A Theory Of Errors in Quantum Measurement

S.G.Rajeev*

University of Rochester. Dept of Physics and Astronomy.

Rochester. NY - 14627

May 3, 2003

Abstract

It is common to model the random errors in a classical measurement by the normal (Gaussian) distribution, because of the central limit theorem. In the quantum theory, the analogous hypothesis is that the matrix elements of the errors of an observable are distributed normally. We obtain the probability distribution this implies for the outcome of a measurement, exactly for the case of traceless 2×2 matrices and in the steepest descent approximation in general. Due to the phenomenon of 'level repulsion', the probability distributions obtained are quite different from the Gaussian.

*rajeev@pas.rochester.edu

In classical physics, there is a well-established 'standard model' of random errors in measurement, the Gaussian or normal distribution [1]: the error is a sum of a large number of more or less independent random contributions, so that the central limit theorem assures us that it is a Gaussian. There is not as yet a similar theory of errors in quantum measurements. In the quantum theory, an observable $\hat{A}$ is represented by a hermitian matrix. A simple model for the error would again be that it is the sum of a large number of independent random additions to $\hat{A}$. If each matrix element is independent of the others (except for the condition of hermiticity), the error would be described by a Gaussian random matrix $\hat{B}$ added to the observable. A measurement of $\hat{A}$ in the presence of this error would yield an eigenvalue of the sum $\hat{R} = \hat{A} + \hat{B}$ instead of $\hat{A}$.

For such a random matrix $\hat{R}$, the analogue of the eigenvalue problem is to ask for the probability distribution of the eigenvalues. Wigner solved such problems in the context of nuclear physics [5, 6], followed by fundamental contributions by Dyson [7], Mehta [8], Itzykson and Zuber [13] and many others. By now this theory has an extensive literature in physics [8] and mathematics [9, 10]. To make this paper more accessible, we will solve the problem of determining the probability distribution of the eigenvalue for a simple example by elementary means first. Then we turn to the more general case, using results from the literature on random matrix theory.

Let us recall in some more detail the theory of errors in classical physics. An observable is a function $A : M \to R$ from the phase space M taking real values. The state of a classical system is given by a point $\xi \in M$ in the phase space. Ideally, a measurement of the observable A on a system in this state will yield the value $A(\xi)$.The standard model of errors in classical measurement is that there are a large number of small, more or less independent, random corrections $B_1, \cdots B_M$ which add to this observable:

$$R = A + \sum_{k=1}^{M} B_r \tag{1}$$

Irrespective (upto mild assumptions) of the distribution of $B_1 \cdots B_M$, the sum

will, in the limit $M \to \infty$ will tend to a Gaussian distribution (the central limit theorem)[1]:

$$P(B) = e^{-cB^2}. \tag{2}$$

We can assume that the mean of this distribution is zero since otherwise it can be absorbed into the definition of A. (In any case such systematic errors cannot be analyzed by statistical methods.) The variance is usually assumed to be the same at all points in the phase-space so that c is a constant. Thus we model the outcome of the measurement by a Gaussian random variable whose mean is the 'true' value $a = A(\xi)$ of the observable and the standard deviation is a measure of the size of the error:

$$p_A(x) \propto e^{-c(x-a)^2}. \tag{3}$$

There is by now a well-developed sampling theory on how best to estimate this mean a and standard deviation σ from repeated measurements of R [1, 2].

If the state of the system is not known exactly, there is a probability density function on the phase space $\rho : M \to R$ which determines the 'instrinsic' probability distribution of A in the absence of errors:

$$\tilde{P}_A(a) = \int \delta(a - A(\xi))\rho(\xi)d\xi \tag{4}$$

If we add in the error, we get the convolution of this with the Gaussian:

$$p_A(x) \propto \int \tilde{P}_A(x - B)P(B)dB \tag{5}$$

Thus there are two separate ways in which statistical considerations enter classical error analysis: the classical observable can have an intrinsic randomness because the state of the system is not completely known; and there can be a random error added to the observable. In the quantum theory, in addition there is another, more fundamental source of randomness: the uncertainty principle. Thus a proper theory of errors has to take into account all three sources of randomness.

In the quantum theory, an observable [3, 4] is represented by a self-adjoint operator on the Hilbert space of states. Let us consider a hermitean matrix $\hat{A}$ of

finite dimension N with eigenvalues $a_1, a_2 \cdots a_N$ and corresponding eigenstates $|u_1> \cdots |u_N>$. Ideally, if a measurement of the observable represented by A is made on a system known to be in state $|\psi>$, the outcome will be one of the eigenvalues a_k, with probability $|<u_k|\psi>|^2$. Another way to state this result (convenient in the following) is that the probability density function of this random variable is

$$\tilde{P}_{\hat{A}}(x) =< \psi|\delta(x-\hat{A})|\psi> . \tag{6}$$

More generally, if the state of the system is only partially known, we have a mixed state described by a density matrix $\hat{\rho}$ which is a positive hermitean matrix with $\text{tr}\ \rho = 1$. We have then the 'intrinsic' probability distribution

$$\tilde{P}_{\hat{A}}(x) = \ \text{tr}\ \hat{\rho}\delta(x-\hat{A}). \tag{7}$$

This formula takes into account two of the sources of randomnes mentioned above: the fundamental uncertainty of quantum mechanics as well as our possible lack of knowledge of the state of the system.

However, this is still an ideal situation that is never realized in practice: there is always some error in the measurement which is a third source of randomness. It is useful to have theory of errors in quantum measurement analogous to the classical theory above, particularly in view of current interest in quantum information theory.

As an elementary example, suppose the observable is a traceless 2×2 matrix $\hat{A} = \begin{pmatrix} a & 0 \\ 0 & -a \end{pmatrix}$. We can think of it as the energy of a spin half particle in a constant external magnetic field along the z-axis. There are different sorts of errors that can affect this energy. The simplest possibility is that the magnetic field fluctuates due, for example, to thermal effects in the currents producing it. We can model this as a sum of small independent additive corrections to the magnetic field. By the central limit theorem, this additional magnetic field $\boldsymbol{B}$ can be represented as a Gaussian random variable.

$$\hat{R} = \hat{A} + \boldsymbol{\sigma}\cdot\boldsymbol{B}. \tag{8}$$

The mean of $\boldsymbol{B}$ can be assumed to be zero. Otherwise it can be absorbed into the definition of the fixed magnetic field; after a rotation we can reduce it to the above diagonal form anyway. It is reasonable, and simplest, to postulate that the error is rotation invariant;i.e., all the components have the same variance:

$$P(\boldsymbol{B}) \propto e^{-c\boldsymbol{B}^2}. \tag{9}$$

We can also think of this as the Boltzmann distribution for the fluctuating magnetic field, as the energy of a magnetic field is proportional to $\boldsymbol{B}^2$. This means that the matrix elements of $\hat{R}$ are also Gaussian random variables, with mean $\hat{A}$.

The eigenvalues of $\hat{R} = \hat{A} + \boldsymbol{\sigma} \cdot \boldsymbol{B}$ will be $\pm r$, where r is a real random variable. Calculating its distribution from that of $\boldsymbol{R}$,

$$P_{\hat{A}}(\hat{R}) \propto e^{-\frac{1}{2}c \text{ tr } [\hat{R}-\hat{A}]^2}, \tag{10}$$

is now the analogue of solving the eigenvalue problem. We will first consider the case where all the states are equally probable: the density matrix is a multiple of the identity. In effect we have to average over all fluctuations that change $\hat{R}$ without affecting its eigenvalues. Representing $\hat{R} = \boldsymbol{\sigma} \cdot \boldsymbol{R}$ as a vector and transforming to spherical polar co-ordinates, we can perform the average over random direction of the vector $\boldsymbol{R}$:

$$\begin{aligned} p_{\hat{A}}(r) &\propto \int \text{ tr } \frac{1}{2}\delta(r-\hat{R})P_{\hat{A}}(\hat{R})d\hat{R} \\ &\propto r^2 \int_{-1}^{1} e^{-c(r^2+2ar\cos\theta)} \sin\theta d\theta \propto r\sinh[2acr]e^{-cr^2}. \end{aligned} \tag{11}$$

Recalling that the observed eigenvalue can take also negative values, we normalize this distribution to get

$$p_{\hat{A}}(r) = \frac{1}{a}\sqrt{\frac{c}{\pi}} r \sinh[2acr]e^{-c(r^2+a^2)}. \tag{12}$$

We can also write this in a way that explicity displays the peaks at $r = \pm a$:

$$p_{\hat{A}}(r) = \frac{1}{2a}\sqrt{\frac{c}{\pi}} r \left[e^{-c(r-a)^2} - e^{-c(r+a)^2}\right] \tag{13}$$

In Fig. 1 we plot this probability distribution.

In summary, the quantum mechanical error is modelled by a Gaussian for the matrix elements which leads to a markedly distribution from the Gaussian for the observed eigenvalue. There is a second order zero for the distribution at the origin, a consequence of the 'level repulsion' of the eigenvalues of a random matrix. Note that the peaks are displaced outwards from the 'true' eigenvalues ± 1 due to this level repulsion. (In the figures we have assumed an unrealistically large error to illustrate the phenomena better.)

The phenomenon of level repulsion is well-known in the theory of random matrices. The set of all traceless hermitean 2×2 matrices with a given spectrum $\pm r$ is a sphere of radius r. The volume of this sphere shrinks as $r \to 0$, so that it is unlikely that the eigenvalues of a random matrix are close together. We can also think of the logarithm of the volume of the set of all matrices with a given spectrum as an entropy [11]. The probability distribution above can be thought of as a compromise between maximizing this entropy and minimizing the energy of a fluctuation in $\boldsymbol{B}$.

So far we dealt with the case when all the states of the system are equally likely;i.e, when the density matrix is proportional to the identity. The probability distribution in general will be given by averaging over the error as well as the states weighted by the density matrix:

$$p_{\hat{A}}(x) = \int \text{ tr } \rho \delta(x - [\hat{A} + \hat{B}]) P(\hat{B}) dB \tag{14}$$

In the special case of a two dimensional Hilbert space, we can expand

$$\hat{\rho} = \frac{1}{2} + \boldsymbol{\sigma} \cdot \boldsymbol{\rho}, |\boldsymbol{\rho}| \leq \frac{1}{2}. \tag{15}$$

The inequality ensures that the density matrix is positive; the trace of $\hat{\rho}$ is normalized to one. We can again evaluate the integral by passing to spherical polar co-ordinates, using the identity

$$\delta(x - \boldsymbol{\sigma} \cdot \boldsymbol{R}) = \frac{1}{2} \left[\delta(x - r) + \delta(x + r)\right] + \frac{\boldsymbol{\sigma} \cdot \boldsymbol{R}}{2r} \left[\delta(x - r) - \delta(x + r)\right] \tag{16}$$

to get

$$p_{\hat{A}}(x) \quad \propto \quad \left\{ x \left[e^{-c(x-a)^2} - e^{-c(x+a)^2} \right] \right.$$

$$+\frac{\eta}{a}\left[\left(x-\frac{1}{2ac}\right)e^{-c(x-a)^2}+\left(x+\frac{1}{2ac}\right)e^{-c(x+a)^2}\right]\Bigg\} \quad (17)$$

where $\eta = \text{ tr } \hat{\rho}\hat{A}$ is the average of A over states. In Fig. 2 we plot this probability density in the extreme case of a system in an eigenstate of A. Note that there is a small peak at the 'wrong' eigenvalue, caused by the errors.

We are ready now to take up the more general case of an $N \times N$ hermitean matrix $\hat{A}$ with an additive random error $\hat{B}$. We can assume that the mean of $\hat{B}$ is the zero matrix: otherwise, we can redefine $\hat{A}$ by absorbing this mean matrix into it. For simplicity we assume that the matrix elements of $\hat{B}$ are statistically independent of each other (except for the constraint of hermiticity). In order for this condition to be true in all choices of basis, the joint probablity distribution should be a Gaussian [8]:

$$P(\hat{B}) \propto e^{-c_1 \text{ tr } \hat{B}^2 - c_2(\text{ tr } \hat{B})^2}, \quad (18)$$

for some positive constants c_1, c_2. Another justification of the normal distribution would come from the central limit theorem: each matrix element of B is the superposition of a large number of small errors.

In the presence of the error, we are really measuring the observable $\hat{R} = \hat{A} + \hat{B}$. The joint probability density function of its matrix elements is

$$P_{\hat{A}}(\hat{R}) \propto e^{-c_1 \text{ tr } [\hat{R}-\hat{A}]^2 - c_2[\text{ tr } \hat{R} - \text{ tr } \hat{A}]^2} \quad (19)$$

We can write $\hat{R} = \hat{U} r \hat{U}^\dagger$ where r is a diagonal matrix and $\hat{U}$ a unitary matrix. By averaging over $\hat{U}$ we can get the joint probability density of the eigenvalues of $\hat{R}$. In this process we must remember that the Jacobian for transforming from the matrix elements of $\hat{R}$ to $r, \hat{U}$ is [8]

$$d\hat{R} \propto \Delta(r)^2 d^N r d\hat{U}, \quad \Delta(r) = \prod_{k<l}(r_k - r_l). \quad (20)$$

Thus the joint distribution for the eigenvalues of $\hat{R}$ is

$$p_{\hat{A}}(r_1, \cdots r_N) \propto \Delta(r)^2 e^{-c_1 \sum_k [r_k^2 + a_k^2] - c_2[\sum_k r_k - \sum_k a_k]^2} \int d\hat{U} e^{2c_1 \text{ tr } r\hat{U}^\dagger \hat{A} \hat{U}} \quad (21)$$

The last integral was evaluated by Harish-Chandra [12] (and rediscovered by Itzykson and Zuber [13] in a more physical context)

$$\int_{U(N)} d\hat{U} e^{2c_1 \text{ tr } r\hat{U}\hat{A}\hat{U}^\dagger} \propto \frac{\det e^{2c_1 r_k a_l}}{\Delta(r)\Delta(a)}. \tag{22}$$

Thus

$$p_{\hat{A}}(r_1, \cdots r_n) \propto \frac{\Delta(r)}{\Delta(a)} \det e^{2c_1 r_k a_l} e^{-c_1 \sum_k [r_k^2 + a_k^2] - c_2 [\sum_k r_k - \sum_k a_k]^2} \tag{23}$$

We can also write this result as

$$p_{\hat{A}}(r_1, \cdots r_N) \propto \frac{\Delta(r)}{\Delta(a)} \sum_{\sigma \in S_N} \text{sgn } (\sigma) e^{-c_1 \sum_k [r_k - a_{\sigma_k}]^2 - c_2 [\sum_k r_k - \sum_k a_k]^2} \tag{24}$$

where the sum is over all permutations of the indices$\{1, 2, \cdots N\}$.

For the simplest example $A = \begin{pmatrix} a & 0 \\ 0 & -a \end{pmatrix}$ we considered earlier, we can write the above as

$$p_{\hat{A}}(r_1, r_2) \propto e^{-[c_2 + \frac{1}{2}c_1][r_1 + r_2]^2} (r_1 - r_2) \sinh[2ac_1(r_1 - r_2)] e^{-\frac{1}{2}c_1 (r_1 - r_2)^2}. \tag{25}$$

This shows that the sum and difference of eigenvalues are independent random variables. The sum is a Gaussian random variable with mean zero. The difference has the distribution we derived earlier. Thus c_2 describes a 'classical' source of error which shifts both eigenvalues the same way (and is a Gaussian) while c_1 is a 'quantum' error that affects their difference.

By integrating over all but one eigenvalue we can get the probability distribution for the outcome of a single measurement of A. This leads to a complicated expression in general but in the steepest descent approximation we can get a simpler form:

$$p_{\hat{A}}(x) \propto \sum_{k=1}^{N} e^{-c(x - a_k)^2} \prod_{k \neq m} \frac{|x - a_m|}{|a_k - a_m|}, \tag{26}$$

where $c = c_1 + c_2$. This probability distribution is peaked near (but not at) the eigenvalues $a_1, \cdots a_N$ as one might expect. However even when the density matrix is proportional to the identity, not all these peaks are of the same height: due to the repulsion of eigenvalues of a random matrix, the extreme values are

more probable. Moreover, the peaks of the probability are displaced towards the edges due to this repulsion. We plot an example in Fig. 3.

The spacing between eigenvalues are of order $\hbar$. In the limit as $\hbar \to 0$ and $N \to \infty$ we get back the classical description. It is possible to develop a semi-classical theory along the lines of Ref.[14]. We hope to return to these issues in a later publication.

It would be interesting to verify our predictions experimentally. Perhaps Superconducting Quantum Interference Devices (SQUIDs) or Spin Resonance will provide such tests.

Acknowledgement I thank A. Agarwal,G. Agrawal, J. Eberly, G. Krishnaswami, E. Langmann, A.C. Melissinos and C. Stroud for discussions.

References

[1] W. T. Eadie, D. Rudjard, F. E. James and B. Sadoulet *Statistical Methods in Experimental Physics* North-Holland,Amsterdam (1971);A. C. Melissinos *Experiments in Modern Physics* Academic Press (San Diego 1966).

[2] C. R. Rao *Linear Statistical Inference and its Applications* Wiley, New York (1965).

[3] J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, (Princeton, 1996).

[4] M. Bell, K. Gottfried, M. Veltmann, *John Bell on the Foundations of Quantum Mechanics* World Scientific (Singapore,2001).

[5] E. P. Wigner, *Ann. Math.* **62**, 548 (1955); **67**, 325 (1958).

[6] C. E. Porter, ed., *Fluctuations of Quantal Spectra*, Academic Press, New York (1965).

[7] F. J. Dyson, *J. Math. Phys.*, **3** 140 (1962).

[8] M. L. Mehta, *Random Matrices*, 2nd ed. (Academic Press, San Diego, CA, 1991).

[9] D. V. Voiculescu, K.J. Dykema and A.Nica, *Free Random Variables* American Mathematical Society, Providence, USA (1992).

[10] N. Katz and P. Sarnak *Random Matrices, Frobenius Eigenvalues, and Monodromy* American Mathematical Society, Providence, USA (1998).

[11] L. Akant, G.S. Krishnaswami and S.G. Rajeev, *Int.J.Mod.Phys.*, **A27** 2413 (2002): hep-th/0111263; A. Agarwal, L. Akant, G.S. Krishnaswami and S.G. Rajeev, *Int.J.Mod.Phys.*, **A18**, 917 (2003): hep-th/0207200

[12] Harish-Chandra, *Am. J. Math.*, **79** ,87(1959).

[13] C. Itzykson and JB Zuber, *J. Math. Phys.*,**21**, 411(1980).

[14] S. G. Rajeev, in the Proceedings of the 70th meeting of Physicists and Mathematicians,IRMA, Strassbourg 2002, ed. by V. Turaev and T. Wurzbacher;hep-th/0210179

Figure Captions

1. Probablity density of the outcome of measuring an observable with A eigenvalues ± 1 in a mixed state with equal weight for the eigenvalues and error parameter $c = 3$. The peaks are shifted to ± 1.15 from ± 1.

2. Probablity density of the same observable in an eigenstate with eigenvalue 1.0 and error parameter $c = 3$. The errors cause a small peak at the 'wrong' eigenvalue.

3. Probability density of an observable with eigenvalues $-1, 0.5, 1.0$ with error parameter $c = 2$. The peaks are shifted to $-1.25, 0.36, 1.41$ respectively.

Figure 1 of S. G. Rajeev

Figure 2 of S. G. Rajeev

Figure 2 of S. G. Rajeev