PROBABILITIES INVOLVING DIRECTIONAL SIMILARITY

JOHN F. NEWELL,*

Abstract

The mathematics behind a method for discovering, comparing, or computing the degrees of similarity or dissimilarity between the orientations of two Euclidean vectors are explored. The method has been in use since the 1800s, but only as a formula whose mathematical origins have been unexplained. The absence of an explanation has led to confusion and speculation regarding causal links. The explanation offered here involves projecting the components of a vector back onto it, thus forming constituent parts which reflect the influence of the components. This allows for the formation of probabilities which reflect degrees of similarity between the vector and its components. In the related experiments, then, the outcomes stochastically reflect changes in the orientations of the detection equipment, and so the search for causal links is shown to be unnecessary.

Keywords: Euclidean vectors; directions; degree of similarity; new probability method; EPR correlations; optics; Law of Malus; polarization; entanglement; quantum mechanics; Bell's theorem; Bell's inequality; CHSH inequality; CH inequality

2010 Mathematics Subject Classification: Primary 51P05

2010 Mathematics Subject Classification: Secondary 60G99; 65C50; 70G55; 78A05; 78A10; 78M25; 81Q99; 81V80; 82D99

* Postal address: 6807 Thomas Blvd., Pittsburgh, PA, 15208, USA

This article is a corrected version of an article first published in *The Mathematical Scientist* Vol. 39(1)

1. Introduction

In quantum mechanics, QM, a problem arises when two verifiably independent polarizers are used to test photons with similar properties. It arises because *(1)* the results for unpolarized photons passing through a given detector show a random distribution of the possible outcomes, *(2)* the two detectors show a correlation rate which varies with an equation whose sole variable is θ, the angle they would form if they were superimposed, and *(3)* 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, Rosen, and Podolsky [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, between the two polarizers. Etienne Malus discovered this in the early 1800s when he noticed that the equation provided a close fit to his data.

These situations suggest *(1)* that there exists a set of hitherto unknown probabilities which are governed by relative rotational orientations, *(2)* that a group of these probabilities are formed from a distinct set of equations which feature the term $\cos^2\theta$, and *(3)* that 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, one 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 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 features 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 QM and optics discussed earlier use an equation which features $\cos^2\theta$, and, for most values of θ, such a system in 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 triangle thus formed (see figure 1). In many cases, the directions of the orthogonal components are known in advance (as, for example, when one seeks the vertical and horizontal components, or when one seeks, 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 triangle.

Since the components form a right triangle, the Pythagorean theorem is applied when one seeks to determine the quantitative values of a, b, or $\mathbf{v}$. Thus, the magnitude of $\mathbf{v}$ (or $||\mathbf{v}||$, 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, $||\mathbf{v}|| = c = 1$, is equal to $a^2 + b^2$.

The Pythagorean theorem is, of course, well known, and there are many variations of 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, realizing 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 follows that $a^2 + b^2 = c^2$.

If we set $c = 1$, and if the angle formed at the intersection, P, of c and a is θ, then

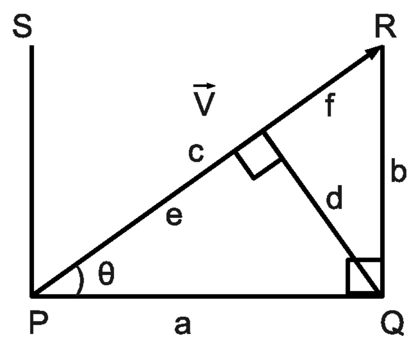

FIGURE 1: Component a projected onto **V** at e, and b projected onto **V** at f

$c^2 = c$, $a = cos\theta$, $b = sin\theta$, 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 the Pythagorean theorem:

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

which one can readily see by simply looking at the diagram 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 one wants to project, say a at Q, onto the line one wants 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. The figure, 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 [equation (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 horizontal, and b with 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 equations (1) and (2).

One might readily see, then, how this system might 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

The analysis above can readily be applied to two equal Euclidean vectors, $\mathbf{v}$ and $\mathbf{w}$ which, placed tail-to-tail (or head-to-head) form an acute angle. Because they are equal, we may readily treat them as (or reduce them to) unit vectors. Working within their common plane (or, where this is not possible or desirable, with transformations or projections of one or both onto a common plane), place the vectors tail-to-tail (or head-to-head), forming the acute angle θ at point P as in figure 2. Selecting one of the vectors, say $\mathbf{v}$, as the vector to be analyzed, form the perpendicular, b, from its head, R, to $\mathbf{w}$ at Q. The segment, a cut off along $\mathbf{w}$, is the component of $\mathbf{v}$ in the direction of $\mathbf{w}$, or the projection of $\mathbf{v}$ onto $\mathbf{w}$, while $b = QR$ is the orthogonal component. Component a reveals that there is similarity or compatibility between $\mathbf{v}$ and $\mathbf{w}$, while b indicates that there is difference or incompatibility between them. By incorporating a and b into $\mathbf{v}$, we arrive at precise measures of the degrees of similarity and difference. This is done, as above, by forming the altitude, d which cuts $\mathbf{v}$ into two segments, e

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

and f; equations (1)-(3) follow, with equation (1) giving the degree of similarity or compatibility, and equation (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 only considered situations involving unit vectors which form an acute angle—that is, situations in which the confining angle, ξ, equals $90°$. 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, then, we must next account for what happens when ξ has any given value, and when vectors $\mathbf{v}$ and $\mathbf{w}$ are unequal.

3.2. Cases involving non-orthogonal reference axes

Let us, then, 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 equation (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 equation (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

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

where A indicates the amplitude of the wave, C and D govern horizontal and vertical displacements, and $B = (m+p)\pi/\lambda$, where λ is the wavelength in radians, and m and p govern any horizontal stretching or compressing of λ. In equation (1) for example, $y = e$, $A = B = m = 1$, $\lambda = \pi$, $n = 2$, $x = \theta$ in radians, and $C = D = p = 0$.

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 the even powers of the cosine 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). Any additional horizontal stretching or compressing of λ is governed by p.

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 period for a standard wave, in contrast, runs, say, from 0 to +1 and back again to 0 when $m = 1$, and, when $m = 2$, the period 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 P 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 equation (1) reads:

$$e = 1 \cos^2\left(\frac{(1+0)\pi}{(2^1)\xi}\theta + 0\right) + 0 \ . \tag{5}$$

This simplifies to:

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

If $\xi = 90°$ or, in radians, $\frac{\pi}{2}$, we get equation (1). If $\xi = 180°$ (or π radians) we get:

$$e = \cos^2\left(\frac{1}{2}\theta\right) \ . \tag{7}$$

For other values of ξ, we could now readily arrive at analogous equations by inserting the correct value for ξ (in radians) into equation (6).

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

We can now consider cases when one or both vectors are not unit vectors.

3.3. Directional similarity for unequal vectors

When vector $\mathbf{v}$ is not a unit vector, the magnitude of a in figure 2 is found by using the following formula:

$$a = ||\mathbf{v}|| \cos\theta \tag{8}$$

where, again, $||\mathbf{v}||$ equals $\mathbf{v}$ divided by a unit vector. Referring to equation (4), we see that $||\mathbf{v}||$ in equation (8) fills the slot for the amplitude, A. When working with probabilities, we have to limit A such that $0 \leq A \leq 1$, thus it becomes necessary to divide equation (8) by a quantity equal or greater than $||\mathbf{v}||$. Additionally, it is sometimes useful to divide equations (1) and (2) so as to reflect the behaviors of subsets within a sample. In such cases, the numerator of A will also contain a variable, s, which counts the number of such subsets governed by the equation, while the denominator features the total number, t of such subsets within the sample. These considerations expand equation (6) to:

$$e = (\frac{||\mathbf{v}||s}{||\mathbf{v}||t}) \cos^2(\frac{\pi}{2\xi}\theta) \ . \tag{9}$$

While we could simplify the amplitude in equation (9) to s/t, its expanded form helps when we turn our attention to cases of similarity between two unequal vectors. In such a case, the larger vector must serve as the unit vector, and the smaller must take on a fractional value in order to keep A in the range between 0 and 1. Thus, if $||\mathbf{w}|| > ||\mathbf{v}||$, equation (9) becomes:

$$e = (\frac{||\mathbf{v}||s}{||\mathbf{w}||t}) \cos^2(\frac{\pi}{2\xi}\theta) \ . \tag{10}$$

Equation (2) can be expanded in the same way simply by substituting b for a, f for e, and sine for cosine in the foregoing discussion.

Of course, in a given context, e and f will have to be written as probabilities, with e indicating the probability for similarity, compatibility, matching outcomes, or sameness of some sort, while f indicates the probability of the opposites of these. Individual contexts will dictate an appropriate notation, for example, $p(\text{same})$ and $p(\text{different})$ or $p(A = B)$ and $p(A \neq B)$ and so forth.

Additional considerations could be discussed at this point (for example, how the equation changes if C or D has an non-zero value, or if A, C or D has an exponent), but these are either trivial matters or lie beyond the scope of this paper.

4. Conclusion

Starting with problems in QM and optics which yielded equations for probabilities which were *(1)* unlike any known equations for probabilities and *(2)* 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 ratio of the size of the angle formed by the axes under consideration to 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 equation (1), or as complex as equation (10).

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 QM since the 1920s. Specifically, the law of Malus is equation (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 QM, equation (7) is featured in a modification—known as the CH Bell test [2]—of an experiment proposed by Bell [1], which served to prove *(1)* that this system of probabilities was independent of other known systems and *(2)* that quantum phenomena followed this system. Equation (1) is featured in a second such modification, known as the $CHSH$ Bell test (named for the variant of the Bell test developed by Clauser, Horne, Shimony and Holt [3]). In both cases, the principal question revolves around the rate that independent detectors (V and W) produce matching or mismatching outcomes when testing similar photons.

Since the detection equipment may either block (b) a photon or allow it to pass (p), the outcomes of both the CH and $CHSH$ tests involve two possible kinds of correlation (V_p, W_p) and (V_b, W_b), and two kinds of non-correlation (V_b, W_p), and (V_p, W_b). Thus the probability of, say, a match at V and W, $p(V = W)$, has two constituent parts, $p(V_p, W_p)$ and $p(V_b, W_b)$. These constituent probabilities will have a form consistent with equation (10). When the probability of getting either kind of match is even, then $p(V_p, W_p) = p(V_b, W_b)$. In this case, since each probability covers a portion of the possible matches, and the probabilities are equal, each of them will feature an s/t term equal to 1/2—that is, for example, for the CH test, $p(V_p, W_p) = (1/2)\cos^2(\theta/2)$. For the cases involving non-correlation, these same relations follow, except that sine is substituted for cosine as indicated in the text above. When large, random samples of unpolarized photons are tested, we should expect the average value for the applicable version of equation (10). In the case of equation (1), the resulting value would be 0.5, a random outcome. Experimentation confirms all the values which would arise from the foregoing discussion.

Consequently, while the 'mysterious' QM probabilities have led to various non-scientific conjectures, we can now see *(1)* that the mystery arose from the fact that probabilities based on directional similarity were unknown, and *(2)* 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 detection equipment). Since this relationship is established as soon as the detectors 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 detectors.

This discovery, then, 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. (1964). On the Einstein-Podolsky-Rosen paradox. *Physics* **1,** 195-200.

[2] Clauser, J., and Horne, M. (1974). Experimental consequences of objective local theories. *Phys.*

Rev. **10,** 526-535.

[3] Clauser, J., Horne, M., Shimony, A., and Holt, R. (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.