

MATSCIENCE REPORT 11

LECTURES ON
FOUNDATIONS OF
QUANTUM MECHANICS AND FIELD THEORY

BY

E. C. G. SUDARSHAN

8662



THE INSTITUTE OF MATHEMATICAL SCIENCES, MADRAS-4, INDIA.

THE INSTITUTE OF MATHEMATICAL SCIENCES

MADRAS - 4 (India)

IMSC
R11
8662

LECTURES ON
FOUNDATIONS OF
QUANTUM MECHANICS AND QUANTUM FIELD THEORY*
(Summer 1962)

By
Professor E. C. G. Sudarshan
University of Rochester, Rochester
New York
(United States of America)

=====
*These lecture notes were prepared by Dr. T. K. Radha,
Dr. Thunga Satyapal and Mr. K. Raman. Institute of Mathematical
Sciences, Madras.

CONTENTS

<u>I.</u>	<u>GENERAL INTRODUCTION</u>	<u>PAGE</u>
	1. Axiomatic Approach to Dynamics	1
	2. Hamiltonian Dynamics	4
	3. Vector Spaces and Lie Algebras	6
	4. Positive Linear Functionals as States	8
	5. Convex Sets and External Elements	11
	6. Spin Algebras	14
	7. Discussion	16
<u>II.</u>	<u>FORMAL FRAMEWORK.</u>	
	1. Introduction	19
	2. The Hilbert space Formulation of Abstract Dynamics	25
	3. Galilean Mechanics in a Hilbert Space Formalism	30
	4. Local Lie Groups	38
	5. Lie Algebras	40
<u>III.</u>	<u>RELATIVISTIC DYNAMICS</u>	
	1. Realizations of the Galilei Group	48
	2. The Lorentz Group and its Realization	52
	3. Interesting Relativistic Systems	58
	4. Trajectory Dynamics	64
	5. Wightman Formulation of Quantum Field Theory	70
	6. Stochastic Dynamics of Quantum Mechanical Systems	74
<u>IV.</u>	<u>STRUCTURE OF INTERACTING SYSTEMS; QUANTUM FIELD THEORY</u>	
	1. Separable Potential Model (S-waves)	79
	2. Unitary Equivalence; Asymptotic Condition	94
	3. The Lee Model	98
	4. The Frame Work of Quantum Field Theory	100
	5. Statistics of Elementary Particles; Parastatistics	108
	6. The Metric Postulate	114

I. GENERAL INTRODUCTION.

1. Axiomatic Approach to Dynamics.

The deduction of the laws of nature from general principles has been the goal of physics from the time of Euclid and Aristotle. Starting with Euclid, Aristotle, Newton, Eddington and Einstein, we can follow it down to the recent axiomatic approaches to physics, axiomatic field theory in particular.

In quantum field theory, the main success has been in quantum electrodynamics, where the basic equations are derived from Maxwell's classical electrodynamics; the success of quantum electrodynamics is in deriving new results from old equations. The emission of radiation, the scattering of light by free and bound electrons, the Lamb shift in the ground state of the hydrogen atom, and the anomalous magnetic moment of the electron are some of the successful predictions of the theory. All the complications of the Hilbert space are reduced to an integral equation for which the solution is written as an infinite series; the above results are all obtained from just the first few terms of this infinite series, i.e. the lowest-order Feynman diagrams, and these results still seem to fit the facts.

However, in quantum electrodynamics, one encountered infinite changes in the effective mass and in the strength of interactions (i.e. the coupling constant) and various tricks had to be used to circumvent these difficulties.

When we come to the Field theory of Strong Interactions, we have no useful way of calculating; in weak interactions, also, the methods of quantum electrodynamics are not applicable, as the

finite result obtained would depend on the recipe used in calculating it. On the other hand, we expect that some features general and universal, common to all types of interactions, should exist in a correct theory of quantized fields. Relativistic invariance in itself seems to be a satisfactory principle (although the particular structure of the dynamics used is not so); we must look for a relativistic dynamical theory for quantal systems.

We start by looking at classical dynamics, where the trajectory of a particle is described by giving the displacement x as a function of the time t . The values of $x(t)$ and $\dot{x}(t)$ may be chosen arbitrarily, but then the value of $\ddot{x}(t)$ would be fixed by the dynamical equation

$$\ddot{x}(t) = \frac{1}{m} F(x(t), \dot{x}(t), t)$$

which is a second order ordinary differential equation -- a parabolic equation. The solution, in terms of two arbitrary constants -- the "initial conditions", may be written

$$x(t) = \mathcal{F}_1(x(t_0), \dot{x}(t_0), t)$$

The generalization of all this to more complicated systems involves two distinct features. First one may try to find out symmetry properties of the system (e.g. the property that for a "central force", the system is invariant under rotations about the centre of force) and relate them to conservation laws (e.g. in the above, the rotational invariance implies the conservation of angular momentum).

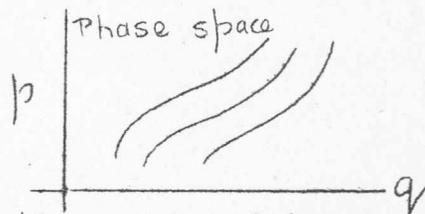
Secondly, one uses generalized coordinates, the equations of motion being given by Lagrange's or Hamilton's equations.

For a system of interacting particles, the total energy contains an additional interaction term, while the total linear momentum \vec{p} and the total angular momentum \vec{J} are still expressed as the sum of the linear momenta and the angular momenta of the individual particles. Dynamical variables of the latter type (i.e. like \vec{p} and \vec{J}) are termed kinematic variables, while variables like the energy, which cannot be expressed as the sum of the individual particle variables when an interaction is present are termed dynamic variables. There are conservation laws for both kinematic and dynamic variables; in this sense they are on an equal footing.

In Hamiltonian (or 'canonical') dynamics, one writes a pair of (simultaneous) first order differential equations in two variables q, p :

$$\left. \begin{aligned} \dot{q}(t) &= \frac{\partial H(q, p, t)}{\partial p} \\ \dot{p}(t) &= -\frac{\partial H(q, p, t)}{\partial q} \end{aligned} \right\}$$

The dynamics may then be visualized in terms of trajectories in a phase space; the dynamics of the system is then contained in the geometry of the trajectories in phase space; this is reminiscent of Aristotle, who emphasized that it is "form" that is important in the description of motion.



Hamiltonian dynamics may be conveniently expressed in terms of Poisson Brackets

$$[A, B] = \frac{\partial A}{\partial q} \frac{\partial B}{\partial p} - \frac{\partial A}{\partial p} \frac{\partial B}{\partial q};$$

the dynamical equations may be written

$$\dot{q} = [q, H] \quad ; \quad \dot{p} = [p, H]$$

One then considers the behaviour of the system under canonical transforms; i.e. one looks at the different possible choices of coordinate systems that leave the general dynamics of the system invariant. What seems to be important is not the form of the trajectories, but the form of the law that expresses their behaviour under canonical transforms.

So far we seem to have said nothing about the forces; we have taken the existence of these for granted. Hertz* started with the principle that one must deduce the "forces" as a manifestation of the existence and motion of particles. Hertz's program was to reduce dynamics to a form in which there would be no arbitrary forces, and no action at a distance. Somewhat as in current elementary particle physics, virtual particles were introduced, using which, all interactions could be expressed in terms of contact interactions. Apart from a few arbitrary constants (to be determined from experiment), all the dynamics was supposed to follow from the theory.

2. Hamiltonian Dynamics.

Consider classical Hamiltonian dynamics which is characterized by the equations

$$\dot{q} = \frac{\partial H}{\partial p} \quad ; \quad \dot{p} = - \frac{\partial H}{\partial q} \quad ;$$

and in general $\dot{\phi}(q, p) = [\phi, H]_{p, q}$

where ϕ is a dynamical variable (and hence a function of q and p)

* H. Hertz : Principles of Mechanics

In the general equation

$$\dot{\phi}(q, p) = [\phi, H]_{P.b.} \quad (1)$$

the Poisson bracket (P.b.) occurring on the R.H.S. is characterized by the following properties:

(i) it is bilinear; i.e. $[\lambda A, B] = \lambda [A, B]$;
 $[A, \mu B] = \mu [A, B]$; $[(A_1 + A_2), B] = [A_1, B] + [A_2, B]$

or more succinctly,

$$[(\lambda_1 A_1 + \lambda_2 A_2), B] = \lambda_1 [A_1, B] + \lambda_2 [A_2, B] \quad (i)$$

(ii) it is antisymmetric (and hence non-commutative):

$$[A, B] = -[B, A] \quad (ii)$$

and

(iii) Instead of obeying the associative law obeyed by ordinary products, it obeys the 'Jacobi identity':

$$[[A, B], C] + [[B, C], A] + [[C, A], B] = 0 \quad (iii)$$

Considering the evolution of a dynamical system in time, we may write

$$\begin{aligned} \phi(q(t), p(t)) &= \phi(0) + t\dot{\phi}(0) + \frac{t^2}{2!}\ddot{\phi}(0) + \dots \\ &= \phi(0) + t\dot{\phi}(0) + O(t^2) \quad \begin{array}{l} \text{to orders} \\ [t^2 \text{ and above}] \end{array} \quad (2) \end{aligned}$$

If $\phi(0) = \psi^2 \geq 0$, then we may expect that

$$\phi(t) \geq 0 \quad (3)$$

From the Jacobi identity it follows that if A and B are constants of motion, then $[A, B]$ is also one, i.e. if $[A, H] = 0$ and $[B, H] = 0$ then it follows that $[[A, B], H] = 0$. Any

transformation on the system that leaves the Poisson bracket $[A, B]_{p, B}$ invariant is called a canonical transformation. Expressed in terms of a parameter λ , it may be written

$$\phi_\lambda = \phi_0 + \lambda [\phi, A] + \lambda^2 [[\phi, H], A] + \dots (4)$$

We shall ask whether we can set up a more general dynamics in which (+) holds formally but which is formulated in terms of a "product" $[A, B]$ which is bilinear, antisymmetric and obeys the Jacobi identity. Examples of such quantities are :

Ex. 1. $[A, B] = \vec{A} \times \vec{B}$ the vector product of two vectors \vec{A} and \vec{B} ; and

Ex. 2. $[A, B] = \lambda (AB - BA)$ where A and B are two matrices.

3. Vector Spaces and Lie Algebras.

In a vector space, the addition of two vectors and the multiplication of vectors by numbers are defined. We may give the vector space an additional structure and make it a Lie Algebra.

To give this additional structure, we define a multiplication between two vectors in such a manner that the product is bilinear, antisymmetric and obeys the Jacobi identity (e.g. we could define the product of vectors \vec{A} and \vec{B} as the Poisson bracket $[A, B]$; with such a definition, the vector space (e.g. one made up of the set of all differentiable functions) becomes a Lie Algebra.

Several other definitions of the product would also make the vector space a Lie Algebra; e.g. we could define the product of two vectors A and B by the Moyal sine Bracket (M.b):

$$(A, B)_{M.b}^{(q, p)} = \text{Sin} \left(\frac{\partial}{\partial q_A} \frac{\partial}{\partial p_B} - \frac{\partial}{\partial p_A} \frac{\partial}{\partial q_B} \right) A(q, p) B(q, p)$$

which obeys the requirements (i) - (iii) above.

[We may note that the P.b. may be written

$$[A, B]_{P.b.}^{(q,p)} = \left(\frac{\partial}{\partial q_A} \frac{\partial}{\partial p_B} - \frac{\partial}{\partial q_B} \frac{\partial}{\partial p_A} \right) A(q,p) B(q,p)$$

We note that this definition (8) of the Moyal bracket assumes infinite differentiability of A and B with respect to q and p.

The Moyal bracket is related to a particular representation of quantum mechanics in terms of classical variables.*

Certain elements A of a Lie Algebra may have vanishing brackets with some other particular element B :

$$[A(q), B(q)]_{L.b.} = 0$$

where L.b. indicates a general Lie bracket.

Ex. $[a, B(q,p)]_{P.b.} = 0$

i.e. a constant a has a vanishing Poisson bracket with every element B(q, p)

An element that has a vanishing bracket with every element of a Lie Algebra is called a neutral element.

For the special case of a Poisson bracket, a neutral element must be a multiple of the identity I.

[The proof is direct. Consider an element X(q, p) such that $[X(q,p), B(q,p)] = 0$ for all B.

a) Take B = q, obtaining $[X, q] = 0$ or $-\frac{\partial X}{\partial p} = 0$

b) Take B = p, obtaining $[X, p]$ or $\frac{\partial X}{\partial q} = 0$

Therefore X = a constant or a multiple of I.

Q.E.D. 7

* Jordan and Sudarshan: Rev. Mod. Phys. 33, 515, (1961)

Ex. 1. When $[A, B] = \vec{A} \times \vec{B}$

no non-trivial neutral element (i.e. besides the null vector) exists.

Ex. 2. When $[A, B] = \lambda (AB - BA)$, a neutral element may or may not exist depending on the collection of matrices A, B, ... When A and B are 'all' finite dimensional matrices, this is essentially the same as Ex. 1. but we may have a subcollection of $n \times n$ matrices which may have neutral members.

4. Positive Linear Functionals as States.

In all the above we have been dealing with equations of motion. We must also concern ourselves with the values of physical quantities.

To a dynamical variable A, we assign a numerical value $\epsilon(A)$ such that the correspondence is linear $A \rightarrow \epsilon(A)$ i.e.

$$\begin{aligned} A + B &\rightarrow \epsilon(A) + \epsilon(B) \\ \lambda A &\rightarrow \lambda \epsilon(A) = \epsilon(\lambda A) \end{aligned}$$

or writing it in a unified way.

$$\epsilon(\lambda_1 A_1 + \lambda_2 A_2) = \lambda_1 \epsilon(A_1) + \lambda_2 \epsilon(A_2)$$

Such a mapping is a mapping of an element of the Lie algebra on to a number $\epsilon(A)$. (By way of contrast, we may note that the mapping induced by multiplication of an element by a number is a mapping of the Lie Algebra ^{onto} itself.

This mapping of elements on to a real number we call a 'Real linear functional or mapping'. We may expect that to the unit element (i.e. $A = 1$) must correspond the number one,

$$\text{i.e. } \epsilon(1) = 1.$$

Thus one restriction that must be imposed on the real linear functional is that it be 'normalized'.

We postulate one further requirement, viz., that it should be non-negative (i.e. ≥ 0) in the sense to be defined later.

Thus we look for real, normalized, non-negative, linear functionals defined over the vector space of all dynamical variables.

The state or 'status'* of a physical system is defined by the rule for assigning numbers to physical quantities. (We may note that in quantum mechanics, an expectation value is defined in terms of both the dynamical variable (i.e. the physical quantity) and the state of the system.)

Consider the requirement that the square of an element (i.e. the 'product' of an element with itself) should be positive-definite,

$$\phi = \psi^2 \geq 0$$

If this 'product' were taken to be the Lie bracket, then it would vanish (since $[A,B]$ is antisymmetric, it follows that $[A,A] = 0$). Thus to express a condition of the form above we must define a 'product' other than the Lie bracket; we take this to be the 'ordinary' product (indicated by a dot). I.e., It is not enough that the vector space be converted into a Lie Algebra by giving it a structure in the form of a vector product or Lie bracket; it must be supplemented by adding the concept of an 'ordinary' product with the properties that:

- a) the multiplication involved must be associative, and
- b) it must be such that we are able to fulfil the requirement

$$\phi(t) = |\psi(t)|^2 \geq 0$$

* We introduce the term 'states' to stress that we are not referring to a state in the sense of Q.M.

This requirement is not satisfied for every Lie algebra; it holds only under very special conditions.

Let us examine the implications of this requirement.

Write $A^2 = A \cdot A$, A being some dynamical variable. For small value of t , we must have

$$A^2(t) = A^2(0) + t[A^2, H] + \dots$$

We also have

$$\begin{aligned} \{A(t)\}^2 &= \{A(0) + t[A, H] + O(t^2)\}^2 \\ &= \{A(0)\}^2 + t\{A(0) \cdot [A, H] + [A, H] \cdot A(0)\} \\ &\quad + O(t^2) \end{aligned}$$

We note that in the evolution with time, the square of a physical quantity would go over into the square of the physical quantity at a later time (i.e. the requirement that $A^2(t) = \{A(t)\}^2$ would be fulfilled) only if we have

$$[A^2, H] = A \cdot [A, H] + [A, H] \cdot A$$

We may generalize this requirement to

$$[(A \cdot B), C] = A \cdot [B, C] + [A, C] \cdot B$$

where A, B, C are arbitrary dynamical variables. An operation of this kind is called a "derivation". [The reason for this is that the characteristic of ordinary derivation of differentiation ∂ is that

$$\partial(A \cdot B) = \partial A \cdot B + A \cdot \partial B$$

Thus we have required that the Lie bracket should have the property of being a derivation with respect to the 'ordinary product'.

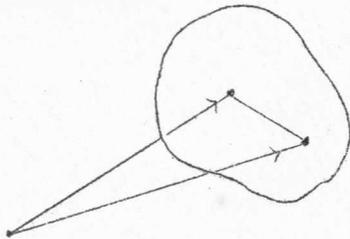
Putting together all the above requirements,
an

we define / 'ABSTRACT DYNAMICS' as a vector space (of a set of dynamical variables) with a structure defined by multiplication in the form of a Lie bracket (which makes the vector space a Lie algebra), supplemented by the concept of the 'ordinary' product, with the property of derivation between the ordinary and Lie-bracket products; further, we have a set of real, non-negative, normalized linear functionals defined over the vector space.

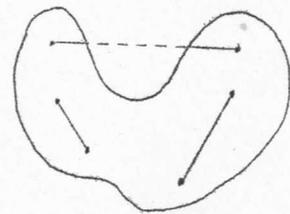
5. Convex Sets and External Elements.

To understand the concept of convex sets, we may consider, for simplicity, the set of points in ordinary 3-dimensional space.

A set of points here is said to be 'convex', if the join of any two points in the set lies completely inside the region bounded by the set



A convex set



A non-convex set.

Fig. 1. Bounded convex sets.

Fig. 2. An unbounded convex region
(or convex cone)

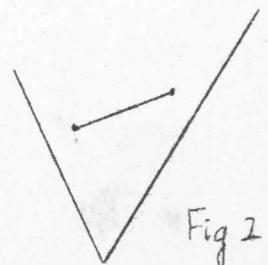


Fig 2

UNBOUNDED

Fig. 1 shows examples of bounded convex and non-convex sets. Fig. 2. gives an example of an unbounded convex region or a convex cone.

In 3-dimensional space, if we specify a set of points by specifying the vectors from some arbitrary origin to the points of the set, we may distinguish convex sets and convex cones as follows:

If \vec{A} , \vec{B} specify any two points of the set, then if the vector $\lambda\vec{A} + \mu\vec{B}$ is also in the set for all $\lambda, \mu \geq 0$, then the set of points is a convex cone.

If the vector $\lambda\vec{A} + \mu\vec{B}$ is in the set for $\mu = 1 - \lambda$, with $0 < \lambda < 1$ then the set of points is a convex set. (Thus a convex cone is necessarily a convex set, but the converse is not true.)

A bounded convex set has a boundary, in addition to its property of being convex; the boundary is defined by the set of points each of which has some neighbourhood not completely contained in the set. Bounded convex sets have "extremal elements". An extremal element of a set is one which does not lie on the segment joining any two points of the set (but may lie at the end-points of the segment), and hence cannot be expressed by a sum of the form $\lambda\vec{A} + \mu\vec{B}$, $\mu = 1 - \lambda$, $0 \leq \lambda \leq 1$, $\vec{A} \neq \vec{B}$ i.e. (DEFINITION) If \vec{A} , \vec{B} specify any two points of a set, then if $\vec{X} = \lambda\vec{A} + (1 - \lambda)\vec{B}$, $0 \leq \lambda \leq 1$ implies that $\vec{A} = \vec{B}$, then the point \vec{X} is said to be an extremal point of the set.

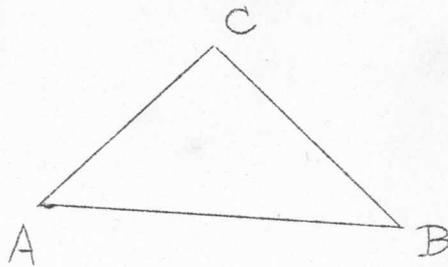
Ex. 1.



Consider the set of points bounded by a circle (the boundary being included).

Every point on the circle is an extremal point.

Ex. 2.



The set of points bounded by a triangle has just three extremal points, viz., the vertices of the triangle.

The real linear functionals (normalized and non-negative) considered previously constitute a convex set.

For if $\epsilon^{(1)}$ and $\epsilon^{(2)}$ are two linear functionals satisfying the conditions required earlier, then the linear functional defined by

$$\epsilon(A) = \lambda \epsilon^{(1)}(A) + (1 - \lambda) \epsilon^{(2)}(A)$$

will also be a linear functional fulfilling the same requirements, provided that λ is a real number $0 \leq \lambda \leq 1$

We may note that in quantum mechanics, the pure states are extremal elements, while the impure states are elements other than the extremal ones and lie on segments joining extremal elements.

Most properties of these other elements can be discussed in terms of the extremal elements, viz., the pure states.

We have in the above, converted the set of all linear functionals into a subset of a new vector space, the subset itself not being the usual type of vector space in that arbitrary linear combinations are not allowed, but only combinations of the form $\lambda \vec{A} + \mu \vec{B}$ with real λ, μ satisfying the condition $\lambda + \mu = 1$; i.e. only if $\lambda + \mu = 1$ will the combination $\lambda \epsilon^{(1)} + \mu \epsilon^{(2)}$ satisfy all the conditions satisfied by $\epsilon^{(1)}$ and $\epsilon^{(2)}$.

The dimension of the vector space has nothing to do with the number of extremal elements.

The approach in terms of linear functionals is due to Segal ('Postulates for quantum mechanics'; Annals of Maths. 1947); it is more general than the approach of Von Neumann in terms of vectors in a Hilbert space, as we shall see later.

Given any associative algebra (i.e. one with the prop.

$A.(B.C) = (A.B).C$ we can always define a bracket $[A, B] = (AB - BA)$

We define:

$$[A, B, C] = (AB).C = C.(AB)$$

Hence we obtain

$$\begin{aligned} [A, B, C] &= (A.B).C - C.(A.B) - A.(C.B) + A.(C.B) \\ &= ABC - ACB + ACB - CAB \\ &= A.(BC - CB) + (AC - CA).B \\ &= A.[B, C] + [A, C].B \end{aligned}$$

i.e. The bracket $[A, B]$ is a derivation with respect to the dot product.

6. Spin Algebras.

Spin systems in Quantum Mechanics are defined by the commutation relations $[J_i, J_j] = J_k$ for the components of the angular momentum operator \vec{J} .

(Classically, $\vec{J} = \vec{q} \times \vec{p}$. Then if q, p satisfy $[q_i, p_k] = \delta_{jk}$ then $[J_i, J_j] = J_k$ where i, j, k are cyclic permutations of 1, 2, 3).

Normally, in Hamiltonian mechanics we deal with Poisson brackets of functions $f(p, q), g(p, q)$ of p and q . We could equally well deal with functions of \vec{J} and P.b's formed from them e.g.

$$[f(\vec{J}), g(\vec{J})] \text{ e.g. we have}$$

$$(i) \quad [J_1^2, J_z] = J_1 \cdot [J_1, J_z] + [J_1, J_z] \cdot J_1$$

which can be evaluated.

$$(ii) \quad [J^2, f(J)] = 0$$

Algebras spanned by functions of \vec{J} with the Poisson bracket defined on these functions we shall designate spin algebras; surprisingly, these do not seem to have been studied in detail. (e.g. we might ask what are the analogues of Liouville's Theorem, etc. for spin systems, this left as an 'exercise for the reader' !)

The equations of motion for spin systems are of the form

$$\dot{J}_1 = [J_1, H(J)] = \phi_1(J),$$

$$\dot{J}_2 = [J_2, H(J)] = \phi_2(J),$$

$$\dot{J}_3 = [J_3, H(J)] = \phi_3(J)$$

$J^2 = J_1^2 + J_2^2 + J_3^2$ is a constant of motion (A, say) and hence $2J_1\dot{J}_1 + 2J_2\dot{J}_2 + 2J_3\dot{J}_3 = 0$ which is a non-linear equation.

As $[H(\vec{J}), H(\vec{J})] = 0$ we note that $H(\vec{J})$ is also a constant of the motion, = B, say.

We can use these constants of motion to solve the non-linear equations of motion of the spin system.

J_1 and J_2 are expressed as functions of A, B and J_3 and we obtain the equation for J_3 as $J_3 = \psi(A, B, J_3)$, the R.H.S. being the same as $\phi_3(J)$

$$\text{The solution of this is } t + c = \int \frac{dJ_3}{\psi(J_3)}$$

in which the R.H.S. is an Abelian integral, as $\psi(J_3)$ is an algebraic function. The equation for most of the familiar spin systems can be solved in terms of elliptic functions.

Brackets other than P.b.'s can also be defined for spin systems. These may be obtained by considering the $(2S + 1) \times (2S + 1)$ dimensional spin S representations of the operators \vec{J} . Consider the symmetrized polynomials in these finite dimensional matrices; there are only $(2S + 1)^2$ such linearly independent polynomials. And their commutators can be expanded in terms of them. We have then a sequence of finite dimensional Lie algebras \mathcal{R}_S with dimension $(2S + 1)^2$ for $S = 0, 1/2, 1, \dots$. These are really Lie algebras defined over equivalence classes of polynomials in a vector; a more detailed discussion is beyond the scope of these lectures.

7. Discussion.

We shall now briefly compare the Segal and Von Neumann formulations of quantum mechanics.

A Hilbert space is a special case of a vector space. It is a vector space in which a product (A, B) is defined which is a (bilinear) mapping of a pair of vectors into a number, the product having the following properties (and the properties of completeness and separability):

$$(\lambda A, \mu B) = \lambda^* \mu (A, B)$$

$$(A_1 + A_2, B) = (A_1, B) + (A_2, B)$$

$$(A, B) = (B, A)^*$$

$$(A, A) \geq 0 \quad \text{the equality holding only when } A = 0.$$

If the last condition is not satisfied, we have a pre-Hilbert space (which can be converted into a Hilbert space of a 'smaller' number of dimensions).

In Von Neumann's formulation, physical states correspond to vectors (with norm = 1) in a Hilbert space. The expectation value $A(f)$ of an operator f in a state A is defined by

$$A(f) = (A, fA)$$

For $f =$ the unit operator $A(f) = (A, A) = 1$. If $f = g^+g$ where g is an operator, then $A(f) = (A, g^+gA) \geq 0$ i.e. the expectation value is positive-definite.

We note that in a Hilbert space, if a vector A is multiplied by a number λ , then the expectation value of a dynamical variable in a state given by the vector A is multiplied by $|\lambda|^2$. This property of the expectation value is different from the properties of the expectation value $E(A)$ in terms of linear functionals.

(It may also be noted that in a Hilbert space, it is not just one vector but a whole family of vectors that corresponds to a given state; for if a vector A determines a certain expectation value, then the vector $e^{i\alpha}A$ determines the same expectation value, α being real, i.e. a phase difference between vectors is irrelevant.)

In Segal's formulation, an expectation value which is linear in the quantity characteristic of the state (i.e. linear in the expression below) is

$$\int f(p, q) \rho(p, q) dp dq$$

where $f(p, q)$ is the dynamical variable of interest.

(Note: In quantum mechanics, in Von Neumann's Hilbert space formulation the expectation value is bilinear in ψ (which characterises the state.) The advantage of

of Segal's formulation is that the expectation value is of the same form as that of a classical distribution.

We ask the question whether one of the two formulations (Von Neumann's and Segal's) is more general than the other. More specifically, we ask:

Q. 1.) Given an arbitrary dynamics, is it possible to express it in Von Neumann's form?

Q. 2.) Given any abstract algebra, can it be, expressed in terms of a Hilbert space?

The answer to these questions seems to be that to a great extent, an arbitrary dynamics can be expressed in terms of linear operators in a Hilbert space, but that a Hilbert space is not sufficient to describe completely all possible forms of dynamics (e.g. A normalized state with a definite momentum in quantum mechanics cannot be expressed in a Hilbert space except as the limit of a sequence of states in that space (however, the limit lies outside the space). However, Segal's formulation includes a specification of such a state in a rigorous manner.

Let A be a vector in Von Neumann's Hilbert space; we define a second Hilbert space by the correspondence

$$A \rightarrow AA^\dagger$$

where we note (i) that AA^\dagger is a matrix, and (ii) that the mapping defined by $A \rightarrow AA^\dagger$ is not one-to-one, as the whole family of vectors $e^{i\alpha}A$, α real, go over to the same operator AA^\dagger .

The expectation value $A(f)$ of a dynamical variable f in a state A is given by $(A, f) \rightarrow \text{Tr} (f AA^\dagger)$

Then, in this form, the Von Neumann formulation goes over into the

II. FORMAL FRAMEWORK.

1. Introduction

In classical mechanics, the equation for a dynamical variable f is

$$\dot{f}(q,p) = [f, H]_{P.O.}$$

where the Poisson bracket $[A, B] = \frac{\partial A}{\partial q} \frac{\partial B}{\partial p} - \frac{\partial A}{\partial p} \frac{\partial B}{\partial q}$

has the properties of (i) being bilinear, (ii) being distributive, and (iii) obeying the Jacobi identity.

We define also a different kind of product, the ordinary product $A \cdot B$, which has the following properties:

- (a) it is distributive, i.e. $(\lambda_1 A_1 + \lambda_2 A_2) \cdot B = \lambda_1 (A_1 \cdot B) + \lambda_2 (A_2 \cdot B)$
- (b) it is associative, i.e. $(A \cdot B) \cdot C = A \cdot (B \cdot C)$
- (c) $A \cdot A \geq 0$ (and $= 0$ only if $A = 0$).
- (d) the Lie bracket defined by the properties (i) - (iii) above should be a derivation with respect to the ordinary product. With these properties, we can define a linear equation

$$\dot{f}(q,p) = [f, H]$$

which has the formal solution

$$f_t = f_0 + t \cdot [f, H] + \frac{t^2}{2} [[f, H], H] + \dots$$

For a positive definite function f , $f = \psi^2$, we have $f_t = (\psi_t)^2$ for all t .

We can now define linear functionals on the system by saying that given A , A is mapped on to $\mathcal{E}(A)$,

$$A \rightarrow \mathcal{E}(A) \quad \text{such that}$$

- (i) $\epsilon(A)$ is normalized, i.e. $\epsilon(1) = 1$.
- (ii) $\epsilon(A)^2 \geq 0$ and $= 0$ only if $A = 0$,
- (iii) $\epsilon(\lambda_1 A_1 + \lambda_2 A_2) = \lambda_1 \epsilon(A_1) + \lambda_2 \epsilon(A_2)$

With a suitably defined $f(q, p)$ we can always write

$$\epsilon(A) = \iint P(q, p) A(q, p) dp dq$$

Ex. If $f(q, p) = \delta(q - q_0) \delta(p - p_0)$, then $\epsilon(A) = \iint \delta(q - q_0) \delta(p - p_0) A(q, p) dp dq = A(q_0, p_0)$ Thus this state ϵ corresponds to the familiar state with p and q 'having the values' q_0, p_0 .

On the other hand given any state $P(q, p)$ we may think of it being made up out of the combination $\iint dq_0 dp_0 P(q_0, p_0) \cdot \epsilon_{p_0 q_0}$ where $\epsilon_{p_0 q_0}$ is the state represented above. We also verify that these states constitute a two parameter infinity of extremal states.

For a time-dependent dynamical variable A_t , we may write $\epsilon_t(A) = \epsilon(A_t)$ with $\epsilon_0 \equiv \epsilon$, where $\epsilon_t(A)$ and $\epsilon_{(t)}$ are two different ways of conceiving of a time-dependent expectation value; i.e. (i) viewed as $\epsilon_t(A)$, it is of the nature of a time-dependent (expectation) functional of an operator A which is independent of time, (ii) viewed as $\epsilon(A_t)$, it is of the form of a time-independent expectation functional of an operator A_t which varies with time. This corresponds to the passage in quantum mechanics from a Schrodinger picture (i) to a Heisenberg picture (ii). In classical mechanics, we have something similar, namely, the passage from the Hamilton-Jacobi picture (corresponding to (i)) to the Hamiltonian picture (corresponding to (ii)).

As mentioned earlier, in quantum mechanics, there are two methods of approach:

- (1) Segal's method of linear functionals,
- (2) The method of state vectors in a Hilbert space.

The observation that in the second method the expectation values are bilinear in the state vector while the dynamical variables are linear operators suggests that we may be able to construct a Hilbert space of a more general character.

Denoting an expectation value in a state ψ (in quantum mechanics) by $\langle A \rangle$, etc., we have

$$\begin{aligned} \langle A_{op} \rangle_{\psi} &= \int \psi^*(x) A_{op}(x, -i \frac{\partial}{\partial x}) \psi(x) dx \\ &= (\psi, A\psi) \end{aligned}$$

We write this as

$$\langle A \rangle_{\psi} = \text{Tr} [A P_{\psi}]$$

We note that $P_{\psi}^2 = \psi\psi^{\dagger} \cdot \psi\psi^{\dagger} = \psi(\psi^{\dagger}\psi)\psi^{\dagger} = \psi\psi^{\dagger} = P_{\psi}$

If P_{ψ} had eigen values, these must be 0 and 1. Comparing with $\text{Tr} (A P_{\psi})$ we make the correspondence

$$\text{Tr} [A P_{\psi}] \rightarrow C(A)$$

Noting that $\text{Tr} [A P_{\psi}]$ has all the properties required of $\epsilon(A)$ viz., it is linear and distributive, and to $\epsilon(1)$ corresponds 1,

$$\text{as } \text{Tr} [1 \cdot P] = \text{Tr} [\psi\psi^{\dagger}] = \text{Tr} [\psi^{\dagger}\psi] = 1 \quad ; \quad P \geq 0$$

(Note: $\text{Tr} P = \sum_{\alpha} P_{\alpha\alpha} = \sum_{\alpha} \psi_{\alpha} \psi_{\alpha}^{\dagger} = \sum_{\alpha} \psi_{\alpha}^{\dagger} \psi_{\alpha} = 1$)

Hence in this special case, the formulation in terms of vectors in a Hilbert space can be made to correspond to the more general approach in terms of linear functionals. In particular, we can make use of the analogous artifices here also. (Incidentally

since $P \geq 0$, $\text{Tr} \{P\} = 1$, P can be always diagonalized.)

To ^{find} the extremal elements, we note that if E_1 and E_2 are admissible linear functionals then E defined as

$$E = \lambda E_1 + (1-\lambda) E_2 ; \quad \lambda < 1,$$

is also an admissible linear functional.

Here we note that $P = \lambda P_1 + (1-\lambda) P_2$ lies on the segment connecting P_1 and P_2 ; hence P is a non-extremal element.

But we note that if $P_1^2 = P_1$ and $P_2^2 = P_2$ are satisfied, then for P as defined above, $P^2 \neq P$ conversely, a quantity P with the property $P^2 = P$ cannot lie on the segment joining P_1 and P_2 .

Hence P^2 is an extremal element.

It may be shown that in finite dimensional spaces this exhausts all the extremal elements but that in spaces of an infinite number of dimensions, this is not true in general. Segal has shown that for these in general the extremal elements do not correspond to vectors in the Hilbert space.

To see this we note that because of the property, $P_\psi^2 = P_\psi$ the eigenvalues of P_ψ can be just 1 and 0; and further, because $\text{Tr} P_\psi = 1$ the eigen value 1 occurs just once and all the other eigen values are 0, i.e. there is just one eigenvector ψ with non-zero eigenvalue and this is given by

$$P_\psi \psi = + \psi$$

Suppose we do not have the above construction for P , but just think of P as a linear functional, then P may not have any eigenvector.

Ex. Consider the special case

$$A_{op}(P, q) = A_{op}(q)$$

where $A_{op}(p, q)$ is the operator corresponding to the dynamical variable $A(p, q)$

Again we have the correspondence

$$A_{op}(q) \rightarrow E(A) = A(x_0)$$

But then, this state corresponds to a state with definite position, i.e. an eigen function corresponding to a definite position. Hence if the wave function existed, it must be proportional to the

δ function, $\psi(x) = \delta(x - x_0)$ and this is not a

vector in the Hilbert space, since it cannot be "normalized"! It can be approximated arbitrarily closely by vectors in the Hilbert space but the "limit" lies outside the Hilbert space.

Thus, as mentioned earlier, the Segal formulation is slightly more general than the Von Neumann formulation. We shall for the time being consider only the Segal formulation.

Consider a Hilbert space with vectors φ and with 'bounded linear operators' (which may be thought of as matrices - with finite numerical elements -- in this first Hilbert space). Label these operators as $A, B, \dots, \rho, \sigma, \dots$ where ρ, σ are operators with special properties, viz.,

(i) $\rho \geq 0$, i.e. if φ is any vector $(\varphi, \rho\varphi) \geq 0$

(ii) $T_{\rho} = 1$

Irrespective of whether these properties are satisfied, we can define the 'ordinary product' $A.B$ and convert the set of all linear operators, or atleast those of some particular type, into vectors spanning a second Hilbert space by defining the quantities $A + B, (\lambda A_1 + \lambda_2 A_2)$ etc, and by defining a scalar product

$$(A, B) = T_{\rho} (A^+ B),$$

which is linear in B and antilinear in A , i.e.

$$(\lambda A, B) = \lambda^* (A, B); \quad (A, \mu B) = \mu (A, B)$$

To get a second Hilbert space we must ensure that

(i) $(A, A) \geq 0$ which follows automatically because

$$(A, A) = \text{Tr}(A^\dagger A) \geq 0$$

(ii) The square of any vector in the Hilbert space is finite; this is ensured by selecting the correct type of operators, i.e. only those with a finite 'length' (For example, we cannot choose the operator 1 , whose length = the dimension of the space, and therefore equal to ∞ for a Hilbert space). Thus, given the first Hilbert space, the second one is defined.

We have an analogue of this second Hilbert space in the phase space of classical mechanics. The numerical value assigned to a dynamical variable $A(q, p)$ depends on a phase space density function $\rho(q, p)$; it is of the form

$$\int A(q, p) \rho(q, p) dq dp$$

in quantum mechanics it takes the form

$$\text{Tr} [A\rho]$$

ρ being the density matrix. Both these can be thought of as scalar products (A, ρ) of vectors A, ρ in the second Hilbert space.

Thus what corresponds to classical mechanics is not the old Hilbert space (of von Neumann) but the second Hilbert space obtained by taking the linear operators as vectors and defining a scalar product as above.

2. The Hilbert Space Formulation of an Abstract Dynamics

Earlier we dealt with bilinear functionals i.e. the mapping of a pair of ordered elements into a number.

Now we deal with a linear mapping of the form

$$A \rightarrow \tau A$$

in an algebra where an antilinear star operation $A \rightarrow A^*$ is defined and $A^*.B^* = (B.A)^*$ i.e. $(\lambda_1 A_1 + \lambda_2 A_2)^* = \lambda_1^* A_1^* + \lambda_2^* A_2^*$ where λ_1^* is the complex conjugate of λ_1 . Generalising the concept of the trace to an arbitrary abstract linear algebra, we consider the mapping

$$A \rightarrow \langle A \rangle \quad \text{where } \langle A \rangle \text{ is a number.}$$

We call this mapping a linear functional. We require for this the following properties:

$$1 \rightarrow \text{a number } \geq 0;$$

$$\lambda_1 A_1 + \lambda_2 A_2 \rightarrow \lambda_1 \langle A_1 \rangle + \lambda_2 \langle A_2 \rangle; \text{ i.e. it is linear.}$$

To get a representation of an abstract dynamics we assert the following:

Assertion: Given any star algebra, and a distinguished positive linear functional over the algebra, we can immediately get a Hilbert space and dynamical operators such that the expectation value comes out in a natural manner.

We consider the general construction for this. Take a set of operators, A, B, C, \dots and associate a set of vectors labelled by A, B, \dots (for clarity, we denote the latter temporarily by ψ_A, ψ_B, \dots) such that the correspondence

between the operators and vectors is linear.

$$\psi_{A+B} = \psi_A + \psi_B \quad (a)$$

Take a set of operators, $A, B, C \dots$ and associate with these the quantities $\psi_A, \psi_B, \psi_C \dots$; then $\lambda A + \mu B \rightarrow \lambda \psi_A + \mu \psi_B$. Define $(\psi_A, \psi_B) = \langle A^* B \rangle$

We note that A stands, not only for an operator but also for a vector, as a scalar product $(A, B) = (\psi_A, \psi_B)$ can be defined. The vectors A form a vector space which becomes a Hilbert space if we have

$$(\psi_A, \psi_A) = \langle A^* A \rangle \geq 0 \quad \text{and if}$$

$\langle A^* A \rangle$ is bounded for all A .

We define a linear operator A on the "second Hilbert Space", i.e., the Hilbert space formed by the vectors A . A is defined by its effect on any vector ψ_B , i.e.

$$A \psi_B = \psi_{A \cdot B} \quad (\text{Definition of } A)$$

which the R.H.S. is a known vector.

We assert that $A(B \psi_c) = (A B) \psi_c$

(Proof. $A(B \psi_c) = A \psi_{B \cdot c} = \psi_{A(B \cdot c)}$ and $(A B) \psi_c = \psi_{(A \cdot B) c}$)

But $A(B \cdot C) = (A \cdot B) C$. The associative property of A, B , etc. follows).

Thus we have exhibited the dynamics by constructing a Hilbert space and exhibiting operators, $A, B \dots$ in this space and showing that these satisfy all the properties of expectation values.

An interesting vector is the vector Ω which corresponds to the identity operator:

$$\begin{aligned} 1 \cdot A &= A; \quad \psi_A = A \psi_1 = A \Omega \\ (\psi_A, \Omega) &= \langle A^* 1 \rangle = \langle A^* \rangle \\ (1, 1) &= 1 \end{aligned}$$

Hence Ω corresponds to the unit matrix in the second Hilbert space: Call Ω the vacuum state or vacuum vector; it is associated with the identity.

We shall refer to the above construction as the Gelfand-Wightman construction: (Note: (i) The zero operator will correspond to the null vector). (ii) The boundedness of A, B, etc. is ensured by considering only appropriate linear functionals which result in bounded operators; (iii) The vacuum vector Ω will be normalized only if the linear functional assigns unity to the identity operator.

(The basic requirement of the Gelfand construction is that of an associative star algebra with a bounded 'distinguished linear functional'. An associative star algebra is one in which given A, there is a corresponding element A^* such that

$$\begin{aligned} A &\rightarrow A^* \\ \lambda A &\rightarrow (\lambda A)^* = \lambda^* A^* \\ AB &\rightarrow (AB)^* = B^* A^* (\neq A^* B^* \text{ in general}) \end{aligned}$$

-- an 'anti-automorphism' (by analogy with 'anti-linear').

We have $(A^*)^* = A$.

The construction: If Ω is the distinguished linear functional which assigns to A the number $\Omega(A)$,

$$A \rightarrow \Omega(A)$$

(which is chosen such that $\Omega(1) = 1$, for convenience) then

$$(\lambda \Omega)(A) = \lambda \Omega(A)$$

If we have A, B, C,, we assign vectors ψ_A, ψ_B, \dots

to A, B, \dots respectively, $A \rightarrow \psi_A$ such that

$$A + B \rightarrow \psi_A + \psi_B$$

$$A \cdot B \rightarrow \psi_A \cdot \psi_B$$

and the scalar product $(\psi_A, \psi_B) = \Omega (A^* B)$

If $\psi_A \dots$ belong to a Hilbert space, then $(\psi_A, \psi_A) > 0$ for any positive operator A , and $(\psi_A, \psi_A) < \infty$ i.e. ψ_A is bounded. A plays a dual role, that of a linear operator A and that of a vector ψ_A in a Hilbert Space.

Consider classical mechanics, where dynamical variables A, B , are $A(q, p), B(q, p)$

We define a distinguished linear functional which assigns

$$A \rightarrow \langle A \rangle = \int A(q, p) dq dp$$

and has the properties that it takes

(i) real quantities to real numbers,

(ii) 0 to 0,

and (iii) non-negative quantities to non-negative numbers.

But unity is not taken to unity, as $\langle 1 \rangle$ is divergent.

(One must consider only functions which fall off rapidly).

Under the canonical transform

$$A \rightarrow A_\lambda = A + \lambda [A, H] + \frac{\lambda^2}{2} [[A, H], H] + \dots$$

$\langle A \rangle$ remains unchanged, as the Jacobian of the transform is 1.

(c.f. The integral for $\langle A \rangle$).

Thus in classical mechanics, we have a Hilbert space, although this Hilbert space is not the same as the one in quantum mechanics. We have functions of the form $\psi(x, y)$ rather than $\psi(x)$. But what matters is that we have a Hilbert space.

We observe that the above set of operators, A, B etc. are not exhaustive. For we have

$$AB \rightarrow AB$$

$$BA \rightarrow BA$$

and in classical mechanics, $AB = BA$ and hence AB and BA are the same; i.e. the operators corresponding to classical dynamical variables are commutative. But in a Hilbert space there always exist non-commuting operators. Hence the above set does not include all possible operators; the operators corresponding to dynamical variables are not irreducible. We can consider for example the operators $\frac{\partial}{\partial q}, \frac{\partial}{\partial p}$ in this Hilbert space; but these are not classical dynamical variables.

Consider operators formed from $\frac{\partial}{\partial q}$ and $\frac{\partial}{\partial p}$, e.g., define an operator \tilde{p} by

$$\tilde{p} A(q, p) = -i \frac{\partial}{\partial q} A(q, p)$$

and \tilde{q} by

$$\tilde{q} A(q, p) = +i \frac{\partial}{\partial p} A(q, p)$$

These are also operators in the Hilbert space but do not commute with the operators considered before. Thus we see that the set of dynamical variables in this scheme is much smaller than the set of all bounded operators.

Von Neumann's formulation postulated that to every operator in the Hilbert space of quantum mechanics there corresponds a dynamical variable and that the operators in quantum mechanics are irreducible. But we know that at least for classical mechanics this is not so; there exist non-Abelian operators not corresponding to any dynamical variable. Therefore the principle that operators

corresponding to dynamical variables are irreducible cannot be taken to be a sacred principle.

3. Galilean Mechanics in a Hilbert Space Formalism.

Consider the behaviour of a dynamical system, under changes in t (i.e. at different instants of time).

To the operator defined by

$$A_t = A_0 + t [A, H] + o(t^2)$$

there corresponds the number

$$\epsilon(A_t) = \epsilon_t(A) ; A \rightarrow \epsilon(A_t) = \epsilon_t(A)$$

We have already noted that this may be interpreted either in (i) The Heisenberg picture, as $\epsilon(A_t)$ i.e. as a time-independent functional ϵ of an operator A_t that varies with time, or (ii) the Schrodinger picture, as $\epsilon_t(A)$ i.e. as a time-dependent functional of a time-independent operator A .

But we also have a third interpretation, which we shall term (iii) The Wigner picture, where the change of a system with time is viewed as a change in the description of the system under a time translation; i.e. the system is observed by different observers (with different clocks); the question of an evolution in time does not enter as t merely labels the observer; the system is described at the same 'intrinsic' time by different observers. (Note: This is similar to the description of a space rotation as the description of a dynamical system fixed in space but viewed by two different observers using different coordinate systems, in contrast to the picture in which there is an actual rotation in space).

In the description of dynamics by Newton's laws of motion, the observer is not specified; the implication is that all observers see the same dynamics. Newton's laws can thus be taken to define the set of allowed observers, or the set of 'canonical observers'.

An equivalent set of observers, for all of whom Newton's principle of inertia holds, is given by a set in which the individual observers are related by

- (i) A time translation, specified by the parameter t
- (ii) a space translation, " " \vec{x}
- (iii) a space rotation, " " $\vec{\theta}$
- (iv) a uniform motion, specified by a velocity \vec{v} .

The restriction to uniform motion in (iv) is necessary, as the motion must be such that a particle that is 'free' for one observer is 'free' for all equivalent observers. (Note; that the concept of 'uniform' motion is a priori; any 'definition' is necessarily circular).

This family of equivalent observers is specified by the values taken by the 10 parameters given by t , \vec{x} , $\vec{\theta}$ and \vec{v} ; these give a coordinate parametrization of the system of all inertial observers.

(In a more general framework, we do not restrict ourselves to free particles; we here assume that there are no external influences but only mutual 'forces' between particles.)

Physical laws are given by equations between dynamical variables and their Poisson brackets. These equations and the P.b's must be preserved under observer \rightarrow observer transformation.

We may take as a fundamental postulate the requirement that the "rule for the conversion of observations" between two frames in the family of equivalent observers must preserve P.b.'s. This is accomplished if we have a canonical transformation connecting the two descriptions; for the continuous group of observers we require that it may be obtained by interaction of infinitesimal canonical transformations.

To verify that the P.b. will be invariant under an infinitesimal canonical transformation, we consider the change in

$$C = [A, B]$$

as a result of the infinitesimal transformations

$$A \rightarrow A + \lambda [A, H] + o(\lambda^2)$$

$$B \rightarrow B + \lambda [B, H] + o(\lambda^2)$$

$$\begin{aligned} \text{Then } C &= [A, B] + [A + \lambda [A, H], B + \lambda [B, H]] \\ &= [A, B] + \lambda \left\{ [[A, H], B] + [A, [B, H]] \right\} \\ &\quad + o(\lambda^2) \\ &= [A, B] + \lambda \left\{ -[B, [A, H]] - [A, [H, B]] \right\} \\ &= [A, B] + \lambda [H, [B, A]] \text{ by the Jacobi identity} \\ &= [A, B] + \lambda [[A, B], H] \end{aligned}$$

$$\text{But } C \rightarrow C + \lambda [C, H] + o(\lambda^2)$$

Therefore, relations between P.b.'s are preserved.

If we enquire what kind of infinitesimal canonical transforms are really acceptable, we obtain the set of all complex infinitesimal canonical transforms. However, we ignore all but real transforms; then we can confine our attention to all transfor-

mation generated by iteration of real infinitesimal transformations. The sequence of two canonical transformations is itself a canonical transformation; and the transformation with generator equal to the negative of the generator of a given transformation is the inverse of the latter.

These form a continuous group. But we are only interested in the sub group formed by those canonical transforms which are related to observer \rightarrow observer transforms. The admissible (or canonical) observers are characterized by ten 'coordinates' (i.e. the different values of the parameters $t, \vec{x}, \vec{\theta}, \vec{v}$); hence the subgroup of observer \rightarrow observer transforms are parameterized by these 10 parameters.

Thus we have a continuous and 'finite' group of canonical transforms ('finite' in the sense that the group is characterized by a finite number of parameters).

We are normally not interested in all values of the parameters, e.g. $\vec{\theta}$ can take only certain values, and in the usual cases, \vec{v} takes on values only in a small interval in real numbers. (slow travellers!) We may say that we are usually interested only in parameters taking on 'sufficiently small' values, i.e. we are usually concerned only with the subset of all observers sufficiently close to some reference observer whom we may term I.

The rule of conversion must be observer-independent; when we pass from a fixed reference observer X to another observer $Y(t, \vec{x}, \vec{\theta}, \vec{v})$ the correlation implied by the rule must not depend on the particular observer, i.e. successive transforms between observers $X \rightarrow Y, Y \rightarrow Z$ must result in a transform $X \rightarrow Z$ that is identical in form with the transform $X \rightarrow Y$.

We then obtain a realization of the 'frame family' (or the family of canonical observers) by a family of canonical transforms.

Ex. Consider the Schrodinger equation

$$i \frac{\partial \Psi}{\partial t} = H \Psi$$

The state vector Ψ_t at time t is given by

$$\Psi_t = U(t) \Psi_0 = U(t, t_0) \cdot \Psi_0$$

where $\Psi_0 \approx \Psi_{t_0}$ and U is a unitary matrix in the Hilbert space of the Ψ 's.

We must require that

$$U(t, t_0) U(t_0, t') = U(t, t')$$

to fulfil the above requirement.

We may note that the realization here by means of matrices in a Hilbert space is a very special type of realization; it is a 'representation'.

In general, in canonical mechanics, we may not be able to talk in terms of representations.

To pass from an observer X to an observer Y , we apply the transformation rules to $t, \vec{x}, \vec{\theta}, \vec{v}$; the transform depends only on the relation between X and Y and not on the parameters of X and Y themselves.

Suppose we characterize the transform

$$X \rightarrow Y \quad \text{by} \quad (t_1, \vec{x}_1, \vec{\theta}_1, \vec{v}_1)$$

$$Y \rightarrow Z \quad \text{by} \quad (t_2, \vec{x}_2, \vec{\theta}_2, \vec{v}_2)$$

and $X \rightarrow Z$ by $(t, \vec{x}, \vec{\theta}, \vec{v})$

We wish to find $(t, \vec{x}, \vec{\theta}, \vec{v})$ in terms of the sets 1 and 2.

A special case is that of pure time translation, for which $t = t_1 + t_2$; for pure space translation, $x \rightarrow x_1 + x_2$. For pure spatial rotations, we cannot write $\vec{\theta} = \vec{\theta}_1 + \vec{\theta}_2$, as finite rotations do not commute unless they are about the same axis.

Combinations of rotations and translations are more complicated.

If we ignore the possibility of relative motion between observers, then the relation between observers will be given by the parameters, t , \vec{x} and $\vec{\theta}$ and we have a 7-parameter family.

Consider the 4 parameters y_0, \vec{y} where y_0 gives the time coordinate and \vec{y} the spatial coordinates of an 'event'. The total effect of a space translation and space rotation is given by the following:

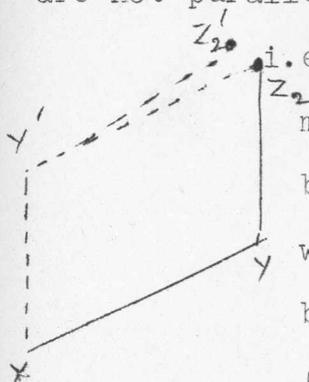
$$\begin{aligned}
 y'_0 &= y_0 \\
 y'_1 &= [R(\theta)y]_1 + x_1 \\
 y'_2 &= [R(\theta)y]_2 + x_2 \\
 y'_3 &= [R(\theta)y]_3 + x_3 \\
 \text{or } \vec{y}' &= [R(\theta)\vec{y}] + \vec{x}
 \end{aligned}$$

This particular family of observers is called the Euclidean family; the set of transformations between them form the Euclidean group in 3 dimensions, which is a 6-parameter group and a 'local group'.

To pass from the 6-parameter group to the 10-parameter group, we must define a law of addition of velocities. In Galilean mechanics, the relative velocity between observers can be arbitrarily large, but in relativity, it has an upper limit, C . The addition

of two velocities, each less than C , results in a velocity that is less than C .

Also, in relativity, the 'addition' of two velocities that are not parallel is not commutative.



i.e. If the transform from observer X to observer Y moving relative to X with a velocity \vec{v}_1 , followed by a transform to observer Z , moving relative to Y with velocity \vec{v}_2 , is given as shown in the figure, by the point Z , then the transform from X to Y' (with a relative velocity \vec{v}_2 with respect to X)

followed by a transform from Y' to Z' (with a relative velocity \vec{v}_1 with respect to Y') will result, in general, in a $z_2' \neq z_2$

An additional spatial rotation is needed to make coincide

with (this is the origin of the 'Thomas precession').

The law of composition of velocities must distinguish the mechanics being studied.

Given an 'inertial' family of frames, the principle of equivalence of dynamical laws in each of these frames is an extension of the principle of Newtonian relativity. We saw that a family of equivalent frames was labelled by 10 parameters, which we shall display by the following notation.

- time translation τ
- space translation a
- space rotation θ
- 'uniform' Relative motion v

We must find the composition rule for two transforms $X \rightarrow Y$ and $Y \rightarrow Z$. For this, we take a reference observer labelled by the values $0, 0, 0, 0$ for $t, \vec{a}, \vec{\theta}, \vec{v}$. A different observer who is not moving relative to the reference observer will be labelled by the coordinates

$$\begin{aligned}\vec{x}' &= \vec{a} + R(\vec{\theta}) \vec{x} \\ t' &= t + \tau\end{aligned}$$

where $R(\vec{\theta})$ is a function describing the spatial rotation. If the observer is moving with respect to the reference observer, an extra principle is required to specify the composition law.

For 'Galilean systems', we have

$$\begin{aligned}\vec{x}' &= \vec{a}_1 + R(\vec{\theta}_1) \vec{x} - \vec{v}_1 t \\ t' &= t + \tau\end{aligned}$$

describing a transformation from an observer at (\vec{x}, t) to one at (\vec{x}', t') . We require that if this is followed by a transformation $(\vec{x}', t') \rightarrow (\vec{x}'', t'')$ then the resultant transformation $(\vec{x}, t) \rightarrow (\vec{x}'', t'')$ must be a transformation of the same general type.

Noting that

$$\begin{aligned}\vec{x}'' &= \vec{a}_2 + R(\vec{\theta}_2) \vec{x}' - \vec{v}_2 (t') \\ &= \vec{a}_2 + R(\vec{\theta}_2) \{ \vec{a}_1 + R(\vec{\theta}_1) \vec{x} - \vec{v}_1 (t) \} - \vec{v}_2 (\tau + t) \\ &= \vec{a}_2 + R(\vec{\theta}_2) \vec{a}_1 - \vec{v}_2 \tau + R(\vec{\theta}_2) R(\vec{\theta}_1) \vec{x} \\ &\quad - [\vec{v}_2 + R(\vec{\theta}_2) \vec{v}_1] t\end{aligned}$$

we obtain $\vec{x}'' = \vec{a} + R(\vec{\theta}) \vec{x} - \vec{v} t$

$$t'' = t + \tau$$

where

$$\begin{aligned}\tau &= \tau_1 + \tau_2 \\ \vec{a} &= \vec{a}_2 + R(\vec{\theta}_2) \vec{a}_1 - \vec{v}_2 \tau_1 \\ R(\vec{\theta}) &= R(\vec{\theta}_2) R(\vec{\theta}_1)\end{aligned}$$

The composition law for the parameters of the transform.

Also $x_n = f_n(e, x) = f_n(0, 0, \dots, 0_n, x_1, \dots, x_n)$
 $e = e \cdot e \dots e_n = 0 = f_n(e, e)$

The requirement (iii) tells us that given x near enough to e we can find a y such that $f_n(x, y) = 0$

The requirement (i) of associativity imposes a further restriction on $f_n(x, y)$; e.g. evaluate the coordinates in two different ways, for small values of x, y .

Assume that $f_n(x, y)$ is analytic at the origin and thus can be expressed as a power series about the origin, write

$$f_n(x, y) = \alpha_n + \sum_s \beta_{ns} x_s + \sum_s \gamma_{ns} y_s + a_{st}^n x_s x_t + b_{st}^n x_s y_t + d_{st}^n y_s y_t + O(3)$$

Using the property $f_n(x, e) = x_n$ we get

$$x_n = \alpha_n + \sum_s \beta_{ns} x_s + b_{st}^n x_s x_t + O(3)$$

or

$$\alpha_n = 0 \quad \beta_{ns} = \delta_{ns}$$

Similarly,

$$f_n(e, x) = x_n \text{ gives}$$

$$b_{st}^n = 0 = d_{st}^n, \quad \gamma_{ns} = \delta_{ns}$$

Therefore, finally, we must have f_n to be of the form

$$f_n(x, y) = x_n + y_n + a_{st}^n x_s y_t + O(3)$$

If the terms of order 3, terms with purely x or purely y must be 0; while there are non-zero terms of the form $x x y, x y y$ etc.

We define $C_{st}^n = a_{st}^n - a_{ts}^n$ these

are termed the structure constants of the group; they contain the essential part of the structure of the group.

The maximum number of non-trivial constants entering the structure of the group is $\frac{n^2(n-1)}{2}$. For an Abelian group $(xy = yx)$ the $C_{s,t}^{\pi}$ are zero; only for a non-commutative group are the structure constants non-zero.

There is a slightly different way of defining the structure constants $C_{s,t}^{\pi}$

Take $xyx^{-1}y^{-1} = (xy)(xy)^{-1} = q$ which is defined as soon as xy is defined; we try to expand this in terms of x, y . (if x, y commute, $q = 1$.) When x, y do not commute, we have for the coordinates of q ,

$$(xyx^{-1}y^{-1})_{\pi} = C_{s,t}^{\pi} x_s x_t + \dots$$

the structure constants are thus defined as the coefficients of the leading terms.

As a result of the associative property of the multiplication, the structure constants satisfy an identity, the Jacobi identity, which is a bilinear identity of the form

$$C_{st}^{\pi} C_{uv}^{\tau} + \dots + \dots = 0$$

Thus not all antisymmetric sets of constants C_{st}^{π} are admissible, but only those obeying the Jacobi identity.

5. Lie Algebras.

We construct a vector space E_n . (we note that in the above we did not use the vector space property.) For this, we take the n -dimensional space near the identity and define an addition

$$(x + y)_{\pi} = x_{\pi} + y_{\pi}$$

and a scalar multiplication

$$(\lambda x)_{\pi} = \lambda x_{\pi}$$

Then we obtain a vector space.

Now define a product $[x, y]$ in this space which maps pairs of vectors into single vectors, with the following properties

$$(i) \quad [x, y] = -[y, x]$$

(ii) It is linear in the first argument, i.e.

$$[\lambda x + \mu y, z] = \lambda [x, z] + \mu [y, z]$$

Then it is automatically linear in the second argument.

(iii) To make the vector space into a Lie algebra, we require the product to satisfy the Jacobi identity

$$[[x, y], z] + [[y, z], x] + [[z, x], y] = 0$$

where '0' is the null vector.

Let e^α , $\alpha = 1, \dots, n$, be a set of linearly independent vectors, then we can expand

$$x = \sum_{\alpha=1}^n x_\alpha e^\alpha$$

Expand the vector $[e^\alpha, e^\beta]$ as

$$[e^\alpha, e^\beta] = \sum c^\gamma e^\gamma$$

and write $c^\gamma = c^\gamma(\alpha, \beta) = C_{\alpha, \beta}^\gamma$

The $C_{\alpha\beta}^\gamma$ have properties similar to the $C_{\delta\epsilon}^\pi$

(i) They are antisymmetric in α, β ,

$$C_{\alpha\beta}^\gamma = -C_{\beta\alpha}^\gamma, \quad \text{by (i) above for } [x, y]$$

(ii) the condition (ii) on $[x, y]$ does not lead to any specific restriction,

(iii) they satisfy the condition

$$\sum_{\delta} \left\{ C_{\alpha\beta}^\delta C_{\delta\gamma}^\epsilon + C_{\beta\gamma}^\delta C_{\delta\alpha}^\epsilon + C_{\gamma\alpha}^\delta C_{\delta\beta}^\epsilon \right\} e^\epsilon = 0$$

(To see this, note that

$$\begin{aligned}
 [[e^\alpha, e^\beta], e^\gamma] &= \left[\sum_{\delta} c_{\alpha\beta}^{\delta} e^{\delta}, e^{\gamma} \right] = \sum_{\delta} c_{\alpha\beta}^{\delta} [e^{\delta}, e^{\gamma}] \\
 &= \sum_{\delta} c_{\alpha\beta}^{\delta} \sum_{\epsilon} c_{\delta\gamma}^{\epsilon} e^{\epsilon}
 \end{aligned}$$

and use the Jacobi identity for the product

$$[e^\alpha, e^\beta]$$

Then we get

$$\sum_{\delta, \epsilon} \left\{ c_{\alpha\beta}^{\delta} c_{\delta\gamma}^{\epsilon} + \dots \right\} e^{\epsilon} = 0.$$

but the e^{ϵ} being linearly independent, this is possible only if the coefficients $\sum_{\delta} \left\{ \dots \right\}$ are each $\equiv 0$.

From this similarity between the $c_{\alpha\beta}^{\gamma}$ and the $C_{\alpha\beta}^{\gamma}$ we may ask whether the two are related in any way.

Given any local Lie group, there exists an associated Lie algebra, i.e. we can define a vector space with a Lie algebra structure defined by the set of constants C (obtained from the local Lie group).

Also, given any Lie algebra, we can determine a local Lie group with structure constants $C_{\alpha\beta}^{\gamma}$ equal to the $c_{\alpha\beta}^{\gamma}$.

This is a remarkable fact, if we recall that the structure constants are determined by the (antisymmetric part of) the bilinear terms in the power series expansion of the composition functions $f_n(x, y)$. The statement made is thus equivalent to the assertion that (apart from "trivial" changes in parametrization) the local Lie group composition function is determined by its structure constants. A demonstration of this assertion is beyond the scope of these lectures; for a thorough treatment, compare the books by Pontrjagin, ("Topological Groups") and Hamermesh ("Group Theory"). We simply state the result:

Theorem:

Given the structure constants, the local Lie group is 'essentially determined'.

Ex. Consider the group defined by

$$\xi^x \rightarrow \xi + x^3$$

Given two transformations x, y , these are commutative,

$$\xi^{(x,y)} \rightarrow \xi + (x^3 + y^3)$$

but the composition law $\eta = (x^3 + y^3)^{1/3}$ is not analytic in x, y at $x = 0, y = 0$.

But this difficulty could be easily removed by defining new parameters

$$x' = x^3, y' = y^3, \text{ so that } \eta' = x' + y'$$

It is because of this freedom of parametrization that the structure constants determine the group only 'essentially'. In fact, the structure constants give the 'normal coordinates' form of the parameters (See Pontrjagin and Hamermesh for more details).

Given the structure constants of the Lie group, we can define a Lie algebra with the same constants and thus deduce a Lie algebra structure for the physical quantities. But there are difficulties in doing this, e.g. the structure constants determine the group only locally and not completely; the group as a whole may have other properties not determined by the structure constants. For example both displacements along a fixed direction and rotations around a fixed axis have the same (trivial) Lie algebra and hence the same local group. But the global groups are clearly different.

The Galilean transformations constitute a local Lie group. A particular transformation defines a point in a 10-dimensional

space; all group properties are satisfied and the composition function is analytic.

To find the structure constants of the Galilean group, we go from the continuous group to a finite-dimensional Lie algebra. We have 1000 coefficients, but most of them are zero !

Consider the simplest non-trivial subgroup of the Galilean group, viz., the Euclidean group in 2-dimension. (Note: the translation subgroup is Abelian, and hence its structure constants are = 0).

We consider the Euclidean group such that \vec{a} lies in the xy plane and $\vec{\theta}$ lies along the z axis. The most general transform is then

$$x' = x \cos \theta - y \sin \theta + a_1,$$

$$y' = x \sin \theta + y \cos \theta + a_2.$$

Take two transforms given by (θ, a_1, a_2) and (θ', a_1', a_2') with small values of θ, a_1, a_2 etc. We compute all the brackets.

We choose the basis

$$e^1 \rightarrow P_1 \quad \text{a displacement along the } x \text{ axis.}$$

$$e^2 \rightarrow P_2 \quad \text{a displacement along the } y \text{ axis.}$$

$$e^3 \rightarrow J \quad \text{a rotation in the } xy \text{ plane}$$

Then the Lie algebra associated with this local Lie group is given by the relations

$$[J, P_1] = -P_2; [J, P_2] = +P_1; \text{ and } [P_1, P_2] = 0$$

Next we consider the 6-parameter Lie group of 3 dimensional spatial rotations and translations; this has a Lie algebra defined

by the following bracket relations:

$$0 = [P_1, P_2] = [P_2, P_3] = [P_3, P_1]$$

$$0 = [J_1, P_1] = [J_2, P_2] = [J_3, P_3]$$

$$[J_1, P_2] = -[J_2, P_1] = P_3 \quad \text{and cyclic permutations}$$

$$[J_1, J_2] = J_3 \quad \text{and cyclic permutations.}$$

The complete 10-parameter group is more complicated: For this, we have the following operators for the Lie algebra;

H corresponding to time translation

\vec{P} corresponding to space translation

\vec{J} corresponding to space rotation,

\vec{G} corresponding to uniform motion

The composition law must be known. The Lie algebra is then given by the Lie bracket relations between the 10 quantities

LIE
ALGEBRA FOR
THE GALILEAN
GROUP.

$$\begin{aligned}
[H, P_j] &= 0, [H, J_j] = 0, [H, G_j] = -P_j \\
[P_j, P_k] &= 0, [J_j, J_k] = \epsilon_{jkl} J_l \quad \begin{matrix} j, k, l \\ \rightarrow 1, 2, 3 \end{matrix} \\
[J_j, P_k] &= \epsilon_{jkl} P_l \\
[J_j, G_k] &= \epsilon_{jkl} G_l; [G_j, G_k] = 0 \\
[G_j, P_k] &= 0
\end{aligned}$$

These bracket relations may be "physically interpreted": e.g. Time translations commute with spatial translations and rotations, but not with transformations to a moving frame, e.g. Time translatio

followed by Movement gives

$$\begin{aligned}
 x &\rightarrow x \rightarrow x - v(t + \Sigma) \\
 t &\rightarrow t + \Sigma \rightarrow t + \Sigma
 \end{aligned}$$

whereas Movement followed by time translation gives

$$\begin{aligned}
 x &\rightarrow x - vt \rightarrow x - vt \\
 t &\rightarrow t \rightarrow t + \Sigma
 \end{aligned}$$

so that there is a net spatial displacement of $v\Sigma$

The other relations can be similarly interpreted. The relation $[J, G] \neq 0$ may be interpreted either as above, by noting that

$$\begin{aligned}
 x &\rightarrow Rx \rightarrow Rx - vt \\
 t &\rightarrow t \rightarrow t
 \end{aligned}$$

is different from

$$\begin{aligned}
 x &\rightarrow x - vt \rightarrow Rx - Rvt \\
 t &\rightarrow t \rightarrow t
 \end{aligned}$$

OR by saying that when we look at the same rotation in space from two frames in relative motion, then we have a relation between the two observed values of the magnitude of the rotation such that the difference is proportional to the movement. The relation $[G_j, G_k] = 0$ tells us that the addition theorem for the velocities must be symmetric in the two frames.

The rule of conversion between equivalent observers corresponds to a canonical transformation which reflects the local Lie group. We must find the infinitesimal generators of the (infinitesimal) canonical transformations i.e. we must find the quantities x such that in the transformation of a dynamical variable

$\phi(q, p)$ between equivalent observers we have

$$\begin{aligned}\phi(q, p) &\rightarrow \phi_v(q, p) \\ &= \phi + v [\phi, x] + \frac{v^2}{2!} [[\phi, x], x] + \dots\end{aligned}$$

which will correspond to the Galilean generator \vec{G} etc.

Instead of studying the rule of conversion of observations, we deal with the structure constants of the Lie group, which then reflect and instead of the latter we study its Lie algebra, i.e. the object is then to obtain a set of dynamical variables $\vec{J}, \vec{P}, \vec{H}, \vec{G}$. Satisfying the bracket relations satisfied by the corresponding elements of the Lie algebra, e.g. In classical mechanics, the Lie bracket is a P.b. and we look for numerical valued functions $g(q, p)$ etc. corresponding to the operators $\vec{G}, \vec{H}, \vec{P}, \vec{J}$.

In quantum mechanics, the Lie bracket is a commutator bracket, and we must look for a set of operators corresponding to $\vec{G}, \vec{H}, \vec{P}, \vec{J}$.

.....

III RELATIVISTIC DYNAMICS1. Realization of the Galilei Group

Consider a dynamics invariant under rotations, where we have an ∞^3 of equivalent observers obtained from a standard observer by a rotation through an angle θ . With any function $f(q,p)$ for the standard observer, we want to associate a function $f_\theta(q,p)$.

$$f(q,p) \rightarrow f_\theta(q,p)$$

such that

$$\lambda f + \mu g \rightarrow \lambda f(\theta) + \mu g(\theta)$$

We require that a relation of the form $h = [f, g]$ should go over into one of the form $h_\theta = [f_\theta, g_\theta]$. This invariance is automatically satisfied by requiring that there must exist a canonical transformation group realising the rotation group.

The composition $\vec{\theta}$ of two transforms $\vec{\theta}_1$ and $\vec{\theta}_2$ must obey the law that the quantity

$$\vec{\theta} = f(\vec{\theta}_1, \vec{\theta}_2)$$

must satisfy the same general transformation law as the rotations through $\vec{\theta}_1, \vec{\theta}_2$.

A copy of the set of transformation by a set of canonical transformations is called a realization. We wish to exhibit a canonical transformation that is of the same nature as the infinitesimal elements of the group. Instead of considering an infinitesimal transformation, it is sufficient if we consider the difference between the transformed quantities and the

original quantities, i.e. if the transformation is given by

$$f_\lambda = f_0 + \lambda [f, A] + \dots$$

the quantity

$$(f_\lambda - f_0) = \lambda [f, A]$$

characterises the rotation group. Next, we consider the bracket appearing in the transformation, i.e. instead of the transformation itself we take the coefficients of the transformation, or in other words, we consider the elements of the group in terms of the elements of the Lie algebra. We then look for dynamical variables with the same composition law as the law for the elements of the Lie algebra (i.e. the bracket relations for the Lie algebra). If we take the elements J_1, J_2, J_3 of the Lie Algebra, then we know the bracket relation

$$[J_1, J_2] = J_3$$

(a)

We find that the association $\vec{J} = \vec{q} \times \vec{p}$

(b)

i.e. $J_1 \rightarrow [q_2 p_3 - q_3 p_2]$ etc. has the property

$$J_3 = [J_1, J_2] \rightarrow [(\vec{q} \times \vec{p})_1, (\vec{q} \times \vec{p})_2] \quad \text{p.b.} = (\vec{q} \times \vec{p})_3$$

i.e. the association (b) is consistent with the bracket relation

(a). For the Galilean group we have the 10 generators

$P_j, H, J_j,$ and $G_j \quad j = 1, 2, 3, \dots$

The composition law for these are

$$\begin{aligned}
 \text{i)} \quad [p_j, p_k] &= [p_j, H] = 0 \\
 \text{ii)} \quad [J_j, p_k] &= \epsilon_{jkl} p_l \\
 \text{iii)} \quad [J_j, J_k] &= \epsilon_{jkl} J_l \\
 \text{iv)} \quad [J_j, G_k] &= \epsilon_{jkl} G_l \\
 \text{v)} \quad [J_j, H] &= 0 \\
 \text{vi)} \quad [p_j, G_k] &= 0 \\
 \text{vii)} \quad [G_j, G_k] &= 0 \\
 \text{viii)} \quad [G_j, H] &= p_j
 \end{aligned}
 \tag{I}$$

We must look for 10 functions of q, p with these composition laws; then we would have a true realization. We find that we cannot get exactly such functions but only something nearly as good.

Let us first ignore \vec{G} . Then the following quantities obey the correct composition laws:

$$p_j(q, p) = p_j, \quad H(q, p) = \frac{1}{2m} \sum_j p_j^2$$

$$J_j(q, p) = \epsilon_{jkl} q_k p_l$$

For G_j we must find functions $G_j(p, q)$ satisfying $[G_j, G_k] = 0$. Therefore G_j must depend only on p or q . But $[G_j, H] \neq 0$ and H is a function of p only; hence G_j must be a function of q only.

We must also have $[J_j, G_k] = \epsilon_{jkl} G_l$. The possible choice $G_j = m q_j$ added to the set (A) would result in practically all the required relations being satisfied. Only instead of having

$$[P_j, G_k] = 0 \quad \text{we have}$$

$$[P_j, G_k] = -m \delta_{jk}$$

where m is a number that is not a function of q and p . Hence, although $[P_j, G_k]$ is not zero, we can achieve the same results as far as generating a transformation is concerned, for the quantity $-m \delta_{jk}$ has vanishing bracket with all $f(q, p)$:

$$[f(q, p), -m \delta_{jk}] = 0$$

The quantity $-m \delta_{jk}$ being a neutral element, has all P.b.'s $= 0$ and is thus 'almost as good as zero'.

$$[P_j, G_k] \neq 0 \quad \text{but} \quad [P_j, G_k], f(q, p) = 0 = [0, f(q, p)]$$

Such a realization is known as a projective realization or "a realization up to an exponent".

If $m = 0$, H is not well-defined and we do not have a realization.

True realizations do exist, although they are not of the type we need. As examples, we may take

$$1) P_j \rightarrow 0; \quad G_j \rightarrow q_j; \quad H \rightarrow 0; \quad J_j \rightarrow \epsilon_{jkl} \cdot q_{kpl}$$

But this does not correspond to a particle; it corresponds to a translationally invariant system with energy $= 0$. Although

This is a 'true realization, it is not one-to-one, i.e. not a 'faithful' realization; we may look for a realization that is faithful and true.

2) All the quantities $\rightarrow 0$; this is a trivial realization.

2) $J_j \neq 0$; all other quantities $\rightarrow 0$.

In quantum mechanics, we consider a special type of realization viz., 'representations' by means of linear operators in a Hilbert space. In classical mechanics, we have fewer methods; for a particular method see A. Loinger, Annals of Physics, We may introduce a classical spin variables S .

$$\vec{J} \rightarrow (\vec{q} \times \vec{p}) + \vec{S}; \quad [S_1, S_2] = S_3; \quad [S_i, q] = 0 = [S_i, p]$$

Dynamical variables here are functions of \vec{q} , \vec{p} and S .

2. The Lorentz Group and its Realizations.

Here, too, we have 10 generators and similar relations, except that G_j is replaced by K_j , the generator of the transform to a moving frame. G_K in the fourth relation in (I) is replaced by K_K , while the relations (vi), (vii), (viii)

of (L) are replaced by $[P_j, K_k] = -\delta_{jk} H$ (vi)

$$[K_j, K_k] = -\epsilon_{jkl} J_l \quad (\text{vii})$$

and $[K_j, H] = P_j$ (viii)

M. Hammermesh: Group Theory.

Inonu and Wigner: Nuovo Cimento.

Bargmann: Ann. Maths. (1947), p. 1

These changes are trivial for the relation (iv) and the last relation.

The Lorentz group may be viewed geometrically as the group of rotations and translations in a 4-space. The relations between P_j , H , J_j , K_j may be interpreted as the following features of the composition laws.

Rotation + rotation = a rotation, by (vii)

Moving transformation + Moving transformation = a general rotation type of transformation by relation (vii)

(iii), (iv) and (vii) are an extension of the relation

$$[J_j, J_k] = \epsilon_{jkl} J_l$$

Similarly the relation rotation + translation = translation, expressed by (ii), is extended into (ii) and (vi). The whole scheme may be viewed geometrically by defining

$$P_\mu = \begin{cases} P_j, & \mu = j = 1, 2, 3 \\ H, & \mu = 0 \end{cases}$$

and

$$J_{\mu\nu} = -J_{\nu\mu} = \begin{cases} J_3, & \text{if } \mu=1, \nu=2 \text{ and cyclic} \\ K, & \text{if } \mu=0, \nu=1 \text{ and cyclic} \end{cases}$$

For the Lorentz group, we can obtain a realization by a particle that is both faithful and true.

$$\text{e.g. } P_j(q, p) = P_j; \quad H = +\sqrt{p^2 + m^2} = \omega$$

$$\text{in general, } H = (\sum p_j^2 + m^2)^{1/2}$$

$$K_j = q_j \sqrt{p^2 + m^2} = \omega q_j$$

$$J_j = \epsilon_{jkl} q_k p_l$$

(As we are dealing with classical mechanics, we have not paid any attention to the order of the q 's and P 's).

We now have

$$\begin{aligned} [P_j, K_k] &= [P_j, \omega q_k] \\ [K_j, K_k] &= [\omega q_j, \omega q_k] \\ &= \omega [q_j, \omega q_k] + q_j [\omega, \omega q_k] \\ &= \omega q_k \frac{\partial \omega}{\partial p_j} - q_j \omega \frac{\partial \omega}{\partial p_k} \\ &= \omega q_k p_j / \omega - q_j p_k \\ &= q_k p_j - q_j p_k = -J_l \end{aligned}$$

In the absence of spin, K_j is like a relativistic moment. For the corresponding quantum mechanical case, refer: Wigner Ann. Math. 1939, (Unitary Representations of the Lorentz group)

Here, we have no projective representations; all representations are faithful and real.

Consider the 3-dimensional rotation group. We have

$$[J_j, J_k] = \epsilon_{jkl} J_l$$

$A = J^2 = \sum J_j^2$ is a constant of the system $[A, J_j] = 0$
 and $[A, f(J)] = 0$ where $f(J)$ is any polynomial in J . In
 classical mechanics, J^2 may be any non-negative number, while in
 quantum mechanics, it is quantized and is equal to $j(j+1)$,
 $J = 0, 1/2, 1, \dots$

Generalizing from this observation, we may consider the
 Casimir invariants. These are not infinitesimal generators;
 they are not even in the Lie algebra; but they can be made up
 out of elements of the Lie Algebra by using an ordinary product;
 and have vanishing brackets with all elements of the Lie Algebra.
 For the Euclidean group characterized by the relations

$$[J_i, J_k] = \epsilon_{ijk} J_l, \quad [J_i, P_k] = \epsilon_{ijk} P_l,$$

$$[P_i, P_k] = 0$$

we have the three Casimir invariants

$$A = \sum_j J_j \cdot J_j, \quad B = \sum_j P_j \cdot P_j \quad \text{and} \quad C = \sum_j C_j \cdot P_j$$

A , B and C commute with everything and must be represented by
 numbers. (Can we prove that these are the only invariants?

This can be shown only for a certain class of groups). For
 the Lorentz group characterised by the relations.

$$[J_i, K_k] = \epsilon_{ikl} K_l, \quad [K_j, K_k] = \epsilon_{jkl} J_l$$

$$[K_i, H] = P_i, \quad [K_j, P_k] = \delta_{jk} H$$

$$[H, P_k] = 0$$

we have the invariants $B = \sum P_j^2 = p^2 - H^2$
 $D = \sum_j R_j \cdot R_j - R_0^2$ where $R_\mu = \epsilon_{\mu\nu\alpha\beta} P^\nu J^{\alpha\beta}$

the outer product of the vector P^ν and the tensor $J^{\alpha\beta}$.

Note: $R_j = H J_j - (\vec{P} \times \vec{J})_j$; $R_0 = \vec{P} \cdot \vec{J}$.

(D is a sort of complicated generalisation of C).

For a particle at rest, $P = 0$, $H^2 = m^2$; $R_j = m J_j$. Thus we obtain the values of the invariants as

$$B = -m^2; \quad D = m^2 S^2 = (\text{mass})^2 \times (\text{Spin})^2$$

where in quantum mechanics, the $(\text{spin})^2$ is quantized, while the first factor, the $(\text{mass})^2$, is always unquantized. For the standard particle-like representations, we have $D \rightarrow m^2 j(j+1)$, $2j$ integral. Besides these, we have a number of "useless representations". e.g. if $m^2 = 0$ and $D \neq 0$, then $S^2 = \text{infinity}$, and we have particles with infinite spin. We tabulate below the "useless representations."

	B	D
Ordinary Particle-like representations	< 0	> 0
	0	(i) $\neq 0 =$ Particles with infinite spin.
	0	(ii) $= 0$, (if $m = 0$)
	0	(i) $S = 0$
	(m pure imaginary)	(ii) $s = \infty$

To see the implications of m being imaginary, consider the relation

$$p = \frac{mv}{(1-v^2)^{1/2}}$$

If p is real and m is imaginary, then (assuming that the particle has a real velocity v) we must have $1-v^2 < 0$ i.e.

$v > 1$ i.e. a particle with imaginary mass can travel only with a velocity $> c$ (For comparison, recall that when m is real and $\neq 0$, the particle velocity must be $< c$ and when $m = 0$, we must have $v = c$).

When $v > c$ its maximum possible value is infinite, The momentum of the particle has then a magnitude which is $|m|$, but may have an arbitrary direction. The energy vanishes since $m/\sqrt{1-v^2} = 0$. If from this extreme value, we increase the particle, energy, keeping m fixed, then v^2 becomes smaller and smaller, its minimum possible value being $= 1$ (or c).

We propose the following nomenclature for particles with different m^2 and s^2 .

$m^2 > 0$	Slugs	1
$m^2 = 0$	Angles	
$m^2 < 0$	Spooks	
$s^2 = \infty$	Goblins	
$P, H = 0$	Fairies.	

Some of the properties of "spooks" are discussed in O.M. Bilanuk, V.K.Deshpande and E.C.G.Sudarshan:

'On Meta-Relativity' American Journal of Physics
30, 718 (1962)

3. Interacting Relativistic Systems:

Instead of having the functions presented above, suppose we have the following:

$$\vec{P} = \vec{P}_{(1)} + \vec{P}_{(2)}; H = \frac{p(1)^2}{2m(1)} + \frac{p(2)^2}{2m(2)}$$

$$J = \vec{q}(1) \times \vec{p}(1) + \vec{q}(2) \times \vec{p}(2)$$

$$\vec{G} = m(1) \vec{q}(1) + m(2) \vec{q}(2)$$

This set of generators also satisfied the required conditions, but is obtained simply from two others (viz., those with $\vec{q}(1)$, $\vec{p}(1)$ and $\vec{q}(2)$, $\vec{p}(2)$). Such representations are 'reducible' and are obtained by taking direct sums of representations. \vec{P} , \vec{H} , \vec{J} , \vec{G} , are all sums of two similar function (1) and (2); in terms of particles one would say there were two particles, \vec{P} , \vec{H} , \vec{J} , \vec{G} all being additive. Such a system is called a non-interacting system or a free system. (Normally, by a free system we mean one in which H is additive; there we require that all the generators be additive).

More interesting are 'interacting' systems, an interacting system being defined as one (with say, two component particles) which has the following properties:

- (1) The total momentum or angular momentum = the sum of the individual momenta or angular momenta; but
- (2) the energy H is not the sum of the individual energies.

On the other hand

$$H = \frac{p(1)^2}{2m(1)} + \frac{p(2)^2}{2m(2)} + V(\vec{q}(1), \vec{q}(2), \vec{p}(1), \vec{p}(2))$$

Similarly G is not in general, the sum of the individual G 's:

$$G = m(1) \dot{\vec{q}}(1) + m(2) \dot{\vec{q}}(2) + W(\vec{q}(1), \vec{q}(2), \vec{p}(1), \vec{p}(2))$$

We further have the intuitive requirement that when the separation of the particles is sufficiently large, they should behave as if they were free, i.e.

$$\text{as } |\vec{q}| \rightarrow \infty, V \rightarrow 0$$

We ask what are the restrictions on the functions V and W .

First we note that the requirements that $[P_j, H] = 0$ implies that

$$[P_j, -V] = 0$$

as we have

$$[P(1)^2/2m, \vec{P}] = 0$$

Further, we must require that

$$[P_j, G_k] = -\delta_{jk} (m(1) + m(2)) \quad \text{but we already have}$$

$$[P_j, m(1) \dot{\vec{q}}(1) + m(2) \dot{\vec{q}}(2)] = -\delta_{jk} (m(1) + m(2));$$

hence it follows that $[P_j, W_k] = 0$

We note that a quantity $F(\vec{q}_1, \vec{q}_2, \vec{p}_1, \vec{p}_2)$ satisfying $[F, P] = 0$ depends only on the difference of the coordinates, $(\vec{q}(1) - \vec{q}(2))$ and not on the total coordinates. Applied to

$V = V(\vec{p}(1), \vec{p}(2), \vec{Q}, \vec{q}(1) - \vec{q}(2))$ where $\vec{Q} = \vec{p}(1) + \vec{p}(2)$, we note that $[P_j, V] = 0$ implies that

$$\frac{\partial V}{\partial p_j^{(1)}} + \frac{\partial V}{\partial p_j^{(2)}} = 0$$

and hence that $V = V(\vec{p}(1), \vec{p}(2), \vec{q}(1) - \vec{q}(2))$

V depends on $(\vec{q}(1) - \vec{q}(2))$ but not on \vec{Q}

Similarly for \vec{W} .

The bracket relation $[J_j, H] = 0$ implies that H must be rotationally invariant; this means that $H(\vec{p}(1), \vec{p}(2), \vec{q}(1), \vec{q}(2))$

must be a function only of

$$p^{(1)2}, p^{(2)2}; \vec{p}^{(1)} \cdot \vec{p}^{(2)}; (\vec{p}^{(1)} \cdot \vec{q}) (\vec{p}^{(2)} \cdot \vec{q}) \\ (\vec{p}^{(1)} \times \vec{p}^{(2)}) \cdot \vec{q}$$

\rightarrow

W must be = (an invariant function (of the above variables)

\times a vector), or the sum of such quantities.

Now consider bracket relations involving \vec{G} and H .

$$[G_k, H] = P_k$$

OR

$$[G_k^0 + W_k, H^0 + V] = [G_k^0, H^0] + [W_k, H^0] \\ + [G_k^0, V] + [W_k, V] \\ = [G_k^0, H^0]$$

OR

$$[W_k, H^0] + [G_k^0, V] + [W_k, V] = 0$$

which is a non-linear partial differential equation connecting \vec{W} and V . Therefore given one of the functions \vec{W} and V , the other is determined. We do not attempt to find the general solution of the equation above, but consider the most interesting particular solution.

Take an interacting system ($V \neq 0$) but with $\vec{W} = 0$.

Then we have the relation $[G_k^{(0)}, V] = 0$ i.e.

$$\left[\left(m^{(1)} q_k^{(1)} + m^{(2)} q_k^{(2)} \right), V \right] = 0$$

$$m^{(1)} \frac{\partial V}{\partial p_k^{(1)}} + m^{(2)} \frac{\partial V}{\partial p_k^{(2)}} = \frac{\partial V}{\partial \left(\frac{p_k^{(1)}}{m^{(1)}} \right)} + \frac{\partial V}{\partial \left(\frac{p_k^{(2)}}{m^{(2)}} \right)} = 0$$

Thus rather than $\vec{p}^{(1)}$ and $\vec{p}^{(2)}$, the velocities $\vec{p}^{(1)}/m^{(1)}$ and $\vec{p}^{(2)}/m^{(2)}$ seem relevant.

Define $\vec{V} = \frac{\vec{p}^{(1)}}{m^{(1)}} - \frac{\vec{p}^{(2)}}{m^{(2)}}$

and $\vec{V}' = \frac{\vec{p}^{(1)}}{m^{(1)}} + \frac{\vec{p}^{(2)}}{m^{(2)}}$

For free particles, these would be the difference and sum respectively, of the velocities. For particles that are not free, we may have velocity-dependent forces. The relation (1) then tells us that V depends only on the difference \vec{v} of the velocities, i.e. on the relative velocity of the two particles and not on \vec{v} (or equivalently, on their absolute velocities).

Thus we have obtained the result that the 'potential' between the particles can depend only on the relative position and the relative velocity of the two particles and that it must be an invariant function of these:

$$V = V(\vec{r}, \vec{q})$$

This is the most general possibility, where we have assumed that only the energy is a dynamic generator and all the others are kinematic generators. (We recall that a generator which is unchanged on going from a free to an interacting system is called a 'kinematic generator', while one that is changed is called a 'dynamic generator').

If we wish to consider the 'moment' \vec{G} also as a dynamic generator, we shall then obtain an infinity of solutions, with W differing from zero.

(Note: In their work on determining all the representations of the Galilean group, Eisenbud and Wigner* implicitly assumed G to be a dynamic generator).

Now consider the Lorentz group. We here have an infinity of interacting particle realisations. As before, while going from a free to an interacting system, we add a function V to H and \vec{W} to \vec{K} , and leave \vec{P} , \vec{J} unchanged. But we cannot take, as a special case, $\vec{W} = 0$, for we have

$$[K_j, P_k] = \delta_{jk} H$$

Now, H is changed, and since \vec{P} is unchanged, \vec{K} must change, i.e. $\vec{K} \neq \vec{K}_0$ or $\vec{W} \neq 0$. Thus we again obtain a nonlinear partial differential equation similar to the one obtained before (). What is the nature of the solution of this?

*L. Eisenbud and E.P. Wigner, Proc. Nat. Acad. Sc. U.S. (1948).

We again find that V is independent of $\vec{Q} = \vec{q}(1) + \vec{q}(2)$ and depends only the relative position.

What can we say about \vec{W} ? We know it must transform as a vector. Can we make any statement about \vec{W} similar to the one we made for the Galilean group (namely, that $\vec{W} = 0$ or G is a kinematic generator).

If, for transformation to a moving system, we add the postulate that

$$\vec{p}^{(i)} \rightarrow \vec{p}^{(i)} + m^{(i)} \vec{v} \quad (\text{true for free particles})$$

then the Galilean generators are determined almost completely. Similarly for the Lorentz group, we could add the postulate that a world line $q(t)$ goes over into another world line $q'(t')$ on transformation to a moving system, $q'(t')$ being obtained from $q(t)$ by a 'Lorentz transformation':

$$q(t) \rightarrow q'(t') \text{ for } \underline{\text{free particles}}$$

We now ask what restrictions would result if we assumed the above for interacting particles also.

The restriction obtained is that we must have

$$V = 0, \vec{W} = 0$$

i.e. the system must be free

Thus if we have (invariant) world lines at all, they must be straight.* A Hamiltonian theory with curved lines would be inconsistent. (Note: All the above is for particles obeying classical mechanics).

*Ref. D.G.Currie: J.Maths. Phys. (to be published)
 D.G.Currie, T.F.Jordan and E.C.G.Sudarshan, Rev. Mod. Phys.
 (in press)

4. Trajectory Dynamics

We consider a third variety of dynamics, descriptive dynamics, e.g. the dynamics defined by Kepler's laws, where we are given the law describing the orbit but not told anything about the dynamical principle that gives rise to the equation of motion.

Consider the conversion of an abstract algebra into an algebra with operators. In addition to the distinguished linear functional defined previously,

$$A \rightarrow A, \quad A \rightarrow \Omega(A):$$

we can have other linear functionals. Here we consider the 'density functional' which is a linear functional with the properties

$$\begin{aligned} A &\rightarrow \rho(A) \\ \rho(A^\dagger A) &\geq 0 \\ \rho(A^\dagger) &= [\rho(A)]^*, \\ \rho(1) &= 1 \end{aligned}$$

A particular state of the system is defined by the density functional in the sense that to every dynamical variable A there corresponds a number $\rho(A)$, $A \rightarrow \rho(A)$, and the set of all the numbers $\rho(A)$ obtained by this mapping may be taken as a description of the state of the system at some particular instant of time. If we know the set $\rho(A)$ as a function of time, i.e. if we know the mappings

$A \rightarrow P_t(A)$ (which we call the trajectory of the system) for all A and all t , when we have a description of the dynamics (considered as a description of the states as a function of time).

To justify such a definition of the dynamics of the system, we must show that if we have a trajectory.

$$A \rightarrow P_t(A)$$

(which is a description in a Schrodinger picture), then this trajectory is equivalent to an equation of motion.

$$A \rightarrow P(A_t)$$

(which is a description in a Heisenber picture).

The method used is a generalization of a demonstration in ordinary quantum mechanics. We recall that to define a Hilbert space H we defined.

$$A \rightarrow \psi_A$$

such that

$$\psi_{A+B} \rightarrow a\psi_A + b\psi_B$$

and

$$[\psi_A, \psi_B] = \Omega(A^\dagger B)$$

The scalar product (ψ_A, ψ_A) is bounded and non-negative (unless $A = 0$) for all A

$$\text{i.e. } \Omega(A^\dagger A) \geq 0 \text{ unless } A = 0$$

Consider the properties of the mapping $A \rightarrow P(A)$; P corresponds to a linear functional on the Hilbert space H . We now make use of the following theorem (stated here without proof).

Theorem:

Any linear functional $L(\psi_A)$ on a Hilbert space H must be of the form

$$L(\psi_A) = (\psi_P, \psi_A)$$

where (ψ_P, ψ_A) is antilinear in ψ_P and linear in ψ_A .

Thus there must exist a ψ_P such that

$$f(A) \equiv (\psi_P, \psi_A)$$

i.e. to every linear functional $f(A)$ there corresponds a vector ψ_P in the Hilbert space.

To get a more specific correspondence, consider the case of quantum mechanics.

We have

$$i \frac{\partial \psi}{\partial t} = H \psi, \quad \psi_t = e^{-iHt} \psi$$

In terms of this, if

$$p = \psi \psi^\dagger$$

then

$$p_t = e^{-iHt} p_0 e^{iHt} = L_t p \text{ (say)}$$

L_t has the properties $L_{t_1+t_2}(p) = L_{t_1}(L_{t_2}(p))$

or

$$L_{(t_1+t_2)} = L_{t_1} \cdot L_{t_2}$$

L_t is a linear transformation on p ; we see that these linear transformations constitute a multiplicative one-parameter group—a trivial Lie group. We may write

$$A \rightarrow p_t(A) = (L_t p_0)(A)$$

The correspondence $P \rightarrow P_t(A)$ being linear, we may instead of considering linear transformation on P , speak in terms of linear transformations on Ψ_P

Given any Ψ_P , we have a Ψ_{P_t} , $\Psi_{P_t} = \Psi_{L_t} P$

Therefore there must be a corresponding operator L_t in the Hilbert space, such that

$$L_t \Psi_P = \Psi_{P_t} = \Psi_{L_t} P$$

Then

$$P_t(A) = (\Psi_{P_t}, \Psi_A) = (L_t \Psi_P, \Psi_A) = (\Psi_P, L_t^* \Psi_A)$$

i.e. we have a correspondence

$$A \rightarrow A_t = \Lambda_t A$$

where $\Lambda_t A$ is such that $(L_t)^* \Psi_A = \Psi_{\Lambda_t A}$

Thus we may either talk of the mappings P_t , or of the mappings A_t

Suppose we have a set of dynamical variables $A_t, B_t, C_t \dots$ with a star operation defined on them which is antilinear.

To define these, we calculate all possible expectation values of $A_t, A_t B_t, A_t B_t C_t, \dots$

For simplicity, suppose we have just one dynamical variable A_t . Take the operators, $A_t, A_{t_1} A_{t_2}, A_{t_1} A_{t_2} A_{t_3} \dots$ and the expectation values

$$\Omega(1), \Omega(A_t), \Omega(A_{t_1}, A_{t_2}), \Omega(A_{t_1}, A_{t_2}, A_{t_3}), \dots$$

If these expectation values are all given, we have a dynamics, for we can then make this particular algebra an algebra of operators $\mathcal{A}(t)$ in a Hilbert space, by defining $\mathcal{A}(t)$ by the

property that

$A(t)\Psi_{A_{t_1}, \dots, A_{t_n}} \equiv \Psi_{A_t, A_{t_1}, \dots, A_{t_n}}$
 provided Ω is a non-negative linear functional.

Thus we have no Poisson bracket and no equation of motion but we still have a dynamics, for we have a trajectory defined by the expectation values.

Ex. The Harmonic Oscillator.

Consider the simple oscillator; in the Heisenber picture this is described by two variables q, p , with $[q, p] = i$ and

$$H = \frac{1}{2} [p^2 + q^2 - 1] \quad (i)$$

The equations of motion are

$$\begin{aligned} \dot{q} &= [q, H] \quad \text{OR} \quad \dot{q} = p, \\ \text{and} \quad \dot{p} &= [p, H] \quad \text{OR} \quad \dot{p} = -q. \end{aligned} \quad (ii)$$

Combining, we get

$$\ddot{q} = -q \quad (iii)$$

We have put the frequency equal to 1. For a frequency ω , we have the solution

$$\begin{aligned} q(t) &= q(0) \cos \omega t - p(0) \sin \omega t \\ p(t) &= -\omega [q(0) \sin \omega t + p(0) \cos \omega t] \end{aligned} \quad (iv)$$

Introduce the variables

$$a = \frac{q + ip}{\sqrt{2}}; \quad a^\dagger = \frac{q - ip}{\sqrt{2}} \quad (v)$$

First construct the state $|0\rangle$, such that

$$\langle 0 | 0 \rangle = 1 \quad (\text{vi})$$

$$a | 0 \rangle = 0 \quad (\text{vii})$$

$|0\rangle$ is the 'ground state' of the oscillator. We can construct the nth excited state, $|n\rangle$,

$$|n\rangle = (a^+)^n / \sqrt{n!} |0\rangle, \quad n = 1, 2, 3, \dots \quad (\text{viii})$$

It has the property

$$a^+ a |n\rangle = n |n\rangle \quad (\text{ix})$$

The definitions of a and a^+ , viz $a = \frac{q + ip}{\sqrt{2}}$, $a^+ = \frac{q - ip}{\sqrt{2}}$

thus define a representation for the operators a , a^+

We also have the expectation value

$$\langle 0 | f(q, p) | 0 \rangle \quad (\text{x})$$

Inserting (iv), we obtain

$$q(t) = \frac{1}{\sqrt{2}} \left[(a + a^+) \cos \omega t + i (a^+ - a) \sin \omega t \right]$$

$$p(t) = -\frac{\omega}{\sqrt{2}} \left[(a + a^+) \sin \omega t - i (a^+ - a) \cos \omega t \right] \quad (\text{xi})$$

and

$$q(0) = \frac{a + a^+}{\sqrt{2}}, \quad p(0) = i \frac{(a - a^+)}{\sqrt{2}} \quad (\text{xii})$$

Thus using (vii), (viii), (ix), (x), (xi), we can calculate the expectation values of all dynamical variables $A_t(q, p)$.

If we synthesize these in terms of the Gelfand construction, we must get back the older results.

This approach to dynamics is due to Wightman and was introduced by him in the formulation of the theory of quantized fields.

5. Wightman Formulation of Quantum Field Theory:

Define a set of field operators

$$\varphi(\vec{x}, t) = \varphi(\underline{x})$$

a vacuum vector $|0\rangle$, and a sequence of expectation values:

$$W^0(t) = 1,$$

$$W^{(1)}(x_1) = \langle 0 | \phi(x_1) | 0 \rangle$$

$$W^{(2)}(x_1, x_2) = \langle 0 | \phi(x_1) \phi(x_2) | 0 \rangle$$

...

$$W^{(n)}(x_1, \dots, x_n) = \langle 0 | \phi(x_1) \dots \phi(x_n) | 0 \rangle$$

(The latter are the 'Wightman functions')

Following Wightman, we can make two statements:

(1) If we act as if the $\phi(\underline{x})$ are really operators, and $|0\rangle$ is a normalizable state vector, then all the vacuum expectation values $W^{(i)}$ can be calculated. Given these vacuum expectation values, we have the dynamics, for we can construct a theory with the vacuum expectation values as the distinguished linear functional.

(In the above, for simplicity we have taken $\varphi^*(x)$
 $= \varphi^\dagger(x)$; else we take the $\varphi^*(x)$ also as operators).

(2) If we are given certain definite properties of the field operators, we have certain definite properties for the expectation values.

Example: The Harmonic Oscillator.

Here we have only one variable viz., the variable $q(t)$ for all t , satisfying a second order differential equation in t .

$$\frac{\partial^2}{\partial t^2} q(t) = -q(t)$$

Then the expectation value $\Omega(A_{t_1}, A_{t_2})$ obeys the equation

$$\frac{\partial^2}{\partial t^2} \Omega(A_{t_1}, A_{t_2}) = -\Omega(A_{t_1}, A_{t_2})$$

Suppose we have two harmonic oscillators, with

$$A_t = q^{(1)}(t) + q^{(2)}(t)$$

Then

$$A_t = \frac{1}{\sqrt{2}} \left\{ a_1 e^{-i\omega_1 t} + a_1^\dagger e^{+i\omega_1 t} \right\} \\ + \frac{1}{\sqrt{2}} \left\{ a_2 e^{-i\omega_2 t} + a_2^\dagger e^{+i\omega_2 t} \right\}$$

We define the vacuum state $|0, 0\rangle$ as the state with the properties that

$$a_1 |0, 0\rangle = 0, a_2 |0, 0\rangle = 0$$

The state $|n_1, n_2\rangle$ is defined by

$$|n_1, n_2\rangle = \frac{(a_1^\dagger)^{n_1} (a_2^\dagger)^{n_2}}{\sqrt{n_1! n_2!}} |0, 0\rangle$$

The vacuum expectation values or Wightman functions are given by

$$\Omega(f(a, p)) = \langle 0, 0 | f(a, p) | 0, 0 \rangle$$

Note: Although we have a system of more degrees of freedom, the Wightman functions look just the same as before we have the same infinite set.

But now the variables A_t , and hence the vacuum expectation values Ω satisfy a differential equation that is of the fourth order in t (and not of the second order as before):

$$\left[\frac{\partial^2}{\partial t^2} + \omega_1^2 \right] \left[\frac{\partial^2}{\partial t^2} + \omega_2^2 \right] A_t = 0$$

(and--similarly for each Wightman function).

Such oscillators which are direct sums of simple oscillators are called Greenberg oscillators. What are the analogues of these in field theory?

To see this, we first note that a simple oscillator corresponds to a free field:

$$H = \int d^3x \mathcal{H}(x) = \frac{1}{2} \int \left\{ \dot{\phi}^2(x) + (m^2 - \nabla^2) \phi^2(x) \right\} d^3x$$

Rewrite this in the form

$$H = \sum_k \frac{1}{2} \left(\frac{\pi^2}{k} + \omega_k^2 \phi_k^2 \right),$$

$$m^2 + k^2 = \omega_k^2$$

For the free field, we can solve the free-field equations and write the free field explicitly as a function of t , and thus we can exhibit the Wightman functions also explicitly.

The Greenberg oscillators correspond to 'Generalized Free Fields' which are direct sums of free fields; in terms of these again, the Wightman functions can be calculated explicitly.

Ref. O.W. Greenberg: Ann. Phys. 16, 158-176 (1961).

In general, we can have a direct sum of any number of oscillators, each with a different weight. The most general case is that of a field that is a direct integral of a set of free fields.

$$\phi(x) = \int d\sigma(m^2) \phi(x, m)$$

for there is no reason why the number of fields should be finite.

We note that just like an interacting field, this field also has a whole spectrum of masses, and is similar to a Heisenberg field, which creates a whole spectrum of particles.

If we have only the above, then the system is very general it can accommodate several other things, besides simple and Greenberg oscillators, e.g. it can accommodate an irreversible dynamics.

Symmetry between the past and the future is an additional prejudice; it is possible to conceive of systems in which it makes sense if we trace it forward in time but not if we trace it back in time; such systems are termed irreversible

e.g. We take a density functional ρ_t that is defined for all $t < 0$ (and hence \mathcal{L}_t is defined only for $t > 0$).

If we try to trace it back, then, for instance, a linear functional that was positive definite may no longer be positive definite, and the theory would be irreversible.

We now ask what are the quantities in quantum mechanics that correspond to density functionals. They are not the wave functions ψ , for these obey the normalization $\sum |\psi_i|^2 = 1$ while we have a linear normalization condition on the linear functionals. Thus the corresponding quantity in quantum mechanics must be a quantity bilinear in the wave functions, viz. a density matrix.

6. Stochastic Dynamics of Quantum-Mechanics Systems.

Consider an algebra of operators A, B, C, \dots with functionals $\rho(A)$ with the properties $\rho \geq 0$, $\rho(1) = 1$. Define \mathcal{L}_t by $\rho_t = \mathcal{L}_t \rho$ we try to express the property that the system can be traced only forwards in time.

Any 2×2 Hermitian matrix A can be written in the form

$$A = \begin{pmatrix} a & c \\ c^* & b \end{pmatrix} + \begin{pmatrix} \sigma & 0 \\ 0 & -\sigma \end{pmatrix}$$

where $\vec{a} \cdot \vec{\sigma} = \sum a_i \sigma_i$

where the σ_i are the Pauli matrices and a_μ are numbers.

If ρ is a Hermitian matrix, with $\text{Tr} \rho = 1$ then ρ may be written

$$\rho = \frac{1}{2} [1 + \vec{P} \cdot \vec{\sigma}]$$

Then we define the functional $f(A)$ by

$$f(A) = \text{tr} (\rho^\dagger A) = \text{tr} (\rho A)$$

Generalize this to

$$\rho_t = \frac{1}{2} [I + \vec{P}_t \cdot \vec{\sigma}]$$

where we choose \vec{P}_t as a vector in the direction of the third axis and with the property

$$\vec{P}_t = \vec{P}_0 e^{-\lambda t}; \quad P_0^2 \leq 1$$

The, for all $t > 0$, the (continuous) sequence of matrices ρ_t has the property that each ρ_t is a density matrix satisfying all the requirements above ($\rho_t \geq 0$ etc.)

But for $t < 0$ the density matrix ρ_t is no longer positive definite, and hence the linear functional, although normalized, is not ≥ 0

$$\rho_t = \rho_0 + \frac{1}{2} \sigma_3 \left\{ \text{tr} (\rho_0 \sigma_3) (e^{-\lambda t} - 1) \right\}$$

Therefore with linear functionals of the type above, we cannot trace the system back in time.

If we have

$$\rho_t = e^{-iHt} \rho_0 e^{+iHt}$$

with ρ and H hermitian, then the property of non-negativeness of ρ is preserved whether $t < 0$ or > 0 , and hence a system with such a time dependence can be traced backwards and forwards in time.

If we write $\rho_t = \mathcal{L}_t \rho_0$ then the \mathcal{L}_t 's have the following composition law:

$$(\mathcal{L}_{t_1}, \mathcal{L}_{t_2})(\rho) = \mathcal{L}_{t_1} [\mathcal{L}_{t_2}(\rho)]; \mathcal{L}_0 = I$$

With this composition law, these quantities constitute a semi-group.

We have

$$\mathcal{L}_{(t_1+t_2)} = \mathcal{L}_{t_1} \cdot \mathcal{L}_{t_2}$$

This has two advantages:

(i) It is the same as the law obtained from density functionals.

(ii) It corresponds to equations of motion of the form

$$\dot{\Pi}(t) = A(t) \Pi(t)$$

with

$$A(t_1+t_2) = A(t_1) + A(t_2)$$

So we restrict ourselves to mappings over linear algebras that satisfy the requirements $\mathcal{L}_{t_1+t_2} = \mathcal{L}_{t_1} \cdot \mathcal{L}_{t_2}$. But this property alone does not guarantee that we can trace back in time for if,

and

$$\rho \rightarrow \frac{1}{2} (1 + \vec{\sigma} \cdot \vec{P})$$

$$\rho_t \rightarrow \frac{1}{2} (1 + \vec{\sigma} \cdot \vec{P} e^{-\lambda t})$$

then the system goes from a state with a positive definite functional back to a state where the density functional is not necessarily positive definite.

If the law $\mathcal{L}_{t_1+t_2} = \mathcal{L}_{t_1} \cdot \mathcal{L}_{t_2}$ is true for positive t and negative t , then we have a 1-parameter, Abelian group, with $\mathcal{L}_0 = \mathcal{L}_{-t} \mathcal{L}_t = 1$. We then have the following theorem due to Jordan and Pinsky.

In Quantum Mechanics, the 1-parameter group must of the form

$$\rho_t = e^{-iHt} \rho_0 e^{+iHt}$$

More precisely, we state the following:

Theorem If we require

(i) a star algebra, with a distinguished linear functional, and (ii) a 1-parameter family of linear mappings on the algebra, which constitute a group, and (iii) the property that the equivalent mapping of operators must leave unity unaltered, then the mapping must be of form

$$\rho \rightarrow e^{-iHt} \rho_0 e^{+iHt}, \quad H^\dagger = H$$

Note: In general, H belongs only to the second Hilbert space and not to the star algebra.

Every linear functional in the star algebra is equivalent to a vector in the second Hilbert space. A 1-parameter mapping of the linear algebra is equivalent to a 1-parameter mapping of vectors of the Hilbert space. This latter mapping of vectors in the Hilbert space must be of a form such that the equivalent mapping of operators of the linear algebra must leave the identity operator unaltered.

Stone's Theorem: asserts that the one-parameter group of mappings in the Hilbert space must be of the form

$$\psi \rightarrow e^{iAt} \psi$$

where A is an operator in the Hilbert space.

Thus there is a correspondence between operators of the star algebra and a vector in the second Hilbert space; to every operator in the star algebra there corresponds a vector in the Hilbert space. The converse is not true, but we can have an extended algebra containing an operator corresponding to every vector in the Hilbert space.

IV. Structure of Interacting Systems.

1. Separable Potential Model (S-waves)

Here also we have something similar to particles.

In the Hilbert space of square integrable functions $f(x)$ of one variable, $\left\{ \left[\int_0^\infty |f(x)|^2 dx \right] < \infty \right\}$, we can define

(i) the operator E which gives $(E f)(x) = f(x)$, or in matrix notation, $E_{x, x'} = \delta(x - x')$ which is a diagonal matrix and a scalar.

(ii) the operator H such that

$$(H f)(x) = X f(x);$$

or written as a matrix, the elements are $H_{x, x'} = x \delta(x - x')$, which again gives a diagonal matrix. Now define

$$H_{x, x'} = x \delta(x - x') + \eta g(x) g(x')$$

with g real, $g^*(x) = g(x)$

This has the property that

$$\begin{aligned} (H f)(x) &= \int H_{x, x'} f(x') dx' \\ &= x f(x) + \eta g(x) \int g(x') f(x') dx' \end{aligned}$$

η is a real number, such that $\eta^2 = +1$.

H is a Hermitian matrix, since it is real and symmetric; hence its "eigenvalues" are expected to be real. We ask what are the eigen functions of

H . i.e. we must find the solutions of

$$\begin{aligned} (H f)_\lambda(x) &= \lambda f_\lambda(x) \\ &= \int H_{x, x'} f_\lambda(x') dx' \\ &= x f_\lambda(x) + \eta g(x) \int g(x') f_\lambda(x') dx' \end{aligned}$$

OR

$$(\lambda - x) f_\lambda(x) = \eta g(x) \int g(x') f_\lambda(x') dx'$$

which is a soluble integral equation.

Write this as

$$f_\lambda(x) = C \delta(\lambda - x) + \frac{\eta g(x)}{\lambda - x + i\epsilon} \int f_\lambda(x') g(x') dx'$$

OR

$$f_\lambda(x) = C \delta(\lambda - x) + \frac{\eta g(x)}{\lambda - x + i\epsilon} J_\lambda ;$$

where J_λ is

independent of x (although it depends on the functional form of

$f_\lambda(x)$)

An eigenvalue equation of the above type will determine the eigenfunctions only upto a multiplicative constant; therefore there is no loss of generality in putting $C = 1$.

$$\begin{aligned} J_\lambda &= \int_0^\infty f_\lambda(x') g(x') dx' \\ &= \int_0^\infty \left\{ \delta(\lambda - x') + \frac{\eta g(x')}{\lambda - x' + i\epsilon} J_\lambda \right\} g(x') dx' \\ &= g(\lambda) + \eta J_\lambda \int \frac{g^2(x') dx'}{\lambda - x' + i\epsilon} \end{aligned}$$

OR

$$J_\lambda = \frac{g(\lambda)}{1 - \eta \int_0^\infty \frac{g^2(x') dx'}{\lambda - x' + i\epsilon}} = \frac{g(\lambda)}{\beta(\lambda + i\epsilon)}$$

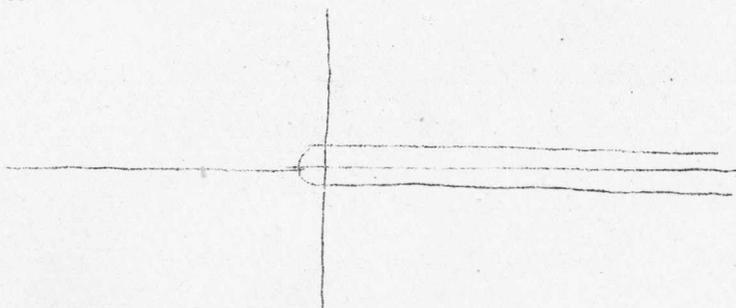
where we have defined the new function

$$\beta(z) = 1 - \eta \int_0^{\infty} \frac{g^2(x) dx}{z - x}$$

This is an analytic function in the z plane cut from 0 to ∞ , with a discontinuity across the cut given by

$$\begin{aligned} \text{Disc. } \beta(z) &= \beta(z_n + i\epsilon) - \beta(z_n - i\epsilon) \\ &= -\eta \int_0^{\infty} g^2(x) dx \left\{ \frac{1}{z_n + i\epsilon - x} - \frac{1}{z_n - i\epsilon - x} \right\} \\ &= 2\pi\eta \int_0^{\infty} dx g^2(x) \delta(z_n - x) \\ &= \begin{cases} 2\pi\eta g^2(z_n) & , \text{ for } z_n > 0 \\ 0 & , \text{ for } z_n < 0 \end{cases} \end{aligned}$$

z plane



The function $\beta(z)$ has no singularities in the left half-plane. Since

$$J_\lambda = \frac{g(\lambda)}{\beta(\lambda + i\epsilon)}, \text{ we have}$$

$$f_\lambda(x) = \delta(\lambda - x) + \frac{\eta g(x) g(\lambda)}{(\lambda - x + i\epsilon) \beta(\lambda + i\epsilon)}$$

If we take $-i\epsilon$ throughout instead of $+i\epsilon$, we would have

$$f_\lambda'(x) = \delta(\lambda-x) + \frac{\eta g(x) g(\lambda)}{(\lambda-x-i\epsilon)(\lambda-i\epsilon)^\beta}$$

and

$$f_\lambda' = e^{-2i\alpha(\lambda)} f(\lambda)$$

where $e^{-2i\alpha(\lambda)} = \frac{f(\lambda+i\epsilon)}{f(\lambda-i\epsilon)}$ = a unimodular number. $\alpha(\lambda)$ is a

real; it is, in fact the phase shift for $\lambda > 0$.

Note: If $\lambda < 0$, the δ function does not contribute to the solution.

To find this solution, we have the condition,

$$J_\lambda' = \eta J_\lambda \int \frac{g^2(x) dx}{\lambda-x+i\epsilon}, \quad \lambda < 0.$$

For $\lambda < 0$,

$$1 - \eta \int \frac{g^2(x) dx}{\lambda-x} = f(\lambda) = 0.$$

(i.e. When $\lambda < 0$ a solution exists if and only if $f(\lambda) = 0$.)

The solution then corresponds to the following choice

$$f_\lambda(x) = \frac{c g(x)}{\lambda-x}$$

We can verify that this is a solution, for

$$\begin{aligned} (Hf_\lambda)(x) &= \int H_{x,x'} f_\lambda(x') dx' \\ &= c \int_0^\infty \{x\delta(x-x') + \eta g(x)g(x')\} \frac{g(x')}{\lambda-x'} dx' \\ &= c \left\{ x \frac{g(x)}{\lambda-x} + \eta g(x) \int \frac{g^2(x') dx'}{\lambda-x'} \right\} \end{aligned}$$

$$\begin{aligned}
 &= C \left\{ \frac{xg(x)}{\lambda-x} + [1-\beta(x)]g(x) \right\} \\
 &= C \left\{ \frac{\lambda g(x)}{\lambda-x} - \beta(x)g(x) \right\} \\
 &= C \frac{\lambda g(x)}{\lambda-x}, \text{ since } \beta(\lambda) = 0 \\
 &= \lambda f_{\lambda}(x)
 \end{aligned}$$

We now ask: Under what conditions is $\beta(\lambda) = 0$.

We have $\beta(-\infty) = 1$

$$\text{Also } \beta'(z_n) = \frac{\partial}{\partial z_n} [\beta(z_n)] = \eta \int_0^{\infty} \frac{g^2(x) dx}{(z-x)^2}$$

Therefore along the negative real axis, $\beta'(z_n)$ always has the same sign.

Therefore, if $\beta(z_n)$ is to decrease from 1 at $z = -\infty$ to 0 at some point, then β' must be a steadily decreasing function; thus

β' must be negative. Therefore we must have $\eta = -1$ (Note $\eta^2 = +1$) —

But this is not sufficient; we must ascertain that $\beta(z_1)$ is negative at $z_1 = 0$.

$$\text{If } \beta(0) = 1 - \int_0^{\infty} \frac{g^2(x) dx}{x}$$

is negative, then we have a "bound state", else we have no bound state.

Thus when $g(x)$ is large there tends to be a bound state, while for small $g(x)$, there tends to be no discrete state. When $\eta = +1$, we have a repulsive interaction as

$$H_{x,x'} = x \delta(x-x') + \eta g(x)g(x')$$

If $\eta = -1$, the interaction is attractive; only then can there be a bound state. — A Hamiltonian of the above type corresponds to a "separable potential"

β varies monotonically with λ , hence, if a bound state exists, there is only one bound state. Also $\beta(z)$ has no complex zeros.

We have

$$\begin{aligned} \text{Im } \beta(z) &= \text{Im} \left\{ 1 - \eta \int \frac{g^2(x') dx'}{z_n + x - x'} \right\} \\ &= -\eta \text{Im} \int \frac{g^2(x') (z_n - i z_i - x)}{(z_n - i)^2 + z_i^2} dx' \\ &= \eta z_i \int \frac{g^2(x') dx'}{(z_n - i)^2 + z_i^2} \end{aligned}$$

Therefore for an imaginary value of z_1 , we have $z_i \neq 0$,

$$\therefore \text{Im } \beta(z) \neq 0.$$

The imaginary part of $\beta(z)$ can have a zero only along the negative real axis, and here there can be at most one zero.

Thus the spectrum of the Hamiltonian spans the real axis from 0 to $-\infty$, and at most one possible negative value (but there may not be such a negative eigen value).

Note: The "eigen functions" $f_\lambda(x)$ do not really belong to the Hilbert space for $\lambda > 0$. For $\lambda < 0$,

The eigen-value spectrum

0
 ∞

We have

$$\int_0^\infty |f_\lambda(x)|^2 dx = |c|^2 \int_0^\infty \frac{g^2(x) dx}{(\lambda - x)^2}$$

If this is to be = 1, we must have

$$|c|^2 = \frac{1}{\int_0^{\infty} \frac{g^2(x) dx}{(\lambda - x)^2}} = \frac{1}{\beta'(\lambda)}$$

OR

$$|c| = \sqrt{1/\beta'(\lambda)}$$

Eigenvectors of a Hermitian matrix belonging to different eigenvalues must be orthogonal. To demonstrate this in our case we proceed as follows: Let

$$H f_\lambda = \lambda f_\lambda ; \quad H f_\mu = \mu f_\mu$$

Then

$$(f_\lambda, H f_\mu) = (f_\lambda, \mu f_\mu) = \mu (f_\lambda, f_\mu)$$

The L.H.S. also

$$\begin{aligned} &= (H^\dagger f_\lambda, f_\mu) = (H f_\lambda, f_\mu) \quad \text{since } H^\dagger = H, \\ &= \lambda (f_\lambda, f_\mu) \end{aligned}$$

$$\therefore (f_\lambda, f_\mu) \text{ must be } = 0 \text{ whenever } \lambda \neq \mu.$$

Note:

$$(f_\lambda, f_\mu) = \int_0^{\infty} f_\lambda^* f_\mu dx$$

$$f_\lambda(x) = \delta(\lambda - x) + \frac{\eta g(\lambda) g(x)}{(\lambda - x + i\epsilon) \beta(\lambda + i\epsilon)}$$

$$\therefore (f_\lambda, f_\mu) = \int \left\{ \delta(\lambda - x) + \frac{\eta g(\lambda) g(x)}{(\lambda - x + i\epsilon) \beta(\lambda + i\epsilon)} \right\}^*$$

$$\left\{ \delta(\mu - x) + \frac{\eta g(\mu) g(x)}{(\mu - x + i\epsilon) \beta(\mu + i\epsilon)} \right\} dx$$

$$\begin{aligned}
 &= \int \left\{ \delta(\lambda-x) + \frac{\eta g(\lambda) g(x)}{(\lambda-x-i\epsilon)(\lambda-i\epsilon)} \right\} \\
 &\quad \left\{ \delta(\lambda-x) + \frac{\eta g(\mu) g(x)}{(\mu-x+i\epsilon)\beta(\mu+i\epsilon)} \right\} \\
 &= \delta(\lambda-\mu) + \frac{\eta g(\mu) g(\lambda)}{(\mu-\lambda+i\epsilon)\beta(\mu+i\epsilon)} \\
 &\quad + \frac{\eta g(\lambda) g(\mu)}{(\lambda-\mu-i\epsilon)\beta(\lambda-i\epsilon)} \\
 &\quad + \frac{\eta^2 g(\lambda) g(\mu)}{\beta(\lambda-i\epsilon)\beta(\mu+i\epsilon)} \int_0^\infty \frac{g^2(x) dx}{(\lambda-x-i\epsilon)(\mu-x+i\epsilon)}
 \end{aligned}$$

Put

$$\begin{aligned}
 &\frac{1}{(\lambda-x-i\epsilon)(\mu-x+i\epsilon)} \\
 &= \frac{1}{(\mu+i\epsilon-\lambda)} \left[\frac{1}{\lambda-x-i\epsilon} - \frac{1}{\mu-x+i\epsilon} \right]
 \end{aligned}$$

and thus note that

$$\eta \int_0^\infty \frac{g^2(x) dx}{(\lambda-x-i\epsilon)(\mu-x+i\epsilon)} = \frac{1}{(\mu+i\epsilon-\lambda)} \left\{ -\beta(\lambda-i\epsilon) + \beta(\mu+i\epsilon) \right\}$$

Substituting these in (1), we find

$$(f_\lambda, f_\mu) = \delta(\lambda-\mu) \quad \text{for } \lambda, \mu \text{ both positive. For } \lambda > 0, \mu < 0 \text{ we can prove this similarly, making use of the condition that } \beta(\mu) = 0.$$

We shall next prove that the set of eigen functions is also complete.

Note: The function

$$f_{\lambda}(x) = \delta(\lambda - x) + \frac{\eta_{\lambda}(\lambda) g_{\lambda}(x)}{(\lambda - x + i\epsilon)\beta(\lambda + i\epsilon)}$$

is the 'in' state in scattering, in which the "spherical" wave part "diverges" from the scattering centre. The 'out' state is obtained from this by time reversal; in this state the "spherical" waves "converge" to the scattering centre.

Instead of picturing these in the configuration space, we could have looked at it in momentum space.

The 'in' and 'out' states have singularities in momentum space.

The plane wave part gives a δ -function; while the spherical waves gives another type of singularity, of the type in

$$\int \frac{g^2(x) dx}{(\lambda - x - i\epsilon)(\mu - x + i\epsilon)}$$

For the 'in' and 'out' state, these singularities are of the type

$$\frac{1}{\lambda - \mu + i\epsilon}$$

respectively; the difference of the 'in' and 'out' states looks a plane wave.

The actual scattered wave has "short-distance" modifications also. One may look at it in terms of either the complete set of 'in' states or the complete set of 'out' states. But if a bound state exists, it is orthogonal to all scattering states; it cannot be expressed in terms of asymptotically plane wave states.

To prove the completeness of the eigen functions of a Hermitian matrix, i.e. to prove that

$$\int_{-\infty}^{\infty} f_{\lambda}^*(x) f_{\mu}(x) dx = \delta(\lambda - \mu)$$

(a)

and

$$\int d\mu f_{\mu}^{*}(x) f_{\mu}(y) = \delta(x-y)$$

(b)

We write the eigen function $f_{\lambda}(x)$ as a matrix:

$$\begin{aligned}\Omega(x, \lambda) &= f_{\lambda}(x) \\ \Omega^{\dagger}(\lambda, x) &= [\Omega(x, \lambda)]^{*}\end{aligned}$$

Then (a) is equivalent to

$$\int_0^{\infty} \Omega^{\dagger}(\lambda, x) \Omega(x, \mu) dx = (\Omega^{\dagger} \Omega)(\lambda, \mu) = \delta(\lambda - \mu)$$

OR

$$\Omega^{\dagger} \Omega = 1$$

Similarly (b) is equivalent to $\Omega \Omega^{\dagger}$. If Ω were a finite-dimensional matrix, it would immediately follow that $\Omega^{\dagger} = \Omega^{-1}$.

But for infinite-dimensional matrices, this is not always true. We ask: under what conditions could this fail to be true?

The product $\Omega \Omega^{\dagger}$ is defined by

$$(\Omega \Omega^{\dagger})(x, y) = \int d\mu \Omega(x, \mu) \Omega^{\dagger}(\mu, y)$$

$$\begin{aligned}\int_0^{\infty} d\mu f_{\mu}^{*}(x) f_{\mu}(y) &= \int_0^{\infty} d\mu \left\{ \delta(\mu-x) + \frac{\eta \bar{g}(x) g(\mu)}{(\mu-x+i\epsilon)\beta(\mu+i\epsilon)} \right\} \\ &\quad \left\{ \delta(\mu-y) + \frac{\eta g(y) \bar{g}(\mu)}{(\mu-y-i\epsilon)\beta(\mu-i\epsilon)} \right\} \\ &= \delta(x-y) + \frac{\eta \bar{g}(x) g(y)}{(y-x+i\epsilon)\beta(y+i\epsilon)} \\ &\quad + \frac{\eta \bar{g}(x) g(y)}{(x-y+i\epsilon)\beta(x-i\epsilon)} \\ &\quad + \eta^2 \bar{g}(x) g(y) \int_0^{\infty} \frac{g^2(\mu)}{\beta(\mu+i\epsilon)\beta(\mu-i\epsilon)} \frac{d\mu}{(\mu-x+i\epsilon)(\mu-y-i\epsilon)}\end{aligned}\tag{2}$$

The last term in (2) is

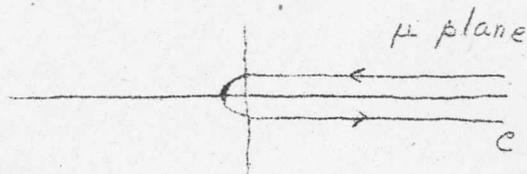
$$\frac{\eta}{2\pi i} g(x)g(y) \int_0^{\infty} \frac{2\pi i \eta g^2(\mu)}{\beta(\mu+i\epsilon)\beta(\mu-i\epsilon)} \frac{d\mu}{(\mu-x+i\epsilon)(\mu-y-i\epsilon)}$$

Notes:

$$\beta(x+i\epsilon) - \beta(x-i\epsilon) = 2\pi i \eta g^2(x)$$

The last term in (2) then becomes

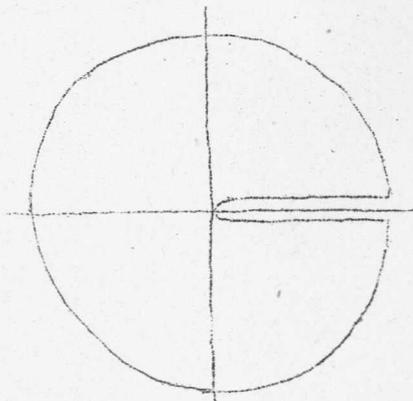
$$\begin{aligned} \frac{\eta}{2\pi i} g(x)g(y) \int_0^{\infty} \frac{d\mu}{(\mu-x+i\epsilon)(\mu-y-i\epsilon)} \left\{ \frac{1}{\beta(\mu-i\epsilon)} - \frac{1}{\beta(\mu+i\epsilon)} \right\} \\ = \frac{\eta}{2\pi i} g(x)g(y) \int_C \frac{d\mu}{(\mu-x+i\epsilon)(\mu-y-i\epsilon)} \frac{1}{\beta(\mu)} \quad (3) \end{aligned}$$



with the contour C as shown.

We wish to write this as an integral along a different contour C'; to do so, we must ask what singularities are crossed in the course of this deformation.

Assume that $\beta(\mu)$ has no zero; then the integrand of (3) will have no singularities on the negative real axis. If we draw small circles round the singularities, then we can show that the integrals round these



circles just cancel the second or third terms in (2). (It can be proved that the contribution of the large circle $\rightarrow 0$ at infinite radius.)

Then we are left with just $\delta(x-Y)$ and thus we see that when $\beta(\mu)$ has no zero, the scattering states alone satisfy the completeness condition.

If $\beta(\mu)$ does have a zero, then drawing a small circle Γ round the corresponding singularity,

we obtain

$$\int_0^{\infty} d\mu f_{\mu}^*(x) f_{\mu}(y) = \delta(x-y) + c$$



plus the contribution from Γ . But then, the definition of Ω and of the product $\Omega \Omega^{\dagger}$ must be changed to

$$\Omega \Omega^{\dagger}(x, y) = \int_0^{\infty} d\mu \Omega(x, \mu) \Omega^{\dagger}(\mu, y) + \Omega(x, \Lambda) \Omega^{\dagger}(\Lambda, y)$$

Thus to (2) is added a fifth term, the bound-state term

$$|c|^2 \frac{g(x)}{\Lambda - x} \frac{g(y)}{\Lambda - y}$$

and the integral in the last term of (2) gets a contribution from the circle Γ .

We again find that, even in the presence of bound states, the completeness condition is satisfied, with the extended definition of Ω which includes Λ .

The properties of Ω :

- (i) $\Omega^{\dagger} \Omega = \Omega \Omega^{\dagger} = 1$
- (ii) $\Omega^{\dagger} H \Omega = H^{\circ}$ or $H \Omega = \Omega H^{\circ}$ or $\Omega^{\dagger} H \Omega = H^{\circ}$

Thus the free Hamiltonian is unitarily equivalent to the Hamiltonian with interaction (when no bound state is present).

Ω is the Moller matrix for the model. What happens when a bound state is present? The spectrum of H is now different from that of H° . But we know that a unitary transformation cannot change the spectrum.

We must calculate $\Omega^\dagger H \Omega = H^{\circ'}$ When there is no bound state $H^{\circ} = H^{\circ}$ where $H^{\circ}(\lambda, \mu) = \lambda \delta(\lambda - \mu)$ when there is a bound state,

$$H^{\circ'}(\lambda, \mu) = \lambda \delta(\lambda - \mu) \text{ for } \lambda, \mu > 0$$

$$= \Lambda \quad \text{when } \lambda = \mu = \Lambda$$

$$= 0 \quad \text{otherwise}$$

i.e.

$$H^{\circ'}(\lambda, \mu) = \begin{bmatrix} \Lambda & 0 \\ 0 & \lambda \delta(\lambda - \mu) \end{bmatrix}$$

Once the spectrum changes, the unitary equivalence is lost.

A separable potential is a simple example; in general, for an arbitrary potential, this kind of explicit computation is not very straightforward! But in general, when there are no bound states, H and H° are unitarily equivalent.

We can use the moral of the story above (Separable potential) to construct solvable interacting relativistic particle systems. For this purpose manufacture a unitary matrix Ω by any expedient means. If \vec{P}° , \vec{J}° , \vec{H}° and \vec{K}° are the generators for the free system, we ask: how can we construct the generators for an interacting systems. We choose the generators $\vec{P} = \Omega \vec{P}^{\circ} \Omega^\dagger$ etc.; we should have $\vec{P} = \vec{P}^{\circ}$, $\vec{J} = \vec{J}^{\circ}$ as these should be unchanged; but we must have $H^{\circ} \rightarrow H \neq H^{\circ}$, $K^{\circ} \rightarrow K \neq K^{\circ}$. Hence Ω must leave \vec{P} and \vec{J} unchanged.

The actual construction of the new generators is a little complicated.*

We thus see that we can have a theory of a scattering system which is relativistically invariant, and has scattering states. But it has certain drawbacks:

* See a forthcoming paper by Jordan, Macforlane and Sudarshan for an explicit construction.

- (i) it has no fundamental basis;
- (ii) in our example above, we have introduced an arbitrary function $g(x)$ which is unsatisfactory;
- (iii) the theory does not automatically furnish any natural explanation of experimental results in general; for to get results that can correlate the results of experiments, it is not enough to consider very general features like relativistic invariance. One must have a theory that gives specific numerical results that can be compared with experiment. To get these, in general, one has to consider the specific features of the interaction. There are certain features of scattering not explained by ordinary particle theory.

Example: Consider electron-photon scattering:

$$e + \gamma \rightarrow e + \gamma$$

The interaction causing the scattering is $\vec{j} \cdot \vec{A}$, where \vec{j} is the charged current and \vec{A} is the vector potential. We expect that for an electron of low momentum and a photon of low frequencies, the associated electric current should be small and should perhaps tend to zero with ω . But in fact the cross-section approaches a finite limit (the Thomson limit) as $\omega \rightarrow 0$.

Similarly consider $\pi^+ + \gamma \rightarrow \pi^+ + \gamma$. In the absence of an electromagnetic field, the pion field is described by the equation

$$(\square^2 + \mu^2) \phi = 0$$

when an electromagnetic field is present, we make the "gauge-invariant" replacement:

$$\partial_\mu \rightarrow \partial_\mu - ie A_\mu$$

and obtain

$$[(\partial_\mu - ie A_\mu)(\partial^\mu - ie A^\mu) + \mu^2] \cdot \phi = 0$$

This gives a term of the form $e^2 A_\mu A^\mu$ which represents an interaction that is independent of the magnitude of the particle momentum. (since it contains no derivative) and depends only on e . For the Dirac equation (which is linear) we have no term of the form $A_\mu A^\mu$. However, such a term is obtained by considering the contribution of negative energy states.

In both the above examples, relativistic invariance has not been sufficient; it has been necessary to invoke specific features of the interaction.

Another example is the requirement of "crossing symmetry" that relates reactions



to



in fact, it is claimed by some people that these two amplitudes are different boundary values of the same analytic function.

In all these, we have requirements going beyond relativistic invariance; a theory arising by making simple canonical transformations cannot automatically exhibit these required features.

By introducing notions like the analyticity of the scattering amplitude, it is possible to rule out particle theories of certain types, e.g. Analyticity would rule out a theory that gave a scattering amplitude of the form

$$\frac{g^2(E)}{\beta(E - i\epsilon)}$$

with an arbitrary real $g(x)$ (which was not analytic!), for example, which strictly vanished for $x > E_{max}$.

2. Unitary Equivalence; Asymptotic Condition.

In the model of a separable potential described in the last lecture, we took a Hamiltonian

$$H_{x,x'} = x \delta(x-x') + \eta f(x) f(x')$$

and found eigenfunctions labelled by x and λ

$$\psi_\lambda(x) = \delta(\lambda-x) + \frac{\eta f(x) f(\lambda)}{(\lambda-x+i\epsilon)\beta(\lambda+i\epsilon)} = \Omega(x,\lambda)$$

where

$$\beta(z) = 1 - \eta \int_0^\infty \frac{f^2(\omega) d\omega}{z-\omega}$$

$$0 < x < \infty \quad ; \quad 0 < \lambda < \infty$$

The model had the interesting property that the free-particle solutions and the solutions in the presence of interaction were both labelled by the same λ and the same x .

The functions ψ had the property

$$(\psi_\lambda, \psi_\mu) = \delta(\lambda-\mu)$$

There existed a certain unitary transform Ω such that

$$H = \Omega H^0 \Omega^\dagger$$

The question of scattering arose not as an intrinsic property of the system but in the relation between the interacting and non-interacting systems.

Consider the definition

$$\psi_\lambda(\omega, t) = e^{i(\lambda - H^0)t} \psi(\omega, 0)$$

i.e. allow the state to evolve with time, not with the time depending of the exact Hamiltonian, $i \frac{\partial \psi}{\partial t} = H \psi$, but with the "wrong" time dependence, that of H_0 .

Then the quantities $\psi_\lambda(\omega)$ above acquire a time dependence.

We have

$$\psi_{\lambda}(\omega) \rightarrow \phi_{\lambda}(t) = \int \psi_{\lambda}(\omega) e^{-i\omega t} d\omega e^{+i\lambda t}$$

This is the wave function in the "interaction representation". We have

$$\begin{aligned} \phi_{\lambda}(t) &= \int d\omega \left\{ \delta(\lambda - \omega) + \frac{\eta f(\omega) f(\lambda)}{(\lambda - \omega + i\epsilon)\beta(\lambda + i\epsilon)} \right\} e^{i(\lambda - \omega)t} \\ &= 1 + \frac{\eta f(\lambda)}{\beta(\lambda + i\epsilon)} \int \frac{f(\omega) e^{i(\lambda - \omega)t}}{\lambda - \omega + i\epsilon} d\omega \end{aligned}$$

The integrand is singular regarded as a function of ω ; such singularities lead to simple asymptotic properties

when $t \rightarrow -\infty$, the integrand $\rightarrow 1$

when $t \rightarrow +\infty$, the integrand $\rightarrow 0$.

We assert the following:

Assertion: $\lim_{t \rightarrow +\infty} \phi_{\lambda}(t) = e^{2i\delta_{\pm}(\lambda)} \phi_{\lambda}^{\circ}(t)$

where $\phi_{\lambda}^{\circ}(t)$ is a free particle wave function with energy λ .

(We are going from the momentum representation to a time representation)

i.e. As $t \rightarrow \pm\infty$, this function behaves as a free particle.

There are two ways of looking at the wave function:

(1) The exact wave function is related to the free wave function by a unitary transformation, OR

(2) We consider the wave function as a time-dependent quantity in the interaction representation; in the asymptotic limit, the wave function is that of a free particle. (For absolutely integrable functions, we have the Riemann-Lebesgue lemma, stating that the Fourier transform with respect to ω of an integrable function of ω tends to zero as $t \rightarrow \pm\infty$.)

that the asymptotic form of this function is the same as that for a free particle. (1) says that the orthogonality and normalization properties are the same as that for a free particle. When the spectrum changes by interaction this is not quite true as we have seen above).

We should now be able to associate particles with the field.

The statement usually made is that if $\phi(x,t)$ is a field operator, then

$$\lim_{t \rightarrow \pm\infty} \phi(\vec{x}, t) = \phi_{in}(\vec{x}, t)$$

Here we are dealing with operators; so the question arises whether the limit implied is the strong limit or the weak limit.

(The usual definition is as follows: Consider the quantity

$\langle B | \phi(x,t) | A \rangle$ with fixed A,B . Then as $t \rightarrow \pm\infty$ if for all A,B these converge to a number, then the convergence is said to be weak; if $\| \phi(x,t) | A \rangle \|$ converges then it is a strong limit.)

OR We take a quantity like $\langle B | e^{-iEt} \phi(x,t) | A \rangle$ and consider the convergence of this quantity.)

For Heisenberg fields of mass m, the time dependence of $\tilde{\phi}(\vec{k}, t)$ is determined by the mass m, (e.g. we have $e^{i\omega t}$ $\omega = \sqrt{k^2 + m^2}$) to the extent that the field $\tilde{\phi}(\vec{k}, t)$ can be split into positive and negative frequency parts.

For a classical field, we would have $\phi(\vec{x}, t) = e^{iHt} \phi(\vec{x}, 0)$ H being the first quantised Hamiltonian.

For quantised field, we have

$$\phi(\vec{x}, t) = e^{-iH_{op}t} \phi(\vec{x}, 0) e^{+iH_{op}t}$$

where H_{op} , the second quantised Hamiltonian, is an operator in Hilbert space; it does not commute with ϕ .

(Note: Given the field, how is the mass of the particle known?)

Given the field, the propagator is known; this gives the mass.)

We consider now how we can label the interacting field by free-field quantities by a limiting procedure; this limit of the Heisenberg fields gives the free fields. (ref. Schweber and Sudarshan). Annals of Physics, 19, 351 (1962).

The asymptotic condition as an explicit condition was first given by Haag; but in perturbation theory, this had always been clearly appreciated. e.g. Feynman always starts off with a free electron, the particle labels are automatically free-particle labels.

However, there is no reason why a satisfactory asymptotic limit should exist; there are several examples of field theories with no useful asymptotic limits.

Ex. 1. $\psi(x) = \phi^2(x)$

where $\phi(x)$ is a free field with mass m . This has several properties of the free field, but not the correct type of asymptotic limit. It contains all possible frequencies from $-\infty$ to $-m$ and from $+m$ to $+\infty$;

$\omega = \pm \sqrt{m^2 + k^2}$; hence \vec{k} can take on all possible real value.

For a fixed $\vec{k} = \vec{k}_1 + \vec{k}_2$ (where \vec{k}_1, \vec{k}_2 refer to the fields $\phi(x)$), \vec{k}_1 and \vec{k}_2 and hence ω , can run over all values; hence the asymptotic field vanishes identically.

Ex. 2. The generalized free field with a continuum weight. For a generalized free field with a discrete weight,

$\psi(x) = \sum \lambda_\alpha \Phi_\alpha(x)$ ($\Phi_\alpha(x)$ referring to a mass m_α)

But when $\psi(x) = \int d\sigma(m^2) \phi_m(x)$

the asymptotic limit = 0.

It is generally taken that the asymptotic condition is more or less 'equivalent' to the particle interpretation (in that it provides particle labels; it takes over half the function of the commutation relations, which is to produce particles out of fields.)

Are there any genuine relativistic interacting quantized fields which satisfy the asymptotic condition? No examples of these are known.

In the above we found that not only were the fields free fields but also the normalization, etc. were the same as the time-dependent limits.

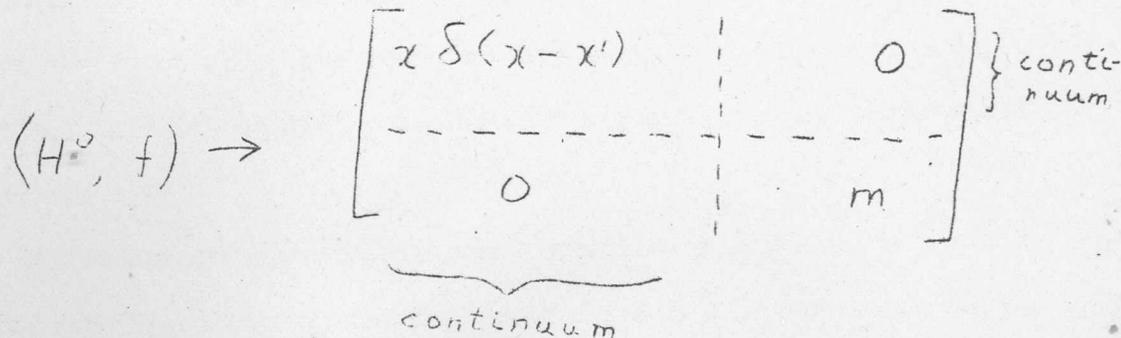
In general, the statements about the normalization, etc. of the asymptotic (time-dependent) limits and the exact wave functions are not the same; this difficulty occurs when we have wave function renormalization. (In general, the time-dependent limit will have a "similar" normalization.)

3. The Lee Model.

Consider a Hilbert space of all square-integrable functions and a single constant f_v (i.e. a space of $\infty + 1$ dimensions) in which the scalar product (f, g) is

$$(f, g) = \int_0^\infty dx f^*(x) g(x) + f_v^* g_v$$

Consider a system whose free Hamiltonian H^0 is defined by



which is partitioned into a continuous part and a discrete sub-matrix (consisting of the single number m). The column vector

$f = \begin{bmatrix} f(x) \\ f_v \end{bmatrix}$; the continuous part represents the state. The Hamiltonian H is given by

$$H \rightarrow \left[\begin{array}{c|c} x \delta(x-x') & \eta(x) \\ \hline \eta(x') & m \end{array} \right].$$

Again we try to find the solutions of

$$H \psi = \lambda \psi$$

H^0 has a continuum of solutions with their corresponding continuous in-

finitude of eigenvalues and a solution $f = \begin{bmatrix} 0 \\ f_v \end{bmatrix}$ with the discrete eigenvalues m . In the presence of interaction, i.e. for H , the continuum states (which are now different) have the same eigenvalues while the discrete state has its eigenvalue shifted to $m - \epsilon$ and its eigenfunction altered from $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ to $\begin{bmatrix} v(x) \\ z \end{bmatrix}$ such that the normalization is preserved.

When we take the limit as $t \rightarrow \pm \infty$ only the singularities contribute; thus only the component z in $\begin{bmatrix} v(x) \\ z \end{bmatrix}$ survives.

(We recollect that there are two different methods of ascribing free-particle labels to a field: (i) we diagonalize and see if we can put free-particle labels; OR (ii) we take the time-dependent limit.

Whenever there is a coupling between a discrete normalizable state and a set of continuum states, one will get different results according as one considers the exact (i.e. time-independent) state or take the time-dependent limit. This is the difference between renormalized and unrenormalized Heisenberg fields.)

A difficulty arises concerning the question of normalization, as in a relativistic field theory, the number of particles is not conserved and

hence the normalization is not conserved; e.g. in perturbation theory, the normalization constant appears to be zero.

4. The Frame Work of Quantum Field Theory.

Any general frame work of quantum theory must involve the following notions:

- (1) A Quantum Theory i.e. a theory - formulated in a Hilbert space.
- (2) Relativistic Invariance.
- (3) Fields
- (4) Stastics
- (5) Positive definiteness of the energy; the spectrum condition.
- (6) Particle interpretation; the asymptotic condition.
- (7) Interacting fields.

We elaborate on these notions as follows:

(1) Hilbert space Formulation.

We must have a theory formulated in a Hilbert space. We have seen that, as far as a finite number of particles are concerned, classical mechanics and quantum mechanics can be so formulated; but in quantum mechanics, we have (square integrable) functions of one variable, $f(x)$.

$$q \rightarrow x, \quad \vec{p} \rightarrow -i \vec{\nabla}, \quad p_x \rightarrow i \frac{\partial}{\partial x}$$

Whereas in classical mechanics, we have functions of two variables,

$$f(x, y), \quad q \rightarrow x, \quad p \rightarrow y$$

In Quantum Mechanics, any quantity that commutes with both q and p must be multiple of the unit operator; but this is not be so in classical mechanics (instead any quantity that has vanishing P.b's with both q and p must be a constant).

A Hilbert space thus admits different types of theories; one way of restricting it is to require that the dynamical variables must be irreducible. In a field theory since we "manufacture" the dynamical variables from the fields we may require that the fields must be "irreducible" i.e., any dynamical variable commuting with all field operators must be a multiple of the identity.

(2) Relativistic Invariance.

In a Lorentz transformation $X \rightarrow \Lambda X + a$, we require that the dynamical variable ϕ must transform according to

$$\phi(x) \rightarrow \phi(\Lambda X + a) \quad (i)$$

This is a considerably stronger restriction than the invariance specified by the requirement

$$\phi(x) \rightarrow U(\Lambda, a) \phi(x) U^{-1}(\Lambda, a) \quad (ii)$$

We require Lorentz invariance in both senses (i) and (ii).

In quantum theory, a canonical transformation is a unitary transformation; the fields must transform according to $U(\Lambda, a)$ is a representation $(\Lambda, a) \rightarrow U(\Lambda, a)$ in the sense that if two transformations (Λ_1, a_1) and (Λ_2, a_2) combine to give (Λ, a) ,

$$(\Lambda_1, a_1)(\Lambda_2, a_2) = (\Lambda, a)$$

then

$$U(\Lambda_1, a_1) U(\Lambda_2, a_2) = U(\Lambda, a) e^{i\eta(\Lambda_1, a_1; \Lambda_2, a_2)}$$

i.e. $U(\Lambda, a)$ is a representation upto a multiplication constant, i.e. a projective representation. For the Lorentz group we can always choose

$$e^{i\eta(\Lambda_1, a_1; \Lambda_2, a_2)}$$

We can have both (i) and (ii) only if

$$\phi(\Lambda x + a) = U(\Lambda, a) \phi(x) U^{-1}(\Lambda, a)$$

We have a 10-parameter group with this invariance property.

(3) Fields.

A free-particles field, or an assembly of free particles, is equivalent to a classical field with P.b.'s replaced by the commutator,

$$\left[\phi(\underline{x}), \frac{\partial \phi(\underline{y})}{\partial y_0} \right]_{x_0 = y_0} = i \delta(\underline{x} - \underline{y})$$

But there are no true operators satisfying this relation. However, we can define

$$\phi^f = \int \phi(x) f^*(x) d^3x$$

$$\pi^g = \int \pi(y) g^*(y) d^3y$$

Then we would have

$$\begin{aligned} [\phi^f, \pi^g] &= i \int \delta(x-y) f^*(x) g(y) d^3x d^3y \\ &= i \int f^*(x) g(x) d^3x \end{aligned}$$

For properly chosen f, g we may take ϕ^f and π^g to be operators; we must think of the "field operators" $\phi(x)$ not as operators but as operator-valued distributions.

4. Particle interpretation.

A collection of free particles is equivalent to a quantized free field, and vice versa; and thus for a free field the particle interpretation is almost immediate. However, when fields begin to "interact", the total energy is not equal to the sum of the energies of the individual fields and there are difficulties in a particle interpretation.

For instance, we may have bound states, so that a proton and an electron can go over into hydrogen atom states that are not of the same

kind as a proton and electron that are "far apart" (ionised hydrogen!). For many purposes, we can label them in term of the free-particle states, but this fails for states in which the particles remain together.

We temporarily ignore such states and assume that all states can be labelled by their asymptotic momenta. We then ask the question: Can we give a particle interpretation of interacting fields also by considering some variety of "asymptotic" fields?

Here we are associating free-particle labels with interacting fields. But free-particle labels can be associated with free fields also.

Thus, by assumption, we are in effect requiring that with the interacting fields we should be able to associate linearly a set of free fields, or put in a different way, that there must exist free fields obtained by considering the "asymptotic limits" of interacting fields. This requirement is the ASYMPTOTIC CONDITION.

But to specify how the "asymptotic limits" are obtained and of what kind they are is a difficult question.

We can use the new "associated free fields" to interpret the interacting fields.

(5) Interacting Fields.

There are two different ways of introducing the concept of an interacting field. The first is to say that it is not a free field. A more useful approach is related to the particle interpretation. We ask the question: what is the amount of scattering (or reaction) that would result when two or more particles come together?

A theory includes a true interaction if it allows the corresponding particles to exchange momentum and energy, or particles to be created or destroyed i.e. for interacting fields, the scattering matrix must not be unity.

Notes: The fields "associated" above are free fields, but the actual Heisenberg fields are not free fields.

(6) Statistics.

For different types of fields, we usually consider one of the two following types of equal-time commutation rules:

- (i) Commutator = an imaginary multiple of a δ -function, associated with Bose statistics, and
- (ii) Anti-commutator = a real multiple of a δ -function, associated with Fermi statistics.

For free fields, we can actually write down the commutation relations for arbitrary times. For equal times, relativistic invariance is generally considered to imply that the commutator or anticommutator vanishes for space-like separations. This "causality" condition is taken to be valid not only for free fields (where it follows by computation from the equal-time commutation relations) but also for interacting fields (where no such computation has so far been completed!).

Thus we can define our statistics by considering whether the "field operators" commute or anticommute at space like points (where the statement about "field operators" is taken as a statement about operator-valued distributions).

Evidence for statistics of different types of particles:-

For electrons, atomic spectra and structure gives evidence for Fermi Statistics; for photons, there is some evidence for Bose statistics from black-body radiation; the statistics for photons could be tested by looking for positive correlations in fluctuations of intensity (or photoelectric counts) in light beams*. For neutrinos, and hyperons, there is little

* See article by L.Mandel in "Progress in Optics", Vol.II, ed. E.Wolf, North-Holland Publishing Company, Amsterdam (1963).

evidence about the statistics. The muon statistics may be inferred (to be Fermi) from muon pair production cross-section. For mesons, there is evidence for its Bose character from the decay $K_1^0 \rightarrow \pi^0 + \pi^0$

Looking at this last in some detail, the line of argument is as follows: K_1^0 has zero spin; thus the two identical π^0 -mesons are in a space-symmetric state and therefore cannot be fermions. Thus the conclusion that the π^0 is a boson is based on the assumption that a particle must be either a boson or a fermion, and that no third type of statistics is admissible.

There seems to be a correlation between spin and statistics, of half-integral spin with Fermi statistics and of integral spin with Bose statistics. This is usually taken as a general postulate, i.e. that integral spin fields (not particles) must obey commutation relations (i.e. Bose Statistics), while half-integral spin fields must obey anticommutation relations (i.e. Fermi statistics).

Pauli proved that under certain conditions the above is true, or rather that the negative of the above is false. As in most proofs, he proved that integral spin fields cannot obey Fermi statistics, and the half integral spin fields cannot obey Bose statistics. But as noted above, concluding from this that integral spin fields are Bose fields etc. overlooks the possibility that a field may obey neither Bose nor Fermi statistics; it may obey some other kind of statistics (para-statistics).

(7) Positive definiteness of the Energy.

In classical physics, we have the second law of thermodynamics which states that one cannot lower the energy of a system indefinitely. The analogue of this in quantum field theory is the statement that the energy of a quantum-mechanical system must have a least value.

In principle, quantum field theory must describe the whole of nature. But a statement about the energy of the whole universe would be without content. To give meaning to it, we take a macroscopic dimension to be $\approx \infty$. Thus for example the measuring apparatus is taken to be of infinite dimensions. Then the requirement about a least possible energy has significance; it is a statement about the interchange of energy between a "laboratory specimen" and an external agency, and tells us that we cannot tap an unlimited amount of energy from the system.

Taking the least possible value of the energy as the zero of the energy, we have the requirement that the energy must be positive in all frames of reference.

But the energy is, relativistically, the time-component of the energy-momentum 4-vector, and the sign of the time-component of a space-like 4-vectors can always be reversed by a Lorentz transform.

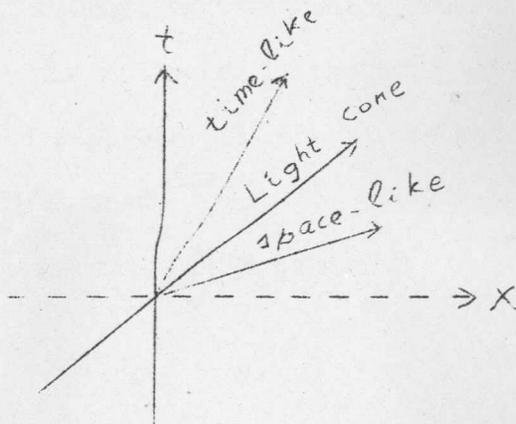
Thus we have the requirement that the energy-momentum 4-vector must not be space-like, as a direct consequence of the condition that the energy must be positive.

The (E, \vec{p}) 4-vectors can be time-like or light-like, for then the sign of E would be left unchanged by a real Lorentz transform. Considering the equivalent mass m defined by

$$m^2 = E^2 - p^2$$

we have the conditions that the spectrum of m^2 must be non-negative.

This is often stated as THE SPECTRUM CONDITION: The Hamiltonian H must be positive definite (or at least non-negative).



Consider the following theorem:

Theorem: In quantum mechanics, if we have a family of time-dependent mappings of density operators which constitute a group, then it is a Hamiltonian system.

Let the operators be A, B, C, \dots such that the functionals $\Omega(A) = \text{Tr } A$ etc. are finite (as the distinguishing linear functionals should be bounded.)

The density operator ρ defines the linear mappings of operators onto numbers; $A \rightarrow \rho(A) = \text{Tr}(\rho^\dagger A) = \text{Tr}(\rho A)$ since ρ is Hermitian.

Then if we have a family of time-dependent mappings $\rho \rightarrow \rho_t = \mathcal{L}_t \rho$ then, if $\mathcal{L}_{t_1} \mathcal{L}_{t_2} = \mathcal{L}_{t_1+t_2}$ for all t_1, t_2 (including negative values), then this group must be of the form

$$\rho_t = e^{-iHt} \rho_0 e^{+iHt}$$

where $H^\dagger = H$

(For proof, see Dynamical mappings of Density operators in Quantum Mechanics, II Time-dependent mappings) by Jordan, Pinsky, and Sudarshan; Journal of Mathematics Physics, 3, 848 (1962).) t need not refer to the time; we need only have some one-parameter group of mappings. If the parameter is not t , then the generator is not the Hamiltonian.

Instead of saying as in the above that the field operators must transform according to

$$\phi(x) \rightarrow U(\Lambda, a) \phi(x) U^{-1}(\Lambda, a) = \phi(\Lambda x + a)$$

we would equally well have postulated a mapping of the density operators, e.g. of the form above

$$\rho_{\Lambda, a} = U^{-1}(\Lambda, a) \rho_{\Lambda, 0} U(\Lambda, a)$$

5. Statistics of Elementary Particles-Para-
Statistics.

We recall that the above property of fields with different commutation relations and spin is termed the spin statistics connection since if we quantise the field, then a many-particle assembly has one or the other statistics, i.e. symmetry or antisymmetry of the wave function under exchange of two particles.

We have seen that for electrons there is good evidence for Fermi statistics for photons, black-body radiation is presumably good evidence for Bose statistics; for muons from photoproduction of muon pairs; but for pions, the evidence is a little more slim, as noted earlier.

In ordinary quantum mechanics, we have

$$\dot{f}(q, p) = [f, H]$$

In field theory, we would like to say that

$$\dot{\phi} = [\phi, H]$$

even in a particle theory, we would like to say that

$$\dot{q} = [q, H]_{p.b.} \rightarrow -i [q, H]$$

Consider the harmonic oscillator;

$$H = \frac{p^2 + q^2 - 1}{2} \quad ; \quad (\omega = 1)$$

Write

$$a = \frac{q + ip}{\sqrt{2}} \quad ; \quad a^+ = \frac{q - ip}{\sqrt{2}}$$

If commutation relations are used, $H = a^+ a$

Even if these are not used, we have

$$H = 1/2 (a^+ a + a a^+) - 1/2;$$

$$\text{for } a a^+ = 1/2 (q^2 + p^2) + i/2 (pq - qp)$$

$$\text{and } a^+ a = 1/2 (q^2 + p^2) - i/2 (pq - qp)$$

In general, can we deduce a commutation or anticommutation relation for a, a^\dagger ?

Note: If $[a, a^\dagger]_- = \text{scalar}$, the relation (a) would be automatically satisfied. But there are an infinity of other solutions besides the one $[a, a^\dagger]_- = \text{scalar}$. (Note: This latter solution is that satisfied by the oscillator).

Reference: H.S. Green, Phys. Rev. 90, 270, (1953).

The other solutions can be divided into the para-Bose and the para-Fermi classes. We now consider a collection of oscillators representing several degrees of freedom.

The Para-Fermi Case.

$$[b_r, b_s]_+ = 0 \quad ; \quad -[b_r^\dagger, b_s]_+ = \delta_{rs} \quad (i)$$

These satisfy the set of relations (b) below:

$$[b_r, [b_s^\dagger, b_t]_-]_- = \delta_{rs} b_t$$

supplement this by one more relation:

$$\left. \begin{aligned} [b_r^\dagger, [b_s, b_t]_-]_- &= \delta_{rs} b_t - \delta_{rt} b_s \\ [b_r, [b_s, b_t]_-]_- &= 0 \end{aligned} \right\} (b)$$

Then we have

There is a whole variety of other solutions (i.e. other than (i)) satisfying (b).

e.g.

$$\begin{aligned} b_r b_s b_t + b_t b_s b_r &= 0 \\ b_r^* b_s b_t + b_t b_s b_r^* &= \delta_{rs} b_t \\ b_r b_s^* b_t + b_t b_s^* b_r &= \delta_{rs} b_t + \delta_{rt} b_r \end{aligned}$$

(ii)

From (i), we can deduce that $b_n^2 = 0$

From (ii), we can deduce that $b_n^3 = 0$

(From $b_n^3 = 0$ one cannot deduce $b_n^2 = 0$)

Example: Take the operator $J_+ = J_1 + iJ_2$

For a spin 0 field, $J_+ = 0$

For a spin 1/2 field, $J_+^2 = 0$

For a spin 1 field, $J_+^3 = 0$

For a spin 3/2 field, $J_+^4 = 0$, etc.

A particular construction for operators of the type (ii) is as follows:

Take (i) with an index $\alpha = 1, 2, \dots$;

i.e.

$$[a_n^\alpha, a_s^\alpha]_+ = 0 ; [a_n^\alpha, a_s^{\alpha+}]_+ = \delta_{n,s}$$

$$\alpha = 1, \text{ OR } \alpha = 2$$

(iii)

But suppose we postulate that

$$[a_n^1, a_s^2]_- = 0 ; [a_n^1, a_s^{2+}]_- = \delta_{n,s}$$

Define $b_n = a_n^1 + a_n^2$

Then when the a_n^α satisfy (iii), then the b_n satisfy (ii). i.e. Para-Fermi statistics = the direct sum of Fermi oscillators which have commutation relations between different classes, i.e. different classes are "relatively Bose".

Similarly, if we take classes of Bose oscillators, in which the classes are "relatively Fermi", then we get para-Bose statistics.

Example: Take

$$[b_n, [b_s^+ b_t]_+]_- = \delta_{n,s} b_t$$

-111-

And take (iii), with $[]_+$ replaced by $[]_-$

Example: Take a system with one degree of freedom that obeys para-Fermi statistics. Then the Hilbert space is quite small.

$$[a, a^+]_+ = 1 \quad ; \quad a^2 = 0$$

$$N = a^+ a \quad ; \quad N' = a a^+ \quad ; \quad N + N' = 1$$

$$N^2 = a^+ a a^+ a = a^+ a = N \quad \therefore N = 0, 1$$

Suppose there exists a state $|\Omega\rangle$ such that

$$a^+ a |\Omega\rangle = 0;$$

$$\text{then } \langle \Omega | a^+ a |\Omega\rangle = 0; \therefore a |\Omega\rangle = 0$$

All operations that can be expressed as a power series in a and a^+ give a linear combination $|\Omega\rangle$ and of $a^+ |\Omega\rangle = |\Omega'\rangle$. Thus a, a^+ have a representation by these two vectors in a 2-dimensional Hilbert space.

Define

$$a^+ a |\Omega'\rangle = |\Omega'\rangle$$

$$|\Omega\rangle = a |\Omega'\rangle$$

(A representation may be obtained by considering

$$a^+ a |\Omega'\rangle = 0 |\Omega'\rangle \quad \underline{\text{OR}} \quad = |\Omega'\rangle$$

a^+ operating on one basis gives the second basis

$$a^+ |\Omega\rangle = |\Omega'\rangle$$

Now consider a system with

$$b = a' + a^2 \quad ; \quad [a', a^2]_- = 0$$

$$[a', a^{2+}]_- = 0$$

A possible representation of these may be obtained by taking the direct sum of the two representations.

Take the state $|0, 0\rangle$ with

$$a^\alpha |0, 0\rangle = 0, \quad \alpha = 1, 2$$

and states $|n_1, n_2\rangle$ such that

$$a^1{}^\dagger a^1 |n_1, n_2\rangle = n_1 |n_1, n_2\rangle \dots$$

Then

$$a^2{}^\dagger a^2 |n_1, n_2\rangle = n_2 |n_1, n_2\rangle \dots$$

$$b^\dagger |0, 0\rangle = |1, 0\rangle + |0, 1\rangle;$$

$$(b^\dagger)^2 |0, 0\rangle = (a^1{}^\dagger + a^2{}^\dagger) \{ |1, 0\rangle + |0, 1\rangle \}$$

$$= 2 |1, 1\rangle$$

$$(b^\dagger)^3 |0, 0\rangle = 0,$$

Since

$$a^1{}^\dagger |1, 1\rangle = 0 = a^2{}^\dagger |1, 1\rangle,$$

it follows that

$$b |0, 0\rangle = 0,$$

$$b^2 |0, 0\rangle = b \{ |1, 0\rangle + |0, 1\rangle \} = 2 |0, 0\rangle,$$

$$\text{as } (a^1 + a^2) |1, 1\rangle = |0, 1\rangle + |1, 0\rangle$$

The three vectors

$$|0, 0\rangle$$

$$\frac{|1, 0\rangle + |0, 1\rangle}{\sqrt{2}}$$

and

$$|1, 1\rangle$$

then constitute an orthonormal representation

in which $b \rightarrow$

(i)	(ii)	(iii)
$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} \sqrt{2} \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ \sqrt{2} \\ 0 \end{bmatrix}$

We started with a direct-product space a four dimensional space, but here we have obtained a three dimensional representation. Therefore there must exist a vector not contained in the above that can by itself constitute a representation. This vector is

$$|x\rangle = \frac{|1,0\rangle + |0,1\rangle}{\sqrt{2}}$$

We have $b^+|x\rangle = 0; b|x\rangle = 0$; thus this vector is a one dimensional representation.

Such a representation is not possible for a Fermi or Bose statistics.

For a para-Bose statistics, the construction is more complicated.

We have $[a, a^+]_- = 1$

It is more difficult to prove the uniqueness of a representation here, i.e. to prove that all representations are unitarily equivalent.

$$|n\rangle = \frac{(a^+)^n}{\sqrt{n!}} |0\rangle$$

The projection operator for the state $|n\rangle$ is $P_n = |n\rangle\langle n|$

To construct it before the statistics are given is quite difficult, this was achieved for ordinary Bose operators by Von Neumann; his work is quoted and simplified by Wightman and Schweber (reference: Phys. Rev. 98.)

We now ask: Given the Wightman system, supplemented by the hypothesis of TCP invariance, how far can we carry through the program if we require that the field obeys a general statistics? It can be shown that a spin-parastatistics connection generalising the spin-statistics connection can be proved under quite general assumptions. See Dell'Antonio, Greenberg and Sudarshan (Rochester preprint) University of Rochester report NYO - and Kamafuchi and Y. Takahashi (Imperial College preprint).

A para-Fermi Statistics of order λ has

$$(a^+)^{\lambda} = 0 \quad ; \quad (a^+)^{\lambda-1} \neq 0$$

At most, $\lambda - 1$ particles can occupy the same state.

A Bose system is thus an infinite-rank para-Fermi system.

For a para-Bose system, we have

$$a_1^+ a_2^+ |\Omega\rangle \neq \pm a_2^+ a_1^+ |\Omega\rangle$$

Thus we have four 2-particle states, viz., the L.H.S. and R.H.S. of the above,

and the states $\frac{(a_1^+)^2}{\sqrt{2!}} |\Omega\rangle$; $\frac{(a_2^+)^2}{\sqrt{2!}} |\Omega\rangle$

Note: For an ordinary Bose system, we have only three 2-particles states.

6. The metric Postulate.

Ordinary perturbation theory leads to infinities, of which the most serious one is the ultra-violet divergence. Convergence at large momenta can be obtained only by changing the metric in the Hilbert space, and taking an indefinite metric. The theory obtained is finite but there is the difficulty of a non-positive-definite probability.

Can a useful model be constructed by relaxing the positive-definiteness condition on the metric? We can still prove many analytic properties; these do not depend on the integral over momenta but only on the Fourier transform. Further, the asymptotic condition is still obtained; this has nothing to do with the metric.

Useful models with several interesting features have been constructed using an indefinite metric; but lack of time prevents us from going into these here.

(for a discussion of these, ref: Sudarshan: Phys. Rev. 123).