bell-like experiments involve separating two quanta (Q_1, Q_2) and 'testing' them by passing each through a magnetic field (if using electrons) or a polarizing filter (if using photons) which is assigned to it (F_1 for Q_1, F_2 for Q_2). One important aspect of the test is to separate F_1 and F_2 far enough apart that there is not enough time (given the constraints of Relativity Theory) for a report of what happened at F_1 to reach F_2 (or vice versa). The key aspect of the test, however, involves rotating F_1 and F_2 relative to each other, since the equations arising from quantum mechanics identify the angle of their relative rotation θ as the *only* important variable for deciding the probability that the outcomes at F_1 and F_2 will correlate or not. Since the probabilities vary with $cos^2\theta$, Bell [1] identified certain angles at which the values predicted by quantum mechanics vary maximally with the values predicted by standard probability theory and so these 'Bell angles' are routinely used for the tests.

Since F_1 and F_2 may be separated by a large distance, determining θ can be tricky business. One way to accomplish it is to start with F_1 and F_2 together, orienting them so that $\theta = 0°$, then marking them at the same spot (e.g., at their topmost part, just as clock faces have a 12 which is usually in the highest position), and then separating them (being careful to always know how to duplicate the original orientation). Since the rotations are a relative relationship, we can reduce potential confusions by keeping, say, F_1 in a fixed position, and rotating F_2 to the desired angle(s), θ.

The experiments proceed by selecting pairs of quanta and sending one to F_1, the other to F_2, and then comparing the results, that is, whether the filters blocked the quanta, allowed both to pass, or let one pass and blocked the other. It should be intuitively plausible that the outcomes should correlate (both blocked or both allowed to pass) to the degree that the orientations of the filters are the same, and that the orientations of the quanta are the same . (For the original EPR experiment, the filters used were magnetic fields which deflected the passing electrons, so the correlations were not a matter of being blocked or not, but of being deflected in one direction or its opposite.) This is, essentially, what the quantum equations have been telling us, and we now have a mathematical basis for understanding why $cos^2\theta$ is the proper measure of this similarity.

More specifically, the CH Bell test [2] uses (6), while (1) is the equation used in the CHSH Bell test [3]. For the original EPR situation [4], since the electrons anti-correlate (e.g., one goes 'up' when the other goes 'down'), it can be useful to multiply A in (4) by -1. When counting only a part of the sample, represent this by a fractional value for A, (s/t); for example, if we only count the photons that pass through a filter (i.e., not those that are blocked), we are missing half of the correlations so s/t = ½.

With the Law of Malus, which is (1) multiplied by the intensity of the incident beam, a photon that passes through F_1 is found to pass through F_2 with a probability of $cos^2\theta$, that is, the degree to which their orientations are the same. In this case, F_1 and F_2 do not have to be separated by a large distance, so the experiment is much easier to run.

Since this probability equation surfaced only in contexts involving quanta, and since it was otherwise unexplained and unexpected, it inspired speculations that quanta behaved by rules that do not apply to more familiar objects. The mathematics, however, does not indicate any reason to restrict its application to quantum objects. Consequently, it should be possible to build macroscopic models that duplicate the behavior of quanta in these circumstances; the difficulty would be in finding a material that behaves like quanta (perhaps dough balls) and in constructing something that approximates the filters or magnetic fields involved. It should also be possible to apply this equation to other circumstances, such as figuring the impact of tilt on a support's bearing capacity.

Another set of speculations that arose in this connection were concerns (or hopes) about action-at-a-distance or other magical or metaphysical relations. These arose because a causal link between Q_1 and Q_2 was hypothesized. We now see, however, that it is the relative orientations of F_1 and F_2 that govern the correlations. Since this relationship is established *as soon as the filters are oriented*, no matter how far apart they are, it *looks like* an instantaneous causal link, but it is not causal: changing the orientation of, say, F_1, does not have any immediate material effect on F_2, but their positions relative to each other (or to any other object) are immediately changed (e.g. , one is now further from the other [or nearer], higher [or lower], tilted, etc.—it takes no time for these relations to be established). These are not causal relations but matters of happenstance. This is not to say that such relations are not useful, but only to clarify that they do not require some elaborate metaphysical adjustment of our understanding of how nature works.

5. Conclusion

Starting with problems in quantum mechanics and optics which yield equations for probabilities which are, firstly, unlike any known equations for probabilities, and, secondly, discovered without any insight into their meaning or geometric origins, we have come to see that they indicate degrees of directional similarity between two axes. In our investigation of directional similarity, we have found two independent systems for forming such probabilities. One involves the ration of the size of the angle formed by the axes under consideration of the size of the central angle subtending a confining arc, that is, θ/ξ, as discussed in Section 1.1. This system underlies the traditional system for labeling the directions around a compass rose. The other system involves finding the component in the direction of the second axis, and then incorporating that component back into the first axis. This creates a part-to-whole relationship between the two axes and yields an equation as simple as (1), or as compex as (4).

This second method is significant because it explains the origins of the 'mysterious' probabilities which have been in use in optics since the early 1800s, and in quantum mechanics since the 1920s. Specifically, Malus' law is (1) multiplied by the intensity of the incident beam, and we may now understand this as indicating that the probability that a photon will pass through both filters depends on the degree of similarity in their alignments. In quantum mechanics, (6) is featured in a modification, known as the CH Bell test [2], of an experiment proposed by Bell [1], which served to prove, firstly, that this system of probabilities was independent of other known systems, and, secondly, that quantum phenomena followed this system. Equation (1) is featured in a second such modification, known as the CHSH Bell test(named for the varient of the Bell Test developed by Clauser *et al.* [3]). In both cases, the principal question revolves around the rate that F_1 and F_2 produce matching or mismatching outcomes when testing similar photons.

Consequently, while the 'mysterious' quantum mechanics' probabilities have led to various non-scientific conjectures, we can now see, firstly, that the mystery arose from the fact that probabilities based on directional similarity were unknown and, secondly, that the correlation rates revealed by experiments (and the resulting violations of Bell inequalities) arise from the degree of similarity in the tests conducted (which results from the similarity in the orientations of the filtering equipment). Since this relationship is established as soon as the filters are oriented, it can create the *illusion* of an instantaneous causal link, but this is not a causal situation: it is a relational circumstance (like comparative height or weight) involving the degree of similarity in the orientations of the filters.

This discovery solves the mathematical puzzle which Bell and others defined, but it leads to some further work in physics, since it may now be possible to construct improved physical or theoretical models of photons, polarizers, and their interactions. These tasks, however, constitute a separate endeavor.

References

[1] Bell, J.S. (1964) On the Einstein Podolsky Rosen paradox. *Physics* **1**, 195-200.

[2] Clauser, J.F. and Horne, M. A. (1974) Experimental consequences of objective local theories *Phys. Rev. D* **10**, 526-535.

[3] Clauser, J.F., Horne, M.A., Shimony, A., and Holt, R.A. (1969). Proposed experiment to test local hidden-variable theories *Phys. Rev. Lett.* **23**, 880 884.

[4] Einstein, A., Podolsky, B. and Rosen, N. (1935). Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* **47**, 777-780.

. -

s ca