

“Special” States in Quantum Measurement Apparatus: Structural Requirements for the Recovery of Standard Probabilities

L. S. Schulman¹

Received February 11, 1991; revised March 11, 1991

In a recently proposed quantum measurement theory the definiteness of quantum measurements is achieved by means of “special” states. The recovery of the usual quantum probabilities is related to the relative abundance of particular classes of “special” states. In the present article we consider two-state discrimination, and model the apparatus modes that could provide the “special” states. We find that there are structural features which, if generally present in apparatus, will provide universal recovery of standard probabilities. These structural features relate to the distribution of certain Hamiltonian matrix elements or interaction times. In particular, those quantities must be asymptotically ($x \rightarrow \infty$) distributed according to the Cauchy law, $C_a(x) = a/\pi(x^2 + a^2)$.

1. INTRODUCTION

In a recent series of publications,⁽¹⁻⁴⁾ it has been suggested that quantum measurements have definite results (rather than superpositions of living and dead “cats”) by virtue of special initial conditions of the measurement apparatus. It is the choice of which of the “special” initial conditions obtains that *determines* the outcome of a particular experiment. For this deterministic theory to recover the usual quantum probabilities, requirements are placed on the relative numbers of apparatus “special” states that lead to one or another outcome. In the present article we report progress in calculating these relative numbers.

¹ Institute for Theoretical Physics, State University of Utrecht, PO Box 80006, Utrecht 3508 TA, The Netherlands, and Physics Department, Clarkson University, Potsdam, New York, 13699-5820.

To follow the forthcoming presentation, familiarity with the above referenced measurement theory is necessary. Reference 4 is an exposition of that theory that I believe provides suitable preparation, and I will henceforth assume knowledge of the first three sections of that article (the substance of which is found dispersed in Refs. 1–3). Implicit in this course of action is a serious caveat: Because that theory contradicts deeply held intuitions on causality and time’s arrow, the reader of the present article who is not familiar with the earlier work will find those intuitions a severe hindrance to comprehension of what now follows.

Our concern here is the relative number of “special” states leading to one or another “definite” outcome in a situation where the experimental system distinguishes between one of two possible values for an observable. To study this, we will focus on a particular example, measurement of the z -component of a spin-1/2 particle by means of the Stern–Gerlach experiment. When we begin to calculate things, however, no specific properties of that setup will be invoked.

The purpose of estimating numbers of special states is to recover the experimentally verified standard probability predictions, namely the relation of probability to $|\psi|^2$, as postulated by Born.⁽⁵⁾ In our forthcoming presentation we will use the expected relation between probability and wave function to establish severe constraints on the properties of measurement apparatus. We will then dramatize the impact of these constraints by suggesting the impossibility of satisfying them. But like a playgoer armed with a telltale review, the reader will already appreciate that the constraints *can* be satisfied. Nevertheless, they can only be satisfied by apparatus with properties to be spelled out below and related to having certain system parameters distributed according to the Cauchy distribution [Eq. (8)]. For our measurement theory to be valid, this would imply the broad occurrence of this distribution, since there is nothing peculiar about the collimator and magnets of a Stern–Gerlach apparatus. It is this requirement on parameter distribution to which the title of the present article refers.

And there is a second surprise. We indicated that satisfying “Probability $\sim |\psi|^2$ ” forced a particular parameter distribution on systems, and, as will be seen below, this forcing arises from the requirements expressed by a certain functional equation. What if we had tried $|\psi|^1$, the first power? It is an amusing historical footnote⁽⁵⁾ that this was Born’s first guess.

For us the first power would not work (nor would any other power). The functional equation would not have a solution. This means that our approach imposes a *theoretical* requirement for “Probability $\sim |\psi|^2$.” From the standpoint of the Copenhagen interpretation I don’t know of any theoretical justification of the fact that the *square* of $|\psi|$ is involved, and

until now my personal view has been to consider this an example of the mysterious meshing of Nature and mathematics: Hilbert space stands out among vector spaces because of its norm (which allows the definition of an inner product) and just this norm is related to physical measurements. But the present approach singles out the *square* of $|\psi\rangle$ for reasons that seem unrelated to Hilbert space structure.

2. ANALYSIS OF APPARATUS

The nominal experimental situation is the Stern–Gerlach spin measurement for an atom with net spin $1/2$, a two-state system. As usual in our theory, no degrees of freedom can be ignored and we must consider a comprehensive Stern–Gerlach apparatus, where “comprehensive” indicates that, at the very least, account must be taken of the source of the atom, its internal degrees of freedom, collimators, sources for the currents giving rise to the magnetic field, walls of the chamber, gases in the chamber, microscopic structure of the film, or whatever it is that registers the ultimate capture of the atom in the spin-up and spin-down regions of the detector. This partial listing should remind the reader that our approach recognizes the measurement as a far more complex procedure than can be analyzed by simply replacing the apparatus by a fixed nonuniform magnetic field and idealizing the rest.

In the face of these myriad degrees of freedom one can simply throw up one’s hands and say, anything goes. Indeed, when one messes with the arrow of time, pretty much anything does go, so long as the microscopic equations of motion are satisfied. However, bearing in mind the two time-boundary condition rationale,⁽⁴⁾ there are still rules, *gradations of the unlikely*. For example, consider the cat map discussed in Ref. 4. This is the map of the unit square into itself: $x' = x + y$, $y' = x + 2y$, both modulo 1. In Ref. 4 we studied two-time boundary conditions in which a pair of subsets of the unit square A_0 and A_T were given. We then considered trajectories² that began in A_0 and after T steps arrived in A_T . At intermediate times (and for large T) the points are widely dispersed even though A_0 and A_T were small. An interesting and perhaps enlightening aspect of this arrangement is that if one watches the system approach the “final” time T , the dispersed points gather in what one might (if unaware of the two-time boundary condition) consider an unlikely way, and show an entropy decreasing convergence.³ But now we ask, having suffered an “unlikely”

² A “trajectory” is a set of pairs $\{(x_0, y_0), (x_1, y_1), \dots, (x_T, y_T)\}$ such that $x_{t+1} = x_t + y_t$, $y_{t+1} = x_t + 2y_t \pmod{1}$ for $t = 0, 1, \dots, T-1$.

³ In more detail: $(x_0, y_0) \in A_0$ and $(x_T, y_T) \in A_T$, so that the trajectories are solutions of a two-time boundary value problem. For any fixed t , call A_t the set of points $\{(x_t, y_t)\}$ that

confluence into A_T , how are the points distributed *within* A_T ? The answer, for most computer simulation runs, is that they are distributed fairly uniformly. Such a “fairly uniform” distribution is of course the *relatively* most likely situation, given the overall conditioning on arrival in A_T . Thus, in satisfying the unlikely two-time boundary condition, one generally satisfies it in the *least unlikely* way. This will be our basic operating assumption in the search for special states.⁴ This procedure is effectively a restatement of our basic probability postulate relating probabilities to dimensions of special state subspaces.

Faced therefore with an enormously complex apparatus, our approach will be to try to identify the *least unlikely* way to accomplish the “specializing,” namely the finding of microscopic apparatus states that cause the final overall quantum state to possess a single macroscopic description.

The macroscopic description of the events we study is as follows: There is a source preparation region from which we have reason to believe there emerges an atom with a spin state oriented at some particular angle with respect to the z -axis. Downstream from this region are collimators, magnets, and finally detectors in which a macroscopic detection event occurs. For many repetitions, the up versus down ratio of detections is the cotangent squared of the angle.

Let us suppose that we know enough about the preparation procedure to assign (by conventional analysis) a wave function of the following sort to the initial atomic spin state

$$u_\theta = \exp(i\sigma_x \theta/2) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (1)$$

After passing through the magnets, a relative amount of atomic wave function proportional to $\cos(\theta/2)$ will arrive in the neighborhood of the spin-up detectors and an amplitude proportional to $\sin(\theta/2)$ will arrive in the neighborhood of the spin-down detectors. Then, by the usual postulates, there will be final up/down probabilities in the ratio $(\cos^2 \theta/2)/(\sin^2 \theta/2)$.

lie on the trajectories for that fixed time. As t increases from 0, A_t spreads (is actually a highly stretched and chopped version of A_0) and its points disperse ever more widely within the unit square, until it is hard to distinguish their distribution from the uniform distribution. As t approaches T , they begin to contract in a way that is substantially a time-reversed playback of the initial dispersal (for A_0 and A_T of the same size and shape).

⁴ There is an important, perhaps philosophical, point here that we do not address: If we are dealing with a single system—the entire universe—how can an ensemble point of view be justified? As remarked in Ref. 4, the same issue arises in justifying the usual initial condition probabilities and I would say that in neither situation is there a complete understanding.

Within the context of our theory, the problem facing us is how to get the wave function of the atom from the u_0 preparation device to evolve *entirely* to the neighborhood of one or the other of the detectors,⁵ and in particular the *least unlikely* way to accomplish this. We have at our disposal the entire collection of microscopic degrees of freedom. Furthermore, once we have found ways to do this, we will want to count the relative numbers for getting up or down spins and see whether we can recover the ratios given above.

As a general principle we will assume that the fewer the number of degrees of freedom that need to be acted on, the better, that is, the less unlikely. Thus, if the “specializing” takes place after the atom has passed through the Stern–Gerlach magnetic fields, then its *translational* degrees of freedom will have to be acted on, which would not have been the case prior to entering the fields (because that’s the function of the fields, to make spin differences also be manifested as positional differences). Our expectation therefore is that the easiest way to get the entire atom into the up counter (say) is to act on it *before* it gets to the Stern–Gerlach fields; at this earlier stage only the single coordinate θ need be affected by whatever microscopic apparatus degrees of freedom are needed to do the job.⁶

At this point it should be clear that our problem has little to do with the specifics of the Stern–Gerlach experiment. Abstractly speaking, we need modes of the system capable of rotating the atomic spin about its x -axis. These are the modes that can affect θ and if we successfully send θ to 0 or π , then the “specializing” will be complete. The normal operation of the Stern–Gerlach apparatus will then carry the atom entirely to one or the other of the exit channels.

To summarize: The least unlikely way to get the atom to go to only one detector area is to rotate it to a spin-up state or a spin-down state *before* it enters the Stern–Gerlach magnetic fields.

What apparatus degrees of freedom can do the job of turning θ ? It is clear (although one usually neglects these effects) that quite a bit of the matter and “stray” fields encountered by the atom on its way to the Stern–Gerlach specific region can interact with the atomic spin. Fluctua-

⁵ Note that there is no reason why an incoming atom with (say) $\theta = 45^\circ$ cannot emerge as entirely spin-up. An irate colleague of mine once insisted that such a rotation was forbidden by conservation laws. Not so. Within the Stern–Gerlach region there are strong, inhomogeneous magnetic fields; having an incoming spin-up state usually emerge as a spin-up state is only a consequence of the adiabatic approximation. The apparatus has no trouble picking up the recoil angular momentum from the rotated atom.

⁶ Note too that as a consequence of having two-time boundary conditions the region where the “specializing” takes place may be distant from where one conventionally thinks the measurement or its registration occurs.

tions in collective magnetic fields, “chance” collisions, and much more, can contribute. For this reason, and also because we expect our considerations to be general, we will not even try to identify the detailed mechanisms. Instead we will focus on only that portion of the vast multiparticle Hamiltonian that is specifically potent for the operation we require. This will involve coupling between the spin raising and lowering operators of the atomic spin and the apparatus degrees of freedom. Various general forms of such couplings can be proposed and we find it most convenient to express this specifically interactive part as

$$H_I = \sum_{\alpha} \phi_{\alpha} P_{\alpha} \sigma_x \quad (2)$$

The operator P_{α} in (2) is built from the apparatus degrees of freedom and we have taken it to be a projection operator on some particular (collective) apparatus mode, labeled by α . The quantity ϕ_{α} is a coupling constant. We also assume that the P_{α} commute with one another, for which a sufficient, but not necessary, condition is that the atoms associated with the respective modes be located in substantially disjoint regions of space. Our picture is that the atom passes through various apparatus regions and if the apparatus mode α is excited in that region (P_{α} has eigenvalue 1) then there is a kick, proportional to ϕ_{α} , to the spin coordinate θ . Clearly this is not the most general possible Hamiltonian and we will comment later on other possibilities.

We remark that our attitude with respect to this “specializing” process is similar to the way one handles Brownian motion. As in many statistical mechanics treatments, the exact nature of the forces is unspecified and one concentrates on the statistics of quantities like ϕ_{α} .

Under time evolution (for a time T), the initial state u_{θ} is subjected to the following operator⁷

$$U = \exp(-iH_I T) = \exp\left(-iT\sigma_x \sum_{\alpha} n_{\alpha} \phi_{\alpha}\right) \quad (3)$$

with $n_{\alpha} = 0$ or 1 and where the apparatus is initially taken to be in an eigenstate of $\{P_{\alpha}\}$. This has the effect of rotating the spin about the x -axis so that

$$\theta \rightarrow \theta - 2T \sum_{\alpha} n_{\alpha} \phi_{\alpha} \quad (4)$$

⁷ In Eq. (2) there is also a suppressed dependence on the atom's position coordinates. Consequently the quantity T appearing in (3) should really depend on the passage time through the region associated with P_{α} . In (3) it might thus be preferable to write $\sum T_{\alpha} n_{\alpha} \phi_{\alpha}$. For simplicity, however, we will leave T 's α dependence inexplicit and absorb this variation into the coefficients ϕ_{α} .

In the language of Refs. 1–4, a “special” state of the apparatus would be one for which the new angle is either 0 or π . That is, a state (characterized by $\{n_\alpha\}$) is “special” if

$$\theta - 2T \sum_{\alpha} n_{\alpha} \phi_{\alpha} = \begin{cases} 0 & \text{mod } 2\pi; \text{ case UP} & \text{or} \\ \pi & \text{mod } 2\pi; \text{ case DOWN} \end{cases} \quad (5)$$

According to Refs. 1–4, the relative number of UP outcomes to DOWN outcomes (in a sequence of experiments with the same initial u_θ) is proportional to the relative number⁸ of apparatus microstates that yield the respective outcomes. In our current formulation this depends on the distribution of the ϕ_α .

It is convenient, and justified by the probability postulate of Ref. 4, to use probabilistic language. Instead of saying “the number of apparatus microstates” accomplishing some end we will speak of the *probability* that the apparatus accomplishes that end. We then combine the product $n_\alpha \phi_\alpha$ into a single random variable, to be called ϕ_α , and which may also take the value zero.⁹ We next define the function $f(\psi)$ to be, alternatively, the number of apparatus states for which $2T \sum n_\alpha \phi_\alpha = \psi$, or the probability that $2T \sum \phi_\alpha = \psi$. These are *not* mod 2π equations. Note that since we only consider ratios of f 's we need to confront neither the possible infinite dimensionality of the overall Hilbert space nor the implicit $d\theta$'s in the definition (5).

The summation in Eqs. (2)–(5) is over α , modes that are in a sense¹⁰ independent of one another. We take the number of such modes (or independent random variables) to be N . The expected large value of N already indicates that an unusual spectrum of fluctuations will be needed in the apparatus. Suppose the probability density for the ϕ_α distribution is $g(\phi)$. Assuming $g(\phi) = g(-\phi)$, one might expect the N -times sampled $g(\phi)$ to go over to a Gaussian $\exp(-\phi^2/2N\sigma^2)$ with σ^2 the single- ϕ second moment. There is no σ that gives the answer we require: Moderate σ will (with large N) lead to a flat distribution when the identification of ϕ with $\phi + 2\pi$, $\phi + 4\pi$, etc., is made. On the other hand, very small σ can at best give a Gaussian, which, when summed, also fails to give a $\cos^2 \theta/2$ distribution. The loophole will turn out to involve σ^2 , the second moment. If we were

⁸ By which is meant⁽¹⁻⁴⁾ the dimension of the appropriate subspace.

⁹ The random variable ϕ_α is a different object from $n_\alpha \phi_\alpha$ and it lumps together the ranges of possibilities open to both n_α and to (the originally defined) ϕ_α . To simplify notation, we also use ϕ_α for the random variable. As noted above, the “ T ” in Eq. (4) should vary with the exposure time to the various perturbing modes. This effect too is included in the total kick administered by the mode α .

¹⁰ That sense alternatively phrased as commutation of $\{P_\alpha\}$ or statistical independence of $\{\phi_\alpha\}$.

to drop its assumed existence, then we would no longer be forced by the central limit theorem to get a Gaussian.

Now we relate $f(\psi)$ to the probabilities that occur in a series of experiments starting with spin states u_θ . The ratio of UP to DOWN observations is $1/\tan^2(\theta/2)$. For this apparatus, not only will $\theta = 2T \sum \phi_\alpha$ yield UP, but so will $\theta + 2k\pi = 2T \sum \phi_\alpha$ for any k , $k = 0, \pm 1, \pm 2, \dots$. Define

$$F(\theta) = \sum_{k=-\infty}^{\infty} f(\theta + 2k\pi) \quad (6)$$

To get a DOWN measurement, not only will $\theta = \pi - 2T \sum \phi_\alpha$ do the job, but adding integer multiples of 2π will accomplish the same result. Using the evenness of $f(\theta)$ [which follows from that of $g(\phi)$], we have that the number of ways to get a DOWN measurement is $F(\theta + \pi)$. It follows that F (and thus implicitly f) must satisfy the functional equation

$$\tan^2 \frac{\theta}{2} = \frac{F(\theta + \pi)}{F(\theta)} \quad (7)$$

To recap, whatever distribution f one gets from g , recovery of the vast body of experience supporting the relation of frequency of observations to absolute value squared of the wave function, requires the satisfying of Eq. (7).

Remark. At this point, pessimism could well be in order regarding the possibility that “special” state abundances could reproduce the probabilities of the Copenhagen interpretation. In the latter, the recipe for experimental prediction is simplicity itself: Take the absolute value squared of the appropriate piece of wave function. According to Refs. 1–4, however, for each apparatus one must look at its complex inner workings and find “special” states whose abundance must recover the standard and universal probabilities. In the case we have discussed, this outrageous requirement takes the form of the functional equation (7) in which F is itself given by a difficult-to-calculate f which is in turn given by an unknown distribution g . How could one hope for success?

3. SATISFYING THE CONSTRAINTS

The key is the observation made earlier that to avoid the straight-jacket imposed by the central limit theorem, $g(\phi)$ should not have a second moment. The function f is the sum of N such random variables. Recalling the Lévy distributions^(6,7) suggests that the large θ dependence of $f(\theta)$

should be of the form $\theta^{-\gamma}$. Furthermore, we expect our “special” states to be unusual, so that $f(0)$ should either be infinite or extremely large.¹¹ We therefore assume the $\theta^{-\gamma}$ dependence to hold down to small θ . From Eq. (6), it then follows that the right-hand side of Eq. (7) will be dominated by $1/f(\theta)$ for θ near 0. This implies that $\gamma=2$. Such a drop-off is characteristic of the Cauchy distribution

$$C_a(x) = \frac{a}{\pi(a^2 + x^2)} \quad (8)$$

For now we assume the microscopic parameter [“ a ” in (8)] to be negligible. The requirement of agreement with experiment, together with our now determined $\gamma=2$, then imposes the demand [a hybrid of (6) and (7)] that for *all* θ

$$\tan^2 \frac{\theta}{2} = \frac{F(\theta + \pi)}{F(\theta)}, \quad \text{with} \quad F(\theta) = \sum_{n=-\infty}^{\infty} \frac{1}{(\theta + 2n\pi)^2} \quad (9)$$

The sum in (9) is well-known expansion of $1/(4 \sin^2 \frac{1}{2}\theta)$. *The first equality in (9) is therefore an identity!*

So the apparently impossible requirements can in fact be satisfied.¹² We now examine issues bypassed earlier.

We first remark that there was little significance to the particular form for the Hamiltonian, Eq. (2). It allowed us to treat the action of the apparatus as a sequence of N independent kicks. For systems in which such a breakup is impossible, the considerations of this paper offer neither contradiction nor confirmation on the matter of “special” states. [An alternative to Eq. (2), leading to the same overall conclusions, replaces P_α by $A_\alpha^\dagger + A_\alpha$, with A_α fermion operators.]

¹¹ This is a reflection of our earlier observation that the effects contributing to “specializing” are usually considered negligible (so that θ not being kicked at all is by far the most likely event). Recall the ice-water example of Ref. 4. Suppose one wished to estimate the one o’clock state of a glass of H_2O from its two o’clock condition, that is, partly ice, partly water. Then, if one neglects coherences and coincidences, one would expect to get a *smaller* ice cube at one o’clock. Of course, because we are working back to an *initial* condition, such coincidences are *not* negligible and the ice cube is in fact larger. However, if one did not include boundary conditions and only enumerated possible states and evolutions, the number of ways to get a smaller ice cube would be vastly larger than that number for any other outcome. This is the analogue of our $f(0)$, which represents the number of ways for the system not to be rotated.

¹² As remarked earlier, for rules other than “Probability $\sim |\psi|^{2n}$ ” we would not have been able to find a physically reasonable solution to the functional equation corresponding to (9). This is shown in the Appendix.

The Cauchy distribution, C_α of Eq. (8), has allowed us to achieve a result that, at least to one schooled in the Gaussian, is counterintuitive. For small a , even after many trials, generally speaking, nothing happens. However, the distribution's lack of both first and second moments means that when large deviations do occur they are without scale. Thus a 10° kick and a 20° kick are both unlikely, but their *relative* degree of unlikeliness is independent of a , for small enough a . The infinite moments also mean that although by some measures (e.g., a) the kicking modes are small, nevertheless their effect can be large.

Another property of the Cauchy distribution that differs sharply from the corresponding finite moment result is the way sums of these variables manage to achieve rare results. The way to get a group of humans whose total height is unusually large is to find a lot of tall people. The way to get a sum of Cauchy distributed random variables to be large is to have *one* be large and all the others small. ("The way" means the *least unlikely way*.)

If X and Y are Cauchy distributed with parameters a and b , then $X + Y$ is Cauchy distributed with parameter $a + b$. Thus if $g(\phi)$ is Cauchy with parameter a , then f is Cauchy with parameter Na . The satisfying of Eq. (7) therefore depended on $Na \ll 1$.

We next explore the consequences of finite Na . Let $\Gamma \equiv Na$. Then

$$\begin{aligned}
 F(\theta) &= \frac{\Gamma}{\pi} \sum_{n=-\infty}^{\infty} \frac{1}{(\theta + 2n\pi)^2 + \Gamma^2} \\
 &= \frac{1}{2\pi} \operatorname{Im} \left(\frac{1}{\tan \frac{1}{2}(\theta - i\Gamma)} \right) = \frac{(1/2\pi) \tanh \frac{1}{2}\Gamma}{\sin^2 \frac{1}{2}\theta + \cos^2 \frac{1}{2}\theta \tanh^2 \frac{1}{2}\Gamma} \quad (10)
 \end{aligned}$$

For an initial state u_θ , the probability of observing UP is

$$\operatorname{Prob}(\text{UP}) = \frac{F(\theta)}{F(\theta) + F(\theta + \pi)} = \frac{\cos^2(\theta/2) + \sin^2(\theta/2) \tanh^2(\Gamma/2)}{1 + \tanh^2(\Gamma/2)} \quad (11)$$

For small Γ this gives

$$\operatorname{Prob}(\text{UP}) \cong \cos^2 \frac{\theta}{2} - \frac{\Gamma^2}{4} \cos \theta \quad (12)$$

This is *not* the standard result. An atom going in with spin up ($\theta = 0$) has probability $\Gamma^2/4$ of being misread as down. It is not clear whether this is measurable. If Γ is microscopic, not only is precision in θ a problem, but there are already *a priori* reasons⁽⁸⁾ for expecting a microscopic level of measurement errors.

There is, however, an aspect of our current considerations that has far reaching consequences and may allow experimental verification. If our

measurement theory is valid and if “special” states occur in the way elucidated by our present considerations, then many systems will have pieces in their Hamiltonians either of the form (2) or of a similar form. In particular they will have parameter distributions or interaction time distributions whose tails are Cauchy distributed. When “special” states are needed, these modes are the most efficient way to provide them, but these dynamical degrees of freedom are present even when no constraints or conditioning exists. Under the latter circumstances they will assume “random” values and share available thermal energy, like any other modes.¹³ Thus material systems are expected to have fluctuations arising from these modes. A question now under investigation is whether these fluctuations are in any way distinguishable from the usual sort. Does the potential for extremely large amplitude fluctuations, in particular the infinite moments in the parameter distribution, lead to any recognizable experimental signature? Although these modes may be rare, their infinite moments may nevertheless have measurable effect outside the specific realm of “special” states.

Having postulated the widespread occurrence of Cauchy noise, we comment on its possible origin and magnitude. In fact, one does not need to look very far to find Cauchy-distributed energy transfers within ordinary matter. Consider a particle scattering from an atom in a metastable state. If in the course of the scattering the atom decays, the range of photon energies involved is of the form (8) (the Lorentz line shape), and by conservation of energy the distribution of energies picked up by the other elements of the system is just the Cauchy distribution. Mechanisms of this sort play a role in the work of Bell *et al.*⁽⁹⁾ in establishing scaling properties in stellar atmospheres.

As to the magnitude of the Cauchy noise, from Eq. (12) one concludes that Γ (and thus $a = \Gamma/N$) is small (or else we are already seeing an experimental disproof of our theory). If one considers scattering from metastable states to be the source of the Cauchy distributed angular kicks, then three physical quantities enter the determination of Γ . These are the transition matrix elements for the metastable state, the passage time for the particle through the region where it interacts with the system in the metastable state, and the interaction energy for the coupling of the atomic spin to the metastable state. Roughly speaking this should give a Cauchy parameter

$$a \sim H_{\text{atomic spin} - \text{metastable system}} \cdot \frac{\text{passage time of atom}}{\text{lifetime of metastable state}}$$

¹³ An analogy can be made to Brownian motion. The same modes that can drive a grain of pollen one centimeter from its starting point also contribute to the fluid specific heat.

Amorphousness, defects, and other circumstances can contribute to extremely long-lived metastable states. That, combined with the weak coupling to the atomic spin, can lead to very small Cauchy parameters. On the other hand, there is reason to believe that the number of such metastable states can be extremely large. (For example, for a spin glass with equal and opposite ferromagnetic and antiferromagnetic bonds there is finite entropy at zero temperature.⁽¹⁰⁾ This suggests that with deviation from perfect equality there will be a large number of metastable near-ground states.)

Although we have just provided a particular source of Cauchy-distributed noise, we suspect that if such noise is present in all possible measurements, then it will arise from considerations even more general than those described. (Gases, for example, are not expected to have the kind of long-lived states that are present in amorphous systems.) Conceivably the long-tailed distribution could result from another property that invariably is associated with quantum measurements, the amplification from microscopic to macroscopic. The associated sensitivity could give rise to the possibility of extremely large deviations. There also seems to me to be an analogy (or perhaps more) to the phenomenon of $1/f$ noise. This noise occurs widely, and in each of its occurrences one can offer theoretical explanations. Nevertheless, the mystery is that despite the particularness of each explanation, there is a common pattern to these phenomena. (There are also universal, if controversial, explanations. Note too that $1/f$ is not really $1/f$, but one over f to some power, near, but generally not exactly equal, to one.) The $1/f$ phenomenon (including its scientific-sociological aspect) provides me at least with an apology for my own inability to exhibit a universal mechanism for Cauchy-distributed noise: The absence of such an explanation in the much more thoroughly observed $1/f$ case makes plausible the difficulties of explaining even ubiquitous phenomena.

As remarked earlier, another way to induce a large change in the spin angle is to have the interaction time be long. There is in fact an extensive literature on *trapping* times, and the passage of electrons through amorphous materials provided one of the first physical examples of the need for distributions without variance.^(6,11)

4. DISCUSSION

We have considered the general problem of apparatus that measures a two-valued observable and for convenience have focussed on an atomic spin measurement by means of a comprehensive Stern–Gerlach apparatus. For a sequence of measurements on an incoming state of the form given

in Eq. (1), laboratory experience gives the ratio of UP to DOWN measurements as $\cos^2 \theta/2 : \sin^2 \theta/2$. The theory of Refs. 1–4 says that “special” states of the apparatus lead to final UP or DOWN states and that the dimensions of the subspaces for the respective outcomes provide the experimentally observed ratio. We here consider apparatus for which the coupling to the incoming states can be treated as a collection of N independent¹⁴ kicks of the angle. The recovery of the universally observed $1/\tan^2(\theta/2)$ then imposes severe restrictions on the distribution of kicks. We find the remarkable result that if the kick distribution coincides asymptotically with the Cauchy distribution, one recovers the universal probabilities. Furthermore, for no simple rule other than “Probability $\sim |\psi|^2$ ” can *any* distribution be found. These results lead to several observations:

1. They suggest the presence of universal noise arising from the Cauchy-distributed elements of the Hamiltonian. We discussed only 2-state systems in which an independent kick assumption could be made, but if Cauchy-distributed noise appears here, it should appear generally as well. Multistate discrimination is discussed in Ref. 4.

2. The scale of the noise is given by the parameter Γ , defined above. Since large Γ would lead to deviations from the $1/\tan^2(\theta/2)$ ratio, information on Γ 's size obtained from noise measurements would allow estimates of whether deviations from $1/\tan^2(\theta/2)$ would be observable. Note that since our demonstration is based on the *asymptotic* requirement for a Cauchy distribution, on scales small compared to a ($=\Gamma/N$), the microscopic distribution function $g(\phi)$ can differ from C_a of Eq. (8). The quantity a is therefore merely a parameter that sets the scale for the asymptotics.

3. Whereas in the not-too-distant past the suggestion of the universal presence of Hamiltonian matrix elements and interaction times with infinite moment distributions may have been deemed unlikely, recently discovered properties of materials make this idea more attractive.^(6,11) Furthermore, we have noted that the Lorentz line shape introduces an element of Cauchy-distributed energy transfers in any system not at absolute zero.

In this article we considered a Hamiltonian in which σ_x appeared. It is of course true that other couplings (e.g., σ_y terms in H_I) can connect UP and DOWN states. Because these are less effective than σ_x at reaching pure UP or DOWN, greater, and thus less likely, kicks of this sort would need to be involved. Confining attention to σ_x is thus a restriction to the *least unlikely* sorts of “special” states.

¹⁴ Meaning, the degrees of freedom effecting the kicks commute.

We close by remarking that although we consider the work described here to be a considerable step forward in the development of the theory of Refs. 1–4, we do not wish to minimize the questions that still exist. In Ref. 4 will be found a discussion of many such open problems.

APPENDIX

In this Appendix we consider the possibility that a power of the (absolute value of the) wave function different from the square should give physical probabilities. Calling this power β , we have, in place of Eq. (7),

$$\left| \tan \frac{\theta}{2} \right|^\beta = \frac{F(\theta + \pi)}{F(\theta)}, \quad \text{with } F(\theta) = \sum_{k=-\infty}^{\infty} f(\theta + 2k\pi) \quad (\text{A1})$$

For θ near zero, we again expect $F(\theta)$ to be dominated by the $k=0$ term, implying that $f(\theta) \sim B\theta^{-\beta}$, with B a constant. Now, although θ has just been described as “small,” it is smallness on the macroscopic level that is intended, since microscopically there will be a cutoff on $f(0)$ (analogous to the physical defusing of the mathematically “infinite” velocities in Brownian motion). On the other hand, the sources of f (and g) are microscopic fluctuations, so that even the small macroscopic angle at which we determined f 's behavior to be $B\theta^{-\beta}$ us already in the asymptotic regime. Since f is the sum of a large number of random variables with no second moments (or even no first moments), in its asymptotic regime it is necessarily of the form of the stable distributions⁽⁷⁾ and therefore has the form of a power. Consequently, for any macroscopic angle we have, not merely similitude, but equality: $f(\theta) = B\theta^{-\beta}$.

Our putative identity therefore takes the form

$$\left| \tan \frac{\theta}{2} \right|^\beta = \frac{F(\theta + \pi)}{F(\theta)}, \quad \text{with } F(\theta) = \sum_{n=-\infty}^{\infty} \frac{1}{|\theta + 2n\pi|^\beta} \quad (\text{A2})$$

(the B cancels). For $\beta \neq 2$, Eq. (A2) is simply not true. First, for the case $\beta \leq 1$, the sum in (A2) diverges and in any reasonable cutoff scheme the ratio [in (A2)] gives 1, not $|\tan \theta/2|^\beta$. Next, for $\beta > 1$, we can see that (A2) is false by expanding near $\theta=0$. To lowest order, if (A2) were true, we would have

$$\left| \frac{\theta}{2} \right|^\beta = \frac{2 \sum_{n=0}^{\infty} [(2n+1)\pi]^{-\beta}}{\theta^{-\beta}} \quad (\text{A3})$$

The θ^β cancels. The right-hand side of Eq. (A3) is related to the Riemann zeta function and were (A3) valid we would have expressed that function

as a simple combination of powers. But it is not valid, and equality in (A3) holds *only* at $\beta = 2$.

We have thus shown that *from our point of view*, that is, when probabilities are gotten by counting special states, no power other than the second (for $|\psi\rangle$) can satisfy the restrictions developed in this article. Thus, without Cauchy-distributed matrix elements (or times) there is *no* way to satisfy those restrictions and with those matrix elements it is only possible to get a simple rule if that rule involves the square.

ACKNOWLEDGMENTS

I thank C. Doering, B. Gaveau, and C. Newman for helpful advice on several of the issues raised here. I have benefitted from discussions with D. Dieks, J. Groeneveld, J. Hilgevoord, D. Hoeksema, M. Shlesinger, J. Uffink, and H. van Beijeren. This work was performed mostly under the auspices of the F. C. Donders Professorship of the Utrecht State University, for which I take this opportunity to express my gratitude. Finally, I also wish to acknowledge partial support of the United States National Science Foundation.

REFERENCES

1. L. S. Schulman, *J. Stat. Phys.* **42**, 689 (1986); *Phys. Lett. A* **102**, 396 (1984); *Phys. Lett. A* **130**, 194 (1988); *Ann. Phys. (N.Y.)* **183**, 320 (1988); *Found. Phys. Lett.* **2**, 515 (1989); in *Path Integrals from meV to MeV*, V. Sa-Yakanit *et al.*, eds. (World Scientific, Singapore, 1989), p. 35.
2. B. Gaveau and L. S. Schulman, *J. Stat. Phys.* **58**, 1209 (1990).
3. L. S. Schulman, C. R. Doering, and B. Gaveau, "Quantum decay in multi-level systems," *J. Phys. A* **24**, 2053 (1991).
4. L. S. Schulman, "Definite quantum measurements," *Ann. Phys. (N.Y.)*, to appear (1991).
5. M. Born, "Zur Quantenmechanik der Stossvorgänge," *Z. Phys.* **37**, 863 (1926); translation printed in J. A. Wheeler and W. H. Zurek, *Quantum Theory and Measurement* (Princeton University Press, Princeton, 1983).
6. M. F. Shlesinger, *Physica D* **38**, 304 (1989); B. D. Hughes, E. W. Montroll, and M. F. Shlesinger, *J. Stat. Phys.* **28**, 111 (1982); E. W. Montroll and B. J. West, in *Fluctuation Phenomena*, E. W. Montroll and J. L. Lebowitz, eds., 2nd edn. (North-Holland, Amsterdam, 1987).
7. W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. II (Wiley, New York, 1971).
8. E. P. Wigner, *Z. Phys.* **131**, 101 (1952); H. Araki and M. Yanase, *Phys. Rev.* **120**, 622 (1960); M. Yanase, *Phys. Rev.* **123**, 666 (1961).
9. T. L. Bell, U. Frisch, and H. Frisch, *Phys. Rev. A* **17**, 1049 (1978).
10. J. E. Avron, G. Roepstorff, and L. S. Schulman, "Ground state degeneracy and ferromagnetism in a spin glass," *J. Stat. Phys.* **26**, 25 (1981).
11. H. Scher, M. F. Shlesinger, and J. T. Bendler, "Time-scale invariance in transport and relaxation," *Phys. Today* **44**(1), 26 (1991).