

ESTI FILE COPY

ESD-TR 66-172

MEMOIRS
OF THE
AMERICAN MATHEMATICAL SOCIETY

ESD RECORD COPY

ESD ACCESSION LIST

ESTI Call No. AL 50878

RETURN TO
SCIENTIFIC & TECHNICAL INFORMATION DIVISION
(ESTI), BUILDING 1211

NUMBER 61 Copy No. 1 of 1 cys.

A NEW FORMULATION
OF PARTICLE MECHANICS

R. T. Prosser

Lincoln Laboratory*
Massachusetts Institute of Technology

*Operated with support from the U. S. Air Force.

ESRL

AD0632968

TABLE OF CONTENTS

	Page
Introduction	1
1. Classical Mechanics	3
2. Quantum Mechanics	4
3. Abstract Mechanics	6
4. Discussion	9
5. Preliminary Consequences	15
6. The Structure of the Moment Algebra	20
7. Representations of the Moment Algebra	25
8. Extensions of the Moment Algebra	34
9. Dynamical Aspects	40
10. Statistical Aspects	50
11. Constraints	53
Bibliography	55

Introduction

The mathematical concept of a mechanical system as a system of mass-points whose behavior is determined by elementary laws of motion has a venerable history. The classical formulation of this concept first appeared in the work of Kepler, Galileo, and Newton, and has engaged the attention of nearly every major mathematician since [2]. By 1900 a formidable body of literature, stretching over some three centuries, bore witness to the intensity of the effort invested in the development of this concept. At that time it was quite generally felt that the foundations were so securely established, both theoretically and experimentally, that all that remained to do was to refine the existing experimental techniques and calculating procedures to bring them into ever closer agreement.

Subsequent refinements in techniques and procedures, however, disclosed discrepancies between theory and experiment in both the macrocosmic and the microcosmic world. These discrepancies could be resolved in the one case by rewriting the laws of motion, but could be resolved in the other only by reconstructing the foundations of the theory. The first part of our century has seen the advent of a quantum formulation of the concept of a mechanical system, whose principal features are by now a matter of general agreement, and whose claim to its throne is as well established within its realm as is that of the classical formulation.

The two formulations of the concept of a mechanical system are quite distinct, and neither of them includes the other. Nevertheless, as one passes from the macrocosmic to the microcosmic world, one passes from one to the other via a form of correspondence principle, which holds that all the relevant elements of structure in the old formulation find corresponding elements in the new. In modern textbooks devoted to these subjects, the correspondence principle is embodied in a single phrase, viz., "replace Poisson brackets with commutator brackets throughout" [19].

To a mathematician this principle suggests immediately that there must be an abstract formulation of the concept of a mechanical system which contains both of the accepted formulations as special cases. The description of such a structure would clarify the relations between the two and illuminate the role of the correspondence principle in passing from one to the other.

It is the purpose of this paper to provide such an abstract formulation and to derive its most elementary properties. In sections 1 and 2 we review the common features of classical and quantum mechanics which are relevant to this purpose. In section 3 we set forth an abstract formulation of these common features in terms of a system of axioms. Section 4 is devoted to a discussion of the axioms, with a

review of possible alternatives. In section 5 we show that the structure described in section 3 always admits a representation in terms of operators on a suitably constructed Hilbert space, and that the Heisenberg uncertainty relations are obtained in consequence. In section 6 we verify that the structure of both classical and quantum mechanics satisfy our axioms. We then go on to show that any other form of mechanics which satisfies our axioms must be reducible either to classical mechanics, or to quantum mechanics, or to a suitable combination of the two; no essentially different structures are possible within this framework.

In section 7 we classify all the representations of the abstract structure in terms of operators on Hilbert spaces, assuming that the associated kinetic energy operator is well behaved in a suitable sense. We show that every such representation must be a direct sum of those commonly adopted for the classical and quantum mechanical structures.

In section 8 we present certain extensions of the structure, and prepare the ground for section 9.

In section 9 we describe briefly the dynamical behavior of the theory. Here we show that every infinitesimal canonical motion of the structure is determined by an appropriate generating function through the canonical bracket operation, and that this generating function may be obtained from an appropriate variational problem. In this way we are able to extend to our abstract mechanics, at least in principle, most of the dynamical properties commonly shared by both classical and quantum mechanics. In sections 10 and 11 we discuss briefly statistical aspects and the problem of constraints in the theory.

The study of the foundations of mechanics has not been neglected in recent literature. From our point of view the best presentation of the structure of classical mechanics is the treatise of Whittaker [25], while its counterpart for quantum mechanics is that of von Neumann [16]. Various reformulations of the structure of quantum mechanics are to be found in the book of Weyl [24] and the papers of Jordan, von Neumann, and Wigner [8], and Segal [20], and Mackey [14]. A formulation which also includes the structure of classical mechanics but which differs considerably from ours appears in the recent papers of Jordan and Sudarshan [9, 25, 27, 30]. The work most nearly akin to ours is a paper on the formulation of quantum field theory by Wightman [23], which provided the nucleus for our development.

It is a pleasure to record here our debt to George Mackey and Irving Segal for their numerous comments and suggestions.*

1. Classical Mechanics

We give here a brief description of the structure of classical mechanics in order to emphasize those features which admit an abstract formulation. Only the simplest properties of the simplest types of systems will be considered. We shall not pretend to include within these few pages the whole content of this formidable subject.

We shall assume that we have a system of mass-points, finite in number, whose behavior is determined by the forces acting upon them according to the laws of classical mechanics. At any instant in time, the state of the system is completely determined by a knowledge of the canonical coordinates, and all quantities of physical interest may, at least in principle, be expressed in terms of these coordinates.

Thus, in order to determine the state of the system at any instant in time, it suffices to determine the values of the canonical coordinates. This amounts to effecting a measuring process. If the measuring process is ideal, in the sense that no uncertainties, ambiguities, or errors are introduced into the values of the coordinates, then the measurement assigns a precise value to each coordinate, and in this way determines the state of the system.

No known measuring process, however, is completely free from uncertainties. No known measuring process can guarantee an accuracy beyond a few significant figures, even in the classical domain, and it seems likely that none ever will. This is perhaps an expression of the limitations of man's capability to know nature. In any case it is desirable to include the effects of uncertain measuring processes in our description of classical mechanics.

An uncertain measuring process will assign, not a single value, but a distribution of values to each of the canonical coordinates of the system. We can interpret this statement in the sense of probability, saying that the measurement of the system in a given state will assign a given value or range of values to each coordinate with a prescribed probability; or in the sense of frequency, saying that repeated measurements of the system in the same given state will assign a given range of values to each coordinate a prescribed fraction of the time. An engineer might say that the measuring process is corrupted with noise. In any case we can incorporate the effects of the uncertainties of the measuring process by agreeing that the measuring process assigns to each canonical coordinate a probability distribution taken over all admissible values.

These considerations lead us to the following arrangement. Let x_1, \dots, x_{2n} denote the canonical coordinates, even in number, of the system, and E_{2n} the Euclidean space of admissible $2n$ -tuples of values for the coordinates. A measuring process will determine a probability distribution, or measure, μ , on E_{2n} ,

which determines the state of the system in the sense described above. The average, or expected, value of the i th coordinate, as determined by the measuring process, is just $m_i = \int_{E_{2n}} x_i d\mu$, while the variance from this average value is just $\sigma_i^2 = \int_{E_{2n}} (x_i - m_i)^2 d\mu$. Moreover, the average value of any quantity of physical interest, whose expression in terms of the canonical coordinates is given by $f = f(x_1, \dots, x_{2n})$, is just $\int_{E_{2n}} f d\mu$. In this way the knowledge of μ determines the state of the system.

The entire framework of classical mechanics can be constructed along these lines. But for our purposes it is more convenient to transcribe these results into a different form. We first recall that every (suitably restricted) measure on E_{2n} is completely determined by the values of its *joint moments*. More precisely, if we define, for each n -tuple $(k_1, k_2, \dots, k_{2n})$ of integers, the joint moment $m(k_1, k_2, \dots, k_{2n})$ of μ via

$$(1.1) \quad m(k_1, k_2, \dots, k_{2n}) = \int_{E_{2n}} x_1^{k_1} x_2^{k_2}, \dots, x_{2n}^{k_{2n}} d\mu$$

and if these moments are all finite and not too badly unbounded, then they completely determine the measure μ [21]. This means that a knowledge of the joint moments determines the state of the system.

To rephrase it again, we define for every polynomial $f(x_1, \dots, x_{2n})$ in the x_i the average value

$$(1.2) \quad \omega(f) = \int_{E_{2n}} f d\mu,$$

and record the following properties:

$$(1.3) \quad \begin{aligned} (1) \quad & \omega(1) = 1, \\ (2) \quad & \omega(\alpha f + \beta g) = \alpha \omega(f) + \beta \omega(g), \\ (3) \quad & \omega(f^* f) \geq 0. \end{aligned}$$

Here α and β are any complex scalars, and f^* is the complex conjugate of f . Thus the average value operation ω is a positive linear functional on the algebra of all polynomials with complex coefficients in the x_i . If ω is finite on all polynomials and not too badly unbounded, it completely determines μ , and hence the state of the system.

2. Quantum Mechanics

We now turn to the structure of quantum mechanics, again emphasizing those features which admit an abstract formulation.

Again we assume that we have a system of mass-points, finite in number, whose behavior in time is determined by the forces acting upon them according now to the laws of quantum mechanics. At any instant in time the state of the system is completely determined by a knowledge of the canonical coordinates, and all

quantities of physical interest may, in principle, be expressed in terms of these coordinates. In order to determine the state of the system at any time, it suffices to determine the values of the canonical coordinates by means of a suitable measuring process.

The measuring processes of quantum mechanics, like their classical counterparts, introduce uncertainties in the values which they assign to the canonical coordinates. But the uncertainties of quantum measurements, unlike those of classical measurements, are constrained by the uncertainty relations of Heisenberg. These relations require, roughly speaking, that the variance in the value of any coordinate, multiplied by the variance in the value of its conjugate coordinate, must always exceed a prescribed lower bound.

From our point of view, these relations are best thought of as a property of the measuring processes of quantum mechanics. We shall not attempt here to explain them in physical terms. Rather we shall accept them as a fundamental property of quantum mechanics, and proceed to show that all the usual results of quantum mechanics are consequences.

The requirements of the Heisenberg uncertainty relations lead us to the following arrangement: Let x_1, \dots, x_{2n} denote the canonical coordinates of the system, now considered as noncommuting variables, and let \mathfrak{H} be a Hilbert space upon which the x_i are represented as noncommuting operators. A measuring process will determine a vector z in \mathfrak{H} , such that the average, or expected, value of the i th coordinate, as determined from the measuring process, is just $m_i = (x_i z, z)$, while the variance from this average value is just $\sigma_i^2 = ((x_i - m_i)^2 z, z)$. In this way the knowledge of z determines the state of the system.

The entire framework of quantum mechanics can now be constructed along these lines. But for our purposes it is again more convenient to transcribe these results into functional form. To do so, we need an analogue of the classical result that a probability distribution μ on E_{2n} is determined by its joint moments. Thus we need to know that a vector v in the Hilbert space \mathfrak{H} is determined by its "joint moments". Under suitable restrictions this is in fact the case, as we shall show below. Thus we may now define for each polynomial f in the coordinates x_i with complex coefficients the average value $\omega(f)$

$$(2.1) \quad \omega(f) = (fv, v)$$

and record the following properties:

- $$(2.2) \quad \begin{aligned} (1) \quad & \omega(1) = 1, \\ (2) \quad & \omega(\alpha f + \beta g) = \alpha \omega(f) + \beta \omega(g), \\ (3) \quad & \omega(f^* f) \geq 0. \end{aligned}$$

Here α and β are any complex scalars, and f^* is the operator adjoint to f on \mathcal{H} . In this way we are lead again to a positive linear functional ω on the algebra of all polynomials with complex coefficients in the canonical coordinates x_i . Moreover, if ω is finite on all polynomials and not too badly unbounded, then it completely determines \mathcal{H} and z , and hence the state of the system (cf. section 7).

3. Abstract Mechanics

The preceding sections have shown that in both classical and quantum mechanics each state of a mechanical system is determined by a distribution of admissible values assigned to the canonical coordinates of the system, and that this distribution is in turn determined by its "joint moments". We now proceed from this observation to construct the framework of an abstract form of mechanics.

Once again we assume that we have a system of mass-points, finite in number, whose behavior in time is determined by the forces acting upon them according to an appropriate generalization of the laws of classical and quantum mechanics. At any instant in time the state of the system is completely determined by a knowledge of the canonical coordinates, and the knowledge of these coordinates is determined by means of a suitable measuring process. The measuring process in fact assigns to each of the canonical coordinates a distribution of possible values, and the properties of this distribution are determined by its "joint moments".

With this description in mind, we denote by x_1, \dots, x_{2n} the canonical coordinates of the system, even in number; and we consider the set A of all polynomials $f(x_1, \dots, x_{2n})$ with these coordinates as generators. We shall not assume that the generators commute in these polynomials, so that the order of the factors is an essential part of each polynomial. The elements of A may be added and multiplied together according to the usual rules for polynomials in noncommutative variables, and with these definitions A forms a noncommutative algebra over the complex scalars, with unit element $f \equiv 1$.

The structure of any such algebra may be completely described in terms of the polynomial relations satisfied by its generators. Rather than specify these relations among the generators directly, however, we shall require instead that the algebra satisfy two additional conditions, necessary for our purpose, from which the structural relations may then be derived.

The first of the conditions we shall require is that A admit a *conjugation* operation, denoted by $*$, which assigns to each polynomial f in A a conjugate polynomial f^* in A . Roughly speaking, the conjugation operation reverses the order of the multiplication and complex-conjugates, the coefficients. More precisely, it must have the following properties:

$$(3.1) \quad (\alpha f + \beta g)^* = \bar{\alpha} f^* + \bar{\beta} g^*,$$

$$(3.2) \quad (fg)^* = g^* f^*,$$

$$(3.3) \quad (f^*)^* = f.$$

These properties determine the conjugation operation uniquely in terms of its action on the canonical generators. For them we stipulate

$$(3.4) \quad x_i^* = x_i.$$

With respect to this conjugation operation we shall say that a polynomial f is *Hermitian* if $f = f^*$, and *positive* if $f = \sum g_i^* g_i$ for suitable polynomials g_i in A . We denote the set of Hermitian polynomials by H , and the set of positive polynomials by P . We note that H forms a linear space over the real scalars, and that P forms a cone in this space. Every polynomial f in A is a linear combination of polynomials in H ; in fact $f = (f + f^*)/2 + i(f - f^*)/2i$. Moreover, every polynomial f in H is a linear combination of polynomials in P ; in fact, $f = (f + 1)^2/4 - (f - 1)^2/4$. Thus the linear combinations of positive polynomials span the algebra A .

The second of the conditions we shall require of our algebra A is that it admit a *bracket* operation, denoted by $\{ , \}$, which assigns to each pair of polynomials f, g in A a third polynomial $h = \{f, g\}$ in A . This operation is to form the natural generalization of both the Poisson bracket of classical mechanics and the commutator bracket of quantum mechanics. In particular, it must have the following properties:

$$(3.5) \quad (\alpha f + \beta g, h) = \alpha \{f, h\} + \beta \{g, h\},$$

$$(3.6) \quad \{fg, h\} = \{f, h\}g + f\{g, h\},$$

$$(3.7) \quad \{f, g\} + \{g, f\} = 0.$$

These properties imply that the bracket operation must be bilinear and skew-symmetric. Moreover, they determine the bracket operation uniquely in terms of its action on the canonical generators. For them we require

$$(3.8) \quad \{x_i, x_j\} = \lambda_{ij} 1$$

where the λ_{ij} are real scalars.

We now stipulate that the real skew-symmetric matrix λ_{ij} , which completes the definition of the bracket operation, be nondegenerate. It is known [24] that this requires that the number of canonical generators be even, and that in this case a new set of generators x'_1, \dots, x'_{2n} may be chosen for A such that

$$(3.9) \quad \{x'_1, x'_j\} = \lambda'_{ij} 1$$

where

$$(3.10) \quad \lambda'_{ij} = \begin{cases} +1 & \text{if } j = i + n, \\ -1 & \text{if } j = i - n, \\ 0 & \text{otherwise.} \end{cases}$$

We shall find it convenient to denote this new basis by $p_1, \dots, p_n, q_1, \dots, q_n$ ($i = 1, \dots, n$), with

$$(3.11) \quad p_i = x'_i, \quad q_i = x'_{n+i}.$$

As we have already indicated, the imposition of these two conditions upon the algebra A forces a particular form upon its structure. Not every polynomial algebra in $2n$ generators admits a consistent combination of the conjugation and bracket operations. The determination of those algebras which do so is the business of section 6.

We now turn to the problem of assigning probability distributions to the canonical coordinates of our system. We have learned from sections 1 and 2 that in both classical and quantum mechanics these distributions may be defined in terms of their "joint moments", i.e., in terms of certain linear functionals on the polynomial algebra generated by the canonical coordinates. Hence in order to assign probability distributions to the canonical coordinates of our abstract mechanics, we shall consider A as the *moment algebra* generated by these coordinates, and consider the positive linear functionals on A as *moment sequences*.

A linear functional ϕ on A is a functional which assigns to each polynomial f in A a complex scalar $\phi(f)$, such that

$$(3.12) \quad \phi(\alpha f + \beta g) = \alpha\phi(f) + \beta\phi(g).$$

We shall say that ϕ is *Hermitian* if

$$(3.13) \quad \phi(f^*) = \overline{\phi(f)},$$

that ϕ is *positive* if

$$(3.14) \quad \phi(f^*f) \geq 0,$$

and that ϕ is *normalized* if

$$(3.15) \quad \phi(1) = 1.$$

We denote by A' the set of linear functionals on A , by H' the Hermitian functionals, by P' the positive functionals, and by Ω the normalized positive functionals. The conjugation operation $*$ on A induces a conjugation operation, also denoted by $*$, on A' , via the relation

$$(3.16) \quad \phi^*(f) = \overline{\phi(f^*)}.$$

In particular, if ϕ is Hermitian, then $\phi^* = \phi$. Every linear functional ϕ in A' is a linear combination of Hermitian functionals; in fact, $\phi = (\phi + \phi^*)/2 + i(\phi - \phi^*)/2i$. It is apparently not known whether every Hermitian functional is a difference of positive functionals, and we shall not assume so in the development presented here.

We now have at hand all of the elements of structure necessary for our formulation of abstract mechanics. We proceed to describe this formulation as an *axiom system*, consisting of the following components:

- (1) A set (x_1, \dots, x_n) of possibly noncommuting variables (the canonical coordinates of the system).
- (2) The algebra A of all polynomials in these variables (the moment algebra of the system).
- (3) The conjugation $*$ and bracket $\{ , \}$ operations on this algebra.
- (4) The set Ω of normalized positive linear functions on this algebra (the state space of the system).

These components are to satisfy the axioms necessary for their various definitions, and nothing more.

We have completed our formulation of abstract mechanics. We shall devote the rest of this paper to an investigation of its consequences.

4. Discussion

In this section we shall examine critically the individual components of our formulation of mechanics and attempt to justify their structure on physical grounds. We shall also attempt to compare our formulation with those others in the literature which are known to us, and to review possible alternatives. Our conclusions here are necessarily tenuous, since they are based, on the one hand, on an imperfect understanding of nature, and on the other, on an imperfect understanding of man.

In the first place, we observe that the individual components of our formulation are few and simple; that each has a well-understood mathematical structure which requires no elaborate elucidation, and that each is essential to the whole. Yet together they are sufficient to embody the entire framework of both classical and quantum mechanics, at least in their statistical aspects, without any further hypothesis. The fact that both classical and quantum mechanics do appear as special cases follows directly from the discussion of sections 1 and 2 and requires no additional argument here.

Next we assert that these two special cases are essentially the only ones: any other mechanical system which admits a description in our formulation is necessarily a suitable combination of these two. Our assertion follows directly from the results of section 6. It implies that our formulation of mechanics essentially *characterizes* the two universally recognized forms of mechanics and admits no others.

Turning now to the individual components of our formulation, we observe that the search for a common description of classical and quantum mechanics led us to

the introduction of the canonical variables and their joint moments. The canonical variables, labeling the states of the system, seem to be inherent in any description of a mechanical system and admit a direct physical interpretation as the coordinates of the system.

The polynomial algebra generated by these variables, on the other hand, is not so easily interpreted. What physical meaning is to be attached to an arbitrary polynomial in the canonical variables? We have taken the position that these polynomials represent linear combinations of the joint moments of the canonical variables, and hence that their interpretation is to be primarily a statistical one.

In order to develop this train of thought a little, let us introduce a new operation, $v^2: A \rightarrow A$, into A : To each polynomial f in A we assign a new polynomial $v^2(f)$ defined by

$$(4.1) \quad v^2(f) = f^* f.$$

Then v^2 satisfies the following consistency relations:

$$(4.2) \quad v^2(\lambda f) = |\lambda|^2 v^2(f),$$

$$(4.3) \quad v^2(f + g) + v^2(f - g) = 2(v^2(f) + v^2(g)),$$

$$(4.4) \quad v^2(v^{2(m)} f + v^{2(n)} f) = v^2(2m) f + 2v^{2(m+n)} f + v^2(2n) f$$

where $v^{2(m)}(f)$ denotes the m -fold iterate of f , with $v^{2(0)}(f) = 1$.

Now v^2 readily admits a physical interpretation: If f is any Hermitian polynomial in A and ω any positive linear functional on A , then $\omega(f)$ is the average value assigned to f by the measuring process when the system is in the state associated with ω , and $\omega(v^2(f - \omega(f)))$ is its variance. For this reason we shall speak of v^2 as the *variance operator* on A .

In order to determine the statistical distributions assigned to the canonical variables, we shall certainly need to know variances as well as averages. This means that we shall need to know the averages of all polynomials obtainable from the canonical variables by repeated application of the operations of addition, scalar multiplication, and variance.

But the least subset of A which contains the canonical variables and is closed under these operations is A itself. In fact, if f and g are in this subset, then so is $f^* g$, since

$$(4.5) \quad 4f^* g = v^2(f + g) - v^2(f - g) - iv^2(f + ig) + iv^2(f - ig).$$

In particular, if f is in this subset, so is f^* (put $g = 1$) and so is fg (replace f by f^*). It follows that all the polynomials of A are in this subset, as asserted.

It is possible to *define* the algebra A as the least complex linear space containing the canonical variables and admitting a variance operator satisfying the

relations (4.1)–(4.4). It can then be shown that A becomes an associative algebra with multiplication and conjugation defined by (4.5). We have chosen a simpler course, and shall not pursue this development here.

A similar objection can be leveled against the conjugation operation, that it has no direct physical interpretation. But we can answer in the same way, that the conjugation operation is necessary in the consistent determination of variances, as shown by the fact that it may be derived from the variance operator via (4.5).

A similar objection can be leveled against the use of the complex numbers as the scalar field. Multiplication by real numbers can be interpreted in terms of changes of scale, but non-real numbers present a special problem. What is the significance of a multiplication by $\sqrt{-1}$? Again we can answer in the same way, by saying that a consistent determination of variances in general requires the use of an algebraically closed scalar field. This is not true in the special case of classical mechanics, where real scalars are sufficient, but it is true in the special case of quantum mechanics, where the commutation relations of the canonical variables involve $\sqrt{-1}$ directly. Attempts to formulate quantum mechanics over the field of real numbers alone have not heretofore been completely successful. The difficulty seems to be that the conjugation operation, already shown to be essential, must act nontrivially on the algebra if the canonical variables do not commute, since it reverses the order of multiplication. In particular, if p and q are canonical variables with $pq - qp = \lambda 1$, then the properties of the conjugation operation require that λ be purely imaginary. This conclusion seems unavoidable.

We turn now to the bracket operation. The bracket operation is the agency through which the development in time of the system is to be expressed. We know that in both classical and quantum mechanics the equations of motion of the system can be completely described in terms of the associated bracket operation. Moreover, these two bracket operations are closely related through the correspondence principle. We have set out to develop a framework which contains both of these operations as special cases, and have therefore included a bracket operation in our formulation.

The properties we have required of our bracket operation are common to both special cases, and are based on the following considerations: We assume that the system admits a complete set of infinitesimal motions of the form $f \rightarrow f + \delta f$, each of which is determined by the canonical coordinates x_i and a prescribed generating function h in A , according to a relation of the form

$$(4.6) \quad \delta f = F(h, f).$$

We assume that this relation is linear in h ,

$$(4.7) \quad F(\alpha h + \beta k, f) = \alpha F(h, f) + \beta F(k, f),$$

and that h is a constant of the motion,

$$(4.8) \quad F(h, h) = 0.$$

From the properties of infinitesimal motions we conclude that

$$(4.9) \quad F(h, \alpha f + \beta g) = \alpha F(h, f) + \beta F(h, g)$$

and that

$$(4.10) \quad F(h, fg) = F(h, f)g + fF(h, g).$$

It follows that $F(h, f)$ is bilinear and skew-symmetric on $A \times A$, and is completely determined in terms of its action on the canonical variables.

For these values we specify that if $h = x_j$, then the resulting motion is an infinitesimal translation:

$$(4.11) \quad \delta x_i = F(x_i, x_j) = \lambda_{ij}1$$

with λ_{ij} a skew-symmetric matrix of real numbers. If all translations are to be admissible as motions, then it follows that the matrix λ_{ij} must be nondegenerate.

These properties suffice to identify F with the bracket operation. Thus we see that the properties of the bracket operation are necessary if the system is to admit a complete set of infinitesimal motions. In section 9 we shall see that these properties are also sufficient.

The same general considerations which gave rise to the moment algebra of all polynomials in the canonical variables have led us to introduce the state space of all positive linear functionals on this algebra. These functionals are best interpreted as assigning expectation values to the moments of the canonical variables, and their properties are all immediate consequences of this role.

One might ask here whether every positive linear functional on the moment algebra can in principle be realized as a state of the system. In order to answer this question, let us consider the set S of all functionals which can be so realized, and try to determine the structure of S . We expect that S must have at least the following properties:

- (1) $\omega, \eta \in S \iff \lambda\omega + (1-\lambda)\eta \in S, \quad 0 < \lambda < 1,$
- (2) $\omega \in S \implies \sigma^*\omega \in S$ for all translations σ of A ,
- (3) $\omega_n \in S$ and $\omega_n \rightarrow \omega \implies \omega \in S$.

The first of these is simply an expression of the strong superposition principle. The second says that S is invariant under all translations $\sigma: x_i \rightarrow x_i + \lambda_i 1$ of A . The third says that S is closed under the taking of (weak) sequential limits.

Now it can be shown that in the two special cases of classical mechanics and quantum mechanics these properties suffice to ensure that S contain every positive linear functional on A . In the general case this result is no longer true.

If we also require, however, that A be *simple*, in the sense that it contains no two-sided ideals which are stable under the conjugation and bracket operations, then it will follow from the results of section 6 that A is isomorphic with the moment algebra of either classical or quantum mechanics, and hence that S must contain all the positive linear functionals on A .

We shall now attempt to compare our formulation of mechanics with some of the others which have appeared in the literature in recent years. We are concerned here only with those formulations which can be presented entirely in mathematical terms, and shall not undertake to examine questions of physics or philosophy.

The recognition that classical mechanics may be formulated in terms of functions of the canonical variables seems to be due to B. O. Koopman [12] who observed that the motion of the classical system is determined by the motion of the Hilbert space of all square-integrable functions on the phase space. This formulation has provided a starting point for much of the subsequent work on classical statistical mechanics, and has made it possible to compare classical and quantum mechanics on the same footing. As we shall see, this formulation is easily derived from ours under the additional assumption that all the canonical variables commute.

The most widely recognized and generally accepted formulation of quantum mechanics is due to von Neumann [16] who observed that the diverse descriptions of Heisenberg and Schrödinger of the motion of a quantum system could be recast in terms of the motion of the Hilbert space of all square-integrable functions on the configuration space. This formulation remains the basis of nearly all subsequent developments, in spite of intensive efforts to find a superior alternative. It, too, is easily derived from ours under the additional assumption that the canonical variables satisfy the familiar Heisenberg commutation relations.

Both of these formulations are usually presented in terms of Hilbert spaces. From our point of view the concept of a Hilbert space is too complicated and too far removed from human experience to serve as a fundamental component of any description of nature. For this reason we have replaced in our formulation the concept of a Hilbert space by that of a statistical distribution, whose properties seem to us more properly fundamental to a statistical theory, and more easily justifiable on physical grounds. The Hilbert space is still intrinsic in our formulation, but now appears as a mathematical construction derived from the fundamental components, rather than as a fundamental component in itself.

Another difference in point of view centers around the concept of observables. In von Neumann's formulation of quantum mechanics, as well as in its classical counterpart, every bounded Hermitian operator on the Hilbert space is taken as an observable entity, on the grounds that every such operator is a function of the canonical variables. From our point of view, only the joint moments—the polynomials—

in the canonical variables, together with certain preassigned functions – the generators of the infinitesimal motions – are necessarily observable, on the grounds that only these functions can be assigned a direct interpretation. All other functions are still intrinsic in our formulation, but do not appear as fundamental.

Several attempts have been made to find a description of quantum mechanics that does not involve the concept of Hilbert space from the outset. In particular, Jordan, Wigner, and von Neumann [8] have tried to base a description upon the Jordan algebra of bounded observable operators. This attempt was vitiated, however, by the discovery of the existence of exceptional Jordan algebras, which do not admit representations as operators on a Hilbert space. No systematic way of excluding the exceptional Jordan algebras from their description has yet been discovered.

Irving Segal [20] has restated their description in terms of C^* algebras, and has developed an extensive and highly ingenious theory which illuminates the darker corners of von Neumann's original description. But Segal's work is open to the objection that the concept of a C^* algebra is even more sophisticated and less familiar than that of a Hilbert space, and moreover excludes the canonical variables themselves from the domain of definition.

George Mackey [14] has devised an alternative formulation of quantum mechanics based on the lattice of all projection operators on the underlying Hilbert space. He has given persuasive arguments in support of his view that the projection lattice is a natural consequence of the physical specifications of quantum mechanics. Yet it seems to us here that the projection lattice is also a complicated structure, and does not characterize the essential features of quantum mechanics. Moreover, no way to exclude those lattices which cannot be represented by projections on a Hilbert space has yet been found.

T. F. Jordan and E. C. G. Sudarshan have recently presented a framework for both classical and quantum mechanics and discussed its various representations [9, 30]. This framework takes as its fundamental component the Lie algebra generated by functions of the canonical variables under the bracket operation. This Lie algebra, however, is infinite dimensional, and its structure has not been worked out. In particular, the authors have not ruled out the possibility of representations which have nothing to do with mechanics. Nevertheless, their formulation has several points in common with ours. (See also [25, 27].)

Finally, we mention that A. S. Wightman has given a description of quantum field theory in terms of the expectation values of the fields, and shown that the theory can be recovered from a knowledge of these expectation values [23]. This description can be rewritten for both the classical and quantum mechanics of particles, and provides an elegant approach to the statistical aspects of both.

Our formulation was first suggested by the work of Wightman, and owes much of its character to his ideas.

In summary, we assert that the formulation of mechanics presented here offers the advantages of simplicity, compactness, and relevance over previous formulations appearing in the literature. Moreover, it encompasses both classical and quantum mechanics, and characterizes them as essentially the only admissible systems. Its greatest strength, however, is at the same time its greatest weakness: Our formulation can account only for those features of either system which are common to both.

5. Preliminary Consequences

In this section we shall derive from the axiom system of section 3 the first consequences. We shall show first that each positive linear functional on A assigns to each Hermitian polynomial in A a distribution of possible values, which may be interpreted as the probability distribution of obtaining those values in a single measurement. We shall show next that each positive linear functional on A determines a representation of A as operators on a Hilbert space, such that the action of the functional on A can be recovered from the structure of the Hilbert space. Finally we shall show that the uncertainty inherent in the joint distribution assigned to any pair of Hermitian polynomials by the functional is bounded below by the modulus of the expected value of their commutator. These results are all an integral part of the statistical aspects of both classical and quantum mechanics, and must be obtainable from the fundamental components of any formulation of mechanics which encompasses both.

Theorem 5.1. *For each $\omega \in \Omega$ and Hermitian $f \in A$, there exists a probability measure μ on the real line E_1 such that*

$$(5.1) \quad \omega(f^n) = \int_{E_1} \xi^n d\mu(\xi), \quad n = 0, 1, 2, \dots$$

Proof. Let $P(\xi)$ be any polynomial with complex coefficients in the one variable ξ on E_1 , and assign to $P(\xi)$ the element $P(f)$ in A . Now observe that this assignment defines a homomorphism ι of the algebra $A(\xi)$ of all polynomials in ξ on E_1 onto the subalgebra $A(f)$ of all polynomials in f in A . The adjoint ι^* of this homomorphism maps the adjoint space of all linear functionals on $A(f)$ into the adjoint space of all linear functionals on $A(\xi)$ via the formula

$$(5.2) \quad (\iota^* \phi)(P) = \phi(\iota P).$$

In particular, it assigns to ω a functional $\iota^* \omega$ which is positive on all polynomials of the form $Q^* Q$, since $(\iota^* \omega)(Q^* Q) = \omega(\iota(Q^* Q)) = \omega((\iota Q)^* (\iota Q)) \geq 0$. Now it is well known that any polynomial $P(\xi)$ which is positive for all values of ξ can always be expressed in the form $Q^* Q$. It follows that if $P(\xi) \geq 0$ for all ξ ,

then $(t^*\omega)(P) \geq 0$.

Thus we are in the situation described by the celebrated Hamburger moment problem [21]. According to its solution, there exists a finite measure μ on E_1 such that

$$(5.3) \quad (t^*\omega)(P) = \int_{E_1} P(\xi) d\mu(\xi)$$

for all polynomials P in $A(\xi)$. If we take $P(\xi) = \xi^n$ and use (5.2) we get (5.1).

In this way ω assigns to each Hermitian f a probability distribution μ of real numbers. This distribution has a simple interpretation: The probability that the value obtained for f in a single measurement will lie in a given interval is just the μ -measure of that interval. In particular, if the μ -measure of that interval is zero, then the probability of obtaining in a single measurement a value for f which lies in that interval is zero. Thus, if the measure μ consists entirely of a sum of isolated point measures, then the values obtained for f in a single measurement must fall at one of the points determined by the point measures. This is a restatement of the conventional requirement, valid for both classical and quantum mechanics, that the values obtained from a single measurement be eigenvalues of the associated operator. In the present context it appears not as an eigenvalue problem but as a property of the statistics.

The measure μ obtained from ω in this way need not be unique, since in general the solution of the moment problem is not unique. A sufficient condition for uniqueness is given in the following corollary:

Corollary 5.2. *The measure μ associated with ω in Theorem 5.1 is unique if ω satisfies*

$$(5.4) \quad \sum_{n=0}^{\infty} \frac{\omega(f^{2n})}{(2n)!} < \infty.$$

Proof. This follows from the usual sufficiency condition for uniqueness of the solution of the moment problem [21].

It is tempting to try to extend this result to obtain joint distributions for several commuting variables. The proof would proceed exactly as in Theorem 5.1 and lead to a moment problem in several dimensions. It is known, however, that a polynomial in several real-valued variables which is positive for all values of the variables is not necessarily expressible as a sum of polynomials of the form Q^*Q [1]. Thus ω is not necessarily positive on such polynomials, and the conditions for the solution of the moment problem are not necessarily satisfied. By requiring a little more of ω , however, we can establish the desired extension of Theorem 5.1.

Definition 5.3. Let $A(f_1, \dots, f_k)$ be a subalgebra of A generated by the k commuting Hermitian elements f_1, \dots, f_k of A . Let ω be a positive linear

functional on A . Then ω is strictly positive on $A(f_1, \dots, f_k)$ if $\omega(P(f_1, \dots, f_k)) \geq 0$ for all polynomials P of k variables for which $P(\xi_1, \dots, \xi_k) \geq 0$.

Theorem 5.4. Let $\omega \in \Omega$ and $f_1, \dots, f_k \in A$ with f_i Hermitian and $f_i f_j = f_j f_i$ for all $i, j = 1, \dots, k$. If ω is strictly positive on $A(f_1, \dots, f_k)$, then there exists a probability measure μ on E_k such that

$$(5.5) \quad \omega(P(f_1, \dots, f_k)) = \int_{E_k} P(\xi_1, \dots, \xi_k) d\mu(\xi_1, \dots, \xi_k)$$

for all polynomials $P(f_1, \dots, f_k)$ in $A(f_1, \dots, f_k)$.

Corollary 5.5. The measure μ is unique if ω satisfies the condition

$$(5.6) \quad \sum_{n=0}^{\infty} \frac{\omega(h^{2n})}{(2n)!} < \infty$$

where $h^2(f_1, \dots, f_k) = f_1^2 + \dots + f_k^2$.

The proofs of these results are exact analogues of the proofs of Theorem 5.1 and its corollary, and will be omitted here.

Not every positive linear functional on A is necessarily strictly positive on $A(f_1, \dots, f_k)$ (see [26], p. 330). It is possible to show, however, that if ω satisfies the condition (5.6), then it must be strictly positive on $A(f_1, \dots, f_k)$ [15].

This condition, while sufficient, is certainly not necessary, and the exact situation remains obscure.

If the k elements f_1, \dots, f_k of A do not commute, then ω can no longer be represented in the form (5.5). It is possible, however, to represent it in terms of the inner product on a certain Hilbert space. The next theorem makes this statement precise.

Theorem 5.6. Let $\omega \in \Omega$ and $f_1, \dots, f_k \in A$, with f_i Hermitian. Then there exists a Hilbert space \mathcal{H} with a distinguished vector v and a representation ρ of $A(f_1, \dots, f_k)$ as an algebra of operators on \mathcal{H} with common dense domain, such that

$$(5.7) \quad \omega(P(f_1, \dots, f_k)) = (\rho P(f_1, \dots, f_k)v, v)$$

for all polynomials $P(f_1, \dots, f_k)$ in $A(f_1, \dots, f_k)$.

For the proof of this theorem we need the following lemmas:

Lemma 5.7. For all $\phi \in P'$ and all $f, g \in A$,

$$(5.8) \quad |\phi(f^*g)|^2 \leq \phi(f^*f)\phi(g^*g).$$

Proof. This is just an unfamiliar form of the familiar Schwarz inequality. The proof runs as follows: For any scalars α and β we have

$$(5.9) \quad \begin{aligned} 0 \leq \phi((\alpha f - \beta g)^*(\alpha f - \beta g)) &= |\alpha|^2 \phi(f^*f) - \bar{\alpha}\beta \phi(f^*g) \\ &\quad - \bar{\beta}\alpha \phi(g^*f) + |\beta|^2 \phi(g^*g). \end{aligned}$$

If we choose $\alpha = (\phi(g^*g)\phi(f^*g))^{1/2}$ and $\beta = (\phi(f^*f)\phi(g^*f))^{1/2}$ then, using the fact that ϕ is Hermitian, we find

$$(5.10) \quad 0 \leq (\phi(f^*f)\phi(g^*g))^{1/2} - (\phi(f^*g)\phi(g^*f))^{1/2}$$

from which (5.8) follows.

Lemma 5.8. *For all $\phi \in P'$, $\phi(1) = 0$ implies $\phi = 0$.*

Proof. For all $f \in A$, we have $0 \leq |\phi(f)|^2 = |\phi(1^*f)|^2 \leq \phi(1^*1)\phi(f^*f) = 0$, which implies $\phi(f) = 0$.

Lemma 5.9. *For any $\phi \in P'$ let $J = \{f: \phi(f^*f) = 0\}$. Then J is a left ideal in A .*

Proof. We must show that if $f, g \in J$ and $h \in A$, then λf , $f + g$, and hf lie in J . That λf lies in J is obvious. For $f + g$, observe that $(f + g)^*(f + g) + (f - g)^*(f - g) = 2(f^*f + g^*g)$, from which it follows that $0 \leq \phi((f + g)^*(f + g)) \leq 2(\phi(f^*f) + \phi(g^*g)) = 0$. Hence $\phi((f + g)^*(f + g)) = 0$, and $f + g$ lies in J . For hf , consider the linear functional θ on A defined by $\theta(h) = \phi(f^*hf)$. Then $\theta(h^*h) \geq 0$, so $\theta \in P'$, and $\theta(1) = \phi(f^*f) = 0$, since $f \in J$, so $\theta = 0$ by Lemma 5.8. In particular, $\phi((hf)^*(hf)) = \theta(h^*h) = 0$, so $hf \in J$, and the proof is complete.

Proof of Theorem 5.6. Let us denote by B the subalgebra $A(f_1, \dots, f_k)$. Given $\omega \in \Omega$, we put $J = \{f \in B: \omega(f^*f) = 0\}$ and form the space $\bar{B} = B/J$ of residue classes of B modulo J . We define on \bar{B} an inner product $(\bar{f}, \bar{g}): \bar{B} \times \bar{B} \rightarrow C$ via

$$(5.11) \quad (\bar{f}, \bar{g}) = \omega(g^*f),$$

for all $f, g \in B$, where $\bar{f} = f + J$, $\bar{g} = g + J$. Since ω is a positive linear functional, this mapping satisfies all the requirements of an inner product save perhaps definiteness. But if $(\bar{f}, \bar{f}) = 0$, then $\omega(f^*f) = 0$ so $f \in J$, and $\bar{f} = 0$. Thus \bar{B} carries an inner product, and its completion under this inner product forms a Hilbert space, which we denote by \mathcal{H} .

We now assign to each element f of B an operator $\rho(f)$ acting on \bar{B} via

$$(5.12) \quad \rho(f)\bar{g} = \overline{fg}.$$

Since J is a left ideal, the definition of $\rho(f)$ is independent of the choice of representative g of \bar{g} . It is easy to verify that $\rho(\alpha f + \beta g) = \alpha\rho(f) + \beta\rho(g)$ and $\rho(fg) = \rho(f)\rho(g)$. Moreover, we have $(\rho(f)\bar{g}, \bar{h}) = (\overline{fg}, \bar{h}) = \omega(h^*fg) = (\bar{g}, \overline{f^*h}) = (\bar{g}, \rho(f^*)\bar{h})$, so that conjugation is represented by a suitable restriction of the adjoint. In particular, the Hermitian elements of B are mapped onto symmetric operators on \mathcal{H} .

Thus ρ represents B as an algebra of operators acting on \mathcal{H} , with common dense domain \bar{B} . We now denote by v the distinguished vector $\bar{1} = 1 + J$ in \mathcal{H} ,

and observe that for any element $f = P(f_1, \dots, f_k)$ of B , we have the relation (5.7). The proof of Theorem 5.6 is complete.

This construction is due originally to Gelfand and Neumark [6], who used it to obtain representations of normed algebras with conjugations. In that case the algebras are represented by algebras of bounded operators on \mathfrak{H} . The construction, however, makes no essential use of the norm, and extends easily to any associate algebra with involution. In this case the representing operators need not be bounded, and are defined only on a common invariant dense domain in \mathfrak{H} .

This construction is sometimes used in the literature of quantum field theory, where it is referred to as "the Gelfand construction". No proof of its validity for algebras without norms, however, is known to us.

Our final result in this section gives a form of the Heisenberg uncertainty relations for this general setting.

Definition 5.10. Let f be any element of A and ω any state in Ω . The variance $(\Delta f)^2$ of f in the state ω is defined by the formula

$$(5.13) \quad (\Delta f)^2 = \omega(f^* f) - \overline{\omega(f)} \omega(f).$$

Theorem 5.11. Let f, g be any pair of Hermitian elements of A , and ω any state of Ω . Then the variances $(\Delta f)^2$ and $(\Delta g)^2$ of f and g in the state ω must satisfy the inequality

$$(5.14) \quad (\Delta f)^2 (\Delta g)^2 \geq |\omega(h)|^2$$

where $h = i(fg - gf)/2$.

Proof. First, suppose $\omega(f) = \omega(g) = 0$. Then $(\Delta f)^2 = \omega(f^2)$ and $(\Delta g)^2 = \omega(g^2)$. So

$$(5.15) \quad (\Delta f)^2 (\Delta g)^2 = \omega(f^2) \omega(g^2) \geq |\omega(fg)|^2$$

by (5.8). Now $fg = (fg + gf)/2 + (fg - gf)/2$, and so

$$(5.16) \quad \begin{aligned} |\omega(fg)|^2 &= \frac{1}{4} |\omega(fg + gf) + \omega(fg - gf)|^2 \\ &= \frac{1}{4} |\omega(fg + gf)|^2 + \frac{1}{4} |\omega(fg - gf)|^2. \end{aligned}$$

The cross terms here vanish, since ω is Hermitian, and $\overline{\omega(fg)} = \omega(gf)$. Combining (5.15) and (5.16), we obtain

$$(5.17) \quad (\Delta f)^2 (\Delta g)^2 \geq \frac{1}{4} |\omega(fg - gf)|^2 = |\omega(h)|^2$$

as required.

If $\omega(f)$ and $\omega(g)$ are not zero, then we replace f by $f - \omega(f)1$ and g by $g - \omega(g)1$, noting that this replacement cancels the mean and does not affect the variance. Thus Theorem 5.11 is proved.

This theorem says that the product of the variances of f and g is bounded below by the square of the mean of their commutator. This means that the probability distributions assigned by ω to f and g cannot be completely independent unless f and g commute. Here, perhaps, is the most direct expression of the statistical significance of the non-commutativity of the moment algebra A .

6. The Structure of the Moment Algebra

We have already indicated in section 3 that the requirement that the moment algebra A carry the conjugation and bracket operations forces certain relations upon the canonical variables. In this section we set out to determine these relations explicitly. We shall show that the canonical variables generate a certain finite dimensional nilpotent Lie algebra. It follows that A is then a homomorphic image of the universal enveloping algebra of this Lie algebra, and the structure of A can be described completely from this fact.

We begin with a simple lemma.

Lemma 6.1. *Let x be any linear combination of the canonical variables x_1, \dots, x_{2n} and f any element of A . Then $\{x, f\} = 0$ implies $[x, f] = 0$.*

Proof. If $x = 0$, the conclusion is trivial. Otherwise, let y be any linear combination of the canonical variables such that $\{x, y\} \neq 0$. By multiplying y by a suitable scalar factor, we may assume that $\{x, y\} = 1$. Now consider the combination $\{x^2, yf\}$. We expand it in two different ways, using the properties of the bracket operation:

$$\begin{aligned} \{x^2, yf\} &= \{x, yf\}x + x\{x, yf\} \\ (6.1) \qquad &= \{x, y\}fx + y\{x, f\}x + x\{x, y\}f + xy\{x, f\} \\ &= fx + xf, \end{aligned}$$

$$\begin{aligned} \{x^2, yf\} &= \{x^2, y\}f + y\{x^2, f\} \\ (6.2) \qquad &= \{x, y\}xf + x\{x, y\}f + y\{x, f\}x + yx\{x, f\} \\ &= xf + xf. \end{aligned}$$

Subtracting (6.2) from (6.1) we find $0 = fx - xf$, as required.

From this lemma we obtain our first structure theorem.

Theorem 6.2. *Assume the canonical variables p_i, q_i are chosen so that $\{p_i, p_j\} = \{q_i, q_j\} = 0$ and $\{p_i, q_j\} = \delta_{ij}1$. Then we have*

$$(6.3) \qquad [p_i, q_j] = [q_i, q_j] = 0,$$

$$(6.4) \qquad [p_i, q_j] = i\delta_{ij}z \text{ for some } z \text{ in the center } Z \text{ of } A,$$

$$(6.5) \qquad [p_i, z] = [q_i, z] = 0.$$

Proof. For (6.3), put $x = p_i$ and $f = p_j$. Then Lemma 6.1 applies and tells us that $[p_i, p_j] = 0$. Similarly, $[q_i, q_j] = 0$ and $[p_i, q_j] = 0$ if $i \neq j$.

Now put $x = p_i$ and $z_i = [p_i, q_i]$, and then observe that $\{x, z_i\} = \{p_i, p_i q_i\} - \{p_i, q_i p_i\} = p_i - p_i = 0$. Hence Lemma 6.1 applies again, and yields $[p_i, z_i] = 0$. Similarly, $[q_i, z_i] = 0$.

It remains to show that the z_i are all equal. For this, put $x = p_i + p_j$ and $f = q_i - q_j$. Then $\{x, f\} = 0$, so $[x, f] = [p_i, q_i] - [p_j, q_j] = z_i - z_j = 0$. The proof is complete.

As an immediate consequence, we have the following corollary.

Corollary 6.3. *If f and g are any two polynomials in A , then*

$$(6.6) \quad [f, g] = iz\{f, g\}.$$

Proof. It suffices to prove this for monomials. If f and g are monomials of degree 1, then the result is a restatement of the relations (6.3)–(6.4). Suppose by induction that it holds for all monomials f of degree d and g of degree 1, and consider the case where degree $f = d + 1$, degree $g = 1$. Write $f = hk$, where degree $h = 1$ and degree $k = d$, and compute: $[f, g] = [hk, g] = [h, g]k + h[k, g] = iz(\{h, g\}k + h\{k, g\})$ (by the induction hypothesis) $= iz\{hk, g\} = iz\{f, g\}$, as required. A similar induction on the degree of g now completes the proof.

The relations (6.3)–(6.5), taken together, say that the variables $p_i, q_i,$ and z form the basis for an $n + 1$ dimensional nilpotent Lie algebra [7]. It seems appropriate to call this Lie algebra the Heisenberg algebra, and the associated Lie group the Heisenberg group, since, as we shall see, its irreducible representations yield the Heisenberg commutation relations. It now follows from the theory of Lie algebras that the algebra A generated by the $p_i, q_i,$ and z must be a homomorphic image of the complex associative enveloping algebra of the Heisenberg algebra [7]. In particular, if all the relations in A are consequences of (6.3)–(6.5), then A is an isomorphic image of, and hence may be identified with, this enveloping algebra.

We shall verify below that this universal enveloping algebra, generated by the p_i and q_i and subject only to the relations (6.3)–(6.5) and their consequences, does indeed admit both a conjugation and a bracket operation, and so satisfies all the requirements of our formulation of a moment algebra. Let us anticipate this result for a moment, and denote this algebra by A_∞ . Then any other algebra which satisfies our requirements must involve the relations (6.3)–(6.5), together with others which are not consequences of these, and hence must be a homomorphic image of A_∞ . The problem of determining these other algebras therefore reduces to that of determining the homomorphisms of A_∞ which preserve the conjugation and bracket operations. This problem reduces in turn to that of determining the

two-sided ideals in A_∞ which are stable under the conjugation and bracket operations.

In order to carry out this program, we must first establish a normal form for the polynomials in A_∞ which eliminates the ambiguity in the form of these polynomials resulting from the relations (6.3)–(6.5).

Definition 6.4. A monomial $f(p_i, q_i, z)$ in A is expressed in *normal form* if each p_i stands to the left of each q_i . A polynomial is in normal form if it is a linear combination of monomials all in normal form.

Thus $p_1^2 q_1^3 z$ is in normal form, but $p_1 q_1^2 p_1 q_1 z$ is not. Since the p_i commute among themselves, as do the q_i , the relative order of the p 's and q 's is immaterial. It is convenient to take the form $z^r p_1^{m_1} \cdots p_k^{m_k} q_1^{n_1} \cdots q_k^{n_k}$ as the normal form for monomials, and we shall do so hereafter.

Lemma 6.5. *Every polynomial in A_∞ has a unique normal form.*

Proof. We proceed by induction on the degree. The statement is trivially true for all polynomials of total degree 0 or 1. If the total degree is 2, the only monomials not already in normal form are the $q_i p_i$, which can be reexpressed as $p_i q_i - iz$.

Suppose the conclusion true for polynomials of total degree d , and consider a monomial f of total degree $d + 1$. If it contains no p_i , it is already in normal form. If it contains p_i , say, then using the relation $q_i p_i = p_i q_i - iz$, we may rewrite it as a polynomial of the form $p_i g + h$, where g and h are polynomials of degree $\leq d$. Now using the induction hypothesis we can rewrite g and h in normal form, which also gives us the normal form for f .

The uniqueness follows from the fact that if two polynomials in normal form are equal, then their difference, which is also in normal form, vanishes. But no polynomials in normal form can vanish in A_∞ , since all relations in A_∞ are consequences of (6.3)–(6.5), which cannot themselves be placed in normal form.

This result is a special form of the Birkhoff-Witt theorem, valid for all Lie algebras [7].

In terms of the normal form we may define the partial derivatives $\partial/\partial p_i$ and $\partial/\partial q_i$ on A . These operators act according to the usual rules for differentiation of polynomials, except that care must be taken with the order of multiplication. If f is in normal form, then clearly so are the $\partial f/\partial p_i$ and $\partial f/\partial q_i$. Moreover, they have the following desirable property.

Lemma 6.6. *Let f be any polynomial in A_∞ , with $\partial f/\partial p_i = \partial f/\partial q_i = 0$. Then the normal form of f is a polynomial in z alone.*

Proof. Write f in normal form, and use the usual rules of differentiation. The derivative $\partial f/\partial p_i$ can vanish only if the normal form of f contains no p_i .

Similarly for q_i . The result now follows from the uniqueness of the normal form.

Next we characterize certain commutators in terms of these derivatives:

Lemma 6.7. *Let f be any polynomial in A_∞ . Then $[p_i, f] = iz(\partial f/\partial q_i)$, and $[q_i, f] = -iz(\partial f/\partial p_i)$.*

Proof. It suffices to prove this for monomials. If f is a monomial of degree 1, then we have again the relations (6.3)–(6.4). Assuming the result true if degree $f = d$, suppose degree $f = d + 1$. Factor f as hk , with degree $h = 1$ and degree $k = d$. Then

$$[p_i, f] = [p_i, hk] = [p_i, h]k + h[p_i, k] = iz \left[\frac{\partial h}{\partial q_i} k + h \frac{\partial k}{\partial q_i} \right] = iz \frac{\partial f}{\partial q_i},$$

as required. Similarly, $[q_i, f] = -iz(\partial f/\partial p_i)$, and this completes the proof.

Corollary 6.8. *Let f be any polynomial in A_∞ , with $[p_i, f] = [q_i, f] = 0$. Then the normal form of f is a polynomial in z alone.*

Proof. Combine Lemma 6.6 and Lemma 6.7.

We now turn to the ideal structure of A_∞ .

Lemma 6.9. *Let J be a non-trivial two-sided ideal in A_∞ . Then J contains a non-zero element whose normal form is a polynomial in z alone.*

Proof. Let f be any non-zero element in J whose total degree in the p_i and q_i is minimal. Then $[p_i, f] = iz(\partial f/\partial q_i)$ and $[q_i, f] = -iz(\partial f/\partial p_i)$ are also elements of J , whose total degree in the p_i and q_i is less by one than that of f . It follows that $[p_i, f] = [q_i, f] = 0$, and hence that the normal form of f is a polynomial in z alone.

Let Z denote the subalgebra of A_∞ consisting of all elements whose normal form is a polynomial in z alone. Then Corollary 6.8 tells us that Z is the center of A_∞ ; i.e., Z consists precisely of all elements of A_∞ which commute with all the canonical variables, and hence with every element of A_∞ . Lemma 6.9 says that every non-trivial two-sided ideal in A_∞ has a non-trivial intersection with Z .

We shall say that a two-sided ideal J of A_∞ is *stable under the conjugation operation* if $f \in J$ implies $f^* \in J$; and *stable under the bracket operation* if $f \in J$ implies $\{f, g\} \in J$ for all $g \in A_\infty$. We now show that those non-trivial ideals of A_∞ which are stable under the conjugation and bracket operations are in fact determined by their intersection with Z .

Theorem 6.10. *Every non-trivial two-sided ideal J of A_∞ which is stable under the conjugation and bracket operations is of the form $J = A_\infty g$, where g is a fixed Hermitian polynomial in z alone. Moreover, every subset of A_∞ of this form is a non-trivial two-sided ideal which is stable under these operations.*

Proof. We shall prove the second statement first. Let g be any Hermitian

polynomial in z alone, and put $J = A_\infty g$. Then J is a left ideal in A_∞ ; and since z commutes with every element in A_∞ , so does $g(z)$, and hence J is also a right ideal. If $fg \in J$, then $(fg)^* = g^* f^* = f^* g \in J$ and J is stable under conjugation. If $fg \in J$, then $\{fg, h\} = \{f, h\}g + f\{g, h\} = \{f, h\} \in J$ (since $\{g, h\} = 0$) and J is stable under the bracket operation.

Now let J be any two-sided ideal of A_∞ which is stable under these operations. Then $J \cap Z$ is an ideal in Z . But since Z is just the algebra of all polynomials in one generator, we know that every ideal of Z is principal, i.e., every ideal is of the form Zg , for some $g \in Z$. If J is stable under conjugation, then so is $J \cap Z$, and it follows that g may be chosen Hermitian. Lemma 6.9 implies that $g \neq 0$ if J is non-trivial.

Put $J' = A_\infty g$. We have already shown that J' is a two-sided stable ideal in A_∞ , and it is obviously contained in J . It remains to show that $J' = J$. Suppose otherwise, and choose an element h , in normal form and of minimal total degree, which lies in J but not in J' . Then $\{p_i, h\} = \partial h / \partial q_i$ is also of normal form and has a total degree less than that of h . Since $\{p_i, h\}$ lies in J , it must also lie in J' . This means that $\{p_i, h\}$ may be expressed in the form $f_i g$, with f_i in normal form. Similarly $\{q_i, h\}$ may be expressed in the form $-g_i g$, with g_i in normal form.

Now we observe that

$$\frac{\partial f_i}{\partial p_j} g = \frac{\partial(f_i g)}{\partial p_j} = \frac{\partial^2 h}{\partial p_j \partial p_i} = \frac{\partial^2 h}{\partial p_i \partial p_j} = \frac{\partial(f_j g)}{\partial p_i} = \frac{\partial f_j}{\partial p_i} g.$$

We know that if $g = 0$, then $J = \{0\} = J'$, which is contrary to our supposition. If $g \neq 0$, then we must have $\partial f_i / \partial p_j = \partial f_j / \partial p_i$. In exactly the same way, we must also have $\partial g_i / \partial q_j = \partial g_j / \partial q_i$, and $\partial f_i / \partial q_j = \partial g_j / \partial p_i$. If all the canonical variables were commutative, these relations would imply that there exists a single polynomial f such that $\partial f / \partial p_i = g_i$, and $\partial f / \partial q_j = f_j$. This same polynomial f , however, will serve the same purpose if the canonical variables are not all commutative, provided only that we write everything in normal form. Thus we have obtained a polynomial f , in normal form, such that $\{p_i, fg\} = \partial(fg) / \partial q_i = f_i g$, and $\{q_i, fg\} = -\partial(fg) / \partial p_i = -g_i g$.

Now consider the difference $h - fg$. We see at once that $\{p_i, h - fg\} = \partial h / \partial p_i - \partial(fg) / \partial p_i = 0$, and similarly, that $\{q_i, h - fg\} = 0$. According to Lemmas 6.7 and 6.9 these relations imply that $h - fg$ lies in $J \cap Z$, and so may be written in the form $h - fg = g'g$, with $g' \in Z$. Thus $h = (f + g')g$, and so $h \in J'$. We have supposed that h is not in J' , and so have arrived at a contradiction.

Thus every stable two-sided ideal J in A_∞ is of the form $A_\infty g$, for some Hermitian polynomial g in the center Z of A_∞ . By rephrasing this result in terms

of the original algebra A which forms the basis of our formulation of mechanics, we have our final result:

Corollary 6.11. *Let A be any algebra of polynomials in the canonical variables which satisfies the requirements of our formulation of mechanics. Then the canonical variables of A necessarily satisfy the relations (6.3)–(6.5), together with one other relation of the form*

$$(6.7) \quad g(z) = 0$$

where g is a Hermitian polynomial in z alone. All other relations among the canonical variables which are valid in A are algebraic consequences of these.

It remains for us to show that these conditions (6.3)–(6.5) and (6.7) are also sufficient.

Theorem 6.12. *The conditions (6.3)–(6.5) and (6.7) are also sufficient for the algebra A to satisfy the requirements of our formulation of mechanics.*

Proof. It follows from Theorem 6.10 that it suffices for us to show that the algebra A_∞ admits both a conjugation and a bracket operation. For the conjugation operation, we observe that A_∞ itself is a homomorphic image of the free algebra of all polynomials in the p_i , q_i , and z , subject to no relations at all. This algebra obviously admits a conjugation relation obtained by reversing the order of multiplication and conjugating coefficients. The kernel of the homomorphism of this algebra onto A_∞ is the two-sided ideal generated by the relations (6.3)–(6.5). It is easy to verify that this ideal is stable under this conjugation, and hence that the homomorphism defines a conjugation on A_∞ .

For the bracket operation, we observe that it must be related to the commutator by (6.6). Hence it suffices to show that, for all f, g in A_∞ , the relations (6.3)–(6.5) imply that $[f, g] = izh$ for some h . The proof of this fact proceeds by induction on degree, exactly as the proof of (6.6). We now define $\{f, g\} = h$. It is easy to verify that this definition satisfies all the requirements of section 3.

The results of this section tell us that it suffices for most purposes to consider the algebra A_∞ as our moment algebra, since every other moment algebra is a homomorphic image thereof. For this reason we may restrict our attention hereafter to the algebra A_∞ without any loss of generality.

7. Representations of the Moment Algebra

In the last section we determined the structure of those algebras which are admissible as moment algebras in our formulation of mechanics. We must now show that every such algebra does in fact arise from a mechanical system. In this section we shall construct a standard representation for each admissible algebra as an algebra of unbounded operators with a common invariant dense domain acting on a Hilbert space, in such a way that the conjugation and bracket operations, as

well as the positive linear functionals, of the algebra are all realized in terms of the structure of the Hilbert space. In the cases of classical and quantum mechanics, this representation reduces to the usual familiar form. In all other cases the representation can be described as a suitable combination of the representations of classical and quantum mechanics. In this sense our formulation of mechanics characterizes the two known systems of mechanics: Under our assumptions, no essentially different system is possible.

We proceed now to the construction of a standard representation for admissible algebras.

We first construct the Hilbert space $\mathcal{L}_2(E_{2n})$ of all complex-valued measurable functions defined on E_{2n} which are square-integrable with respect to the usual Lebesgue measure. We shall denote this Hilbert space by \mathcal{H}_0 . At the same time we construct the dense subspace of \mathcal{H}_0 consisting of all the Hermite functions, and denote it by \mathcal{D}_0 .

On \mathcal{D}_0 we now construct the family of unbounded operators $p_i(0)$, defined by multiplication by the i th coordinate, and $q_i(0)$, defined by multiplication by the $(n+i)$ th coordinate, with $i = 1, \dots, n$. Then \mathcal{D}_0 is a common dense invariant domain for these operators, and on \mathcal{D}_0 they are all essentially self-adjoint, and they all commute. Hence we may construct on \mathcal{D}_0 the commutative algebra A_0 of unbounded operators consisting of all polynomials with complex coefficients in the operators $p_i(0)$ and $q_i(0)$. Each element of A_0 is defined by multiplication by the corresponding polynomial in the $2n$ coordinates.

We introduce into A_0 the conjugation and bracket operations as follows: The conjugate f^* of any polynomial A_0 is defined by multiplication by the complex conjugate polynomial in the coordinates. The bracket $\{f, g\}$ of any two polynomials f and g in A_0 is defined by multiplication by the derived polynomial $\{f, g\}_0$ in the coordinates, where $\{, \}_0$ denotes the classical Poisson bracket.

It is easy to verify that with these definitions A_0 becomes an algebra which satisfies all the requirements of Corollary 6.6, with $g(z) = z$. In view of the discussion given in section I, it is natural to identify this algebra with the structure of classical mechanics.

We next construct the Hilbert space $\mathcal{L}_2(E_n)$ of all complex-valued functions defined on E_n which are square-integrable with respect to the usual Lebesgue measure. We form one copy of this Hilbert space for each nonzero real number λ , and denote it by \mathcal{H}_λ . At the same time we construct the dense subspace of \mathcal{H}_λ consisting of all the Hermite functions, and denote it by \mathcal{D}_λ .

On \mathcal{D}_λ we now construct the family of unbounded operators $p_i(\lambda)$, defined if $\lambda < 0$ by $-\sqrt{\lambda}$ times differentiation by the i th coordinate, and defined if $\lambda > 0$

by $\sqrt{\lambda}$ times multiplication by the i th coordinate. We also construct the family $q_i(\lambda)$, defined by the relation $q_i(\lambda) = p_i(-\lambda)$. Then \mathcal{D}_λ is a common invariant dense domain for the $p_i(\lambda)$ and $q_i(\lambda)$, and on \mathcal{D}_λ they are all essentially self-adjoint and satisfy the relation $[p_i(\lambda), q_i(\lambda)] = i\lambda 1$. Hence we may also construct on \mathcal{D}_λ the algebra of unbounded operators consisting of all polynomials with complex coefficients in the $p_i(\lambda)$ and $q_i(\lambda)$.

We now introduce into A_λ the conjugation and bracket operations as follows: The conjugation f^* of any polynomial f in A_λ is the polynomial obtained from f by reversing the order of the operators and conjugating the coefficients. The bracket $\{f, g\}$ of any two polynomials f and g in A_λ is the polynomial obtained from f and g by forming the combination $[f, g]/i\lambda$.

It is easy to verify that with these definitions A_λ becomes an algebra which satisfies all of the requirements of Corollary 6.6, with $g(z) = z - \lambda$. In view of the discussion given in section 2, it is natural to identify this algebra with the structure of quantum mechanics, with the parameter $2\pi\lambda$ playing the role of Plank's constant [19].

Now let m be any Borel measure on E_1 . We construct the Hilbert space \mathcal{H}_m , consisting of all measurable functions v defined on E_1 with values in the \mathcal{H}_λ , such that $v(\lambda)$ lies in \mathcal{H}_λ , and such that $\int \|v(\lambda)\|^2 d\mu(\lambda)$ is finite. At the same time we construct the dense subspace \mathcal{D}_m , consisting of those functions v in \mathcal{H}_m such that $v(\lambda)$ lies in \mathcal{D}_λ .

On \mathcal{D}_m we construct a family of unbounded operators p_i , defined by the relation

$$(7.1) \quad (p_i v)(\lambda) = p_i(\lambda)v(\lambda),$$

and $q_i(\lambda)$, defined by the relation

$$(7.2) \quad (q_i v)(\lambda) = q_i(\lambda)v(\lambda).$$

It is easy to verify that \mathcal{D}_m is an invariant dense domain for these operators, and that on \mathcal{D}_m each is essentially self-adjoint, and together they satisfy the relations

$$(7.3) \quad \begin{aligned} [p_i, p_j] &= [q_i, q_j] = 0, \\ [p_i, q_j] &= iz\delta_{ij} \end{aligned}$$

where z is the operator defined by

$$(7.4) \quad (zv)(\lambda) = \lambda v(\lambda).$$

It follows that we may construct on \mathcal{D}_m the algebra A_m , consisting of all polynomials with complex coefficients in the p_i and q_i . Every polynomial f in A_m operates on elements in \mathcal{H}_m according to the formula

$$(7.5) \quad (fv)(\lambda) = f(\lambda)v(\lambda),$$

where $f(\lambda)$ is the corresponding polynomial in A_λ . We introduce into A_m the conjugation and bracket operations in the now obvious way:

$$(7.6) \quad (f^* v)(\lambda) = f^*(\lambda) v(\lambda),$$

$$(7.7) \quad (\{f, g\} v)(\lambda) = \{f(\lambda), g(\lambda)\}_\lambda v(\lambda)$$

where $\{, \}_\lambda$ denotes the bracket operation in A_λ .

The properties of A_m are summarized in the following theorem:

Theorem 7.1. *The algebra A_m satisfies all the requirements of Corollary 6.9 with $g(z)$ taken as the Hermitian polynomial of least degree such that*

$$(7.8) \quad \int_{E_1} g^2(\lambda) dm(\lambda) = 0$$

if one exists, and $g(z) \equiv 0$ otherwise.

Proof. The fact that the canonical variables p_i and q_i in A_m satisfy the relations (6.3)–(6.6) follows immediately from the construction. For the last statement it suffices to observe that every polynomial $g(z)$ in Z acts on \mathcal{H}_m according to the formula

$$(7.9) \quad (gv)(\lambda) = g(\lambda) v(\lambda).$$

This follows from (7.7). Now $g(z) = 0$ in A_m if and only if $gv = 0$ for every v in \mathcal{H}_m . This holds if and only if

$$(7.10) \quad \begin{aligned} \|gv\|^2 &= \int_{E_1} \|(gv)\lambda\|^2 dm(\lambda) \\ &= \int_{E_1} g^2(\lambda) \|v(\lambda)\|^2 dm(\lambda) = 0 \end{aligned}$$

for every v in \mathcal{H}_m . The usual argument now shows that (7.10) is equivalent with (7.8). We conclude that any polynomial $g(z)$ in z alone vanishes in A_m if and only if it satisfies (7.8). In particular any Hermitian polynomial of least degree in z alone which vanishes in A_m must satisfy (7.8), and all others are multiples thereof. If no polynomial in z alone satisfies (7.8), then no such polynomial can vanish in A_m , and the relation (6.6) holds only for the polynomial $g(z) \equiv 0$.

We now see that the precise form of the relation (6.6) which holds in A_m depends on the form of the measure m . If m is not purely a point measure, or if m is purely a point measure supported on an infinite number of distinct points, then clearly (7.8) can hold for no polynomial other than $g(z) \equiv 0$. Hence every such measure gives us a representation of A_∞ , and incidentally provides the verification required in the last section that A_∞ is actually an admissible algebra.

If m is a point measure supported on a finite number of distinct points, say $\lambda_1, \dots, \lambda_k$, then clearly (7.8) holds for every polynomial which vanishes at each

of the points λ_i . The polynomial of least degree with this property is $g(\lambda) = \prod_{i=1}^k (\lambda - \lambda_i)$. Hence every such measure gives us a representation of an admissible algebra, with the relation (6.6) given by

$$(7.11) \quad g(z) = \prod_{i=1}^k (z - \lambda_i) = 0.$$

In particular, if m is support at just one point λ_1 , then A_m reduces to A_{λ_1} , and the representation reduces to the one previously constructed for A_{λ} .

In this way we have found representations for all algebras A satisfying the relations (6.3)–(6.6) as well as for A_{∞} . For A_{∞} , we may choose m to be any measure such that (7.8) holds only for the polynomial $g(z) \equiv 0$; and for any other A we may choose m to be any measure such that (7.8) holds for all polynomials $g(z)$ which vanish in A .

If the polynomial $g(z)$ which defines the relation (6.6) in A is of the special form (7.11), then m may be chosen so that the representations of A obtained in this way are *faithful*, in the sense that every polynomial which is nonzero in A is represented by an operator which is nonzero on \mathcal{H}_m . This follows from the fact that every representation is a homomorphism of A . If this homomorphism annihilates any polynomial in A , then by the results of the last section it must also annihilate a polynomial in z alone in A . Hence (7.8) must hold for this polynomial, and so this polynomial must vanish on the support of m . If m is chosen to have the largest support consistent with (7.8), then it follows that this polynomial must vanish wherever $g(z)$ vanishes, and hence must be a multiple of $g(z)$. This means that it vanishes in A , and our conclusion follows.

If the polynomial $g(z)$ is not of the form (7.11), however, then the representations obtained in this way cannot be faithful. This follows from the fact that (7.8) must hold for $g(z)$. If $g^2(\lambda)$ never vanishes, then the only measure m consistent with (7.8) is the zero measure, $m \equiv 0$, and the representation maps every element of A into zero. If $g^2(\lambda)$ does vanish for some λ , then every measure consistent with (7.8) is supported on the zeros $\lambda_1, \dots, \lambda_k$ of $g^2(\lambda)$, and hence the polynomial $\prod_{i=1}^k (z - \lambda_k)$ also satisfies (7.8). Since $g(z)$ is not of this form, this polynomial must be a proper factor of $g(z)$ of degree less than that of $g(z)$. It therefore vanishes on \mathcal{H}_m , but not in A .

This situation raises the question of whether there exist any faithful representations of admissible algebras A satisfying the relations (6.3)–(6.6) but with $g(z)$ not of the form (7.11). Our next result shows that the answer is negative.

Theorem 7.2. *Let A be any algebra of polynomials in the canonical variables which satisfies the relations (6.3)–(6.7) of Corollary 6.9. Then A admits a faithful representation as an algebra of operators on a Hilbert space if and only if the polynomial $g(z)$ has the form (7.11).*

Proof. In any case $g(z)$ admits a factorization of the form $g(z) = \beta \prod_{i=1}^k (z - \alpha_i)$ where the α_i are complex scalars, not necessarily distinct.

Suppose first that for some α_i we have $\alpha_i = \lambda_i + i\mu_i$, with λ_i and μ_i real and $\mu_i \neq 0$. Then, since $g(z)$ is Hermitian, we must have $\alpha_j = \lambda_i - i\mu_i$ for some α_j . Then $g(z)$ can be written in the form

$$(7.12) \quad g(z) = ((z - \lambda_i)^2 + \mu_i^2)g'(z),$$

where $g'(z)$ a polynomial of degree less than that of $g(z)$. Hence

$$(7.13) \quad g^*(z)g(z) = ((z - \lambda_i)^4 + 2(z - \lambda_i)^2\mu_i^2 + \mu_i^4)g'^*(z)g'(z).$$

Now the right-hand side of (7.13) is a sum of polynomials of the form f^*f . Any positive linear functional ϕ on A must vanish on g^*g , and so must vanish on this sum, and therefore must vanish on each term separately. In particular, it must vanish on $\mu_i^4 g'^*(z)g'(z)$, and hence on $g'^*(z)g'(z)$. If A admits a full set of states, then we must have $g'(z) = 0$ in A . But this relation cannot be an algebraic consequence of (6.6), since the degree of $g'(z)$ is less than that of $g(z)$.

Thus we see that the α_i must all be real. Suppose now that $\alpha_i = \alpha_j$. Then $g(z)$ can be written in the form

$$(7.14) \quad g(z) = (z - \alpha_i)^2 g'(z)$$

where again $g'(z)$ is a polynomial of degree less than that of $g(z)$. Any positive linear functional ϕ on A must vanish on g^*g , and hence must vanish on $(z - \alpha_i)^2 g'^*g'(z)$, since, by Lemma 5.4,

$$(7.15) \quad \begin{aligned} \phi((z - \alpha_i)^2 g'^*g') &\leq \phi((z - \alpha_i)^4 g'^*g')\phi(g'^*g') \\ &= \phi(g^*g)\phi(g'^*g'). \end{aligned}$$

Again, if A admits a faithful representation, then we must have that $(z - \alpha_i)g'(z) = 0$. But again, this relation cannot be an algebraic consequence of (6.6), since the degree of $(z - \alpha_i)g'(z)$ is less than that of $g(z)$. The proof is complete.

We shall now examine the problem of determining for each admissible algebra A the class of all representations of A as an algebra of operators on a Hilbert space.

We first observe that any representation of A is also a representation of A_∞ , since every admissible algebra A is a homomorphic image of A_∞ . Conversely, every representation of A_∞ is also representation of A provided only that it annihilates the polynomial $g(z)$ defining A . Hence it suffices to examine the representations of A_∞ .

We next observe that any attempt to classify the representations of A_∞ must hinge on a definition of equivalence which determines when and in what sense two representations of A_∞ are equivalent. For this purpose we shall adopt the

following definitions:

Definition 7.3. Let ρ_1 and ρ_2 be two representations of A_∞ as algebras of operators acting on the same Hilbert space \mathcal{H} . Then ρ_2 is an (proper) *extension* of ρ_1 if the common dense domain of definition \mathcal{D}_2 of $\rho_2(A_\infty)$ (properly) contains the common dense domain of definition \mathcal{D}_1 of $\rho_1(A_\infty)$, and if $\rho_2(A_\infty)$, when restricted to \mathcal{D}_1 , equals $\rho_1(A_\infty)$. ρ_2 is a *maximal* representation if it admits no proper extension.

Lemma 7.4. *Every representation ρ of A_∞ as an algebra of operators acting on a Hilbert space \mathcal{H} admits a maximal extension.*

Proof. This result follows from a straightforward application of Zorn's lemma.

Hence it suffices to deal with maximal representations. For these we adopt the following definition of equivalence:

Definition 7.5. Let ρ_1 and ρ_2 be two maximal representations of A_∞ as algebras of operators acting on possibly different Hilbert spaces \mathcal{H}_1 and \mathcal{H}_2 . Then ρ_1 and ρ_2 are equivalent if there exists a unitary operator W mapping \mathcal{H}_1 onto \mathcal{H}_2 such that W maps the common dense domain of definition \mathcal{D}_1 of $\rho_1(A_\infty)$ onto the common dense domain of definition \mathcal{D}_2 of $\rho_2(A_\infty)$, and for every element f of A_∞ we have $W\rho_1(f) = \rho_2(f)W$.

We recall here that every representation ρ of A_∞ as an algebra of operators acting on a Hilbert space \mathcal{H} gives rise to a set of positive linear functionals ϕ on A_∞ via the relation $\phi(f) = (\rho(f)v, v)$, one for each vector v in the common dense domain of definition \mathcal{D} of $\rho(A_\infty)$. It follows from the Gelfand construction (Theorem 5.6) that every positive linear functional on A_∞ can be realized in this way. It is easy to verify that if two maximal representations of A_∞ are equivalent in the sense of Definition 7.5, then they give rise to precisely the same positive linear functionals of A_∞ .

We now turn to the standard representations of A_∞ , of the type constructed at the beginning of this section. We shall need to know whether any pair of standard representations are equivalent in the sense of Definition 7.5. To answer this question we must first establish an important property of these standard representations.

Lemma 7.6. *Let ρ be any standard representation of A_∞ , and let h be the polynomial*

$$(7.16) \quad h(p_i, q_i, z) = \left[\sum_{i=1}^n p_i^2 + q_i^2 \right] + z^2$$

in A_∞ . Then $\rho(h)$ is essentially self-adjoint.

Proof. Let m be the measure associated with ρ . Then if m is the point measure supported at the origin, then ρ is the standard representation of A_∞ ,

with $z = 0$, and $\rho(h)$ is the operation of multiplication by the sum of the squares of the coordinates on the domain \mathfrak{D}_0 , consisting of all Hermite functions in $L_2(E_{2n})$. It is well-known that this operator is essentially self-adjoint on \mathfrak{D}_0 .

If m is the point measure supported at $\lambda > 0$, then ρ is the standard representation of A_λ , with $z = \lambda 1$, and $\rho(h)$ is the operator $\lambda(\sum_{i=1}^n -\partial^2/\partial \xi_i^2 + \xi_i^2) + \lambda^2$ acting on the domain \mathfrak{D}_λ consisting of all Hermite functions in $L_2(E_n)$. It is well-known that this operator is also essentially self-adjoint.

Now if m is arbitrary, then ρ_m is a direct integral of the ρ_λ with respect to m . In particular, $\rho_m(h) = \int_{E_1} \rho_\lambda(h) dm(\lambda)$. It is known from the theory of direct integrals that a direct integral of essentially self-adjoint operators is again essentially self-adjoint.

Lemma 7.7. *Every representation ρ of A_∞ satisfying the conclusion of Lemma 7.6 determines a unique representation of the associated Heisenberg group G , whose infinitesimal form is a maximal extension of ρ .*

Proof. This result is an application of a recent result of Nelson [15], according to which the conclusion of Lemma 7.6 implies the conclusion of Lemma 7.7. We shall not attempt a proof here, but shall rather refer to his paper [15] for details.

Corollary 7.8. *Every representation ρ of A_∞ satisfying the conclusion of Lemma 7.6 admits precisely one maximal extension $\bar{\rho}$. Two such representations have equivalent maximal extensions if and only if they induce equivalent unitary representations of G .*

In this way we have reduced the problem of classifying the representations of A_∞ satisfying the conditions of Lemma 7.6 to the problem of classifying the unitary representations of G . Fortunately this problem has been completely solved. Restating the solution in terms of A_∞ , we have our principal result:

Theorem 7.9. *Every maximal representation ρ of A_∞ as an algebra of operators on a Hilbert space \mathcal{H} , such that $\rho(h)$ is essentially self-adjoint on \mathcal{H} , is equivalent to a direct sum of (maximal extensions of) standard representations of A_∞ , and nonstandard representations of A_0 . The maximal extensions of two standard representations of A_∞ are equivalent if and only if the associated measures have the same null sets. The maximal extensions of the nonstandard representations of A_0 are given in Corollary 7.10.*

Proof. For a proof of this result for the group G , we refer to [11]. It follows from Lemmas 7.6–7.8 that the result for G implies the result for A_∞ .

Thus we see that every maximal representation of A_∞ is either (the maximal extension of) a standard representation of A_∞ , or (the maximal extension of) a nonstandard representation of A_0 , or a direct sum of such representations.

This leaves us with just two points to clear up. First, we must determine the

nonstandard representations ρ of A_0 for which $\rho(h)$ is essentially self-adjoint.

The standard representation of A_0 was obtained by forming $\mathcal{L}_2(E_{2n})$ with respect to Lebesgue measure, and representing every polynomial in A_0 by the operation of multiplication by that polynomial in the coordinates of E_{2n} . This suggests that other representations of A_0 might be obtained in the same way. We proceed as follows:

We let μ be any Borel measure on E_{2n} , all of whose moments are finite. We form $\mathcal{H} = \mathcal{L}_2(E_{2n})$ with respect to μ . We let \mathcal{D} be the dense subspace of \mathcal{H} consisting of all polynomials in the coordinates of E_{2n} . Now we represent A_0 on \mathcal{D} , by defining $\rho(f)$ as that operator on \mathcal{D} obtained by multiplying every polynomial in \mathcal{D} by the polynomial f in the coordinates of E_{2n} . It is easy to verify that this prescription does define a representation ρ of A_0 on \mathcal{H} , such that $\rho(h)$ is essentially self-adjoint on \mathcal{D} . Our last result says that these representations suffice.

Corollary 7.10. Every maximal representation ρ of A_0 for which $\rho(h)$ is essentially self-adjoint is equivalent to the maximal extension of a representation of the form described above, or a direct sum of such representations. The maximal extensions of two such representations are equivalent if and only if the associated measures have the same null sets. The maximal extension of such a representation is equivalent to the maximal extension of the standard representation of A_0 if and only if the associated measure has the same null sets as Lebesgue measure.

Proof. As before, the condition that $\rho(h)$ be essentially self-adjoint tells us the ρ induces a unitary representation of the Heisenberg group G . Moreover, since $\rho(A_0)$ is a commutative algebra, the image of G must be a commutative group. It is known that the commutative representations of G are precisely those of the form described in the conclusion [13]. Since ρ is the infinitesimal form of this unitary representation of G , the same conclusion also holds for ρ .

Finally, we turn to the problem of classifying those representations ρ of A_∞ for which $\rho(h)$ is not essentially self-adjoint. Examples of such representations may be found in the paper of Nelson [15], who showed that they do not always determine unitary representations of the Heisenberg group G . Moreover, those that do so may not do so uniquely. The detailed classification of such representations therefore remains an open problem.

Even so, one might hope that the positive linear functionals on A_∞ all determine positive definite functions on G and hence could be classified in terms of the unitary representations of G . But it is possible to construct positive linear functions on A_0 , and hence on A_∞ , which do not determine any positive definite function on G [26, pp. 232–236]. We have succeeded in showing in [29], however, that every positive linear functional on A_∞ which is positive on a cone somewhat

larger than P (i.e., which is "strictly positive" in the sense of Definition 5.3) determines a positive definite function on G , and hence can be obtained from either a standard representation of A_∞ or a non-standard representation of A_0 . We do not yet know, however, how to specify this larger cone solely in terms of the data in A_∞ .

In any case, we obtain directly from Theorem 5.3 and Theorem 7.9 the following Corollary.

Corollary 7.11. *Every positive linear functional ω on the moment algebra A which admits an extension to a positive definite function on the Heisenberg group G is either of the form*

$$(7.17) \quad \omega(f) = (\rho f v, v),$$

(where ρ is a standard representation of A or a non-standard representation of A_0), or is a countable sum of such functionals.

If a state ω is a countable sum of functionals of the form (7.17), where each ρ is a standard representation, then we shall say that ω is a *standard state* of A . It follows from Corollary 7.10 that the non-standard states of A all arise from singular measures on the phase space of classical mechanics. From our point of view, no singular measure can arise as the result of any practical measuring process, because of the nature of the uncertainties introduced by the measurements. For this reason only the standard states of A can be of any physical interest in our formulation of mechanics.

8. Extensions of the Moment Algebra

In most problems of mechanics it is necessary to introduce functions of the canonical coordinates which are not polynomials, and probability distributions which are not determined by their joint moments. This is particularly true in dynamical problems involving motions of the system which transform polynomial functions into non-polynomial functions and states with finite moments into states with infinite moments. Once the structure of the moment algebra has been determined, it is then an easy matter to construct a larger algebra of all (reasonably defined) functions of the canonical variables, in such a way that the standard states of the moment algebra admit natural extensions as states of the function algebra. In this section we shall provide the details of this extension process.

We begin by introducing the *algebra of exponential polynomials*. Let B_∞ be the set of all polynomials with complex coefficients in the variables $e^{i\xi_0 x_0}$, $e^{i\xi_1 x_1}, \dots, e^{i\xi_{2n} x_{2n}}$, where the ξ_i are real scalars, and the x_i are the generators of A_∞ , with $x_0 = z$. The elements of B_∞ can be added and multiplied according to the usual rules for polynomials, and with these definitions B_∞ becomes a

non-commutative algebra with a $(2n + 1)$ -parameter family of generators.

The generators are subject to the relations

$$(8.1) \quad e^{i\xi_i x_i} e^{i\xi_j x_j} = e^{-\xi_i \xi_j [x_i, x_j]} e^{i\xi_j x_j} e^{i\xi_i x_i},$$

$$(8.2) \quad e^{i(\xi_i + \eta_i) x_i} = e^{i\xi_i x_i} e^{i\eta_i x_i},$$

where the commutator $[x_i, x_j]$ is given by (6.3)–(6.5). In particular, since the x_i span a Lie algebra L , the $e^{i\xi_i x_i}$ span the Lie group G associated with L , and the relations (8.1)–(8.2) are just the exponential forms of the relations (6.3)–(6.5).

B_∞ admits a conjugation operation, $*$, which conjugates coefficients and reverses the order of multiplication. It satisfies the relations (3.1)–(3.3) and thus is completely determined by its action on the generators. For them we have

$$(8.3) \quad (e^{i\xi_i x_i})^* = e^{-i\xi_i x_i}.$$

In terms of the conjugation operation we shall define in B_∞ the space of Hermitian polynomials and the cone of positive polynomials just as we did in A_∞ .

Moreover, we shall also define the dual space B'_∞ of B_∞ , consisting of all linear functionals on B_∞ , as well as the space of Hermitian functionals and the cone of positive functionals in B'_∞ . In this way we obtain for B_∞ the same elements of structure, apart from the bracket operation, that we have developed for A_∞ . The role of the bracket operation is now taken over by the commutation relations (8.1)–(8.2).

The correspondence between A_∞ and B_∞ is best thought of as implemented by exponentiation. The generators x_i of A_∞ span a Lie algebra L , and their exponentials, the $e^{i\xi_i x_i}$, span the associated Lie group G . The theory of Lie groups now tells us that A_∞ may be identified with the universal enveloping algebra of L , while B_∞ may be identified with the algebraic group algebra of G . From these facts all necessary relations between the two may be obtained.

Every positive linear functional on B_∞ gives rise to a representation of B_∞ as an algebra of operators acting on a Hilbert space. The proof of this result is exactly the same as for A_∞ . But because the generators of B_∞ satisfy (8.3), they must be represented by unitary operators on the Hilbert space, and it follows that *the elements of B_∞ are always represented by bounded operators*. This fact distinguishes B_∞ from A_∞ , and is a decided advantage in some applications.

It is known from the general theory of Lie groups that every such representation of B_∞ determines an associated representation of A_∞ [15], in which the kinetic energy operator h introduced in section 7 is self-adjoint; and that conversely, every such representation of A_∞ in which the kinetic energy operator h is self-adjoint is determined in this way by a representation of B_∞ [15]. Hence the

classification of representations of B_∞ is the same as that of the representations of A_∞ in which h is self-adjoint, as given in Theorem 7.9.

It follows that every standard state of A_∞ induces a positive linear functional, which we shall also call a standard state, on B_∞ , via the correspondence between standard states and standard representations. In particular, if ω is a standard state of A_∞ , then the induced state ω on B_∞ satisfies

$$(8.4) \quad \omega(e^{i\xi_i x_i}) = \sum_{n=0}^{\infty} i^n \frac{\omega(x_i^n)}{n!}$$

whenever the series converges.

There are positive linear functionals on B_∞ , however, which do not arise in this way from states on A_∞ . This situation is a result of the fact that the representations of B_∞ involve bounded operators, while those of A_∞ involve unbounded operators. Thus the states of A_∞ are determined by vectors lying in the invariant domain \mathfrak{D} of the representation, while the positive functionals of B_∞ may be determined by vectors lying outside the domain \mathfrak{D} , and if so, are not induced by states of A_∞ . For this reason we may regard these positive functionals of B_∞ as states of the system which are not determined by their "joint moments".

We can arrive at this same conclusion by another route. Let us recall that in the case of classical mechanics the states of the system are represented by probability distributions on the phase space E_{2n} . Heretofore we have considered only those probability distributions which are determined by their joint moments. We now want to include in our discussions those probability distributions which are not determined by their joint moments, and in particular, those whose joint moments are not all finite.

It is well-known from the theory of probability that every probability distribution μ , whether its moments are finite or not, is always completely determined by its *characteristic function* [3]

$$(8.5) \quad \hat{\mu}(\xi_1, \dots, \xi_{2n}) = \int_{E_{2n}} e^{i(\xi_1 x_1 + \dots + \xi_{2n} x_{2n})} d\mu(x_1, \dots, x_{2n})$$

or, what is the same thing, by its values on the exponential monomials over E_{2n} . In our present terminology, μ determines, and is completely determined by, a positive linear functional on the algebra B_∞ of exponential polynomials in the x_i . It is easy to see from (8.5) that the joint moments of μ are all finite if and only if $\hat{\mu}$ is infinitely differentiable at the origin, and that in this case the value of μ on the polynomial f is given by

$$(8.6) \quad \int_{E_{2n}} f(x_1, \dots, x_{2n}) d\mu = \lim_{\xi \rightarrow 0} f\left[\frac{1}{i} \frac{\partial}{\partial \xi_1}, \dots, \frac{1}{i} \frac{\partial}{\partial \xi_{2n}}\right] \hat{\mu}(\xi_1, \dots, \xi_{2n}).$$

These observations carry over to the case of quantum mechanics. There are states of the system are represented by vectors in the Hilbert space $\mathcal{L}_2(E_n)$. Heretofore we have considered only those vectors which lie in the domain \mathcal{D} of the operators representing the canonical coordinates. We now want to include in our discussions those vectors which lie outside of \mathcal{D} . It is known from the theory of quantum mechanics [16] that every vector v in $\mathcal{L}_2(E_n)$, whether or not it lies in \mathcal{D} , is completely determined by its *characteristic function*

$$(8.7) \quad \hat{v}(\xi_0, \xi_1, \dots, \xi_{2n}) = (e^{i\xi_0 x_0} e^{i\xi_1 x_1} \dots e^{i\xi_{2n} x_{2n}} v, v)$$

i.e., by the values it assigns to the exponential monomials in the canonical operators. Thus v determines, and is determined by, a positive linear functional on the algebra B_∞ of exponential polynomials in the x_i . Moreover, the "joint moments" of v are all finite if and only if \hat{v} is infinitely differentiable at the origin, and in this case the value that v assigns to a polynomial f in normal form in the x_i is given by

$$(8.8) \quad (f(x_0, \dots, x_{2n})v, v) = \lim_{\xi \rightarrow 0} f\left[\frac{1}{i} \frac{\partial}{\partial \xi_0}, \dots, \frac{1}{i} \frac{\partial}{\partial \xi_{2n}}\right] \hat{v}(\xi_1, \dots, \xi_{2n}).$$

There are other positive linear functions on the algebra B_∞ of exponential polynomials which are not of this form. They have the property that their values on the monomials $e^{i\xi_0 x_0}, \dots, e^{i\xi_{2n} x_{2n}}$ are not continuous functions of the arguments ξ_i . It is known that their structure is quite pathological, and that they do not give rise to any representation of A_∞ . For this reason we shall restrict our attention to positive linear functionals whose values on the generating monomials are *continuous* in the arguments ξ_i .

These observations together suggest that we regard as *extended states* of the system those normalized positive linear functionals on the algebra B_∞ which are continuous as functions of the arguments ξ_i . We shall denote the space of all extended states by Ω^- . Those extended states on B_∞ which determine states on A_∞ we shall continue to call states, and we shall regard the state space Ω as part of the extended state space Ω^- . (See the remark at the end of section 7.)

Every extended state ω may be expressed in terms of its *characteristic function* $\hat{\omega}$, defined by

$$(8.9) \quad \hat{\omega}(\xi_0, \xi_1, \dots, \xi_{2n}) = \omega(e^{i\xi_0 x_0} e^{i\xi_1 x_1} \dots e^{i\xi_{2n} x_{2n}}).$$

Thus $\hat{\omega}$ is a continuous function on the Heisenberg group G , here parametrized by the ξ_i . Since ω is a positive linear functional, $\hat{\omega}$ is a positive definite function on G , in the sense that

$$(8.10) \quad \sum \alpha_i \bar{\alpha}_j \hat{\omega}(g_i g_j^{-1}) \geq 0$$

for all scalars α_i and group elements g_i in G . Thus the theory of positive definite functions on Lie groups applies [13]. In particular, it can be shown that, as in the two special cases, an extended state is a state if and only if its characteristic function is infinitely differentiable at the origin of G , and that, in this case, the value assigned to a polynomial f in normal form is just

$$(8.11) \quad \omega f(x_0, x_1, \dots, x_n) = \lim_{\xi \rightarrow 0} f \left[\frac{1}{i} \frac{\partial}{\partial \xi_0}, \frac{1}{i} \frac{\partial}{\partial \xi_1}, \dots, \frac{1}{i} \frac{\partial}{\partial \xi_{2n}} \right] \hat{\omega}(\xi_0, \dots, \xi_{2n}).$$

Finally, we learn from section 7 that the extended states all arise from vectors in the Hilbert spaces described there just as the states do. Moreover, every vector in those spaces defines an extended state, while it defines a state if and only if it lies in the maximal dense domain of the operator representatives of the polynomials in the canonical variables. We shall say that an extended state is a *standard extended state* if it arises from a standard representation. As we pointed out at the end of section 7 we are primarily interested in the standard representations, and hence in the standard extended states.

In this way we obtain a satisfactory extension of the definition of a state of the system. It is now an easy matter to extend the definition of a function of the canonical variables. We shall restrict our attention to functions which yield an expected value in every extended state of the system, since only these functions can be "measured" in every extended state. We shall further suppose that two functions which yield the same expected values in every extended state are the same, since no measurement can distinguish them. Finally, we shall require that the function be bounded, in the sense that its expected values are bounded. Unbounded functions may be treated in the same way, but are technically more difficult to manage, and the bounded functions will suffice for our purposes.

These considerations suggest that we define an extended function of the canonical variables in terms of its expected values in the extended states. Thus an *extended function* f of the canonical variables assigns to each extended state ω in Ω^- a scalar $f(\omega)$ such that

$$(8.12) \quad f(\alpha\omega + (1-\alpha)\eta) = \alpha f(\omega) + (1-\alpha)f(\eta),$$

$$(8.13) \quad \sup \{ |f(\omega)| : \omega \in \Omega^- \} < \infty.$$

We denote by F_∞ the set of all extended functions of B_∞ . Note that every exponential polynomial g may be regarded as an extended function by putting

$$(8.14) \quad g(\omega) = \omega(g).$$

In this sense B_∞ is included in F_∞ , and so F_∞ is not empty. A_∞ , on the other hand is not included in F_∞ , since the polynomials in A_∞ do not admit expected values in every extended state, and the expected values they do admit are not bounded.

It is obvious that complex linear combinations of extended functions are again extended functions, and so F_∞ is a linear space over the complex scalars. It is not obvious that two extended functions can be multiplied together in any consistent way to obtain a third, but it is also true. Our next result says that F_∞ forms a \mathbb{W}^* algebra [17].

Theorem 8.1. *The space F_∞ of extended functions admits a multiplication and a conjugation operation in such a way that under these operations it becomes a \mathbb{W}^* algebra.*

Proof. Let us define a norm for the elements g of B_∞ via

$$(8.15) \quad \|g\|^2 = \sup\{\omega(g^*g) : \omega \text{ in } \Omega^-\}.$$

It is known that this definition does indeed provide a norm, and that the completion C_∞ of B_∞ in this norm is a C^* algebra. Moreover, the dual space N_∞ of this C^* algebra contains all the extended states of B_∞ , and is spanned by linear combinations of these states. It is also known that the second dual \mathbb{W}_∞ of C_∞ is a \mathbb{W}^* algebra which includes B_∞ in a natural way as a weak- $*$ -dense subalgebra. By definition the elements of \mathbb{W}_∞ are determined by their values on N_∞ , and hence by their values on Ω^- . Thus \mathbb{W}_∞ consists of extended functions, and is included in F_∞ [17].

It remains to show that \mathbb{W}_∞ includes F_∞ . Any extended function f in F_∞ is defined as a bounded convex functional on Ω^- , and since Ω^- spans N_∞ , f is defined as a linear functional on N_∞ . We must show that it is bounded. Let ϕ be any Hermitian functional of norm 1 in N_∞ , and recall that ϕ may be expressed as $\phi = \alpha\omega - 1(1 - \alpha)\eta$, where ω and η are in Ω^- . Then

$$|f(\phi)| \leq \alpha|f(\omega)| + (1 - \alpha)|f(\eta)| \leq \sup\{|f(\omega)| : \omega \in \Omega^-\}$$

which we have supposed finite [17]. Thus f is bounded on N_∞ , and therefore lies in \mathbb{W}_∞ . We conclude that F_∞ may be identified with the \mathbb{W}^* algebra \mathbb{W}_∞ .

The computation of extended functions is facilitated by the following corollary.

Corollary 8.2. *Every extended function f of F_∞ may be approximated by exponential polynomials g in the sense of*

$$(8.16) \quad |f(\omega_i) - g(\omega_i)| < \epsilon$$

for any ϵ , and any finite set $\{\omega_i\}$ of extended states.

Finally, we need to know what happens to the extended functions under representations.

Theorem 8.3. *Let ρ be any representation of B_∞ as an algebra of bounded operators on a Hilbert space. Then ρ may be lifted to a representation of F_∞ as the \mathbb{W}^* algebra of operators generated by $\rho(B_\infty)$.*

Proof. See [17]. Thus we see that in any representation of the exponential polynomials as bounded operators, the extended functions automatically appear as bounded operators which can be weakly approximated by the exponential polynomial operators. In this way we obtain a satisfactory extension of the exponential polynomial functions of the canonical variables.

We have carried out this development only for the algebra A_∞ . There is no difficulty in doing the same for the other possible moment algebras, since they are all homomorphic images of this one. Details will be omitted here.

9. Dynamical Aspects

The efforts of the previous sections have been devoted entirely to the descriptive aspects of mechanics, i.e., to the determination of the possible states of the system. In this section we shall consider briefly the dynamical aspects of mechanics, i.e., the description of the possible motions of these states of the system.

We shall first define a motion of the system as a transformation of the states of the system. We shall require on physical grounds that the transformation be one-to-one and invertible, and that it preserve the convex structure of the state space and the commutation relations of the moment algebra. We shall then show that every such transformation induces either an automorphism, or an anti-automorphism, or a suitable combination of the two, of the algebra of extended functions, and that it can be represented by either a unitary operator, or an anti-unitary operator, or a suitable combination of the two in any standard representation.

It follows that every differentiable one-parameter group of motions of the system can be realized as a differentiable one-parameter group of unitary operators in any standard representation. Moreover, if k is the Hermitian generator of this group, then the equations of motion of the canonical variables can be written simply as

$$(9.1) \quad \frac{dx_j}{dt} = i[k, x_j].$$

Now in the (essentially unique) standard representation of the algebra A_λ , where $\lambda \neq 0$, these equations of motion may be rewritten, using (6.6), as

$$(9.2) \quad \frac{dx_j}{dt} = \{h, x_j\}$$

where h is an (unbounded) extended function of the canonical variables, chosen so that $k = \lambda h$. If $\lambda = 0$, then the bracket operation is no longer determined by (6.6), and need not be preserved by the motion; but if we assume in addition that the motion does preserve the bracket operation, then we obtain the same result: The equations of motion (9.1) may be rewritten as (9.2), where now h is an

extended function of the canonical variables chosen so that

$$k = \frac{1}{i} \sum_j \left[\frac{\partial h}{\partial p_j} \frac{\partial}{\partial q_j} - \frac{\partial h}{\partial q_j} \frac{\partial}{\partial p_j} \right].$$

It follows that every differentiable one-parameter group of canonical motions is necessarily determined by the Hamiltonian equations of motion in the form (9.2).

We shall next observe that the class of motions defined in this manner is in some ways too restrictive. A closer investigation reveals that in both classical and quantum mechanics there are physically interesting motions which do not preserve the state space, but transform some of the states into distributions whose joint moments are not all finite, i.e., into extended states.

Thus we are led to define an extended motion of the system as a transformation of the extended states of the system. Again we require that it be one-to-one and invertible, and that it preserve the convex structure of the extended state space and the commutation relations of the exponential polynomial algebra. Every motion determines uniquely an extended motion, but not all extended motions are determined in this way. We shall then show that every extended motion induces either an automorphism, or an anti-automorphism, or a suitable combination of the two, of the algebra of extended functions, and can be represented by either a unitary operator, or an anti-unitary operator, or a suitable combination of the two in any standard representation. Moreover, continuous one-parameter groups of these extended motions are also determined by the equations of motion (9.2) whenever the right-hand side can be properly defined.

Finally, we shall show that the equations of motion (9.2) are consequences of a suitable generalization of the Hamilton variational principle. In this sense the equations of motion in our formulation of mechanics are always determined by a Hamiltonian function of the canonical variables via the Hamilton variational principle.

We begin with a formal definition of a motion of the system. Intuitively, we expect that any motion of the system will transform each state of the system into another state of the system. Moreover, it will transform different states into different states, and superpositions of states into superpositions of states. It will preserve the expected values of the commutation relations. Finally, it will admit an inverse motion, which will restore the system to its original configuration. These considerations lead us to

Definition 9.1. A *motion* of a mechanical system is a transformation $\pi: \Omega \rightarrow \Omega$ of the space of all states of the system onto itself such that, if ω, η lie in Ω and $0 \leq \alpha \leq 1$, then

$$(9.3) \quad \pi(\omega) = \pi(\eta) \text{ if and only if } \omega = \eta,$$

$$(9.4) \quad \pi(\alpha\omega + (1 - \alpha)\eta) = \alpha\pi(\omega) + (1 - \alpha)\pi(\eta),$$

$$(9.5) \quad \pi(\theta) = \omega \text{ for some } \theta \text{ in } \Omega,$$

$$(9.6) \quad (\pi\omega)(z) = \omega(z).$$

Examples of motions can be constructed as follows: For $\lambda = 0$, let T be any one-to-one measure-preserving transformation of E_{2^n} onto itself, and U_0 the unitary operator induced by T on \mathcal{H}_0 . For $\lambda \neq 0$, let U_λ be a unitary operator on \mathcal{H}_λ , chosen in such a way that the operator-valued function U_λ is a measurable function of λ . Then U_λ determines a unitary operator U_m on every standard representation of A_∞ on \mathcal{H}_m , given by

$$(9.7) \quad U_m v = \int U_\lambda v(\lambda) dm(\lambda)$$

and hence determines a transformation on the state space of A_∞ . It is easy to see that this transformation satisfies all the requirements of our definition of a motion, and that this motion preserves the subset of standard states, associated with the standard representations.

This is not the only possibility, however. Other motions of A_∞ arise as follows. Let J_λ be the anti-unitary operator defined in \mathcal{H}_λ by the formula $J_\lambda v = \bar{v}$, i.e., by ordinary conjugation of the functions in \mathcal{H}_λ . Let $k(\lambda)$ be any measurable function of λ whose only values are 0 and 1. Put $V_\lambda = J_\lambda^{k(\lambda)}$. Then V_λ is a measurable operator-valued function of λ which is the identity operator where $k(\lambda) = 0$ and the conjugation operator where $k(\lambda) = 1$. Then V_λ induces an isometric operator on every standard representation of A_∞ , and hence determines a transformation on the state space of A_∞ . It is easy to see that this transformation is also a motion which preserves the standard states.

We now want to show that every motion of the system which preserves the standard states must be a combination of motions of these two special forms; no other motions are possible. Our first result characterizes motions in terms of extended functions of A_∞ .

Lemma 9.2. *Let π be any motion of the state space Ω . Then π induces a linear isometry π^* on the space F_∞ of extended functions, such that*

$$(9.8) \quad (\pi^*f)(\omega) = f(\pi\omega).$$

Proof. We recall that Ω is a subset of the extended state space Ω^- , which spans the space N_∞ . It is known that every Hermitian functional ϕ of norm 1 in N_∞ has a unique decomposition as a convex difference of positive functionals of norm in Ω^- [17]:

$$(9.9) \quad \phi = \lambda\phi^+ - (1 - \lambda)\phi^-$$

where ϕ^+ and ϕ^- have norm 1, and $0 \leq \lambda \leq 1$. If ϕ^+ and ϕ^- lie in Ω , then we

may define $\pi\phi$ by

$$(9.10) \quad \pi\phi = \lambda\pi\phi^+ - (1 - \lambda)\pi\phi^-.$$

Then π is defined as a linear transformation on a subspace of N_∞ . The norm of $\pi\phi$ is given by

$$(9.11) \quad \|\pi\phi\| \leq \lambda\|\pi\phi^+\| + (1 - \lambda)\|\pi\phi^-\| = 1.$$

Hence π decreases the norm in N_∞ . Since the same is true for π^{-1} , we conclude that π preserves the norm in N_∞ . Since Ω is norm-dense in Ω^- , it spans a norm-dense subspace of N_∞ . It follows that π may be uniquely extended to an isometry of N_∞ .

Let π^* be the adjoint mapping defined on the dual F_∞ of N_∞ . Then π^* is also an isometry and by definition satisfies (9.4).

Our next result says that the isometries of F_∞ must have a special form.

Lemma 9.3. *Let π^* be any isometry of F_∞ , with $\pi^*(1) = 1$. Then there exists an orthogonal projection e in the center of F_∞ , with $\pi^*(e) = e$, such that π^* is an isomorphism of $F_\infty e$ and an anti-isomorphism of $F_\infty(1 - e)$.*

Proof. This result, valid for isometries of any W^* algebra, has been obtained by Kadison in [10].

So far we have made no use of the requirement that $\pi\omega(z) = \omega(z)$. We now show that under this requirement, we must have $\pi^*f = f$ for all functions f of the form $e^{i\xi z}$, and hence for all functions f in the W^* subalgebra of F_∞ generated by the $e^{i\xi z}$.

Lemma 9.4. *Let π be any motion of the system. Then for all states ω in Ω , we have $\pi\omega(z^n) = \omega(z^n)$, $n = 1, 2, 3, \dots$*

Proof. From Theorem 5.1 we know that, for any ω in Ω ,

$$(9.12) \quad \omega(z^n) = \int_{E_1} \lambda^n d\mu(\lambda)$$

for a suitably chosen probability measure μ on E_1 . It follows that, as far as the polynomials in z are concerned, the motion π on Ω induces a motion π on the set of all finite measures on E_1 which satisfies the requirements of Definition 9.1.

Let I denote any half-open interval $\alpha \leq \lambda < \beta$ in E_1 , and let μ be any probability measure on E_1 . We assert that $(\pi\mu)(I) = \mu(I)$.

To see this, let μ' be the restriction of μ to the interval I , and μ'' the restriction of μ to $E - I$. Then $\mu'(I) = \mu(I)$, $\mu''(I) = 0$ and $\mu' + \mu'' = \mu$.

Now $(\pi\mu')(I) = \mu'(I)$. Otherwise, there is a half-open interval J , disjoint from I , with $(\pi\mu')(J) > 0$. If so, let ν' be the restriction of $\pi\mu'$ to J and ν'' be the restriction of $\pi\mu'$ to $E - J$. Then $\nu' + \nu'' = \pi\mu'$, so $\pi^{-1}\nu' + \pi^{-1}\nu'' = \pi^{-1}(\nu' + \nu'') = \pi^{-1}(\pi\mu') = \mu'$. It follows that the support of $\pi^{-1}\nu'$ is contained

in the support of μ' , and hence in I . Thus $\pi^{-1}\nu'$ and ν' have supports in I and J , respectively. This means that $\int \lambda d(\pi^{-1}\nu') \neq \int \lambda d\nu'$, which contradicts our assumption on π . We conclude that $\pi\mu'(I) = \mu'(I) = \mu(I)$.

It now follows that $(\pi\mu)(I) = (\pi\mu')(I) + (\pi\mu'')(I) \geq (\pi\mu')(I) = \mu(I)$. The same argument, applies to π^{-1} , shows that $\mu(I) \geq (\pi\mu)(I)$, and we have $(\pi\mu)(I) = \mu(I)$.

Now we see that if $f(\lambda)$ is any step function made up of half-open intervals, then we must have $(\pi\mu)(f) = \mu(f)$.

If we approximate λ^n uniformly by a step function $f(\lambda)$ so that $|\lambda^n - f(\lambda)| < \delta$, then we see that $|(\pi\mu)(\lambda^n) - \mu(\lambda^n)| \leq |(\pi\mu)(\lambda^n - f)| + |\mu(f - \lambda^n)| < 2\delta$, and it follows that $(\pi\mu)(z^n) = \mu(z^n)$ for all n .

In the same way we can show that $(\pi\mu)(f) = \mu(f)$ for all bounded continuous functions f of λ . It follows that, under the assumptions, $\pi\mu = \mu$ for all measures μ on E_1 . This gives us immediately

Lemma 9.5. *Let π be any motion of the system. Then for all states ω in Ω^- , we have $(\pi\omega)(e^{i\xi z}) = \omega(e^{i\xi z})$.*

Proof. The preceding proof has established this result for all ω in Ω . Since Ω is dense in Ω^- , the rest follows from the uniform continuity of π .

Lemma 9.6. *Let π be any motion of the system, and π^* the associated isometry of the extended functions. Then for all f which are extended functions of z , (i.e., which lie in the W^* subalgebra of F_∞ generated by the exponential polynomials $e^{i\xi z}$ in z) we have $\pi^*(f) = f$.*

Proof. Approximate f by exponential polynomials in z and use Lemma 9.5.

Lemma 9.6 contains the essential consequences of the requirement that $\pi\omega(z) = \omega(z)$, and can be used to sharpen considerably the results of Lemma 9.3. We first observe that the algebra F_∞ of extended functions may be split into two parts, one of which is essentially classical, (i.e., commutative) and the other of which is essentially quantal (i.e., non-commutative).

Lemma 9.7. *The algebra F_∞ of extended functions of A_∞ contains a central orthogonal projection e , such that $F_C = F_\infty e$ is commutative, and $F_Q = F_\infty(1 - e)$ has center generated by the exponential polynomials in z .*

Proof. From Theorems 7.9 and 8.3 we learn that the irreducible representations of F_∞ are all described in terms of the algebras B_λ of all bounded operators on the spaces \mathcal{H}_λ , $\lambda \neq 0$, and the algebra B_0 of all bounded measurable functions on the phase space E_{2n} . Let F_Ω be the largest subalgebra of F_∞ whose irreducible representations are all of the form B_λ , for $\lambda \neq 0$, and let F_C be the largest subalgebra of F_∞ whose only irreducible representations are of the form B_0 . Then the theory of operator algebras tells us that F_∞ is the direct sum of F_C and F_Q [17].

Moreover, there exists a projection e in the center of F_∞ , such that $F_C = Fe$ and $F_Q = F(1 - e)$. Finally, F_C obviously commutative; and F_Q is non-commutative, and has center generated by the exponential polynomials in Z .

We are now ready to state our principal result.

Theorem 9.8. *Let π be a motion of the system which preserves the standard states, and π^* the associated isometry of the space F of extended functions. Then $\pi^*F_C = F_C$ and $\pi^*F_Q = F_Q$. On F_C , π^* is an automorphism induced by a Lebesgue measure-preserving transformation of the phase space E_{2n} . F_Q is the direct sum of two subalgebras, F'_Q and F''_Q , such that on F'_Q , π^* is an automorphism induced by a unitary transformation u^* of F'_Q , and on F''_Q , π^* is the product of such an automorphism and a conjugation of the form described after Definition 9.1.*

Proof. Since π^* commutes with the exponential polynomials in z (Lemma 9.5), we see that π^* takes F_C into F_C and F_Q into F_Q . Since $F_C = \mathcal{L}_\infty(E_{2n})$ is commutative, π^* acting on F_C is an automorphism of $\mathcal{L}_\infty(E_{2n})$, and we know that every such automorphism is induced by a measure-preserving transformation of E_{2n} [28]. Since F_Q has center generated by the exponential polynomials in z , and since every irreducible representation of F_Q is of the type B_λ , with $\lambda \neq 0$, we know that every isometry of F_Q is of the form required by the theorem [10].

Corollary 9.9. *If π^* is an automorphism of F_∞ , then on F_C it is induced by a measure-preserving transformation of E_{2n} , and on F_Q it is induced by a unitary function in F_Q .*

After this rather lengthy analysis of the individual motions of the system, we are now prepared to discuss the one-parameter groups of motions which will determine the development of the system in time.

Definition 9.10. *A one-parameter group of motions of the system is a one-parameter family $\{\pi(t)\}$ of motions such that for all ω in Ω , and all f in A ,*

$$(9.13) \quad \pi(0)\omega = \omega,$$

$$(9.14) \quad \pi(s)\pi(t)\omega = \pi(s+t)\omega,$$

$$(9.15) \quad (\pi(t)\omega)(f) \text{ is a continuous function of } t.$$

If $\pi(t)$ is a one-parameter group of motions of Ω , then clearly $\pi^*(t)$ is a one-parameter group of isometries of F_∞ . Moreover, condition (9.15) implies that these isometries must all be automorphisms, since no anti-automorphism can be obtained from the identity by a continuous deformation. This observation, together with Theorem 9.8, enables us to characterize a one-parameter group of motions as follows.

Theorem 9.11. *Let $\pi(t)$ be a one-parameter group of motions on Ω which preserve the standard states, and $\pi^*(t)$ the associated group of isometries of F_∞ .*

Then on F_C , $\pi^*(t)$ is induced by a one-parameter group of measure-preserving transformations on E_{2n} , and on F_Q , $\pi^*(t)$ is induced by a one-parameter group of unitary functions in F_Q .

The weak continuity of $\pi(t)$ defined by (9.15) forces a weak continuity upon $\pi^*(t)$. The weak continuity of $\pi^*(t)$ in turn implies that $\pi^*(t)$ is weakly differentiable. In particular, we have

Theorem 9.12. *Let f be any extended function in F , and let $f(t) = \pi^*(t)f$. Then there exists an unbounded Hermitian extended function k , defined on a dense subset of the extended states, such that for those extended states ω for which k is defined, $f(t)(\omega)$ is differentiable in t , and*

$$(9.16) \quad \lim_{t \rightarrow 0} \frac{d}{dt} f(t)(\omega) = i[k, f](\omega).$$

Proof. Suppose f lies in F_Q . Then $f(t) = u^*(t)fu(t)$ for some weakly continuous one-parameter group of unitary operators in F_Q . We know that every such group of unitary operators has the form $u(t) = e^{ikt}$, for some unbounded Hermitian operator k , defined at least on a dense subset of Ω^- . For all states ω in this subset, then, $f(t)(\omega) = e^{-ikt}fe^{ikt}(\omega)$ is differentiable in t , and its derivative at the origin is given by (9.16).

Suppose now that f lies in F_C . Then again $f(t) = u^*(t)fu(t)$, where now $u(t)$ is the one-parameter group of unitary operators on $\mathcal{L}_2(E_{2n})$ induced by the group of transformations of the phase space E_{2n} which determine $\pi^*(t)$. Again we know that $u(t) = e^{ikt}$, where k is an unbounded Hermitian operator on $\mathcal{L}_2(E_{2n})$, and again this relation leads to (9.16).

The general case is an obvious combination of these two.

If we assume, in addition, that the original group of motions $\pi(t)$ of the state space Ω is differentiable, in the sense that, for all f in A ,

$$(9.17) \quad \pi(t)\omega(f) \text{ is differentiable in } t$$

then it can be shown that the Hermitian generator k of (9.16) is defined on all states ω in Ω , and (9.16) holds for all states.

We shall regard (9.16) as the *equation of motion* for the extended function f . It is by now obvious that the behavior of f under the group of motions $\pi(t)$ is completely determined by equation (9.16), and in particular, this holds true in every standard representation.

Next we notice that the definition of the bracket operation in each of the irreducible standard representations B_λ of F given in section 7 can be used to define a bracket operation on a suitably chosen dense subalgebra of F , such that the relation (6.6) holds. In F_Q , this bracket operation has the form

$$(9.18) \quad \{f, g\} = \frac{1}{iz} [f, g]$$

for all extended functions f and g which admit division by z , while in F_C it is given by

$$(9.19) \quad \{f, g\} = \sum_j \frac{\partial f}{\partial p_j} \frac{\partial g}{\partial q_j} - \frac{\partial f}{\partial q_j} \frac{\partial g}{\partial p_j}$$

for all differentiable functions in $B_0 = \mathcal{L}_\infty(E_{2n})$.

In terms of this bracket operation, the equation of motion (9.16) can be rewritten in Hamiltonian form. In F_Q , (9.16) becomes

$$(9.20) \quad \frac{df}{dt} = \{h, f\}$$

where h is an unbounded Hermitian extended function chosen so that

$$(9.21) \quad k = zh.$$

In F_C , (9.16) can be put in Hamiltonian form only if the motion preserves the bracket operation. This requirement is automatically satisfied in F_Q because of (9.18), but is not necessarily satisfied in F_C . It is known, however, that if the motion does preserve the bracket operation in F_C , then the equation of motion has the form

$$(9.22) \quad \frac{df}{dt} = \{h, f\}$$

where now h is an unbounded function on E_{2n} chosen so that

$$(9.23) \quad k = \frac{1}{i} \left[\sum_j \frac{\partial h}{\partial p_j} \frac{\partial}{\partial q_j} - \frac{\partial h}{\partial q_j} \frac{\partial}{\partial p_j} \right]$$

and conversely, if the equation of motion has the form (9.22), then the motion necessarily preserves the bracket operation in F_C [25].

If we call *canonical* all motions which preserve the bracket operation in F , then we can state our final result as follows:

Theorem 9.13. *Every differentiable one-parameter group of canonical motions of the state space is completely determined by the Hamiltonian equations of motion*

$$(9.24) \quad \frac{df}{dt} = \{h, f\}$$

where the Hamiltonian h is a suitably chosen (unbounded) Hermitian extended function of the canonical variables.

It must be admitted here that this seemingly fortuitous result is a consequence of a long series of carefully chosen definitions. Nevertheless, it implies that any reasonable motion of the system can always be cast in Hamiltonian form.

In some problems our definition of motion is too restrictive to allow all

motions of physical interest. In the case of classical mechanics, for example, the motion of a single point particle moving along a single axis under a prescribed force is a motion in our sense only if the potential function is sufficiently well-behaved. Otherwise the motion may transform states with all joint moments finite into states with some joint moments no longer finite.

A similar situation prevails in the quantum mechanical analogue of this one-particle system. A motion of this system is a motion in our sense only if the potential operator is sufficiently well-behaved. Otherwise the motion transforms some of the vectors in the common dense domain \mathcal{D} of the canonical variables into vectors which lie outside this domain.

We can include motions of this type within our framework by defining an *extended motion* of the system as a transformation of the space of extended standard states. Then the whole preceding development can be adjusted to cover extended motions as well, with the result that Theorem 9.13 holds whenever the right-hand side of the Hamiltonian equations of motion are properly defined. Thus extended motions behave just like motions, apart from questions of domain.

We shall conclude our investigation of the dynamical aspects of our formulation of mechanics by showing that the equations of motion may also be obtained from a variational principle. For this purpose, we let ρ be any standard representation of A on \mathcal{K} , and h any Hermitian polynomial in A such that $\rho(h)$ is self-adjoint on \mathcal{K} . We define for any f in A the operators $f(t)$ and $\dot{f}(t)$ by means of

$$(9.25) \quad f(t) = \exp(-ikt)\rho(f)\exp(ikt),$$

$$(9.26) \quad \dot{f}(t) = \exp(-ikt)\rho(-i[h, f])\exp(ikt).$$

We then form the action integral

$$(9.27) \quad I = \int_0^t (L(\tau)v, v) d\tau$$

where v is an arbitrary vector chosen from the domain \mathcal{D} in \mathcal{K} , and $L(t)$ is the Lagrange function defined by

$$(9.28) \quad L(\tau) = \sum_{i=1}^n p_i(\tau) \dot{q}_i(\tau) - h(\tau).$$

By a variation of the canonical coordinates, we shall mean simply a translation of the form

$$(9.29) \quad \begin{aligned} p &\rightarrow p + \delta p = p + \delta\alpha 1, \\ q &\rightarrow q + \delta q = q + \delta\beta 1. \end{aligned}$$

We now introduce a variation of this form into each of the canonical coordinates at each time τ in the interval $0 < \tau < t$, and try to estimate its effect on the action integral.

$$(9.30) \quad \delta I = \delta \int_0^t (L(\tau) v, v) d\tau.$$

If the variation vanishes at the end points, then (9.30) may be rewritten as

$$(9.31) \quad \delta I = \int_0^t \delta(L(\tau)v, v) d\tau.$$

Using (9.28), we obtain for the integrand

$$(9.32) \quad \begin{aligned} \delta(L(\tau)v, v) = & \left[\sum_{i=1}^n \delta p_i(\tau) \dot{q}_i(\tau) v, v \right] + \left[\sum_{i=1}^n p_i(\tau) \delta \dot{q}_i(\tau) v, v \right] \\ & - \left[\sum_{i=1}^n \frac{\partial h}{\partial p_i} \delta p_i v, v \right] - \left[\sum_{i=1}^n \frac{\partial h}{\partial q_i} \delta q_i v, v \right]. \end{aligned}$$

Since the variation of each coordinate is made at each time τ , the variation commutes with the time derivative. Hence the integral of the second term in (9.32) may be rewritten as

$$(9.33) \quad \begin{aligned} \int_0^t \sum_{i=1}^n p_i(\tau) \delta \dot{q}_i(\tau) d\tau &= \int_0^t \sum_{i=1}^n p_i(\tau) \frac{d}{d\tau} \delta q_i(\tau) d\tau \\ &= - \int_0^t \sum_{i=1}^n \dot{p}_i(\tau) \delta q_i(\tau) d\tau. \end{aligned}$$

Here we have used the fact that all variations vanish at the end points. Now (9.31) becomes

$$(9.34) \quad \begin{aligned} \delta I = & \int_0^t \left[\sum_{i=1}^n \left[\dot{q}_i - \frac{\partial h}{\partial p_i} \right] \delta p_i v, v \right] \\ & - \left[\sum_{i=1}^n \left[\dot{p}_i + \frac{\partial h}{\partial q_i} \right] \delta q_i v, v \right] d\tau. \end{aligned}$$

Now we observe that the variations δp_i and δq_i are actually scalar multiples of the identity in A , and hence may be taken out of the inner product. Hence

$$(9.35) \quad \begin{aligned} \delta I = & \int_0^t \left[\sum_{i=1}^n \left[\dot{q}_i - \frac{\partial h}{\partial p_i} \right] v, v \right] \delta \alpha_i d\tau \\ & - \int_0^t \left[\sum_{i=1}^n \left[\dot{p}_i + \frac{\partial h}{\partial q_i} \right] v, v \right] \delta \beta_i d\tau. \end{aligned}$$

If we require that the action integral be stable under variations of the form (9.29), in the sense that $\delta I = 0$ for all such variations, then the usual arguments of the calculus of variations will show that we must have

$$(9.36) \quad \left[\left[\dot{q}_i - \frac{\partial h}{\partial p_i} \right] v, v \right] = \left[\left[\dot{p}_i + \frac{\partial h}{\partial q_i} \right] v, v \right] = 0$$

for all vectors v in the domain \mathcal{D} . Thus we must have

$$(9.37) \quad \dot{q}_i = \frac{\partial h}{\partial p_i} ; \dot{p}_i = - \frac{\partial h}{\partial q_i}$$

or, in other form

$$(9.38) \quad \begin{aligned} \dot{q}_i &= \{q_i, h\}, \\ \dot{p}_i &= \{p_i, h\} \end{aligned}$$

as the equations of motion of the system. These equations, of course, imply those of (9.25).

In this way we see that the equations of motion (9.25) are, just as in the classical case, consequences of the variational principle $\delta I = 0$ for the action integral I of (9.27). With a little care, the same result can be established in the same way for a more general class of Hamiltonian operators $h(p, q)$.

It is now easy to show that the equations of motion (9.16) always admit unique solutions of the form (9.25) provided only that the Hamiltonian operator is essentially self-adjoint in the representation of A under consideration. Thus the equations of motion determine the development in time of any mechanical system described by our formulation of mechanics.

The relations here derived for the equations of motion and the development of the system in time are equally valid for any one-parameter group of symmetries of the system.

§10. Statistical Aspects

Since our formulation of mechanics is primarily statistical in origin, it seems natural to suppose that it must incorporate the elements of both classical and quantum statistical mechanics. We shall show here that such is indeed the case. For this purpose it suffices to show that the standard states of the system in our formulation are all determined by an appropriate form of density matrix [16].

We must first introduce into each standard representation of A_∞ a functional which is invariant under all motions.

For $\lambda \neq 0$, let τ_λ denote the functional which assigns to each bounded operator on \mathcal{H}_λ its trace. For $\lambda = 0$, let τ_0 denote the functional which assigns to each bounded function on E_{2n} its Lebesgue integral. Now for any probability measure m on E_1 , let τ_m be the functional defined on the bounded operator on \mathcal{H}_m be the formula

$$(10.1) \quad \tau_m(b) = \int_{E_1} \tau_\lambda(b_\lambda) dm(\lambda).$$

It is clear that τ_m is invariant under any motion; since if $\lambda \neq 0$, $\tau_\lambda(u^*bu) = \tau_\lambda(b)$, for all unitary operators u , and if $\lambda = 0$, τ_0 is invariant under each motion of the phase space E_{2n} .

Now suppose r is a positive bounded operator on \mathcal{H}_m such that $\tau_m(rf)$ is finite for all representatives f of polynomials in A_∞ . (For example, r might be a

projection on a one-dimensional subspace of the invariant domain \mathfrak{D}_m .) Then it is easy to verify that the formula

$$(10.2) \quad \omega(f) = r_m(rf)$$

defines a positive linear functional on A_∞ which is a standard state. If π is any motion of the standard states, then

$$(10.3) \quad (\pi\omega)(f) = r_m(ru^*fu) = r_m(uru^*f)$$

where u is the isometric operator on \mathfrak{H}_m defined by π .

Our next result says that every standard state on A must be of this form.

Theorem. *Let ω be any standard state of A_∞ . Then there exists a measure m on E_1 and a bounded positive operator r on \mathfrak{H}_m such that, for all f in A_∞ ,*

$$(10.4) \quad \omega(f) = r_m(rf).$$

The measure m may be normalized so that

$$(10.5) \quad \omega(z^n) = \int_{E_1} \lambda^n dm(\lambda)$$

and in this case both m and r are essentially unique.

Proof. First, suppose ω has the special form

$$(10.6) \quad \omega(f) = (fv, v)$$

for some vector v in a standard representation. If we choose m so that (10.5) holds, then (10.6) may be written as

$$(10.7) \quad \omega(f) = (fv, v) = \int_{E_1} (f_\lambda v_\lambda, v_\lambda) dm(\lambda)$$

where v_λ lies in \mathfrak{H}_λ , and $\|v_\lambda\| = 1$.

Let r_λ be the orthogonal projection on v_λ in \mathfrak{H}_λ . Then the inner product $(f_\lambda v_\lambda, v_\lambda)$ may be written as $r_\lambda(r_\lambda f_\lambda)$, and (10.7) becomes

$$(10.8) \quad \omega(f) = \int_{E_1} r_\lambda(r_\lambda f_\lambda) dm(\lambda) = r_m(rf)$$

which is the same as (10.4).

We know from Corollary 7.11 that every standard state of A_∞ must be a countable convex combination of standard states of the special form (10.6). Hence, the usual techniques of Hilbert space theory will now show that (10.4) must also hold for all standard states of A_∞ [16]. The uniqueness is a consequence of the uniqueness of (10.7).

Theorem 10.1 says that every standard state on A_∞ is determined by a density operator in a suitably chosen standard representation of A_∞ . Moreover, we see from (10.3) that every motion of this standard state is determined by a motion of the associated density operator of the form $\pi^* : r \rightarrow uru^*$.

Thus we conclude that the usual forms of both classical and quantum statistical mechanics are derivable from our formulation of mechanics, and that a suitable generalization of these two special cases can be established for every mechanical system which admits a description in our terms.

We must point out here that because of the commutation relations (6.3)–(6.5) the statistics of the different pairs (p_i, q_i) of conjugate variables are independent. This means that the different pairs are distinguishable by a suitable measuring process. It follows that the statistics of a many-particle system obtained from the formalism described so far distinguishes different particles, and thus is a form of Maxwell-Boltzman statistics. This statement is true for both the classical case and the quantum case, as well as the general case.

It is not hard, however, to incorporate other forms of statistics into the same framework. If different particles are not distinguishable, then different pairs (p_i, q_i) of conjugate variables will not be statistically independent, and this fact must be reflected in the structure of the moment algebra.

If we require that the statistics remain invariant under the exchange

$$(10.9) \quad (p_i, q_i) \leftrightarrow (p_j, q_j)$$

then we must simply restrict the moment algebra to polynomials which remain invariant under this exchange. The entire development goes through exactly as before, but now every state of the system is invariant under the exchange. This means that the statistics obtained from such a formalism is a form of Bose-Einstein statistics.

Similarly, if we require that the statistics reverse polarity under the exchange (10.9) then we must simply restrict the moment algebra to polynomials which reverse polarity under this exchange. Again the entire development goes through exactly as before, but now every state of the system reverses polarity under the exchange. This means that the statistics obtained from such a formalism is a form of Fermi-Dirac statistics. All these possibilities are available in every mechanical system which admits a description in our terms.

So far we have considered only the mechanics of a system of point particles, whose states we assume to be determined completely by the canonical positions and momenta. If the particles of the system exhibit other degrees of freedom (e.g., spin or charge) so that other coordinates are necessary to determine completely the states of the system, then these other coordinates must also appear in the moment algebra.

If, for example, it is necessary to include the spin coordinates $\sigma_1, \dots, \sigma_n$, then these variables must also appear in the moment algebra, and their relations with the other generators determined. If we assume, for instance, that the spin

variables are statistically independent of the position and momenta variables and of each other, then we know from Theorem 5.3 that they must commute with the position and momenta variables as well as with each other. Furthermore, if only a finite set of values is available for each spin variable σ_i , then each σ_i must also satisfy a polynomial identity of the form $g(\sigma_i) = 0$. If, for example, the spin variable σ_i can assume only the values ± 1 (spin up or spin down) then σ_i must satisfy the polynomial identity $\sigma_i^2 - 1 = 0$. In this way the additional variables required to describe additional degrees of freedom may be incorporated into the framework of our formulation of mechanics.

We emphasize here that, at least in principle, any combination of commutation relations (classical or quantum), statistics (Maxwell-Boltzmann, Bose-Einstein, or Fermi-Dirac), and spin (zero, half-integer or integer) can be built into our framework. There is no difficulty in principle which forbids any combination. In particular, the connection between spin and statistics derived in quantum field theory is not a consequence of our formulation of particle mechanics, nor of any formulation which describes only systems with a fixed finite number of particles.

11. Constraints

Up until now we have assumed that the system whose mechanics we are describing is free of constraints. This assumption appears operationally in the statement that every normalized positive linear functional on the moment algebra A determines a state of the system.

If the system is subject to constraints, then only those functionals on the algebra which satisfy the conditions of constraint can determine states of the system. If, for example, the particles of the system are constrained to lie within the unit cube (i.e., $0 \leq q_i \leq 1$), or on the surface of the unit sphere (i.e., $\sum q_i^2 = 1$), or below a fixed energy level (i.e., $h \leq 1$), then clearly all values assigned to the coordinates by any measuring process must satisfy the conditions of these constraints. It follows that we can account for the presence of constraints in the system by suitably restricting the space of admissible states.

We shall consider here only those constraints which can be given in terms of inequalities on the values assigned to a finite set of prescribed Hermitian functions of the canonical variables. (Note that every equality can always be described by a pair of inequalities!) Such constraints can always be put in the form

$$(11.1) \quad f_i(x_1, \dots, x_{2n}) \geq 0.$$

We shall then take as the space of admissible states those elements ω of Ω which assign to each of the f_i a probability distribution of possible values, via Theorem 5.1, which vanishes on the negative real axis. In terms of moments, this

condition says that if $g(\xi)$ is any polynomial in the single real variable ξ , such that $\xi \geq 0$ implies $g(\xi) \geq 0$, then

$$(11.2) \quad \omega(g(f_i)) \geq 0.$$

If there exist no states in Ω satisfying (11.2), then we say that the constraints are incompatible with the system. Otherwise the constraints are compatible with the system, and the states which satisfy (11.2) are compatible with the constraints.

It is a straightforward matter to verify that the space S of states compatible with the constraints has the properties

$$(1) \quad \omega, \eta \in S \Rightarrow \lambda\omega + (1 - \lambda)\eta \in S, \quad 0 \leq \lambda \leq 1,$$

$$(2) \quad \{\omega_n\} \in S \text{ and } \omega_n \rightarrow \omega \Rightarrow \omega \in S,$$

but not the property

$$(3) \quad \omega \in S \Rightarrow \sigma^*\omega \in S \text{ for all translations } \sigma \text{ of } A.$$

Thus S is closed under strong superpositions and weak limits, but not under translations (see section 4). Properties (1) and (2), taken together, imply that S is a *face* of Ω [17]. The same statement holds true for their extensions: S^- is a (weakly closed) face of Ω^- .

If the moment algebra A is commutative (the classical case), then so are the algebras B of exponential polynomials and F of extended functions. In this case it is known that the set J of extended functions which vanish on the compatible states forms an ideal in F , and the quotient algebra F/J can be realized as the algebra of extended functions defined only on the compatible states [17]. In this way we recover the familiar result that the constraints of the system define a subset of the phase space from (11.1), and the states and extended functions compatible with the constraints are all defined relative to this subset.

If the moment algebra A is not commutative, however, then neither are the algebras B and F , and the preceding argument no longer holds. In this case there exists in F an orthogonal projection e , such that the algebra eFe can be realized as the algebra of extended functions defined only on the compatible states. If the functions f_i defining the constraints lie in the center of F , then so does the orthogonal projection e , and we have an analogue of the commutative case. The only functions which lie in the center of F , however, are extended functions of z alone. In the non-commutative case, then, only constraints involving the commutation relations (i.e., involving z alone) can be realized in terms of an ideal J in the algebra of extended functions.

We note here that the process of specializing A_∞ to A_λ (i.e., of specializing our abstract framework to classical or quantum mechanics) may be viewed as the process of placing a constraint of the form $z = \lambda$ on the abstract system which

defines the ideal in A_∞ generated by $z - \lambda$. If $\lambda = 0$, then further compatible constraints may be adjoined, defining ideals in A_0 , but if $\lambda \neq 0$, then A_λ is simple, and no further compatible constraints can be adjoined which will define ideals in A_λ . For this reason the placing of constraints on the system is a more natural procedure in classical than in quantum mechanics; classical constraints have no quantum mechanical counterparts in terms of the moment algebra, though they always do in terms of the state space.

What about motions? If π is a motion of the system such that the defining functions f_i are constants of the motion (i.e., $\pi^* f_i = f_i$), then it follows that π takes states compatible with the constraints into states again compatible with the constraints, and hence leaves the space of compatible states invariant. In this case we shall say that the motion is compatible with the constraints. Conversely, it can be shown, using the machinery of section 9, that if the motion is compatible with the constraints, then the constraints may be expressed in the form (11.1), where the f_i are constants of the motion.

It follows that, in the presence of constraints on the system, the only motions of the system compatible with these constraints are those defined on the space of compatible states, and for such motions the constraints may be expressed in terms of constants of the motion.

BIBLIOGRAPHY

- [1] E. Artin, *Über die Zerlegung definiter Funktionen in Quadrate*, Abhandlungen aus dem Math. Seminar der Hamburgischen Universität, Bd. 5 (1926), pp. 100–115.
- [2] E. T. Bell, *Men of Mathematics*, Simon and Schuster, New York, 1937.
- [3] S. Bochner, *Vorlesungen über Fouriersche Integrale*, Leipzig, 1932.
- [4] J. Dixmier, *Sur les Représentations Unitaires des Groupes de Lie Nilpotent. I*, Amer. J. Math., vol. 81 (1959), pp. 160–170.
- [5] I. M. Gelfand, *Normierte Ringe*, Recueil Math. (Mat. Sbornik) N.S., vol. 9 (1941), pp. 3–24.
- [6] I. M. Gelfand and M. Neumark, *On the Imbedding of Normed Rings into the Ring of Operators on Hilbert Space*, Recueil Math. (Mat. Sbornik) N.S., vol. 12 (1943), pp. 197–213.
- [7] N. Jacobson, *Lie Algebras*, Interscience, New York-London, 1962.
- [8] P. Jordan, J. von Neumann and E. Wigner, *On an Algebraic Generalization of the Quantum Mechanical Formalism*, Ann. Math. vol. 34 (1934), pp. 29–64.
- [9] T. F. Jordan and E. C. G. Sudarshan, *Lie Group Dynamical Formalism and the Relation between Quantum Mechanics and Classical Mechanics*, Rev. Modern Phys. vol. 23 (1961), pp. 515–524.

- [10] R. V. Kadison, *Isometries of Operator Algebras*, Ann. Math., vol. 54 (1951), pp. 325–338.
- [11] A. A. Kirillov, *Unitary Representations of Lie Groups* (Russian), Uspehi. Math. Nauk., vol. 17 (1962), no. 4 (106), pp. 57–110.
- [12] B. O. Koopman, *Hamiltonian Systems and Transformations in Hilbert Space*, Proc. Nat. Acad. Sci., vol. 17 (1931), pp. 315–318.
- [13] L. Loomis, *An Introduction to Abstract Harmonic Analysis*, van Nostrand, New York, 1953.
- [14] G. Mackey, *Quantum Mechanics and Hilbert Space*, Math. Monthly, vol. 64 (1957), part II, no. 8, pp. 45–57.
- [15] E. Nelson, *Analytic Vectors*, Ann. Math., vol. 70 (1959), pp. 572–615.
- [16] J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton Univ., Princeton, 1955.
- [17] R. T. Prosser, *On the Ideal Structure of Operator Algebras*, Memoirs Amer. Math. Soc., no. 45 (1963), pp. 1–28.
- [18] C. E. Rickart, *General Theory of Banach Algebras*, van Nostrand, New York, 1960.
- [19] L. Schiff, *Quantum Mechanics*, McGraw-Hill, New York-London, 1955.
- [20] I. Segal, *Postulates for General Quantum Mechanics*, Ann. Math., vol. 43 (1947), pp. 930–948.
- [21] J. A. Shohat and J. D. Tarmarkin, *The Problem of Moments*, Amer. Math. Soc., Providence, 1943.
- [22] M. H. Stone, *Linear Transformations in Hilbert Space and Their Applications to Analysis*, Amer. Math. Soc., Providence, 1932.
- [23] A. Wightman, *Quantum Field Theory in Terms of Vacuum Expectation Values*, Phys. Rev., vol. 101 (1956), pp. 860–866.
- [24] H. Weyl, *The Theory of Groups and Quantum Mechanics*, Dover, New York, 1930.

Added in Proof:

- [25] D. G. Currie, T. F. Jordan and E. C. G. Sudarshan, *Relativistic Invariance and Hamiltonian Theories of Interacting Particles*, Rev. Mod. Phys. 35 (1963), pp. 350–375.
- [26] I. M. Gelfand and N. Ya. Vilenkin, *Generalized Functions*, Vol. 4, Academic Press, New York, 1964.
- [27] T. F. Jordan, M. A. Pinsky and E. C. G. Sudarshan, *Dynamical Mappings of Density Operators in Quantum Mechanics*, II. Time Dependent Mappings, J. Math. Phys. 3 (1962), pp. 848–852.
- [28] J. von Neumann, *Einige Satze Ueber Messbare Abbildungen*, Ann. Math. 33 (1932), pp. 574–586.

- [29] R. T. Prosser, *The Moment Problem for Lie Groups*, submitted to Bull Amer. Math. Soc.
- [30] E. C. G. Sudarshan, *Structure of Dynamical Theories*, Lectures in Theoretical Physics, Brandeis Summer Institute of 1961, Benjamin, Inc., New York, 1962; pp. 143-199.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R&D		
<i>(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)</i>		
1. ORIGINATING ACTIVITY (Corporate author) Lincoln Laboratory, M.I.T.	2a. REPORT SECURITY CLASSIFICATION Unclassified	2b. GROUP None
3. REPORT TITLE A New Formulation of Particle Mechanics		
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Journal Article		
5. AUTHOR(S) (Last name, first name, initial) Prosser, Reese T.		
6. REPORT DATE 1966	7a. TOTAL NO. OF PAGES 57	7b. NO. OF REFS 30
8a. CONTRACT OR GRANT NO. AF 19(628)-5167	9a. ORIGINATOR'S REPORT NUMBER(S) JA-2278	
b. PROJECT NO.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) ESD-TR 66-172	
10. AVAILABILITY/LIMITATION NOTICES Distribution of this document is unlimited		
11. SUPPLEMENTARY NOTES Memoirs of The American Mathematical Society, Number 61	12. SPONSORING MILITARY ACTIVITY Air Force Systems Command, USAF	
13. ABSTRACT It is the purpose of this paper to provide an abstract formulation of the concept of a mechanical system containing both the classical and quantum formulations, and to derive its most elementary properties.		
14. KEY WORDS Classical mechanics Quantum mechanics Abstract mechanics		

UNCLASSIFIED

Security Classification