PRE 27827

NEGATIVE PROBABILITY

Richard P. Feynman
Professor of Theoretical Physics
California Institute of Technology

Some twenty years ago one problem we theoretical physicists had was that if we combined the principles of quantum mechanics and those of relativity plus certain tacit assumptions, we seemed only able to produce theories (the quantum field theories) which gave infinity for the answer to certain questions. These infinities are kept in abeyance (and now possibly eliminated altogether) by the awkward process of renormalization. In an attempt to underrstand all this better, and perhaps to make a theory which would give only finite answers from the start, I looked into the "tacit assumptions" to see if they could be altered.

One of the assumptions was that the probability for an event must always be a positive number. Trying to think of negative probabilities gave me cultural shock at first, but when I finally got easy with the concept I wrote myself a note so I wouldn't forget my thoughts. I think that Prof. Bohm has just the combination of imagination and boldness to find them interesting and amusing. I am delighted to have this opportunity to publish them in such an appropriate place. I have taken the opportunity to add some further, more recent, thoughts about applications to two state systems.

Unfortunately I never did find out how to use the freeedom of allowing probabilities to be negative to solve the original problem of infinities in quantum field theory!

It is usual to suppose that, since the probabilities of events must be positive, a theory which gives negative numbers for such quantities must be absurd. I should show here how negative probabilities might be interpreted. A negative number, say of apples, seems like an absurdity. A man starting a day with five apples who gives away ten and is given eight during the day has three left. I can calculate this in two steps: $5 - 10 = -5$ and $-5 + 8 = 3$. The final answer is satisfactorily positive and correct although in the intermediate steps of calculation negative numbers appear. In the real situation there must be special limitations of the time in which the various apples are received and given since he never really has a negative number, yet the use of negative numbers as an abstract calculation permits us freedom to do our mathematical calculations in any order simplifying the analysis enormously, and permitting us to disregard inessential details. The idea of negative numbers is an exceedingly fruitful mathematical invention. Today a person who balks at making a calculation in this way is considered backward or ignorant, or to have some kind of a mental block. It is the purpose of this paper to point out that we have a similar strong block against negative probabilities. By discussing a number of examples, I hope to show that they are entirely rational of course, and that their use simplifies calculation and thought in a number of applications in physics.

First let us consider a simple probability problem, and how we usually calculate things and then see what would happen if we allowed some of our normal probabilities in the calculations to be negative. Let us imagine a roulette wheel with, for simplicity, just three numbers: 1, 2, 3. Suppose however, the operator by control of a switch under the table can put the wheel into one of two conditions A, B in each of which the probability of 1, 2, 3 are different. If the wheel is in condition A, the probabilities of 1, $p_{1A} = 0.3$ say, of 2 is $p_{2A} = 0.6$, of 3 is $p_{3A} = 0.1$. But if the wheel is in condition B, these probabilities are $p_{1B} = 0.1$, $p_{2B} = 0.4$, $p_{3B} = 0.5$ say as in the table. We, of course, use the table in this way:

	Cond. A	Cond. B
1	0.3	0.1
2	0.6	0.4
3	0.1	0.5

Suppose the operator puts the wheel into condition A 7/10 of the time and into B the other 3/10 of the time at random. (That is the probability of condition A, $P_A = 0.7$, and of B, $P_B = 0.3$.)

Then the probability of getting 1 is Prob. 1 = 0.7 (0.3) + 0.3 (0.1) = 0.24, etc. In general, of course, if α are conditions and $p_{i\alpha}$ is a conditional probability, the probability of getting the result i if the condition α holds, we have ($p_{i\alpha}$ = Prob (if α then i))

$$P_i = \sum_\alpha p_{i\alpha} \cdot P_\alpha \tag{1}$$

where P_α are the probabilities that the conditions α obtain, and P_i is the consequent probability of the result i. Since some result must occur in any condition, we have

$$\sum_i p_{i\alpha} = 1 \tag{2}$$

where the sum is that over all possible independent results i. If the system is surely in some one of the conditions, so if

$$\sum_\alpha P_\alpha = 1$$

then

$$\sum_i P_i = 1 \quad , \tag{3}$$

meaning we surely have some result, in virtue of (2).

Now, however, suppose that some of the conditional probabilities are negative, suppose the table reads so that, as we shall say, if the system is in condition B the probability of getting 1 is -0.4. This sounds absurd but we must say it this way if we wish that our way of thought and language be precisely the same whether the actual quantities $p_{i\alpha}$ in our calculations are positive or negative. That is the essence of the mathematical use of negative numbers -- to permit an efficiency in reasoning so that various cases can be considered together by the same line of reasoning, being assured that intermediary steps which are not readily interpreted (like -5 apples) will not lead to absurd results. Let us see what $p_{1B} = -0.4$ "means" by seeing how we calculate with it. We have arranged the numbers in the table so that $p_{1B} + p_{2B} + p_{3B} = 1$ in accordance with eq. (2). For example, the condition A has probability 0.7 and B has probability 0.3,

	Cond. A	Cond. B
1	0.3	-0.4
2	0.6	1.2
3	0.1	0.2

we have for the probability of result 1,

$$P_1 = 0.7\ (0.3) + 0.3\ (-0.4) = 0.09$$

which would be all right. We have also allowed p_{2B} to exceed unity. A probability greater than unity presents no problem different from that of negative probabilities, for it represents a negative probability that the event will not occur.

Thus the probability of result 2 is, in the same way,

$$P_2 = 0.7\ (0.6) + 0.3\ (1.2) = 0.78 \quad .$$

Finally, the probability of result 3 presents no problem for

$$P_3 = 0.7\ (0.1) + 0.3\ (0.2) = 0.13 \quad .$$

The sum of these is 1.00 as required, and they are all positive and can have their usual interpretation.

The obvious question is what happens if the probability of being in condition B is larger, for example, if condition B has probability 0.6, the probability of result 1 is negative 0.4 (0.3) + 0.6 (-0.4) = -0.12. But suppose nature is so constructed that you can never be sure the system is in condition B. Suppose there must always be a limit of a kind to the knowledge of the situation that you can attain. And such is the limitation that you can never know for sure that condition B occurs. You can only know that it may occur with a limited probability (in this case less than 3/7 say). Then no contradiction will occur in the sense that a result 1 or 2 or 3 will have a negative probability of occurrence.

Another possibility of interpretation is that results 1, 2, 3 are not directly observable but one can only verify by a final observation that the result had been 1, 2 or 3 with certain probabilities. For example, suppose the truly physically verifiable observations can only distinguish two classes of final events. Either the result was 3 or else it was in the class of being either 1 or 2. This class has the probability $P_1 + P_2$ which is always positive for any positive P_A, P_B. This case corresponds to the situation that 1, 2 3 are not the finally observed results, but only intermediaries in a calculation.

Notice that the probabilities of conditions A and B might themselves be negative (for example, $P_A = 1.3$, $P_B = -0.3$) while the probabilities of the results 1, 2, 3 still remain positive.

It is not my intention here to contend that the final probability of a verifiable physical event can be negative. On the other hand, conditional probabilities and probabilities of imagined intermediary states may be negative in a calculation of probabilities of physical events or states.

If a physical theory for calculating probabilities yields a negative probability for a given situation under certain assumed conditions, we need not conclude the theory is incorrect. Two other possibilities of interpretation exist. One is that the conditions (for example, initial conditions) may not be capable of being realized in the physical world. The other possibility is that the situation for which the probability appears to be negative is not one that can be verified directly. A combination of these two, limitation of verifiability and freedom in initial conditions, may also be a solution to the apparent difficulty.

The rest of this paper illustrates these points with a number of examples drawn from physics which are less artificial than our roulette wheel.

Since the result must ultimately have a positive probability, the question may be asked, why not rearrange the calculation so that the probabilities are positive in all the intermediate states? The same question might be asked of an accountant who subtracts the total disbursements before adding the total receipts. He stands a chance of going through an intermediary negative sum. Why not rearrange the calculation? Why bother? There is nothing mathematically wrong with this method of calculating and it frees the mind to think clearly and simply in a situation otherwise quite complicated. An analysis in terms of various states or conditions may simplify a calculation at the expense of requiring negative probabilities for these states. It is not really much expense.

Our first physical example is one in which one usually uses negative probabilities without noticing it. It is not a very profound example and is practically the same in content as our previous example. A particle diffusing in one dimension in a rod has a probability of being at x at time t of $P(x,t)$ satisfying $\partial P(x,t)/\partial t = -\partial^2 P(x,t)/\partial x^2$. Suppose at $x = 0$ and $x = \pi$ the rod has absorbers at both ends so that $P(x,t) = 0$ there. Let the probability of being at x at $t = 0$ be given as $P(x,0) = f(x)$. What is $P(x,t)$ thereafter? It is

$$P(x,t) = \sum_{n=1}^{\infty} P_n \sin nx \exp(-n^2 t) \qquad (4)$$

where P_n is given by

$$f(x) = \sum_{n=1}^{\infty} P_n \sin nx \qquad (5)$$

or

$$P_n = \frac{2}{\pi} \int f(x) \sin nx \, dx \quad . \qquad (6)$$

The easiest way of analyzing this (and the way used if P(x,t) is a temperature, for example) is to say that there are certain distributions that behave in an especially simple way. If f(x) starts as sin nx it will remain that shape simple decreasing with time, as e^{-n^2t}. Any distribution f(x) can be thought of as a superposition of such sine waves. But f(x) cannot be sin nx if f(x) is a probability and probabilities must always be positive. Yet the analysis is so simple this way that no one has really objected for long.

To make the relation to our previous analysis more clear, the various conditions α are the conditions n (that is, the index α is replaced by n). The a priori probabilities are the numbers P_n. The conditions i are the positions x (the index i is replaced by x) and the conditional probabilities* (if n then x at time t) are

$$p_{i\alpha} \rightarrow p_{x,n} = e^{-n^2t} \sin nx \quad .$$

Equation (4) is then precisely Eq. (1), for the probabilities p_i of having result n is now what we call P(x,t). Thus Eq. (4) is easily interpreted as saying that if the system is in condition n, the chance of finding it at x is exp $(-n^2t)$ sin nx, and the chance of finding it in condition n is P_n.

No objection should be made to the negative values of these probabilities. However, a natural question is what are the restrictions which insure that the final probability for the event (finding a particle at x at time t) are always positive. In this case they are simple. It is that the a priori probabilities, although possibly negative, are restricted by certain conditions. The condition

* These do not satisfy Eq. (2) for we have particles "lost" off the end of the rod, and the state of being off the rod is not included among the possibilities i.

is that they must be such that they could come from the Fourier analysis of an everywhere positive function. This condition is independent of what value of x one wishes to observe at time t.

In this example, the restrictions to insure positive probabilities can be stated once and for all in a form that does not depend on which state we measure. They are all positive simultaneously.

Another possibility presents itself. It can best be understood by returning to our roulette example. It may be that the restrictions on the conditions A, B which yield a positive probability may depend on what question you ask. In an extreme example, there may be no choice for the P_α that simultaneously make all P_i positive at once. Thus, although certain restrictions may make probability of result 1 positive, result 3 under these circumstances would have a negative probability. Likewise, conditions ensuring that P_3 is positive might leave P_1 or P_2 negative. In such a physical world, you would have such statements as "If you measure 1 you cannot be sure to more than a certain degree that the condition is A, on the other hand it will be alright to think that it is certainly in condition A provided you are only going to ask for the chance that the result is 3." For such a circumstance to be a viable theory, there would have to be certain limitations on verification experiments. Any method to determine that the result was 3 would automatically exclude that at the same time you could determine whether the result was 1. This is reminiscent of the situation in quantum mechanics in relation to the uncertainty principle. A particle can have definite momentum, or a definite position in the sense that an experiment may be devised to measure either one. But no experiment can be devised to decide what the momentum is, to error of order Δp, which at the same time can determine that the position x is within Δx unless $\Delta x > \hbar/\Delta p$.

It is possible, therefore, that a closer study of the relation of classical and quantum theory might involve us in negative probabilities, and so it does. In classical theory, we may have a distribution function $F(x,p)$ which gives the probability that a particle has a position x and a momentum p in dx and dp (we take a simple particle moving in one dimension for simplicity to illustrate the ideas). As Wigner has shown, the nearest thing to this in quantum mechanics is a function (the density matrix in a certain representation) which for a particle in a state with wave function $\psi(x)$ is

$$F(x,p) = \int \psi^*(x-y/2)\, \exp -i(py) \psi(x+y/2) dy \quad , \qquad (7)$$

(if the state is statistically uncertain we simply average F for the various possible wave functions with their probabilities).

In common with the classical expression, we have these properties:

(a) $F(x,p)$ is real.

(b) Its integral with respect to p gives the probability that the particle is at x:

$$\int F(x,p) dp/(2\pi) = \psi^*(x)\psi(x) \quad . \qquad (8)$$

(c) Its integral with respect to x gives the probability that the momentum is p

$$\int F(x,p) dx = \phi^*(p)\phi(p) \quad , \qquad (9)$$

where $\phi(p)$ is the usual Fourier transform of $\psi(x)$. $\phi(p) = \int e^{-ipx}\psi(x) dx$.

(d) The average value of a physical quantity M is given by

$$\langle M \rangle = \int w_M(x,p) F(x,p) dx\, dp \quad , \qquad (10)$$

where w_M is a weight function depending upon the character of the physical quantity.

The only property it does not share is that in the classical theory $F(x,p)$ is positive everywhere, for in quantum theory it may have negative values for some regions of x,p. That we still have a viable physical theory is ensured by the uncertainty principle that no measurement can be made of momentum and position simultaneously beyond a certain accuracy.

The restriction this time which ensures positive probabilities is that the weight functions $w_M(x,p)$ are restricted to a certain class -- namely, those that belong to hermitian operators. Mathematically, a positive probability will result if w is of the form

$$w(x,p) = \int X(x-Y/2) e^{+ipy}\, X^*(x+Y/2) dY \quad , \qquad (11)$$

where X is any function and X^* is its complex conjugate. Generally, if $w(x,p)$ is the weight for the question, "what is the probability that the physical quantity M has numerical value m?" w must be of the form (11) or the sum of such forms with positive weights. With this limitation, final probabilities are positive.

To make the analogy closer to those previously used, we can take two systems a, b, in interaction, such that measurements on b can provide predictions of probabilities for a. Thus, using the one-dimensional case again, we have a two-point correlation function $F(p_a,x_a;\ p_b,x_b)$ defined via an obvious generalization of Eq. (7) to two variables. This corresponds to the conditional probability $p_{i\alpha}$. Then if a quantity M is measured in b, the a priori probabilities for various p_b,x_b are given by an appropriate $w_M(x_b,p_b)$ (the analogue of P_α in Eq. (1)). The probability that system "a" has position and momentum x_a,p_a is (the analogue of P_i), then

$$P(x_a,p_a) = \int F(x_a,p_a;\ x_b,p_b)w_M(x_b,p_b)dx_b\ dp_b \quad ,$$

the analogue of Eq. (1). As an example, we may take the strong correlation possible arising from the two-particle wave function $\delta(x_a-x_b)$ which is

$$F(x_a,p_a;\ x_b,p_b) = \delta(p_a+p_b)\delta(x_a-x_b) \quad ,$$

which means that the particles a, b, have the same position and opposite momenta so that a measurement of b's position would permit a determination of a's and a measurement of b's momentum would determine a's (to be the opposite). This particular F is entirely positive and classical in its behavior so that letting $w_M(x_b,p_b)$ be $\delta(x_a-b)\delta(p_a-Q)$ would not lead to negative probabilities directly, for (1) gives $P(x_a,p_a) = \delta(x_a-b)\delta(P_a+Q)$ in this case, but further use of such a P in subsequent interactions has the danger of producing negative probabilities. We have become quite used to the rules of thought and limitations of an experiment, which ensures that they never arise in quantum mechanics.

It is not our intention to claim that quantum mechanics is best understood by going back to classical mechanical concepts and allowing negative probabilities (for the equations for the development of F in time are more complicated and inconvenient than those of ψ)*. Rather we should like to emphasize the idea

---

* The classical equations for F for a particle moving in a potential are

$$\partial F(x,p,t)/\partial t = -p/m \cdot \partial F/\partial x + V'(x)\partial F/\partial p$$

while the quantum equations are

$$\partial F(x,t)/\partial t = -p/m \cdot \partial F/\partial x + \int G(x,Q)F(x,p+Q)dQ$$

so instead of the momentum changing infinitesimally during an infinitesimal time, Δt, it may jump by an amount Q with probability when it is at x,

$$\Delta t\ G(x,Q) = \Delta t \cdot 2 \text{ Imay part} \int e^{iQ\cdot Y}\, V(x+Y/2)dY$$

which is a real, but possibly negative probability.

that negative probabilities in a physical theory does not exclude that theory, providing special conditions are put on what is known or verified. But how are we to find and state these special conditions if we have a new theory of this kind? It is that a situation for which a negative probability is calculated is impossible, not in the sense that the chance for it happening is zero, but rather in the sense that the assumed conditions of preparation or verification are experimentally unattainable.

We may give one more example. In the quantum theory of electrodynamics, the free photon moving in the z direction is supposed to have only two directions of polarization transverse to its motion x,y. When this field is quantized, an additional interaction, the instantaneous Coulomb interaction, must be added to the virtual transverse photon exchange to produce the usual simple

$$(j_x j_x + j_y j_y + j_z j_z' - j_t j_t')e^2/q^2 \quad , \tag{12}$$

virtual interaction between two currents, j and j'. It is obviously relativistically invariant with the usual symmetry of the space j_x, j_y, j_z and time j_t components of the current (in units where the velocity of light, is c=1). The original starting Hamiltonian with only transverse components does not look invariant. Innumerable papers have discussed this point from various points of view but perhaps the simplest is this. Let the photon have four directions of polarization of a vector x,y,z,t no matter which way it is going. Couple the time component with ie instead of e so that the virtual contribution for it will be negative as required by relativity in Eq. (12). For real photons, then, the probability of a t-photon emission is negative, proportional to $-|\langle f|j_t|i\rangle|^2$ the square of the matrix element of j_t between initial and final states, just as the probability to emit an x photon is $+|\langle f|j_x|i\rangle|^2$. The total probability of emitting any sort of photon is the algebraic sum of the probabilities for the four possibilities,

$$|\langle f|j_x|i\rangle|^2 + |\langle f|j_y|i\rangle|^2 + |\langle f|j_z|i\rangle|^2 - |\langle f|j_t|i\rangle|^2 \quad . \tag{13}$$

It is always positive, for by the conservation of current there is a relation of j_t and the space components of $j, k_\mu j_\mu = 0$ if k_μ is the four-vector of the photon. For example, if k is in the z direction, $k_z = \omega$, and $k_x = k_y = 0$ so $j_t = j_z$ and we see Eq. (13) is equal to the usual result where we add only the transverse emissions. The probability to emit a photon of definite polarization e_μ is (assume e_μ is not a null vector)

$$-|\langle f|j_\mu\, e_\mu|i\rangle|^2/(e_\mu\, e_\mu) \quad .$$

This has the danger of producing negative probabilities. The rule to avoid them is that only photons whose polarization vector satisfies $k_\mu\, e_\mu = 0$ and $e_\mu\, e_\mu = -1$ can be observed asymptotically in the final or initial states. But this restriction is not to be applied to virtual photons, intermediary negative probabilities are not to be avoided. Only in this way is the Coulomb interaction truly understandable as the interchange of virtual photons, photons with time-like polarization which are radiated as real photons with a negative probability.

This example illustrates a small point. If one t photon is emitted with a negative probability $-\alpha(\alpha > 0)$, and another t photon is emitted say independently with probability $-\beta(\beta > 0)$, the chance of emitting both is positive $(-\alpha)(-\beta) = \alpha\beta > 0$. Should we not expect then to see physical emission of two such photons? Yes, but (if these photons are moving in the z direction) there is a probability to emit z photons α and β also, and there are four emission states: two t photons with probability $+\alpha\beta$; two z photons with probability $+\alpha\beta$; the first z and second t probability $(+\alpha)(-\beta) = -\alpha\beta$ and the first t second z with probabilities $-\alpha\beta$ so again, for total emission rate only the transverse photons contribute.

Although it is true that a negative probability for some situations in a theory means that that situation is unattainable or unverifiable, the contrary is not true, namely a positive probability for a situation does not mean that that situation is directly verifiable. We have no technique for detecting t photons which is not similarly sensitive to z photons so that we can only always respond to a combination of them. Likewise, no direct test can be made that the two t photons are indeed present without including the additional probabilities of having z photons. The fact for example, that $F(x,p)$ is everywhere positive $\left(\exp\left(-\frac{p^2/m+m\omega^2x}{2\hbar\omega}\right)\right)$ for the ground state of an oscillator does not mean that for that state we can indeed measure both x and p simultaneously.

As another example we will give an analogue of the Wigner function for a spin $\frac{1}{2}$ system, or other two state system. Just as the Wigner function is a function of x and p, twice as many variables as in the wave function, here we will give a "probability" for two conditions at once. We choose spin along the z-axis and spin along the x-axis. Thus let f_{++} represent the "probability" that our system has spin up along the z-axis and up along the x-axis simultaneously. We shall define the quantity f_{++} for a pure state to be the expectation of $\frac{1}{4}(1+\sigma_z+\sigma_x+\sigma_y)$, where σ_x,σ_y, and σ_z are the Pauli matrices. For a mixed state we take an average over the pure state values. Likewise f_{+-} is the expectation of $\frac{1}{4}(1+\sigma_z-\sigma_x-\sigma_y)$, f_{-+} is the expectation of $\frac{1}{4}(1-\sigma_z+\sigma_x-\sigma_y)$ and f_{--} is the expectation of $\frac{1}{4}(1-\sigma_z-\sigma_x+\sigma_y)$.

Understanding that this "probability" can be negative, we shall train ourselves to deal with it otherwise as a real probability and thus dispense with the warning quotes hereafter. Analogously f_{+-} is the probability that the spin is up along the z-axis and down along the x-axis (that is pointing in the negative x direction). Likewise f_{-+} and f_{--} give the probability that the spin is along the negative z-axis and along the x-axis in the positive or negative sense, respectively. These are all the possible conditions so we have $f_{++}+f_{+-}+f_{-+}+f_{--}=1$. As an example, we might have $f_{++}=0.6$, $f_{+-}=-0.1$, $f_{-+}=0.3$ and $f_{--}=0.2$.

Now the probability that the spin is up along z is simply the sum of the probability that it is up along z and up along x, and the other possibility, that it is up along z but down along x; that is simply $f_{++}+f_{+-}$ or $0.6+(-0.1)=0.5$ in our example. The probability the spin is down along z is $f_{-+}+f_{--}$,also 0.5. In the same way the probability that the spin is along the positive x-axis, independent of its value along z is $f_{++}+f_{-+}$ or 0.9. We, of course, cannot measure simultaneously the spin in the z and in the x direection, so we cannot directly determine f_{+-} and there is no difficulty with its negative value.

These four numbers give a complete expression of the state of the system, and the probability for any other question you can ask experimentally is some linear combination of them. For example, the probability that a measurement of spin along the y-axis gives 'up' is $f_{++}+f_{--}$ or 0.8, and that it gives 'down' is $f_{+-}+f_{-+}$ or 0.2. In fact, for a two state system any question is equivavent to the question vis the spin up along an axis in some direction. If that direction is defined by the unit vector V with components V_x,V_y,V_z then we can say the probability that the spin is up along this direction if the condition of the electron is ++ is $p_{++}(V)=\frac{1}{2}(1+V_z+V_x+V_y)$. For the other conditions we have $p_{+-}(V)=\frac{1}{2}(1+V_z-V_x-V_y)$, $p_{-+}(V)=\frac{1}{2}(1-V_z+V_x-V_y)$, and $p_{--}(V)=\frac{1}{2}(1-V_z-V_x+V_y)$. In the general case then where the f's give the a priori probabilities of each condition the probability of finding the spin up along V is the sum on a of $p_a(V)f_a$ or $\frac{1}{2}((1+V_z+V_x+V_y)f_{++}+(1+V_z-V_x-V_y)f_{+-}+(1-V_z+V_x-V_y)f_{-+}+(1-V_z-V_x+V_y)f_{--})$. In order that this always gives positive results, in addition to the condition that the sum of the f's is unity, there is the restriction that the sum of the squares of the four f's be less than $\frac{1}{2}$. It equals $\frac{1}{2}$ for a pure state.

If there are two electrons in a problem we can use classical logic, considering each of them as being in one of the four states, $++,+-,-+,--$. Thus suppose we have two electrons correllated so their total spin is zero moving into two detectors, one set to determine if the spin of the first electron is in the direction V, and the other set to measure whether the second electron has its spin in the direction U. The probability that both detectors respond is $\frac{1}{4}(1 - U.V)$. Thus if one is found up along any axis, the other is surely down along the same axis. This situation usually causes difficulty to a hidden variable view of nature. Suppose the electron can be in one of a number of conditions a for each of which the chance of being found to be spinning up along the V-axis is $p_a(V)$. If the second electron is in condition b it's probability of being found along U is $p_b(U)$. Suppose now that the chance of finding the two electrons in conditions a, b, respectively, is P_{ab}. This depends on how the electrons were prepared by the source. Then the chance of finding them along the V and U axes is $\sum_{a,b} P_{ab}p_a(V)p_b(U)$ which is to equal $\frac{1}{4}(1 - U.V)$. This is well known to be impossible if all the "probabilities" P_{ab} and p are positive. But everything works fine if we permit negative probabilities and use for a our four states with the $p_a(V)$ as defined previously. The probabilities for the correllated states in the case that the total spin is zero are P_{ab} equals $\frac{1}{8}$ if a and b are different states, and $-\frac{1}{8}$ if they are the same.

For another example of a two state sytem consider an electron going thru a screen with two small holes to arrive at a second screen(see figure 1). We can say there are four ways or conditions by which the electron can go thru the holes, corresponding to the $++, +-, -+$, and $--$ conditions. If we take up spin to correspond to going thru hole number 1 and down spin to represent going thru hole 2, then the other variable corresponding to spin in the x direction means going thru the two holes equally in phase. Ordinarily we cannot say which hole it goes thru and what the phase relation is (just as ordinarily we do not say which way the z-spin is and which way the x-spin is) but now we can and do. For example, f_{--} gives the probability of going thru hole 2 but 180 degrees out of phase (whatever that could mean). For each of these conditions we can calculate what the chance is that the electron arrives at a point x along the screen. For example, $P_{++}(x)$, the probability for arrival at x for the condition $++$ (thru 1 in phase) and $P_{+-}(x)$, the probability for $+-$ (thru 1 but out of phase) are sketched roughly in figure 1 as the curves b and c respectively. The independent probabilities are negative for some values of x. The functions thru hole 2 are these reflected in x; $P_{-+}(x) = P_{++}(-x)$ and $P_{--}(x) = P_{+-}(-x)$. The total chance to go thru hole 1 , $P_{++} + P_{+-}$, the sum of the two irregular curves shown in the figure is just the smooth bump, the solid line at a, with its maximum under hole 1, not showing inteference effects. But the total probability to arrive with holes out of phase, $P_{+-} + P_{--}$ shows the typical inteference pattern at the bottom of the figure at d.

Obviously the particular choice we used for the two state system is arbitrary, and other choices may have some advantages. One way that generalizes to any number of holes or of states, finite or otherwise, is this. Suppose an event can happen in more than one way, say ways A, B, C, etc. with amplitudes a, b, c, respectively, so that the probability of occurring is the absolute square of $a + b + c + ...$ This can be described by saying the event can happen in two ways at once. For example we can say that the event happens by "coming" in way A and "going" in way B (or, if you prefer, by "looping" via A and B) with a "probability" $P(A, B) = \frac{1}{2}(1 + i)a^*b + \frac{1}{2}(1 - i)b^*a$ where a^* stands for the complex conjugate of a. The probability of "coming" and "going" by the same way

A is $P(A,A) = a^*a$ and is the conventional positive probability that the event would occur if way A only were available to it. The total probability is the sum of these P for every pair of ways. If the two ways in P, "coming", and "going" are not the same, P is as likely to be negative as positive.

The density matrix, ρ_{ij}, if the states are i is then represented instead by saying a system has a probability to be found in each of a set of conditions. These conditions are defined by an ordered pair of states "coming" in i and "going" in j with "probability" $p(i,j)$ equal to the real part of $(1+i)\rho_{ij}$. The condition that all physical probabilities reamain positive is that the square of $p(i,j)$ not exceed the product $p(i,i)p(j,j)$ (equality is reached for pure states).

Finally, suppose that, because of the passage of time, or other interaction, or simply a change in basis, the state i has an amplitude S_{mi} of appearing as state m, where S is a unitary matrix (so the new density matrix ρ' is given by $S^{-1}\rho S$). We then discover vwe can find the new probabilities $p'(m,n)$ by summing all alternatives i,j of $p(i,j)$ times a factor that can be interpreted as the probability that the state "coming" in i, "going" in j turns into the state "coming" in m, "going" in n. This "probability" is $\frac{1}{2}(S^*_{im}S_{jn} + S^*_{jn}S_{im}) + \frac{i}{2}(S^*_{jm}S_{in} - S^*_{in}S_{jm})$.

With such formulas all the results of quantum statistics can be described in classical probability language, with states replaced by 'conditions' defined by a pair of states (or other variables), provided we accept negative values for these probabilities. This is interesting, but whether it is useful is problematical, for the equations with amplitudes are simpler and one can get used to thinking with them just as well.

My interest in this subject arose from many attempts to quantize electrodynamics or other field theories with cutoffs or using advanced potentials, in which work apparently negative probabilities often arose. It may have applications to help in the study of the consequences of a theory of this kind by Lee and Wick.

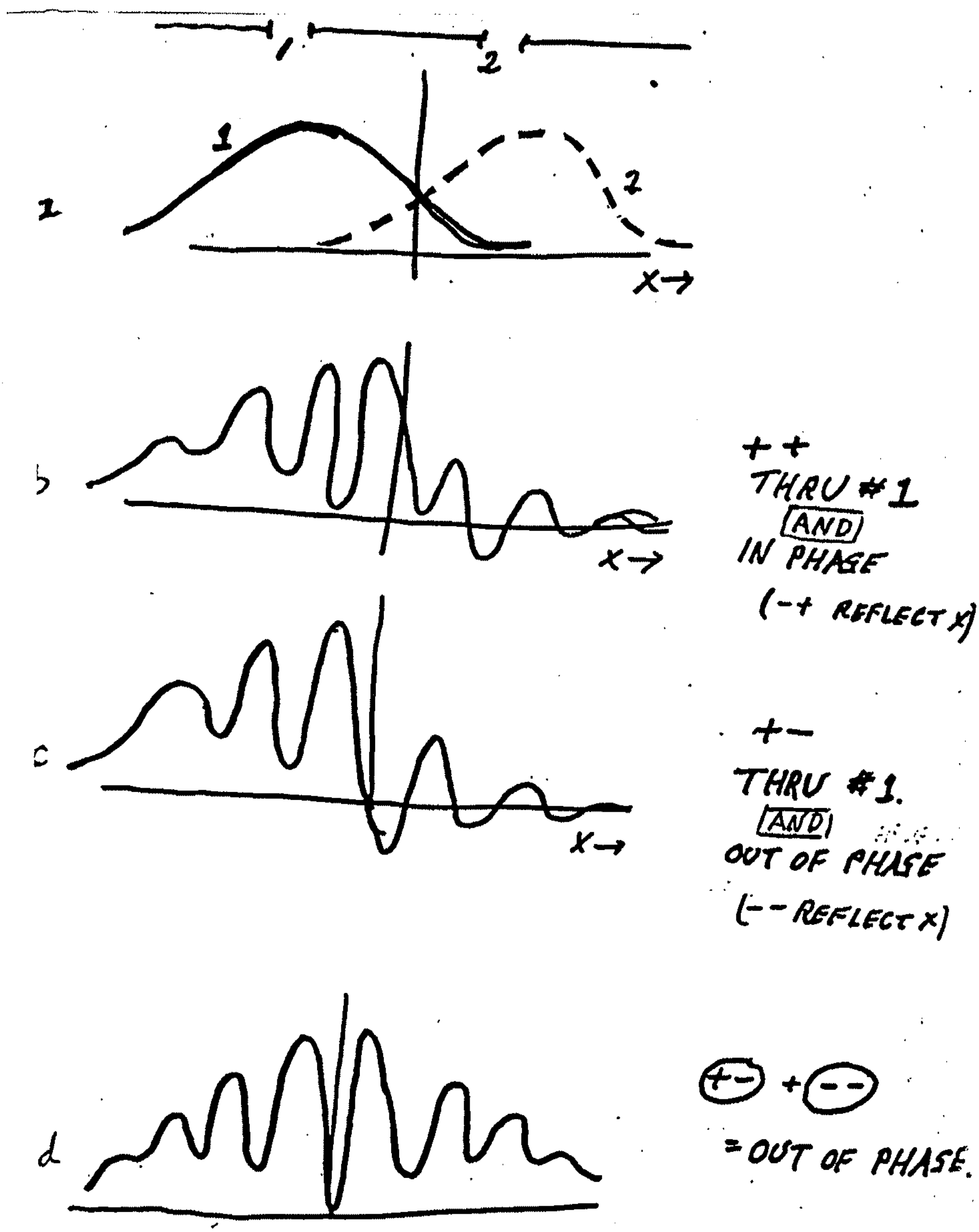
1
2
x→
b
++
THRU #1
AND
IN PHASE
(-+ REFLECT x)
c
+-
THRU #1.
AND
OUT OF PHASE
(-- REFLECT x)
d
+- + --
= OUT OF PHASE.

Figure 1.