This new many-body quantum formalism retains all correlations between particles in a correlation function, whereas the conventional formalism puts part of the correlation in the choice of symmetries or antisymmetries of the wave functions and part in the form of the interaction potential. The possibilities opened by this new format for the treatment of physical problems are now being explored.

In summary, the developments presented here give further support to our contention that much of the formal structure of physical theory can be deduced directly from the statistical properties of experimental data. Perhaps one of the most exciting possibilities opened up by this is that it indicates the way for applying the formal structures of physical theory to nonphysical problems as well.

PHYSICAL REVIEW D VOLUME 2, NUMBER 4 15 AUGUST 1970

Statistical Basis for Quantum Theory: The Embedding Formalism*

R. E. COLLINS
The Department of Physics, The University of Houston, Houston, Texas 77004
(Received 8 December 1969)

The formulation of quantum theory on the basis of the statistical definition of a duplicated experiment, as initiated in a recent paper, is used to reexamine the embedding formalism introduced in earlier papers. It is shown that the familiar relations between momentum and coordinate representations are given a statistical basis by the embedding theory. It is also shown that a complete quantum theory includes three fundamental sets of operators and that one of these seems to represent an operator formalism for thermodynamics.

INTRODUCTION

IN the development of any new way of looking at old ideas, one progresses in an irregular fashion and must continuously reexamine each step. A recent paper[1] presented a derivation of a generalized quantum formalism directly from the statistical definition of a duplicated experiment, but two earlier papers[2,3] presented derivations of some parts of the formalism on the same basis by a very different argument; these employed an embedding of one space in another.

These embedding arguments were rather crude and, in the case of the first paper[2] particularly, they now seem rather contrived and definitely improperly presented. After finding the more direct path from the statistical description of an experiment to a quantum formalism, as presented in the most recent paper,[1] the embedding formalism has been reexamined. This has yielded what seems to be a more coherent picture of the statistical basis of quantum theory. In particular, it provides a purely statistical formulation of representation theory as used in conventional quantum mechanics. However, this reexamination has also yielded a better understanding of the operator structure of the recent paper[1] as well, and points to an enlarged operator formalism for thermodynamics. This is presented here beginning with the statistical definition of a duplicated experiment.

* Supported in part by a Frederick Gardner Cottrell Grant in aid from the Research Corporation.
[1] R. E. Collins, Phys. Rev. D 1, 379 (1970).
[2] R. E. Collins, Phys. Rev. 183, 1081 (1969).
[3] R. E. Collins, preceding paper, Phys. Rev. D 2, 658 (1970).

DUPLICATED EXPERIMENTS, STATISTICS AND TIME

The results of a duplicated experiment can in principle be summarized as a spectrum of configurations of the variables determined by measurement in the experiment (the observables) together with the probability for each such configuration (frequency of occurrence). The spectrum of values for any observable is in general determined by values of parameters which must be fixed in order to duplicate the conditions of the experiment; also the probability for the occurrence of a configuration is conditioned by these same parameters.

Thus let

$$r=(x,y,z,\dots) \tag{1}$$

be the set of all observables whose values are determined in each duplication and suppose r contains K independent variables. Then each observable can be viewed as a function of the elements of a vector of K independent components,

$$\lambda=(\lambda_1,\lambda_2,\dots,\lambda_K), \tag{2}$$

and a distinct configuration is labeled by a configuration of λ. Thus we represent any observable as

$$x(\lambda,\eta,t)=x_\lambda(\eta,t), \tag{3}$$

where $\eta_1, \eta_2, \dots, \eta_M$ and t label values of conditioning parameters that must be the same in every duplication of the experiment. Here $\eta_1, \eta_2, \dots, \eta_M$, denoted by η, are those parameters that are individually identified and fixed in each duplication, while t is introduced as a single parameter label for all other parameters which

condition the experiment, but which are not identified and controlled by the experimenter. Thus t, "the time," is a one-parameter label for the configuration of that part of the "external universe" which conditions the outcome of the experiment.

We tacitly assume the distinct configurations of the observables in the experiment to form a *denumerable* set and then denote the probability for a configuration as

$$\pi=\pi(\lambda|\eta,t), \quad \lambda\in D$$

where

$$\pi\geqslant 0 \tag{4}$$

and

$$\sum_{\lambda\in D}\pi(\lambda|\eta,t)=1 \tag{5}$$

is an *identity* valid for all configurations of η and t. However, we may also admit a continuum of distinct results, in which case this sum becomes an integral.

This description of a duplicated experiment admits, as a particular special case, a completely *deterministic* outcome for each duplication of the experiment, namely,

$$\begin{aligned}\pi(\lambda|\eta,t)&=1, \quad \lambda=\lambda'(\eta,t)\\ &=0, \quad \lambda\neq\lambda'(\eta,t),\end{aligned} \tag{6}$$

but in our general discussion we will assume a non-deterministic experiment. Of course, if we considered a continuum of distinct results as possible, we would have a Dirac δ function defining a deterministic form for the distribution function.

Here the role of time indicates that if our preparation of the experiment includes identification and fixing of *all* parameters that condition the outcome of the experiment, then the spectral functions, such as $x(\lambda,\eta)$, and the probability $\pi(\lambda|\eta)$, will *not* depend upon "time." That is, there are then no unidentified, or unfixed, parameters to be labeled by t that condition the experiment. Quite obviously, whether the experiment is time dependent depends upon our *choice* of observables as well as our ingenuity in preparing the experiment.

MATRIX FORMALISM

Since the probability π, now to be referred to as the *state probability*, is non-negative, we can write

$$\pi(\lambda|\eta,t)=c_\lambda^*(\eta,t)c_\lambda(\eta,t), \tag{7}$$

where $c_\lambda(\eta,t)$ is a complex function with complex conjugate $c_\lambda^*(\eta,t)$; this determines $c_\lambda(\eta,t)$ to within an *arbitrary phase*. Then, as in our earlier paper,[1] we see Eq. (5) as the matrix form

$$C^\dagger C=1=\sum_{\lambda\in D}c_\lambda^*(\eta,t)c_\lambda(\eta,t), \tag{8}$$

where C is a column matrix with elements $c_\lambda(\eta,t)$. Also, the expectation value of any observable appears as

$$\langle x\rangle=\sum_{\lambda\in D}x_\lambda(\eta,t)c_\lambda^*(\eta,t)c_\lambda(\eta,t)=C^\dagger XC, \tag{9}$$

where X is the real diagonal matrix with elements $x_\lambda(\eta,t)$.

Then, since any other observable, such as y, for example, has a corresponding real diagonal matrix representation, and diagonal matrices commute, we have

$$XY-YX=0 \tag{10}$$

for any two observables in the experiment. Of course, this commutation is invariant under all unitary transformations, as are the matrix forms in Eqs. (8) and (9), but these real diagonal matrices then become Hermitian matrices. Thus the observables in the experiment are represented by a collection of commuting Hermitian matrices, just as pointed out in our earlier paper.[1]

Our earlier paper[1] presented this matrix formalism and then proceeded to examine how one might describe changes in these matrices as the parameters η_j, $j=1, 2, \ldots, K$ and t are altered, but here our primary interest is in the relation of this matrix structure to an embedding formalism.[2,3] However, we point out that if t is a *continuous* parameter, then we may differentiate Eqs. (8) and (9) with respect to t to obtain

$$C^\dagger\frac{\partial C}{\partial t}+\frac{\partial C^\dagger}{\partial t}C=0 \tag{11}$$

and

$$\frac{\partial\langle x\rangle}{\partial t}=\frac{\partial C^\dagger}{\partial t}XC+C^\dagger X\frac{\partial C}{\partial t}+C^\dagger\frac{\partial X}{\partial t}C, \tag{12}$$

provided that these derivatives exist. Then, if we *assume* the existence of a matrix H such that

$$\partial C/\partial t=iHC, \tag{13}$$

we find that Eq. (11) *can* be satisfied if H is Hermitian, while Eq. (12) becomes

$$\frac{\partial\langle x\rangle}{\partial t}=C^\dagger\left[\frac{\partial X}{\partial t}-i(H^\dagger X-XH)\right]C. \tag{14}$$

This defines an operator $\delta/\delta t$ acting on X given by

$$\frac{\delta X}{\delta t}=\frac{\partial X}{\partial t}-i(H^\dagger X-XH), \tag{15}$$

and a similar equation exists for each observable, i.e., X replaced by Y, for example. This looks much like the structure of ordinary quantum theory, but we point out that Eq. (11), with the derivatives expressed in terms of H using Eq. (13), *can* be satisfied for a *non-Hermitian* matrix H.[4] Another alternative is explored at length in our earlier paper[1]; here, we merely wish to indicate that the kinematics of quantum theory arises from an examination of the evolution of the matrices C, X, etc.

[4] That is, the matrix $(H-H^\dagger)$ operating on the vector C produces a new matrix (vector) $(H-H^\dagger)C$ that is orthogonal to C, $C^\dagger(H-H^\dagger)C=0$. This possibility is explored at length in a paper in preparation.

VECTOR FORMALISM AND REPRESENTATION THEORY

The vector λ was introduced in Eq. (2) to label the distinct configurations of all observables and is defined to have the set D of distinct configurations. This set D is *not* conditioned by the η_j and t, and is taken to be the direct product set of the set of all integers, that is,

$$D=I_1\otimes I_2\otimes\cdots\otimes I_K, \tag{16}$$

where I is the set of all integers. If some or all of the observables in the experiment have a continuum for the range of their values, then some or all of the I_i here are replaced by R, the set of all real numbers. Here the notation $\lambda\in D$ implies $\lambda_i\in I_i$ for $i=1, 2, \ldots, K$.

If for some configuration of the η_j and t we must describe a duplicated experiment in which only a *finite* set of distinct results are actually observed, then $\pi(\lambda|\eta,t)$ has the value zero for all $\lambda\in D$ except for some subset $D_0\subset D$. However, we allow the spectral functions $x(\lambda,\eta,t)$, $y(\lambda,\eta,t)$, ... to be defined over the entire domain D.

Then in the matrix formalism all matrices are infinite dimensional and, as carried through in our previous paper,[1] can be represented on an infinite-dimensional abstract vector space which must be a Hilbert space. This is a separable Hilbert space, having both discrete and continuous sets of basis vectors.

Thus, since the $c_\lambda(\eta,t)$ are elements of a square summable sequence, as in Eq. (8), they have a representation on a particular set of basis vectors, $|\lambda\rangle$, $\lambda\in D$, of this vector space as

$$\langle\lambda|\eta,t\rangle=c_\lambda(\eta,t), \tag{17}$$

and are the representatives on this basis of the vector

$$|\eta,t\rangle=\sum_{\lambda\in D}\langle\lambda|\eta,t\rangle|\lambda\rangle, \tag{18}$$

which, according to Eq. (8), has unit norm

$$\langle\eta,t|\eta,t\rangle=1. \tag{19}$$

This representation on a discrete basis is to be identified as the "momentum representation" in quantum mechanics.

Furthermore each observable, for example x, is represented by an abstract operator, say $\mathbf{x}(\eta,t)$, acting in this space, and the particular basis vectors $|\lambda\rangle$ must be simultaneous eigenvectors of *all* the compatible observables in the experiment. That is, for example,

$$\mathbf{x}(\eta,t)|\lambda\rangle=x_\lambda(\eta,t)|\lambda\rangle. \tag{20}$$

Then matrix representatives such as X, for example, with elements

$$X_{\lambda\lambda'}=\langle\lambda|\mathbf{x}|\lambda'\rangle, \tag{21}$$

will be *diagonal* on this particular basis.

But the separable Hilbert space also possesses continuous sets of basis vectors such as, say, $|\xi\rangle$, $\xi\in G$, where $\xi=(\xi_1,\xi_2,\ldots,\xi_K)$ is a K-dimensional vector like λ but has a continuous domain, say

$$G\subset R_1\otimes R_2\otimes\cdots\otimes R_K. \tag{22}$$

Thus, whereas a discrete basis has the orthonormal and completeness properties

$$\begin{aligned}\langle\lambda|\lambda'\rangle=\delta_{\lambda\lambda'}=1, &\quad \lambda=\lambda'\\ =0, &\quad \lambda\neq\lambda'\end{aligned} \tag{23}$$

and

$$\sum_{\lambda\in D}|\lambda\rangle\langle\lambda|=1, \tag{24}$$

the continuous basis has

$$\begin{aligned}\langle\xi|\xi'\rangle=\delta(\xi-\xi')=\infty, &\quad \xi=\xi'\\ =0, &\quad \xi\neq\xi'\end{aligned} \tag{25}$$

and

$$\int_G|\xi\rangle\langle\xi|d^K\xi=1. \tag{26}$$

Thus the vector $|\eta,t\rangle$ may be given a "coordinate representation" as

$$\langle\xi|\eta,t\rangle=\sum_{\lambda\in D}\langle\lambda|\eta,t\rangle\langle\xi|\lambda\rangle, \tag{27}$$

or

$$\psi(\xi,\eta,t)=\sum_{\lambda\in D}c_\lambda^*(\eta,t)\psi_\lambda(\xi), \tag{28}$$

and in the usual language of quantum theory we call $\psi(\xi,\eta,t)$ the wave function of the system and interpret $\psi^*\psi$ as the probability density for the vector ξ with the values of the η_j and t as contingent events. Similarly, we call $\psi_\lambda^*\psi_\lambda$ the probability density for ξ contingent upon the system "being in the state" represented by λ, as well as particular values of the η_j and t being given. Furthermore, we call $c_\lambda^*c_\lambda$ the probability for the state λ.

While these interpretations of quantities formed by introducing a coordinate representation for the state vector $|\eta,t\rangle$ have proved to be extremely successful in the application of quantum mechanics to physical problems, they have not been given a clear statistical basis. Now we can show that the embedding formalism introduced in our earlier papers[2,3] provides such a statistical basis.

EMBEDDING FORMALISM

Given any conditional probability $\pi(\lambda|\eta,t)$ with λ on some domain D (not necessarily that defined above), we can always introduce another set of random variables

$$\xi=(\xi_1,\xi_2,\ldots,\xi_N), \quad \xi\in G \tag{29}$$

with

$$N\geqslant K, \tag{30}$$

such that with functions

$$\lambda_i=\lambda_i(\xi), \quad i=1, 2, \ldots, K \tag{31}$$

every function of the λ_i, such as $x(\lambda,\eta,t)$, for example, becomes a function of the ξ_j. Furthermore, with $N \geqslant K$, we are assured that the λ_i may remain as functionally independent quantities, but unless $N=K$ holds, these relations have no inverse.

Furthermore, we can define a probability density $\rho(\xi|\eta,t)$ for the ξ_j:

$$dP(\xi|\eta,t)=\rho(\xi|\eta,t)d^N\xi, \quad \xi\in G \tag{32}$$

with

$$\int_G \rho(\xi|\eta,t)d^N\xi=1, \tag{33}$$

such that

$$\pi(\lambda|\eta,t)=\int_{G_\lambda}\rho(\xi|\eta,t)d^N\xi, \tag{34}$$

where

$$\xi\in G_\lambda\subset G$$

and a G_λ is defined for each configuration of λ by

$$\lambda_i(\xi)=\lambda_i, \quad i=1, 2, \ldots, K. \tag{35}$$

This is an *embedding* of the λ_i in the space of the ξ_j.

It is evident that the subsets $G_\lambda\subset G$ are disjoint,

$$G_\lambda\bigcap G_{\lambda'}=\varnothing, \quad \lambda\neq\lambda' \tag{36}$$

where $\varnothing$ is the null set, and also, in view of Eqs. (5) and (33), that we must have

$$\int_G d^N\xi=\sum_{\lambda\in D}\int_{G_\lambda} d^N\xi, \tag{37}$$

or

$$G=\bigcup_{\lambda\in D} G_\lambda, \tag{38}$$

this being the union of all these disjoint sets.

In terms of such an embedding, we find that with the notation

$$x(\lambda(\xi),\eta,t)=x'(\xi,\eta,t), \tag{39}$$

the expectation value of any function of the λ_i becomes

$$\langle x\rangle=\int_G x'(\xi,\eta,t)\rho(\xi|\eta,t)d^N\xi. \tag{40}$$

It should be quite evident that beyond the constraint $N\geqslant K$, the choice of embedding space is almost completely arbitrary; we only need to be assured that all integrals defining expectation values exist, i.e., every quantity defined in terms of the λ_i must also be properly represented in the embedding picture. In this regard, note that all observables such as $x(\lambda,\eta,t)$, for example, become sectionally continuous step functions $x(\lambda(\xi),\eta,t)$ in the new variables.

There is a particular representation of the embedding distribution $\rho(\xi|\eta,t)$ that can always be introduced, namely,

$$\rho(\xi|\eta,t)=\sum_{\lambda\in D}\pi_\lambda(\eta,t)\rho_\lambda'(\xi|\eta,t), \tag{41}$$

where the $\rho_\lambda'(\xi|\eta,t)$ are required to have the properties

$$\begin{aligned}\rho_\lambda'(\xi|\eta,t)&>0, \quad \xi\in G_\lambda\subset G\\ &=0, \quad \xi\notin G_\lambda\subset G\end{aligned} \tag{42}$$

and

$$\int_{G_\lambda}\rho_\lambda'(\xi|\eta,t)d^N\xi=1. \tag{43}$$

With this form for $\rho(\xi|\eta,t)$, Eq. (34) is an identity, and the expectation value for any quantity $f(\xi)$ will be given by

$$\langle f\rangle=\sum_{\lambda\in D}\pi_\lambda(\eta,t)\int_{G_\lambda} f(\xi)\rho_\lambda'(\xi|\eta,t)d^N\xi, \tag{44}$$

and we must obviously require that

$$f_\lambda=\int_{G_\lambda} f(\xi)\rho_\lambda'(\xi|\eta,t)d^N\xi \tag{45}$$

be the value of f corresponding to the distinct result labeled by λ. In fact, as pointed out above, all proper observables are represented by step functions having uniform value over each of the subsets $G_\lambda\subset G$. If f is such a quantity, it could be brought outside the integral here and Eq. (45) would then be an identity.

But now other features of this representation for $\rho(\xi|\eta,t)$ emerge. Note that in view of Eq. (42), we may write the integral in Eq. (45) as

$$f_\lambda=\int_G f(\xi)\rho_\lambda'(\xi|\eta,t)d^N\xi, \tag{46}$$

that is, the integral over *all* of G instead of just the subset $G_\lambda\subset G$. However, the subset $G_\lambda\subset G$ may not be simply connected and, in fact, may be "scattered" over G like "freckles," and therefore the function $\rho_\lambda'(\xi|\eta,t)$ may be an extremely discontinuous function over G. Hence this integral must be taken in the Lebesgue sense. But then we can introduce another, *continuous*, distribution function $\rho_\lambda(\xi|\eta,t)$ that is equal to $\rho_\lambda'(\xi|\eta,t)$ *almost everywhere* in G, that is,

$$\int_G f(\xi)[\rho_\lambda'(\xi|\eta,t)-\rho_\lambda(\xi|\eta,t)]d^N\xi=0, \tag{47}$$

where $f(\xi)$ is *any* sectionally continuous, measurable function over G, i.e., the points of discontinuity of $f(\xi)$ form a set of zero measure in G.[5]

Thus, if all observables are restricted to be sectionally continuous, measurable functions of the ξ_j on G, we see $\rho(\xi|\eta,t)$ in the form

$$\rho(\xi|\eta,t)=\sum_{\lambda\in D}\pi_\lambda(\eta,t)\rho_\lambda(\xi|\eta,t), \quad \xi\in G \tag{48}$$

[5] Jacob Korevaar, *Mathematical Methods* (Academic, New York, 1968), pp. 373 ff.

with the equality interpreted as being valid almost everywhere in G, meaning at all points $\xi\in G$ except possibly on some set of zero measure.

ANOTHER VIEWPOINT ON EMBEDDING

The embedding formalism was introduced above on the basis of a probability $\pi(\lambda|\eta,t)$ being given for λ on a discrete domain D, with observables as discrete-valued functions of λ, say, $x(\lambda,\eta,t)$; but another viewpoint is possible.[2]

Suppose we have given a probability distribution $\rho(\xi|\eta,t)$ with ξ defined on a continuous domain G, together with a collection of observables, such as $x'(\xi,\eta,t)$, which are continuous-valued functions, *not* step functions as considered above. Then suppose that we partition G into disjoint subsets by specifying a discrete set of values for a vector-valued function $\lambda(\xi)$, say, $\lambda\in D$. This is essentially the procedure of our earlier paper.[2] However, it should be quite evident here that we could then define $\pi(\lambda|\eta,t)$ as in Eq. (34), and in fact all of the embedding formalism would be unchanged except on one key point, namely, we would then have f_λ in Eq. (45) actually as a particular average value of $f(\xi)$ over the subset $G_\lambda\subset G$. This point was given prominence in our earlier paper,[2] where we noted that every "observed value" is an average value in some sense.

PROBABILITIES AND L^2 SPACE

Since $\rho(\xi|\eta,t)$ is non-negative, it may be represented in the form

$$\rho(\xi|\eta,t)=\psi^*(\xi,\eta,t)\psi(\xi,\eta,t)\,, \tag{49}$$

where ψ is a complex function with complex conjugate ψ^* and the phase of ψ is arbitrary. Then Eq. (33) appears as

$$\int_G|\psi(\xi,\eta,t)|^2d^N\xi=1 \tag{50}$$

and we see that $\psi(\xi,\eta,t)$ must be Lebesgue square integrable on G. Then $\psi(\xi,\eta,t)$ belongs to the space of Lebesgue square integrable functions on G; $\psi\in L^2$ on G. Thus with the orthonormal basis $U_\gamma(\xi)$, $\gamma\in\Gamma$ such that

$$\int_G U_\gamma^*(\xi)U_{\gamma'}(\xi)d^N\xi=1\,,\quad \gamma=\gamma'$$
$$=0\,,\quad \gamma\neq\gamma'\,, \tag{51}$$

we have

$$\psi(\xi,\eta,t)=\sum_{\gamma\in\Gamma}b_\gamma^*(\eta,t)U_\gamma(\xi)\,,\quad \xi\in G \tag{52}$$

which must be read as an integral if the basis is continuous rather than discrete. That is, if the domain G is unbounded in any one of the coordinates ξ_j, then the corresponding sum on the left may be replaced by an integral. We note that if the specification of the domain G involves any of the η_j or t, then the $U_\gamma(\xi)$ *must* have these as parameters, in view of Eq. (51).

With this form for $\psi(\xi,\eta,t)$, we have

$$\rho(\xi|\eta,t)=\sum_{\gamma\in\Gamma}\sum_{\gamma'\in\Gamma}b_\gamma^*b_{\gamma'}U_\gamma U_{\gamma'}^* \tag{53}$$

and the norm condition, Eq. (33), becomes, in view of the orthonormal property of the U_γ,

$$\sum_{\gamma\in\Gamma}b_\gamma^*b_\gamma=B^\dagger B=1\,, \tag{54}$$

where B is the column matrix with elements b_γ. Also, the expectation value of any observable such as x, for example, becomes

$$\langle x\rangle=B^\dagger X'B\,, \tag{55}$$

where the matrix X' is the Hermitian matrix with elements given by

$$X_{\gamma\gamma'}=\int_G x'(\xi,\eta,t)U_\gamma(\xi)U_{\gamma'}^*(\xi)d^N\xi\,, \tag{56}$$

i.e., this form results simply by inserting Eq. (52) into Eq. (40).

Since X' is Hermitian, it may be diagonalized by a unitary transformation, say by the unitary matrix T. Thus we insert $T^\dagger T$ into Eqs. (52), (54), and (55) to obtain

$$\psi(\xi,\eta,t)=B^\dagger U=B^\dagger T^\dagger TU=C'^\dagger\Psi$$
$$=\sum_{\gamma\in\Gamma}c_\gamma'^*\psi_\gamma(\xi,\eta,t)\,, \tag{57}$$

$$B^\dagger B=B^\dagger T^\dagger TB=C'^\dagger C'=1\,, \tag{58}$$

and also

$$\langle x\rangle=B^\dagger X'B=B^\dagger T^\dagger TX'T^\dagger TB=C'^\dagger XC'\,. \tag{59}$$

Here

$$C'=TB \tag{60}$$

and

$$\Psi=TU\,. \tag{61}$$

Also,

$$X=TX'T^\dagger \tag{62}$$

is a diagonal matrix. Here we have introduced the notation of column matrices U and Ψ with elements $U_\gamma(\xi)$ and $\psi_\gamma(\xi,\eta,t)$, respectively. Quite obviously, since T diagonalized X', which depends on the η_j and t, the $\psi_\gamma(\xi,\eta,t)$ generated as a new basis for the L^2 space from the $U_\gamma(\xi)$ by T *must* be functions of the η_j and t through the elements of T. Here the elements of X are

$$X_{\gamma\gamma'}=\delta_{\gamma\gamma'}\int_G x'(\xi,\eta,t)|\psi_\gamma(\xi,\eta,t)|^2d^N\xi=\delta_{\gamma\gamma'}\bar{x}_\gamma(\eta,t)\,. \tag{63}$$

Examining Eqs. (58) and (59) we see these as

$$\sum_{\gamma\in\Gamma} c_\gamma'^*(\eta,t)c_\gamma'(\eta,t)=1 \tag{64}$$

and

$$\sum_{\gamma\in\Gamma} \bar{x}_\gamma(\eta,t)c_\gamma'^*(\eta,t)c_\gamma'(\eta,t)=\langle x\rangle, \tag{65}$$

which have the same structure as Eqs. (8) and (9). Thus, if γ consists of the same number of independent components as λ, and the domain Γ is identical to the domain D, it would seem that we could choose the embedding such that $\gamma\equiv\lambda$, $\Gamma\equiv D$, $\bar{x}_\gamma=x_\lambda$, and the c's here are identical to those in Eqs. (8) and (9).

Certainly this is only possible if the set D of distinct configurations is at least countably infinite, because every basis for L^2 on G is at least countably infinite, but, as we pointed out above, we can always choose D to be an infinite set. On the other hand, the number of independent components in the label γ is always identical to the dimensionality N of the ξ space; therefore, in order to have the same number of components γ_i making up γ as we have λ_i making up λ, we must have

$$N=K. \tag{66}$$

That is, *the dimensionality of the ξ space must be the same as that of the λ space.*

Thus, now assuming this condition imposed so that $\gamma\equiv\lambda$, $\Gamma\equiv D$, we assume the $c_\gamma'(\eta,t)$ to differ from the $c_\lambda(\eta,t)$ by at most some phase function. Recall that in the construction of the matrix structure above, the $c_\lambda(\eta,t)$ each carried an unspecified phase; also, recall that here $\psi(\xi,\eta,t)$ carries an arbitrary phase and since, by the orthonormal property,

$$c_\lambda'^*(\eta,t)=\int_G \psi(\xi,\eta,t)\psi_\lambda^*(\xi,\eta,t)d^N\xi, \tag{67}$$

we see that the $c_\lambda'(\eta,t)$ must be partially arbitrary. The phases do not appear in Eqs. (8) and (9), and thus we may require that

$$c_\lambda'^*c_\lambda'=c_\lambda^*c_\lambda, \tag{68}$$

and then demand that

$$\bar{x}_\lambda(\eta,t)=x_\lambda(\eta,t). \tag{69}$$

But this last equation then suggests that

$$\rho_\lambda(\xi|\,\eta,t)=\psi_\lambda^*(\xi,\eta,t)\psi_\lambda(\xi,\eta,t), \tag{70}$$

with $\rho_\lambda(\xi|\,\eta,t)$ defined as in Eq. (48). However, writing $\psi^*\psi$ for $\rho(\xi|\,\eta,t)$, and employing the expansion of ψ in the ψ_λ as in Eq. (57), we have

$$\rho(\xi|\,\eta,t)=\sum_{\lambda\in D}\sum_{\lambda'\in D} c_\lambda'^*c_{\lambda'}'\psi_\lambda\psi_{\lambda'}^*, \tag{71}$$

and this has the structure of Eq. (48) only if the off-diagonal terms sum to zero.

Our earlier paper[2] erroneously implied that a choice for the arbitrary portions of the c_λ' could be made so that these off-diagonal terms would indeed sum to zero *everywhere* in G. This was based on a theorem given in the Appendix of that paper. However, we see from our discussion of embedding above that any representation for $\rho(\xi|\,\eta,t)$ need only be valid *almost everywhere* in G, not everywhere. The diagonalization of the X' matrix above yielding Eq. (65) implies that

$$\int_G x'(\xi,\eta,t)\sum_{\lambda\neq\lambda'}\sum_{\lambda'\in D} c_\lambda'^*c_{\lambda'}'\psi_\lambda(\xi,\eta,t)\times\psi_{\lambda'}^*(\xi,\eta,t)d^N\xi=0, \tag{72}$$

and if this can be shown to hold for *every* observable, then indeed this sum of off-diagonal terms vanishes almost everywhere in G.

If we construct the expectation value of any other observable, say y, for example, using $\psi^*\psi$ for $\rho(\xi|\,\eta,t)$ and the expansion of ψ on the U_λ basis, we will obtain the form

$$\langle y\rangle=B^\dagger Y'B, \tag{73}$$

with the elements of Y' defined just like those of X' in Eq. (56), but with the function $x'(\xi,\eta,t)$ replaced by $y'(\xi,\eta,t)$. Then we may ask whether the *same* unitary transformation will diagonalize both X' and Y'; if this is the case, we obtain $\langle y\rangle$ in a diagonal form just like $\langle x\rangle$ in Eq. (65) and this then is equivalent to Eq. (72) with $x'(\xi,\eta,t)$ replaced by $y'(\xi,\eta,t)$.

The necessary and sufficient condition for both X' and Y' to be diagonalized by T is for these matrices to commute:

$$X'Y'-Y'X'=0. \tag{74}$$

In fact, the same T matrix will diagonalize the matrix representatives of *all* observables, formed like X' and Y' here, provided only that these commute. Our earlier paper[2] demonstrated that such matrices will indeed commute if the spectral functions $x'(\xi,\eta,t)$, $y'(\xi,\eta,t)$, ... are such that all functions such as $x'(\xi,\eta,t)U_\lambda(\xi)$, $\lambda\in D$, belong to L^2 on G. Here the $U_\lambda(\xi)$ may be *any* orthonormal basis of L^2 on G. It was also shown[2] that this property of the spectral functions implied a finite variance for every such observable. The collection of all these spectral functions must then include all members of L^2 on G plus many other functions that are not members of L^2 on G, and it forms a nondenumerable set of functions.

Thus we can assert that for the basis set $\psi_\lambda(\xi,\eta,t)$, $\lambda\in D$, of L^2 on G, in terms of which the expectation values of all the compatible observables have the diagonal form as in Eq. (65), we have[6]

$$\rho(\xi|\,\eta,t)=\sum_{\lambda\in D} c_\lambda^*c_\lambda\psi_\lambda\psi_\lambda^* \tag{75}$$

[6] This reduction of the double sum in Eq. (71) to this single sum is usually introduced as the *reduction postulate* in conventional treatments of quantum theory [see, for example, K. Gottfried, *Quantum Mechanics* (Benjamin, New York, 1966), Vol. I, pp. 178 ff], but here such an *ad hoc* postulate is not required.

almost everywhere in G. Then the identification of $\rho_\lambda(\xi|\eta,t)$ as $\psi_\lambda^*\psi_\lambda$ in Eq. (70) is valid almost everywhere in G. However, we must remember that this requires all observables to be represented in the embedding by spectral functions having the square integrable property just described above, and we recall that these are already required to be sectionally continuous.

COORDINATE REPRESENTATION AS EMBEDDING

With the results just given, we are in a position to comprehend not only the statistical meaning of a coordinate representation of the state vector $|\eta,t\rangle$, but several other aspects of quantum mechanics as well.

First observe that the statistical description of a duplicated experiment given earlier does *not* imply that the λ_i are themselves operationally defined quantities that are directly observed or measured in the experiment, although these may be inferred (calculated) from any K of the observables (pointer readings) $x(\lambda,\eta,t)$, $y(\lambda,\eta,t)$, ... that form an independent set. Thus, the introduction of the coordinate representation, or the embedding with $N=K$, is just a "change of variables" $\lambda \rightarrow \xi$, where again the ξ_j are not directly observed. The new statistical description implies only the *same* observables as before.

Of course, we can construct embeddings with $N>K$, as indicated in our earlier paper. The statistical description in terms of the larger number of independent variables is then capable of describing not only the experiment at hand, but also experiments in which a larger number (N) of independent variables are measured. In quantum mechanics we usually refer to this situation as one of *degeneracy*, meaning that the mathematical description of the system has surplus labels available for distinguishing the results of measurement.

By exploiting the representation of the embedding distribution $\rho(\xi|\eta,t)$ as $\psi^*\psi$ and then the square integrable property of ψ, we were able to prescribe the conditions which assured the existence of all integrals in the embedding formalism. However, other lines of development are possible and are indicated, although not altogether properly, in our earlier papers,[2,3] i.e., with

$$\rho(\xi|\eta,t)=e^{-\sigma(\xi,\eta,t)}/B(\eta,t)\,, \tag{76}$$

with $\sigma(\xi,\eta,t)$ a real function and $B(\eta,t)$ the integral of the exponential over G.

FUNDAMENTAL OPERATORS AND QUANTUM FORMALISM

We have seen that the statistical definition of a duplicated experiment provides us with a matrix description of a system (an experiment) corresponding to fixed values of the η_j and t, and this was shown to have a representation on an abstract vector space which is a Hilbert space. We then showed that the change of basis from the discrete basis $|\lambda\rangle$, $\lambda\in D$, to a continuous basis $|\xi\rangle$, $\xi\in G$, had a statistical interpretation in terms of an embedding. Thus the underlying mathematical framework of the matrix structure can be formed by supplying three fundamental sets of operators, These are described as follows.

A basis for the N-dimensional Hilbert space can be defined by specifying a complete set of commuting operators whose eigenvectors constitute a basis; these *generate* the basis. The identification of such a set of *fundamental operators* is based on the fact that the orthonormal property, Eq. (23) or (25), is invariant under unitary transformation. Thus the unitary operator

$$\mathbf{U}=e^{i\xi\cdot\mathbf{P}}\,, \tag{77}$$

where the component operators $\mathbf{P}_n$, $n=1, 2, \ldots, N$ are a set of commuting Hermitian operators, maps an initial base vector $|0\rangle$ into the basis vector $|\xi\rangle$, and therefore $\mathbf{U}$ generates a basis for the space. If we interpret the ξ_j as space coordinates, as is customary, then the $\mathbf{P}_n$ correspond to the conjugate momenta.

Thus we see that *every* duplicated experiment having N independent observables must admit a description in terms of N space coordinates and N conjugate momenta. Of course, these operators may be given various representations.

But the N operators which generate a basis for the space are not the only operators of interest. There is also a Hermitian operator such as $\mathbf{x}$, for example, corresponding to every observable in the experiment. The operators $\mathbf{P}_n$ may have no special relation to the particular experiment in question and may be the *same* operators for *every* experiment having the same number of independent variables. On the other hand, all of the commuting operators $\mathbf{x}$, $\mathbf{y}$, ... are special to the system being considered, and the discovery of their forms is a matter of experiment.

However, it is not the operators of arbitrary observables such as $\mathbf{x}$, $\mathbf{y}$, ... which constitute the only other fundamental operators mentioned above; another set of *fundamental operators* consists of those which generate the evolution of the state vector $|\eta,t\rangle$ as the parameters η_j and t are changed, and which also generate the evolution of the operators $\mathbf{x}(\eta,t)$, ... of all observables. It was these operators which were considered in our previous paper.[1] However, we did *not* recognize the role of the η_j properly in that paper, where they were shown to be coordinates conjugate to momenta, and for a very good reason—the formalism of operators generating translation of the η_j is just like that for the space coordinates. In fact, we there denoted those quantities here symbolized by the η_j as q_j, having the canonical coordinates in mind.

We have shown[1] that if one postulates a linear mapping of spectral matrices such as, say,

$$X_\delta=VXV^\dagger\,, \tag{78}$$

where the δ subscript denotes translated values,

$\eta_j \rightarrow \eta_j+\delta\eta_j$, $t \rightarrow t+\delta t$, for the conditioning parameters, then every spectral matrix must evolve by the same matrix V, and we must have

$$V^\dagger V = I \tag{79}$$

in order to preserve commutation of spectral matrices. Furthermore, if *continuous* translations of the η_j and t are admitted, we must consider $X_\delta \rightarrow X$ as all $\delta\eta_j$ and δt approach zero, and hence we must have $V \rightarrow I$. Then V may be unitary and, as we showed,[1] all spectral matrices must then commute with V. This was then shown to imply that none of the compatible observables in the experiment are represented by operators having the η_j and t as parameters. However, if any of the η_j were discrete-valued parameters, then we would have only an isometric V instead of a unitary V, and these conclusions would not follow.

For example, if we consider a particle in a box, the volume of the box is a parameter, say v, that is fixed in the preparation of the system, and it is a perfectly meaningful question to ask how the probability for a particular momentum of the particle is conditioned by this parameter. In this classic problem, this probability is independent of time but is conditioned by the volume of the box. Since the volume is considered as continuously variable, none of the compatible observables have operators involving the volume as a parameter, but the state vector must change under translation of this parameter.

Thus, in the matrix format we could have, for example,

$$\partial C/\partial v = -i\mathcal{P}C\,, \tag{80}$$

in analogy to Eq. (13), where $\mathcal{P}$ is in general a *non-Hermitian* matrix.

Similarly, if no compatible operator in the experiment has v as a parameter, the analog of Eq. (14) is

$$-\partial\langle x\rangle/\partial v = iC^\dagger(\mathcal{P}^\dagger X - X\mathcal{P})C\,, \tag{81}$$

and if x here were the energy, we would interpret the left-hand member as the pressure in the usual thermodynamic sense. Of course, $\mathcal{P}$ could be Hermitian in some circumstances, but even so it would not in general correspond to an observable in the experiment if x is observable, i.e., observables commute.

Thus, not only does this statistical formulation provide a basis for the familiar operators of quantum theory, but also for another class of operators such as that represented here by $\mathcal{P}$.

It seems now that all of the operators which generate translation of fixed parameters, like that here corresponding to $\mathcal{P}$ for the volume v, might constitute an operator formalism for thermodynamics, all, that is, except the one which generates translation of the time t. It was these operators which were treated in our most recent paper.[1]

The idea, expressed in our recent paper,[1] that phase transitions might be represented by isometric operators corresponding to finite translations of fixed parameters, as the volume v here, gains more credence in this recognition of a thermodynamic operator formalism. However, much work remains to develop these ideas fully.

CONCLUSION

In this paper we have achieved a specific objective, namely, exhibiting the proper connection of the embedding formalism to the formal operator structure of quantum theory. Specifically, we have shown that an embedding is the statistical equivalent of a change in representation. But in establishing this relation of embedding to representation theory, we have also clarified the role of operators associated with translations of conditioning parameters; this has indicated the possibility of an operator formalism for thermodynamics.