ON SOME NEW CONCEPTS IN PROBABILITY THEORY

R. Jagannathan
The Institute of Mathematical Sciences
Madras - 600113, India

I shall first discuss the new concepts 'disparity', 'activity' and 'duality' introduced recently in the study of some stochastic processes by Alladi Ramakrishnan (cf. [1-5] and references therein) and in conclusion advance arguments based on our work [6-8] to indicate the interesting possibility of a fundamental role for classical probability in the quantum theory of subnuclear phenomena.

For a physical system with n possible states let $\pi(j;t)$ denote the probability that it is in the state j at time t and $\pi(j/k;t)$ denote the probability that the system is in the state j at time t given that it is in the stake k at time t = 0. Obviously

$$\sum_j \pi(j;t) = 1, \quad \sum_j \pi(j/k;t) = 1 \tag{1}$$

Then the Chapman-Kolmogorov equation

$$\pi(j;t_2) = \sum_k \pi(j/k;t_2-t_1)\,\pi(k;t_1) \tag{2}$$

or equivalently the matrix equation

$$\vec{\pi}(t_2) = [\pi(t_2-t_1)]\,\vec{\pi}(t_1) \tag{3}$$

with $\pi(j;t)$ as the j-th element of the probability vector $\vec{\pi}(t)$ and $\pi(j/k;t)$ as the jk-th element of the matrix $[\pi(t)]$ representing the probability transition operator, gives the temporal evolution of the system.

Now if,

$$[\pi(t_2-t_1)] \approx I + R\,(t_2-t_1) \quad \text{for } t_2-t_1 = \Delta \approx 0 \tag{4}$$

where I is the identity matrix and R defines the fundamental transition probability matrix such that

$$\pi(j/k;t) \rightarrow R(j/k)\Delta \quad \text{for } j \neq k \text{ as } t \rightarrow \Delta \approx 0$$

$$R(j/k) \geqslant 0, \quad \text{for } j \neq k, \quad R(k/k) = -\sum_{j\neq k} R(j/k) \tag{5}$$

then the evolution equation for the process, (2) or (3), can be written as

$$\frac{d\vec{\pi}(t)}{dt} = R\vec{\pi}(t) \tag{6}$$

with the formal solution

$$\vec{\pi}(t) = e^{Rt}\vec{\pi}(0) \tag{7}$$

i.e. in (3) $[\pi(t)] = e^{Rt}$. In terms of matrix elements (6) becomes

$$\frac{d\pi(j;t)}{dt} = \sum_{k\neq j} R(j/k)\,\pi(k;t) - \left[\sum_{k\neq j} R(k/j)\right]\pi(j;t) \tag{8}$$

When the process evolves according to (8) it is obvious that for given j, t and Δt, $\pi(j;t+\Delta t) - \pi(j;t)$ can assume any value between -1 and 1, <0, >0 or $= 0$. Then, the following question arises. Is there any sense of 'direction' in the evolution of the stochastic process governed by the above equations or in other words does there exist any quantity associated with the above process which strictly decreases (or increases) with time? Recently this question has been answered by Ramakrishnan in [1-3] through the introduction of the new concepts 'disparity' and 'activity' as follows.

Let two identical physical systems evolve according to (8) with the same fundamental R matrix but starting from two different initial distributions $\vec{\pi}_1(0)$ and $\vec{\pi}_2(0)$ at t = 0. According to (7) at time t

$$\vec{\pi}_1(t) = e^{Rt}\,\vec{\pi}_1(0), \qquad \vec{\pi}_2(t) = e^{Rt}\,\vec{\pi}_2(0) \tag{9}$$

Now let the 'disparity' between the two distributions at time t be defined by

$$D(t) = \sum_j |\pi_2(j;t) = \pi_1(j;t)| \tag{10}$$

which is always nonnegative with an upper bound 2. Then Ramakrishnan's theorem states that the disparity D(t) as defined by (10) decreases with time t.

Proof of the above theorem can be given easily as follows. Let at a given time t

$$D_+(t) = \sum_{\substack{\text{sum over all}\\ j_+ \text{ such that}\\ \pi_2(j_+;t)\geqslant\pi_1(j_+;t)}} \left[\pi_2(j_+;t) - \pi_1(j_+;t)\right] > 0 \tag{11}$$

and

$$D_-(t) = \sum_{\substack{\text{sum over all}\\ j_+ \text{ such that}\\ \pi_2(j_-;t)<\pi_1(j_-;t)}} \left[\pi_2(j_-;t) - \pi_1(j_-;t)\right] < 0 \tag{12}$$

It is to be noted that for any given t the sets (j_+) and (j_-) do not have any common element. Hence obviously for all t

$$D_+(t) + D_-(t) = 0 \tag{13}$$

so that

$$D(t) = D_+(t) - D_-(t) = 2D_+(t) = 2\,|D_-(t)| \tag{14}$$

and as long as $D(t) > 0$ neither of the sets (j_+) and (j_-) can be empty. Thus it is enough if it is proved that $D_+(t)$ decreases strictly with time. This can be shown as follows. Using (6) and the fact $\sum_{\text{all } j} R(k/k) = 0$ lead to

$$\frac{dD_+(t)}{dt} = \sum_{j_+} \Big\{ \sum_{j'_+} R(j_+/j'_+)[\pi_2(j'_+;t) - \pi_1(j'_+;t)]$$

$$+ \sum_{j'_-} R(j_+/j'_-)[\pi_2(j'_-;t) - \pi_1(j'_-;t)] \Big\}$$

$$= \sum_{j'_+} [\sum_{j_+} R(j_+/j'_+)][\pi_2(j'_+;t) - \pi_1(j'_+;t)]$$

$$+ \sum_{j'_-} [\sum_{j_+} R(j_+/j'_-)][\pi_2(j'_-;t) - \pi_1(j'_-;t)]$$

$$= \sum_{j'_+} [-\sum_{j_-} R(j_-/j'_+)][\pi_2(j'_+;t) - \pi_1(j'_+;t)]$$

$$+ \sum_{j'_-} [\sum_{j_+} R(j_+/j'_-)][\pi_2(j'_-;t) - \pi_1(j'_-;t)] \tag{15}$$

Now using the definition of the elements of R in (5) and the classification of j's into (j_+) and (j_-) it is seen that

$$\frac{dD_+(t)}{dt} < 0 \tag{16}$$

which proves Ramakrishnan's theorem.

When the states of the system are to be labelled by a continuous parameter x instead of the discrete index j, the equations (1), (2),, (5) and (8) become

$$\int dx\, \pi(x;t) = 1, \quad \int dx\, \pi(x/x';t) = 1 \tag{17}$$

$$\pi(x;t_2) = \int dx'\, \pi(x/x';t_2-t_1)\, \pi(x';t_1) \tag{18}$$

$$\pi(x/x';t) \longrightarrow R(x/x')\Delta \qquad \text{for } x \neq x' \quad \text{as } t \longrightarrow \Delta \approx 0 \tag{19}$$

and

$$\frac{d\pi(x;t)}{dt} = \int dx' R(x/x')\, \pi(x';t) - \pi(x;t)\Big(\int dx' R(x'/x)\Big) \tag{20}$$

Consequently the 'disparity' between the distributions of two identical physical systems governed by the same evolution equation (20) becomes

$$D(t) = \int dx\, |\pi_2(x;t) - \pi_1(x;t)| \tag{21}$$

where $\pi_2(t)$ and $\pi_1(t)$ are the distributions reached at time t from the initial distributions $\pi_2(0)$ and $\pi_1(0)$ respectively at $t = 0$ according to

$$\pi_1(x;t) = \int dx' \, \pi(x/x';t) \, \pi_1(x';0)$$
$$\pi_2(x;t) = \int dx' \, \pi(x/x';t) \, \pi_2(x';0) \qquad (22)$$

Now if $\pi_2(x;0)$ is taken as $\pi_1(x;\Delta t)$ then $\pi_2(x;t) = \pi_1(x;t+\Delta t)$. Then the corresponding expression for $D(t) = \int dx \, | \pi_1(x;t+\Delta t) - \pi_1(x;t) |$ can be associated with system 1 itself meaningfully. Thus Ramakrishnan introduced in [1] the quantity

$$A(t,\Delta t) = \begin{cases} \sum_j | \pi(j;t+\Delta t) - \pi(j;t) | & \text{for discrete state-space} \\ \int dx \, | \pi(x;t+\Delta t) - \pi(x;t) | & \text{for continuous state-space} \end{cases} \qquad (23)$$

called the 'activity' of the given system in the time interval $(t,t+\Delta t)$. Looked at as a special case of 'disparity' it is clear that the probability 'activity' of the system also decreases with t for any chosen time interval Δt.

Having thus explained the meaning of the 'sense of direction' in the case of the above type of Markovian stochastic processes through the introduction of the concepts 'activity' and 'disparity' in [1-3] Ramakrishnan analyses in [4] the basic problem of inverse probability in such cases which involves the question : Is it possible to make statements about the past given the present state of the system? He shows that a 'duality' is inherent in this respect in the mathematical theory of stochastic processes as follows.

In conventional treatments (2) or (3) is usually written only for $t_2 > t_1$. Ramakrishnan has demonstrated as early as 1955 in [5] that the equation

$$\vec{\pi}(t_2) = [\pi(t_2-t_1)] \, \vec{\pi}(t_1) = e^{R(t_2-t_1)} \, \vec{\pi}(t_1) \qquad (24)$$

is valid even $t_2 < t_1$. In such a case $[\pi(t_2-t_1)]$ can be used meaningfully as an 'operator' to relate $\vec{\pi}(t_2)$ to $\vec{\pi}(t_1)$ for $t_2 < t_1$ though $[\pi(t_2-t_1)]$ will have negative elements without any probability significance. However $\vec{\pi}(t_2)$ can be obtained from $\vec{\pi}(t_1)$ for $t_2 < t_1$ in a different manner also. The joint probability that the system is in state j at time t_2 and in state k at time t_1 for $t_1 > t_2$ can be written in two ways as

$$\pi(j,t_2;k,t_1) = \pi(j;t_2) \, \pi(k/j;t_1-t_2) = \pi(k;t_1) \, P(j/k;t_2-t_1) \qquad (25)$$

where the 'inverse probability'

$$P(j/k;t_2-t_1) = \pi(j;t_2) \, \pi(k/j;t_1-t_2) \, \pi(k;t_1)^{-1} \qquad (26)$$

is nonnegative by definition, of course assuming that for any k $\pi(k;t_1) > 0$. Then $\vec{\pi}(t_2)$ can be obtained from $\vec{\pi}(t_1)$ for $t_2 < t_1$ through the relation

$$\vec{\pi}(t_2) = [P(t_2-t_1)] \, \vec{\pi}(t_1) \qquad (27)$$

also instead of through (24) where the matrix $[P(t_2-t_1)]$ with $P(j/k;t_2-t_1)$ is the typical element can be recognised to be given by

$$[P(t_2-t_1)] = D(t_2) \, [\pi(t_1-t_2)]^T \, D^{-1}(t_1) \, , \qquad t_2 < t_1$$

$$[\pi(t_1-t_2)]^T_{jk} = \pi(k/j;t_1-t_2) \quad , \quad D(t)_{jk} = \pi(j;t)\,\delta_{jk} \quad . \tag{28}$$

Thus in the case of $t_2 < t_1$ Ramakrishnan calls $[\pi(t_2-t_1)]$ as the 'analytic inverse' and $P(t_2-t_1)$ as the 'Bayes inverse' and notes that $\vec{\pi}(t_2)$ can be related to $\vec{\pi}(t_1)$ in a 'dual' manner either through $[\pi(t_2-t_1)]$ using (24) or through $[P(t_2-t_1)]$ using (27). This 'duality' in relating $\vec{\pi}(t_2)$ to $\vec{\pi}(t_1)$ is not surprising since more than one matrix can transform one vector to another. However the 'analytic inverse' is fundamental since it is independent of $\vec{\pi}(t_1)$ and $\vec{\pi}(t_2)$ whereas the 'Bayes inverse' transforming $\vec{\pi}(t_1)$ to $\vec{\pi}(t_2)$ depends on $\vec{\pi}(t_1)$ and $\vec{\pi}(t_2)$ (cf. [4] for more details).

I shall now conclude by mentioning an interesting possibility in the quantum theory of subnuclear phenomena. (cf. [6-8] for more details). So far all known particles are described by wavefunctions $\psi(q)$ with $|\psi(q)|^2 dq$ representing the probability of locating the particle somewhere in the position interval (q, q+dq) and it is assumed that q can take all values continuously from $-\infty$ to ∞. This assumption is implicit in prescribing the position operator $\hat{q}$ and the conjugate momentum operator $\hat{p}$ to obey the Heisenberg relation

$$[\hat{q}, \hat{p}] = i\hbar \tag{29}$$

which can have only infinite dimensional realizations as is well known. As pointed out by Weyl in [9] the Heisenberg relation (29) can be written equivalently in the form

$$\exp(i\alpha\hat{p}/\hbar)\exp(i\beta\hat{q}/\hbar) = \exp(i\alpha\beta/\hbar)\exp(i\beta\hat{q}/\hbar)\exp(i\alpha\hat{p}/\hbar) \quad , \; -\infty \leqslant \alpha\,,\ \beta \leqslant \infty \tag{30}$$

and hence $\hat{q}$ and $\hat{p}$ can be thought of as the generators of the continuous group of unitary operators

$$G_{\alpha\beta\gamma} = \exp(i\alpha\hat{p}/\hbar)\ \exp(i\beta\hat{q}/\hbar)\ \exp(i\gamma/\hbar) \quad , \quad -\infty \leqslant \alpha,\beta,\gamma \leqslant \infty$$

$$G_{\alpha\beta\gamma}G_{\alpha'\beta'\gamma'} = G_{\alpha+\alpha',\,\beta+\beta',\,\gamma+\gamma'-\alpha'\beta} \tag{31}$$

Thus Weyl associated the kinematical basis of Schrödinger-Heisenberg-Dirac quantum mechanics with the above continuous unitary group and expressed the prophetic hope : "..... the field of discrete groups offers many possibilities which we have not yet been able to realize in Nature ; perhaps these holes will be filled by applications to nuclear physics".

A fundamental role for discrete analogues of the above unitary group G, as envisaged by Weyl, ofcourse with the natural requirement that the continuous group G underlying the current quantum mechanics of atomic and nuclear phenomena should be the asymptotic limit of such discrete groups (correspondence principle of Bohr), can arise at a deeper, say subnucleonic, level if in that realm the phase-space formed by all eigenvalues of q and p or equivalently α and β has a lattice structure given by, for instance,

$$\alpha = n\epsilon \;,\; \beta = m\eta \;,\; \epsilon\eta = 2\pi\hbar/(2J+1)\,,\; n,m = -J,-J+1,\ldots,-1,0,1,\ldots,J-1,J \tag{32}$$

where J is a positive integer. Then the new position operator Q and the momentum operator P can be looked upon as the 'generators' of the finite group G defined by

$$G_{nmk} = A^n B^m \omega^k, \quad AB = \omega BA, \quad A^{2J+1} = B^{2J+1} = I, \quad \omega^{2J+1} = 1$$

$$G_{nmk} G_{n'm'k'} = G_{(n+n'),(m+m'),(k+k'-n'm)}, \quad (x) = x.\mathrm{mod}(2J+1)$$

$$A = \exp(i\epsilon P/\hbar), \quad B = \exp(i\eta Q/\hbar), \quad \omega = \exp(i\epsilon\eta/\hbar) = \exp(i2\pi/(2J+1)) \qquad (33)$$

and with the faithful and unique irreducible representation

$$\langle n | Q | n' \rangle = n\epsilon\delta_{nn'}$$
$$\langle n | P | n' \rangle = \begin{cases} 0 & \text{for } n = n' \\ (\eta/2i)(-1)^{n-n'} \operatorname{cosec}[(n-n')\pi/(2J+1)] & \text{for } n \neq n' \end{cases}$$
$$n, n' = -J, -J+1, \ldots, -1, 0, 1, \ldots J-1, J \qquad (34)$$

(cf. [6-12] for details)

A formalism of quantum mechanics in which the usual infinite-dimensional Schrodinger-Heisenberg realizations of q and p are replaced by finite-dimensional matrices Q and P respectively should be, in my opinion, an ideal candidate for the quantum mechanics of quark phenomena if it is true that these subnucleonic constituents have permanent confinement or in other words have only a finite spectrum for position as observed in the rest frame of the composite nucleon. Such Weylian Finite-Dimensional Quantum Mechanics (WFDQM) is being developed by us (cf. [6-8]) to describe the quantum mechanics of quarks with reference to the rest frame of the nucleons, based on the early pioneering works of Weyl [9] and Schwinger [10] on the group structure of quantum kinematics and the recent works of Ramakrishnan and collaborators [11,12] on generalized Clifford algebras of the type defined in (33). Gudder and Naroditsky also have analysed in [13] similar forms of 'finite-dimensional quantum mechanics'. A form of quantum mechanics, analogous to our WFDQM, is also implicit in the work of Drell, Weinstein and Yankielowicz [14] on the lattice field theory of fermions wherein the derivative operation on the lattice is taken as the action of the matrix operator $iP/\hbar$ on the concerned vector defined on the lattice with the matrix P defined exactly as in (34) above.

So far in our work [6-8] quarks have been considered as belonging to a formalism of WFDQM based on some particular group G as defined in (33) and (34) with a fixed value for J. For any given J of course the Heisenberg uncertainty principle holds in the sense that any state of the quark can not belong to a single eigenpoint in the phase-space lattice since still the commutator [Q,P] is nonzero. But the question arises whether the phase-space lattice itself can have a sharply defined structure as described above. In other words, can the value of J be taken as a fixed integer for a given type of quark? This may not be so. Consequently J may have characteristic classical-type (in my opinion) of probability distributions so that the actual quantum mechanics of quarks may be a linear superposition of close-lying forms of Weylian quantum mechanics, in the sense described above, with classical probability weights. It should be interesting to explore this possibility further.

References

1. Alladi Ramakrishnan, 'A mathematical feature of evolutionary stochastic processes', Sonderdruck aus Methods of Operations Research 36, 239 (1980) (paper presented at the Oberwolfach Conference in November 1978)
2. Alladi Ramakrishnan, 'Approach to stationarity in stochastic processes', Proc. Tamil Nadu Acad. Sciences, 2, 189 (1979)
3. Alladi Ramakrishnan, 'A new concept in probability theory', J. Math. Anal. Appl. 83, 408 (1981)
4. Alladi Ramakrishnan, 'Duality in stochastic processes', J. Math. Anal. Appl. 84, 483 (1981)
5. Alladi Ramakrishnan, 'Inverse probability and evolutionary Markov stochastic processes', Proc. Indian Acad. Sciences A41, 145 (1955)
6. R. Jagannathan, T.S. Santhanam and R. Vasudevan, 'Finite-dimensional Quantum Mechanics of a particle', International J. Theor. Phys. 20, 755 (1981)
7. R. Jagannathan and T.S. Santhanam, 'Finite-dimensional quantum mechanics of a particle-II', International J. Theor. Phys. 21, 351 (1982)
8. R. Jagannathan, 'Finite-dimensional quantum mechanics of a particle-III:Weylian quantum mechanics of confined quarks' International J. Theor. Phys. 22, (1983) to appear
9. H. Weyl, Theory of Groups and Quantum Mechanics, Dover, New York (1932) Section 4.14
10. J. Schwinger, Quantum Kinematics and Dynamics, Benjamin, New York (1970) Sections 2.12 and 6.18
11. Alladi Ramakrishnan, L-Matrix Theory or Grammar of Dirac Matrices, Tata-Mc Graw Hill, New Delhi, (1972)
12. R. Jagannathan and N.R. Ranganathan, 'Generalized Clifford groups-I and II' Reports on Mathematical Physics, 5, 131 (1974), 7, 229 (1975)
13. S. Gudder and V. Naroditsky, 'Finite-dimensional quantum mechanics', International J. Theor. Phys. 20, 619 (1981)
14. S.D. Drell, M. Weinstein and S. Yankielowicz, 'Strong-coupling field theories II : Fermions and gauge fields on a lattice', Phys. Rev. D14, 1627 (1976)