© *Proc. Indian natn. Sci. Acad.*, **50**, A, No. 4, 1984, pp. 386–396.
Printed in India.

Quantum Theory

ON INFORMATION ASSOCIATED WITH AN OBJECT

SUBHASH C KAK

Department of Electrical and Computer Engineering, Louisiana State University, Baton Rouge, LA 70803

(*Received 29 August 1983; after revision 16 December 1983*)

This paper presents an exploratory study of the significance of the notion of information and that of the fundamental uncertainty in a quantum description for object classification. Assuming that the correspondence principle applies to information, we have been able to relate this uncertainty to the number of symmetries in a quantum description. If space-time is considered a discrete lattice the number of interaction symmetries turns out to be finite and close to the currently accepted value.

Key Words : Quantum Description; Uncertainty Principle; Elementary Particles

INTRODUCTION

IF one should ask the question why are quantum and classical phenomena different there are three ways one could attempt an answer. First, one could ascribe the difference to the uncertainty principle becoming operative at the atomic distances. Second, one may attribute the difference to the existence of two new forces in the nuclear domain, viz. the strong and the weak forces. Third, one may look at classical phenomena as a limit of quantum phenomena. It is also reasonable to view the above three reasons as being related. If that assumption is correct then the short-range forces (strong and weak) should be variants of the long-range forces (electromagnetic and gravitational) resulting from the mediation by the uncertainty principle.

In recent years a unification of the short-range and the long-range forces has been sought in the above manner and achieved to a certain degree. All the four forces are now described by theories that have the same general form. It has also been found how the weak force and electromagnetism can be understood in the context of a single theory. This development has proceeded by studying aspects of quantum electrodynamics as well as properties of elementary particles. While it is clear that this is how a detailed theory should evolve it seems justified to ask if any limits or bounds are associated with the interaction symmetries in a quantum description. In other words, is it possible to predict the amount of information in strong and weak interaction symmetries using the uncertainty principle alone? The fact that we have taken the uncertainty relations to represent the divide between classical mechanics and quantum mechanics suggests that such a prediction should be possible provided we extend the applicability of the correspondence principle to information as well. Such an extension would imply the following equation (Kak, 1976, 1977 and 1982)

$$I\,(\text{classical m.}) = I\,(\text{q.m.}) + \text{uncertainty} \qquad \ldots(1)$$

where I denotes information. Now since uncertainty implies negative information, the above equation would require a positive information to be associated with a quantum description.

Section 2 reviews some elementary properties of information. In section 3 we indicate the combinatorial aspects of particle statistics. The significance of attributes of an object has been taken up in section 4. Different approaches to the study of particle structure are discussed. Thus, an object could be taken to be a system consisting of different subjects or a specific structure (topologically different from the structures for other objects) built out of a collection of the same subobject. Section 5 presents a calculation on the amount of information associated with an object under the assumptions of eqn. (1) and a discrete space-time lattice.

Measure of Information

We shall briefly describe a measure of information frequently used in communication theory that seems appropriate for our purpose. Let us consider a box that is filled with 2 kinds of particles in equal number. Supposing the particles are identical excepting for their charge, clearly a single measurement of charge suffices to determine the kind of the particle being observed. An information of 1 bit will be associated with such a measurement of the object. In general if the probability of the measurement x is $P(x)$, the information $I(x)$ is

$$I(x) = \log_2 [1/P(x)] \qquad \ldots(2)$$

Hence when $P(x) = 1/2$, for example, the information is 1 bit.

The average information with the system of measurements is called the entropy $H(x)$ and given by:

$$H(X) = \Sigma P(x) \log [1/P(x)] \qquad \ldots(3)$$

What significance could our information measure have in object classification? If a total of n bits were necessary to identify an object, quite clearly each of these bits represents an attribute. Furthermore, the total number of distinguishable elementary objects would then equal 2^n.

Let us now consider an object with spin $\frac{1}{2}$. Any measurement of spin for such an object will yield one of two states: parallel or anti-parallel to the object's motion. While such a measurement will not aid in the object classification it defines an internal attribute of the object.

Every measurement requires expenditure of some energy. According to statistical considerations the entropy corresponding to M number of equally likely states of the system is

$$k \log_e M$$

where k is the Boltzmann's constant. At the same time the energy input E brings about an entropy change of E/T, where T is the temperature in degrees Kelvin. Clearly, then, the energy required to obtain 1 bit of information is $kT \log_e 2$.

Let us now determine the probability density functions that maximize average information subject to appropriate constraints.

Case 1. Let us maximize the expression

$$I = \int_{-\infty}^{\infty} p(x) \log p(x) \, dx$$

subject to the conditions that

$$\int_{-\infty}^{\infty} p(x) \, dx = 1 \quad \text{...(4)}$$

$$\int_{-\infty}^{\infty} x^2 p(x) \, dx = \sigma^2 \quad \text{...(5)}$$

σ^2 is a given constant.

We form the expression

$$F_0 = F + \lambda_1 F_1 + \lambda_2 F_2 = p \log p + \lambda_1 p + \lambda_2 x^2 p$$

where λ_1 and λ_2 are Lagrangian multipliers. Using the Euler-Langrage equation on the above :

$$\frac{\partial F_0}{\partial p} = 0 = 1 + \log p + \lambda_1 + \lambda_2 x^2$$

or

$$P = \exp(-(\lambda_1 + 1)) \exp(-\lambda_2 x^2) \quad \text{...(6)}$$

The multipliers λ_1 and λ_2 are evaluated by substituting equation 6 into equations 4 and 5, from which we obtain finally.

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-x^2/2\sigma^2), \; -\infty < x < \infty \quad \text{...(7)}$$

In words, therefore, for fixed variance the normal distribution has the largest entropy. An easy calculation shows that its value equals

$$H = \tfrac{1}{2} \log_2 2\pi e \sigma^2 \quad \text{...(8)}$$

Case 2. If we maximize $H(X)$ for a limited peak value x, the constraint is

$$\int_{-M}^{M} p(x) \, dx = 1$$

This can be shown to lead to the uniform distribution

$$p(x) = \begin{cases} \frac{1}{2M}, & |x| \leqslant M \\ 0 & |x| > M \end{cases} \quad \text{...(9)}$$

for which the associated entropy is log $2M$.

Case 3. Let us maximize $H(X)$ for x limited to non-negative values and a given average value. The constraints are

$$\int_0^\infty p(x)\,dx = 1$$

$$\int_0^\infty xp(x)\,dx = u$$

The density function $p(x)$ is found from

$$\frac{\partial}{\partial p}[-p \log p + \lambda_1 p + \lambda_2 xp] = 0$$

or $$p(x) = \exp(\lambda_1 - 1)\exp(\lambda_2 x)$$

As before λ_1 and λ_2 may be eliminated by substituting into the constraint equations to give us

$$p(x) = \begin{cases} 0 & , x < 0 \\ \frac{1}{u}\exp(-x/u), & x \geqslant 0 \end{cases} \qquad ...(10)$$

The entropy associated with this distribution is $1n\,(ue)$.

The choice of a specific distribution to compute maximum entropy would therefore depend on the physical problem. If an object in free space is being studied one should consider the Gaussian distribution. An object in a potential well would be governed by the uniform distribution. Processes characterized by an asymmetric variable like time would, on the other hand, be governed by the exponential distribution.

The cases governed by the uniform and the exponential distributions will not be considered further in this paper.

When the variable x is discrete: $0, \pm x_0, \pm 2x_0, \ldots$, the determination of the probability distribution which maximizes the entropy subject to appropriate constraints can be carried out similarly. As example for the discrete analog of Case 1.

$$\Sigma\, P(nx_0) = 1$$

$$\Sigma\, n^2 x_0^2\, P(nx_0) = \sigma^2$$

the probability distribution maximizing entropy is discrete Gaussian:

$$P(nx_0) = \exp(-(1 + \lambda_1))\exp(-\lambda_2 n^2 x_0^2) \qquad ...(11)$$

where λ_1 and λ_2 are solutions of:

$$\exp(-(1 + \lambda_1))\,\Sigma\exp(-\lambda_2 n^2 x_0^2) = 1$$

$$\exp(-(1 + \lambda_1))\,\Sigma\, n^2 x_0^2 \exp(-\lambda_2 n^2 x_0^2) = \sigma^2 \qquad ...(12)$$

The analogs for Case 2 and Case 3 can be derived similarly.

Probability Distribution of States

This section is a review of the different probability functions that arise owing to the different physical assumptions about the objects. Consider a system charaterized by N states numbered 1, 2, ..., N. Consider further n objects ($n < N$) numbered 1, 2, ..., n that occupy some of the states of the system. Assuming that all the states are equally likely we wish to determine the probability that states numbered 1 through n will have one object each. Such a pattern will be denoted by E.

The probability of E depends on two things:

(i) Are the objects distinguishable or indistinguishable?

(ii) Can more than one object be placed in the same state?

This gives rise to four possibilities, which will be considered one by one.

Objects distinguishable; Any number in a state

Total number of patterns $= N^n$

Number of patterns of $E = n!$

$$p(E) = \frac{n!}{N^n} \qquad \text{...(13)}$$

This distribution is the Maxwell-Boltzmann statistics.

Objects distinguishable; Only one in each state

Total number of patterns $= {}_NP_n = \dfrac{N!}{(N-n)!}$

Number of patterns of $E = n!$

$$p(E) = \frac{n!(N-n)!}{N!} = \frac{1}{\binom{N}{n}} \qquad \text{...(14)}$$

This case is assumed not to arise in any physical process.

Objects indistinguishable; Any number in a state

Total number of patterns $= \binom{N+n-1}{n}$

Patterns of $E = 1$

$$P(E) = \frac{1}{\binom{N+n-1}{n}} \qquad \text{...(15)}$$

This is the Bose-Einstein statistics.

Objects indistinguishable; Only one in a state

Total number of patterns $= \binom{N}{n}$

Patterns of $E = 1$

$$P(E) = \frac{1}{\binom{N}{n}} \qquad \ldots(16)$$

This is the Fermi-Dirac statistics. Note that the result is the same if only one object is allowed to occupy a state no matter whether the objects are distinguishable or indistinguishable.

When N is very large and the states are approximately equally likely (as at high temperatures) the Fermi-Dirac and Bose-Einstein statistics give results that are essentially the same as for the classical Maxwell-Boltzmann statistics. At low temperatures, the low-energy states are more likely than the high-energy states and therefore the expressions given above must be modified.

The determination of entropy associated with an object would depend on the statistics obeyed by the object.

On the Structure of an Object

Consider that we are given a variety of objects each of which is characterized by k attributes. If each of these attributes defines a choice between two possibilities the total number of different objects, without counting compounds composed of more than one object, will be $2^k - 1$. These objects can be represented as k-long binary strings:

$$\begin{array}{l} 0\ 0\ \ldots\ 0\ 0\ 1 \\ 0\ 0\ \ldots\ 0\ 1\ 0 \\ 0\ 0\ \ldots\ 0\ 1\ 1 \\ \quad\ \ldots \\ \quad\ \ldots \\ 1\ 1\ \ldots\ 1\ 1\ 1 \end{array} \qquad \ldots(17)$$

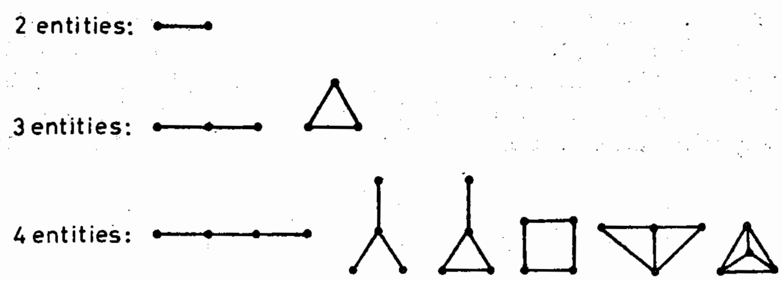
An object with k-attributes can be viewed, in its structural aspects, in three different ways which we describe briefly.

Model 1

Each object is considered as being formed out of upto k different sub-objects. The locations of ones in the binary string representation shown above would represent the sub-objects constituting the object. The vacuum state may be represented by the null string 00 ... 000.

Model 2

Each object is formed out of a single fundamental entity which combines with itself to form different patterns, with each topological category representing the object and its excitation states. To illustrate this we consider patterns formed out of 2, 3, and 4 entities as shown below:



These patterns have been formed subject to the condition that at best only one branch can connect any 2 nodes, and topologically equivalent patterns have been placed in a single pattern category. A branch between two nodes represents interaction. More than one branch may, therefore, represent enhanced interaction or, equivalently, excited object state. The number of different patterns depends also on the dimensionality of the basis space: the patterns sketched above are valid for a 2-dimensional space.

The graphs shown above are only to illustrate the idea of this construction and are not meant as serious models. A different set of construction rules will doubtless generate other families of graphs.

Model 3

Each object is a complex structure built out of different entities, some of which may occur more than once. In other words, this model combines aspects of both of the preceding models. It also defines varieties of structures much richer than in models 1 and 2.

As an illustration of this model we consider the following hierarchical structure.

$$
\begin{array}{l}
0\ 0\ \ldots\ 0\ 1 \\
0\ 0\ \ldots\ 1\ 0 \\
\vdots \\
1\ 1\ \ldots\ 1\ 1 \\
\hline
0\ 0\ \ldots\ 0\ 1^2 \\
0\ 0\ \ldots\ 1^2\ 0 \\
0\ 0\ \ldots\ 1^2\ 1^2 \\
\hline
\text{etc.}
\end{array}
$$

where 1^i represents that the entity in question occurs i times. More complex hierarchical structures where the topology of the system itself is considered (as in model 2) can also be furnished (Bastin *et al.*, 1979).

In any of the above models a choice of a suitable grammer can be used to furnish more complex strings (sentences) starting from the primitives of the basic $2^k - 1$ strings (words). Such a procedure would then define the potentially infinite variety in matter. Each complex super-object would be a grammatically correct statement in this language. This grammar can, however, be found only by means of axioms that must be empirically validated. It is possible, furthermore, that for systems where observations can only be indirect, as in sub-atomic physics, theories based on any of the three models could be shown to be isomorphic and thus equally effective.

Uncertainty Relations and Information

State preparation and measurement in quantum mechanics have a statistical basis. In fact no matter what care is taken there is a limitation to the accuracy with which the state of an object can be prepared. As this limitation is probabilistic it is natural to investigate the significance of information (entropy) associated with it.

It is also clear that if we assumed the space variable to be continuous then the entropy associated with an object in a quantum mechanical description would be infinite. This would imply that the energy to be expended in order to extract all the information about the object would also be infinite. This is clearly unsatisfactory; we, therefore, postulate that the fineness of the measurement of two conjugate variables y_1 and y_2 satisfies the constraint $y_1 y_2 = \hbar/2$. This postulate has the following pleasing aspects. First, the information associated with an object is finite and one may extend the correspondence principle to equate information in the quantum numbers with the uncertainty inherent in a quantum description. Second, the spacetime variables are defined in terms of discrete lattices, which also implies that products like px (angular momentum) should vary in steps of $\hbar/2$.

As we are interested in a statistical average of the state function, the non-statistical variables shall be of no interest to us. In physical terms this implies that entropy does not depend on variables that are deterministic, because one can, in principle, design experiments to evaluate them. However, once again, for reasons that are related to a continuous variable having infinite entropy, all variables like spin, charge and mass should also vary in steps.

For convenience of our analysis we take the state function in one dimension only. The state function may be represented as $\psi(x, p)$. That x and p are random variables is evident by the fact that the uncertainty principle relates to state preparation and the maximum uncertainty, for a fixed variance, will be defined when x and p have Gaussian distributions. The variance of x and p will also be defined by our discreteness postulate through relations of the type $\sigma_x \sigma_p = \hbar/2$.

The entropy will be calculated for $|\psi|^2$ and not ψ because it is the former that is of measurable significance besides being a probability density function. Other random variables associated with the object can be generated by appropriate (linear, measure-preserving) transformations on the random variables defined by the state function, and hence the entropy of these variables will be the same value. This follows from the fact that a hermitian transformation on a random variable does not change its entropy.

Let us now consider a probabilistic discrete model which follows when one replaces the expression for the entropy of the state function represented by H_0 as given below:

$$H_0 = -\Sigma P(x) \log P(x) \quad \text{...(18)}$$

by the modified expression

$$H_0 = -\int dx\, \omega(x) \log \omega(x) + \log \frac{1}{x_0} \quad \text{...(19)}$$

where $P(x) = \omega(x)\, x_0$. The logarithms are to the base 2 so that entropy of n bits can be directly interpreted as uncertainty between 2^n equally likely alternatives. Thus $\omega(x)$ is a probability density function and x_0 is the elemental distance. Note that x_0 can, a priori, be taken either as the fineness of measurement or as a fundamental attribute of space. We opt for the latter meaning, however, owing to the limitations of the measurement explanation. In fact, from the symmetry of spacetime:

$$x_0 = y_0 = z_0 = ct_0. \quad \text{...(20)}$$

Let us consider the state function $\psi(x)$ of the object to be such that it leads to maximum entropy. As shown in the previous section the probability density for such an unrestrained particle will be Gaussian*. Thus

$$\psi(x) = a \exp\,[-x^2/2\sigma_x^2] \quad \text{...(21)}$$

where a is an appropriate constant. The expression for $\omega(x)$ is therefore

$$\omega(x) = |\psi(x)|^2 = (2\pi\sigma_x^2)^{-1/2} \exp\,(-x^2/\sigma_x^2) \quad \text{...(22)}$$

If a measurement is made of the mometum, the probabitity of a result p is given in terms of the overlap integral $\phi(p)$:

$$\phi(p) = \int_{-\infty}^{\infty} \exp\,(-ipx/\hbar)\, \psi(x)\, dx, \quad \text{...(23)}$$

and the probability of momentum p is

$$\omega(p) = |\phi(p)|^2. \quad \text{...(24)}$$

Calculating (14) we obtain for $\phi(p)$.

$$\phi(p) = b \exp\,[-p^2\sigma_x^2/2\hbar^2]. \quad \text{...(25)}$$

Defining σ_p by the relation

$$\sigma_x\sigma_p = \hbar/2$$

*We have made an approximation that can be avoided. One can use the probability distribution function that maximizes entropy for a discrete space instead of the Gaussian function of eqn. (7). The change is so small that there is no significant effect on the nature of our conclusions, however.

we have

$$\phi(p) = b \exp[-p^2/8\sigma_p^2] \quad ...(26)$$

and

$$\omega(p) = (4\pi\sigma_p^2) \exp[-p^2/4\sigma_p^2] \quad ...(27)$$

We assume that the momentum variable also has an elemental value which equals p_0, such that $x_0 p_0 = \hbar/2$.

We now calculate the uncertainty (negative information) associated with the distribution of the x and p variables and take its mean to represent the particle information H_0.

$$H_0 = \tfrac{1}{2}[\text{entropy of } x + \text{entropy of } p] \quad ...(28)$$

Using (8) and (10) in (13) and (18) this results in

$$H_0 = \tfrac{1}{2}\left[\tfrac{1}{2}\log \pi e\sigma_x^2 + \tfrac{1}{2}\log 4\pi e\sigma_p^2 + \log \frac{1}{x_0} + \log \frac{1}{p_0}\right]$$

$$= \tfrac{1}{2}[\log 2\pi e + \log \sigma_x\sigma_p - \log x_0 p_0] \quad ...(29)$$

Since $\sigma_x\sigma_p = x_0 p_0 = \hbar/2$, we finally obtain

$$H_0 = \tfrac{1}{2}\log 2\pi e \quad ...(30)$$

as the entropy due to the uncertainty along the x coordinate. The fact that the uncertainty due to the cell size cancels out with a corresponding term owing to the Gaussian probability of the variable may appear unsatisfactory at first sight, yet such a connection is inherent in mathematical information theory. Whether a continuous distribution can be used without approximation when the variance and the discrete variable separation are of the same order is a question that needs to be explored further, however.

The entropy, considering all the three spatial directions, is now

$$H_0 = \tfrac{3}{2}\log 2\pi e \simeq 6.38$$

It may also be argued that a particle with spin 1/2 has another degree of freedom, since the spin polarization is unknown. This would cause the total degrees of freedom to increase to about 7.38 for a spin 1/2 particle.

If the time variable is taken on the same footing as the space variables, as one may using special theory of relativity, then this variable would also have the same uncertainty as each space variable and therefore H_0 shall equal $2\log 2\pi e \simeq 8.4$. And as spin is a consequence of reconciling relativity and quantum theory, one may not consider it separately anymore.

As the quantum-mechanical object has an entropy owing to the uncertainty relation, it follows on application of the correspondence principle that at most 8.4 (or about 9) bits, or quantum numbers, will have to be associated with a particle in a quantum-mechanical description. These may be identified as the internal symmetry attributes that remain invariant in all interactions.

To build a complete theory along the above lines, however, requires clarification of several issues. We shall enumerate some of them.

1. Is charge a quantum effect? If it is indeed so then the eight quantum numbers due to uncertainty include it.

2. Is it valid to include quantum numbers as being due to spin polarization uncertainty?

3. What is the significance of the space-time lattice with respect to the reality criterion? Does the lattice exist prior to measurement or is it a result of the interaction of the observer with the object?

Concluding Remarks

We have made an exploratory study of the significance of the fundamental uncertainty inherent in a quantum description for object classification. We have argued that it is, in principle, possible to relate the number of symmetries in quantum interactions to uncertainty if the correspondence principle were assumed to apply to information as well. The most significant result of the study is that the object classification numbers (quantum numbers) turn out to be finite, which indicates that as energies in the elementary particle experiments go up, we should not keep on getting new symmetries. It is also significant that the number of object attributes obtained corresponds to the currently accepted value. The calculation in section 5 is based on several assumptions, however, and is meant to merely illustrate the central thesis of the paper, viz., the principle embodied in eqn. (1).

The programme outlined in this paper can be extended along several directions. Some of these are :

1. Determine if the measure of information used in this paper is the most appropriate.

2. Explore further the object structures using the different models of section 4.

3. Make the calculation of section 5 rigorous.

Acknowledgement

The author wishes to thank Professor C W Kilmister of the University of London for his comments on an earlier version of this paper.

References

Bastin, T, Noyes, H P, Amson, J and Kilmister, C W (1979) "On the physical interpretation and the mathematical structure of the combinatorial hierarchy". *Int. J. theor. Phys.*, **18**, 445.

Kak, S C (1976) "On quantum numbers and uncertainty". *Nuovo Cimento*, **34B**.

——— (1977) "On quantum numbers and uncertainty II". *Nuovo Cimento*, **41B**.

——— (1982) "On information in a quantum description". *Tech. Report EE* 603, *LSU* June 1981. Also presented at the IEEE International Symposium on Information Theory, Les Arcs, France.