Observables and Fields

Rudolf Haag
II. Institut für Theoretische Physik
der Universität Hamburg
Hamburg, Germany

CONTENTS

Introduction

The following story was reported to me. A few years ago Klaus Hepp gave some lectures in the Brandeis summer school. At some stage he praised the beauty of axiomatic field theory. Next day he found the note on the blackboard:
"Axiom 1: Axiomatic Field Theory is beautiful in an empty sort of way."

Presumably this note expresses also pretty accurately the feelings of the majority of today's audience and indeed there is an element of truth in it. Specifically, after about 18 years of hard efforts, the principal objective of this enterprise has not yet been achieved. This objective was to find out whether an adequate framework for the description of elementary particle physics could be developed within the conceptual structure provided by the principles of quantum physics, special relativity theory and locality (= "Nahwirkungsprinzip"). Of course, from the point of view of the development of basic physical theory, this question has been one of the central themes for the past 35 years and is by no means a monopoly of axiomatic field theorists. What distinguishes various groups is not the question itself, but rather the attitude towards it. There are three major ideologies:

1) The answer to the above question will certainly be no. We need a radical change of our concepts, some brilliant new idea. It is just as futile to approach elementary particle physics with the conceptual structure of 1930 as it was to attack atomic physics within the frame of classical mechanics. Therefore we should look for daring new approaches. Some examples: modification of geometry by assumption of a fundamental length, elimination of concepts which are far removed from experimental possibilities or, on the formalistic level, non associative algebra, indefinite metric in Hilbert spaces, etc.

2) The answer may be yes if we are sufficiently careful. It is worthwhile to develop a framework which incorporates the old principles, formulating them precisely, separating the essential and the peripheral features of traditional Quantum Field Theory, recognizing the numerous mathematical pitfalls. One should then demonstrate that this framework is internally consistent and study whether it leads to any consequences which are in disagreement with experience.

3) The time is not ripe for any assessment of the fundamental principles. The most fruitful task for the theoretician at present is to analyze experiments, looking for regularities and for phenomenological models which describe the essential features.

It is unfortunately in the nature of ideologies that they tend to crystallize. One has to make a determined effort to keep the channels of communication between the different camps open. I think that the organizers of this summer school had this need in mind when they asked me to lecture here and therefore I do not feel apologetic for exposing you to some ideas and problems in axiomatic field theory.

Our first concern will be with the "axioms" themselves, the formulation of the input assumptions. In the course of the years there has been some development both in the direction of simplification by recognizing the essential elements and in the direction of enriching the structure. Let me give a brief sketch of this development.

I. Axiomatic Quantum Field Theory in Various Formulations

A. The Simplest Kind of Field Theory

In the years 1953-56 the motivation was provided by the divergence difficulties of standard Lagrangian field theory models and the wish to see whether the renormalization procedures could be welded into a mathematically well defined scheme independent of a perturbation expansion. For this purpose it seemed adequate to consider the simplest type of field theory, namely that of a single neutral scalar field A

describing a single type of particle (of course with interaction). The generalization to several fields with more complicated transformation properties seemed obvious and straight forward. In laying out the framework one was guided then by the experience gained from the perturbation treatment of a Lagrangian field theory keeping those structural features of the renormalized perturbation solution which could be precisely expressed mathematically. This led to the following principal assumptions:

1) Principles of Quantum Physics. Essentially the mathematical and conceptual structure outlined in the books by Dirac and von Neumann. It may suffice here to say that the mathematics deals with a Hilbert space $\mathfrak{H}$ whose vectors correspond to physical states; observables are represented by self adjoint operators acting in $\mathfrak{H}$ and there are the well known rules for calculating probabilities for the results of observations.

2) Poincaré invariance

The Poincaré group (inhomogeneous Lorentz group) consists of translations in space time and homogeneous Lorentz transformations. A general element is denoted by (a, Λ) where a is a 4-vector of translation, Λ a homogeneous Lorentz transformation and the notation suggests that Λ is applied first, a later. We assume that the Poincaré

group is represented by unitary operators in $\mathcal{H}$

$$(a,\Lambda) \rightarrow U(a,\Lambda).$$

The action of the unitary operator $U(a,\Lambda)$ on a vector in $\mathcal{H}$ shall have the obvious physical interpretation; i.e., the image vector corresponds to the state which is prepared by the same intrinsic apparatus as the original state but shifted in its space-time placement and motion by (a,Λ). We shall write $U(a)$ instead of $U(a,1)$ (pure translation) and $U(\Lambda)$ instead of $U(0,\Lambda)$ (Lorentz transformation).

Actually the situation is slightly more complicated. One does not need (and in general does not have) a true representation of the Poincaré group but rather a representation of its "covering group." This corresponds to the well known replacement of a Lorentz-matrix Λ by a complex 2 x 2 matrix with determinant 1. The correspondence of Λ to such a matrix $\alpha(\Lambda)$ is determined only up to a sign, see [2],[3].

3) The vacuum state and stability

Writing

$$U(a) = e^{-iP_\mu a^\mu} \qquad (1.1)$$

the infinitesimal generators P_μ of the translations may be interpreted as observables. They correspond to the total

energy and linear momentum of the system. We assume:

a) There is a vector Ω in $\mathfrak{H}$ which is invariant under all $U(a,\Lambda)$. It is the only discrete eigenvector of any $U(a)$. It corresponds to the physical vacuum state. Clearly it has zero energy and momentum.

b) Ω is the ground state of the system. Frequently one makes stronger assumptions on the energy-momentum spectrum in order to introduce particles and to exclude the additional complications associated with the occurrence of zero mass particles.

For instance instead of b)

b') The simultaneous spectrum of the operators P_μ is as in Fig. 1. It consists of the single point $P_\mu = 0$ corresponding to the vacuum state, then the hyperboloid $p^2 = -m^2$, $p^0 > 0$ corresponding to the states of a single particle (mass m) and then the continuum above the 2-particle threshold.

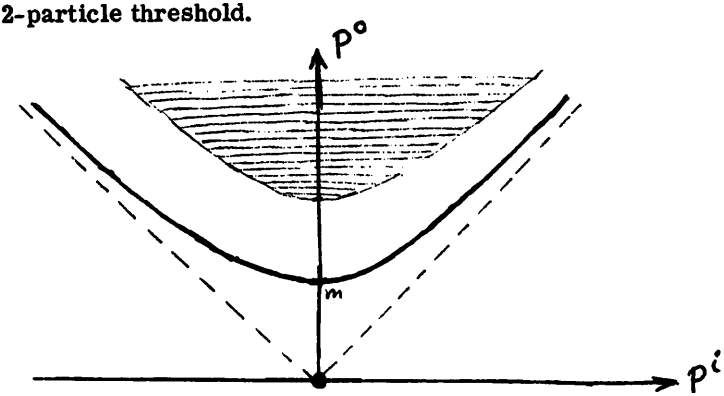

Figure 1.

4) The Field

It had been recognized very early in the development of Quantum Electrodynamics that the field at a point cannot be a proper observable (e.g., the analysis of idealized measurements of electromagnetic field strengths by Bohr and Rosenfeld[1]). One has to consider averages of the field (denoted in our case by A) over space-time regions, such as $\frac{1}{VT}\int_0^T dt \int_V d^3x\, A(x)$ or, in general, weighted averages with smooth weight functions $f(x)$,

$$A(f) = \int d^4x\, A(x) f(x) . \qquad (1.2)$$

In a handwaving way the mathematical nature of the field can be understood if we have the physical picture that $A(x)$ shall represent an operation on the physical system at the point x. One may anticipate then that such an operation must transfer an unlimited amount of energy-momentum to the system.

To express this expectation more precisely we define some notation (corresponding to a direct integral decomposition of the Hilbert space with respect to energy-momentum). Let $|p,n\rangle$ be a (improper = continuous) basis of state vectors, where p denotes the total energy-momentum of the state and the discrete index n is used to distinguish the states with the same p. A general state may be written

$$\Psi = \sum_n \int d\mu(p)\, \psi_n(p)\, |p,n\rangle\,, \qquad (1.3)$$

where the "spectral measure" $d\mu(p)$ may be chosen without loss of generality as (1.4)

$$d\mu(p) = d^4p \left\{ \delta^4(p) + \theta(p^0)\delta(p^2+m^2) + \theta(p^0)\theta(-p^2-4m^2) \right\},$$

the first term corresponding to the vacuum state, the second to the single particle states, and the last to the states above the 2-particle threshold. The normalization of Ψ is given by

$$(\Psi, \Psi) = \sum_n \int d\mu(p)\, |\psi_n(p)|^2 \qquad (1.5)$$

corresponding to orthogonality relations

$$\langle p',n' | p,n\rangle\, d\mu(p) = \delta_{n'n}\, \delta^4(p'-p) \qquad (1.6)$$

The integrand in (1.3), i.e., the object

$$\Psi_p = \sum_n \psi_n(p)\, |p,n\rangle \qquad (1.7)$$

for fixed p can be considered as a vector in a Hilbert space $\mathcal{H}_p$, the metric in $\mathcal{H}_p$ being given by

$$(\Psi_p, \Psi_p)_p = \sum_n |\psi_n(p)|^2 \qquad (1.7')$$

We expect now that the matrix elements $\langle p',n'|A(\omega)|p,n\rangle$ are finite. **In fact, we may even assume that the restriction** of $A(x)$ **which maps from** $\mathcal{H}_p$ **to** $\mathcal{H}_{p'}$ is a bounded operator with a norm $N_{p'p}$. **But even for fixed** p the norm $N_{p'p}$ will not decrease sufficiently for large p' to make $N^2_{p'p}$ integrable with respect

to $d\mu(p')$. In fact as a consequence of covariance and locality of A (assumptions 5 and 6) one would guess that for large p' and fixed p

$$N^2_{p'p} \to \| A(x)\Omega \|^2_{p'} \equiv \rho(p'^2) \qquad (1.8)$$

because all states with bounded energy-momentum should be equivalent in their aspect around a single point x (since such a state cannot give a singular preferential emphasis to any point).

If we define $\mathcal{D}$ as the set of states with fast enough decrease of $\| \Psi_p \|_p$ so that

$$\int d\mu(p') d\mu(p) \; \| \Psi_{p'} \|_{p'} \; N_{p'p} \; \| \Psi_p \| \; < \infty \, ,$$

then the matrix elements

$$\langle \Phi | A(x) | \Psi \rangle \qquad \text{with both } \Phi \text{ and } \Psi \text{ from } \mathcal{D}$$

will be finite. Thus the field at a point may be regarded as a bilinear form over the domain $\mathcal{D}$. Alternatively, if instead of $A(x)$ we take $A(f)$ we get due to (1.1) and the covariance of A under translations

$$\langle p', n' | A(f) | p, n \rangle = \langle p', n' | A(0) | p, n \rangle \tilde{f}(p'-p) \qquad (1.9)$$

with $\tilde{f}(q) = \int d^4x \, e^{-iq\cdot x} f(x)$.

If $f(x)$ is very smooth then $\tilde{f}(q)$ will have a fast decrease for large q. In particular, if $f(x)$ is infinitely often differentiable with respect to the four coordinates x_μ then $\tilde{f}(q)$ will decrease faster than any power in any direction in q-space. Hence for such f and for $\Psi \in \mathcal{D}$ the vector $A(f)\Psi$ will exist, provided that the growth of $N_{p'p}$ is bounded by a polynomial. Under these circumstances the domain $\mathcal{D}$ will also be stable under the application of $A(f)$.

To summarize: The field A at a point may be regarded as a bilinear form on the domain $\mathcal{D}$ but not as an operator. Alternatively, the field averaged with a sufficiently smooth weight function is an operator on $\mathcal{D}$.

Consequences: a) In traditional field theories the dynamics is given by field equations which involve nonlinear functions of the field at a point. The above remarks indicate that it is not at all clear how such nonlinear functions of $A(x)$ can be defined or, in fact, whether they can be defined at all. The answer is known for free fields (where we do not need it) see e.g. [4], [5], and it has been studied for the renormalized perturbation series in some models [6], [7]. It has, however, so far not been possible to incorporate a formulation of a specific dynamical law into the framework of axiomatic field theory.

b) The kinematics of traditional field theories, i.e. the specification of the "degrees of freedom", is given by canonical commutation relations between the field at a fixed time but at different points in space. For the formulation of such relations we do not need the field at a point but only at a sharp time. We must ask therefore whether

$$A(\sigma,t) = \int d^3x\ \sigma(\underset{\sim}{x})\, A(\underset{\sim}{x},t)$$

is an operator, when σ is a sufficiently smooth function in 3-dimensional space. The averaging with σ provides a cut-off in spatial momentum transfer but not in energy. Looking at the expected behavior (1.8) of $N_{p'p}$ one sees that such a cut-off will be sufficient if $\rho(\varkappa^2)$ decreases sufficiently fast for large $\varkappa^2$ and thus furnishes a cut-off in the mass transfer. This is the case for a free field, where $\rho(\varkappa^2) = \delta(\varkappa^2 - m^2)$ but probably not in any model with non-vanishing interaction. The expectation that $\rho(\varkappa^2)$ does not have a fast decrease comes originally from the experience with renormalized perturbation expansions. In recent years, however, some progress has been made towards an understanding of the general reasons (independent of consideration of specific models) [8], [9]. If we accept the fact that $\rho(\varkappa^2)$ does not decrease fast enough to make $A(\sigma,t)$ into a well defined operator then also the formulation of kinematical relations becomes a difficult problem. In

fact it is questionable whether the separation between kinematics and dynamics makes any sense at all.

5) Covariance of the Field

The field shall have a simple transformation property under the Poincaré group. In the example considered here it is

$$U(a,\Lambda)\, A(x)\, U^{-1}(a,\Lambda) = A(\Lambda x + a) \quad (1.10)$$

6) Locality: Einstein's Causality Principle

Whatever the complete physical interpretation of $A(f)$ may be, one wants to assert that it corresponds to an observable whose measurement involves only the part of space-time in which the function $f(x)$ does not vanish. This space-time region is called the "Support of f." If we accept Einstein's postulate that no physical effect can propagate faster than light then the measurement of $A(f)$ cannot perturb that of $A(g)$ whenever the supports of f and g lie space-like to each other. Under these circumstances $A(f)$ and $A(g)$ are compatible observables; the operators $A(f)$ and $A(g)$ should commute:

$$[A(f), A(g)] = 0 \quad \text{if supp } f \text{ is spacelike to supp } g, \quad (1.11)$$

or symbolically

$$[A(x), A(y)] = 0 \quad \text{for } (x-y)^2 > 0\,. \quad (1.11')$$

7) Primitive Causality

Einstein's principle covers only one aspect of causality, the one which is added when we pass from a non-relativistic to a relativistic theory. Common to both is the requirement that the knowledge about the state of a system which is obtainable by measurements at one time suffices to determine the state. In view of the comments made under 4) about the field at a sharp time, we should allow finite time intervals for the measurement. The requirement of primitive causality is then the following: The set of operators $A(f)$ for all functions f with support in a time interval $t-\epsilon < x^0 < t+\epsilon$ should generate a complete system of observables. This requirement may be mathematically expressed in another way: If an operator Q commutes with the $A(f)$ for all functions f with support in a time interval, then Q is a multiple of the identity operator.

B. Adaptation to More Realistic Situations

From a formalistic point of view this is a straight forward matter. The guiding ideas came originally from the study of free fields (linear field equations).

There one has a very direct connection between the nature of the field and the types of particles described by the theory. This connection is at the root of the famous particle-

wave dualism. Its overemphasis has been responsible for widespread and long lasting misunderstandings about the role of Quantum Fields. According to this view each field should be associated with a certain type of particle. The transformation properties of the field are related to the spin of the particle; the equal time commutation relations of the field are determined by the statistics of the particle. Thus one generalizes the assumptions 5) to

5B) The theory deals with several fields ϕ_ρ^i, the upper index distinguishing the different fields, the lower index the components of one field which go over into each other under Lorentz transformations. For each i one has a finite dimensional representation of the homogeneous Lorentz group by matrices $D^i(\Lambda)$. The transformation law of the field is

$$U(a,\Lambda)\,\phi_\rho^i(x)\,U^{-1}(a,\Lambda) = \sum_\sigma D^i_{\rho\sigma}(\Lambda^{-1})\,\phi_\sigma^i(\Lambda x + a). \tag{1.12}$$

To take care of the occurrence of Fermi statistics the causality principle 6) is generalized to

6B) If ϕ^i is associated with a Bose particle then one has

$$[\phi^i(x), \phi^i(y)] = 0 \quad \text{for } x-y \text{ spacelike} \tag{1.13}$$

if ϕ^i is associated with a Fermi particle then one has instead

$$\left[\phi^i(x), \phi^i(y)\right]_+ \equiv \phi^i(x)\phi^i(y) + \phi^i(y)\phi^i(x) = 0$$
$$\text{for } x-y \text{ spacelike} \qquad (1.14)$$

The requirement (1.14) can be put in accordance with Einstein's causality principle if one assumes that every observable contains only even polynomials in the Fermi Fields. The commutation relations between different fields at spacelike distances are usually assumed to be: all Fermi fields anticommute with each other, all Bose fields commute, any Bose field commutes with any Fermi field.

Comments: Replacing 5) and 6) by 5B), 6B) gives a framework which appears flexible enough to accommodate elementary particle phenomenology. The predictive power derived from the bare framework has so far not been spectacular but not entirely void either. The three celebrated successes (all till now in agreement with experience) are the PCT - theorem, the connection between spin and statistics and a few quantitative statements about scattering amplitudes which follow from analytic properties, for instance dispersion relations for π-N scattering. The first two of these successes are of a qualitative nature, and it is somewhat disconcerting that with our present understanding of their roots they depend absolutely crucially on the detailed assumptions listed

above. Thus these two conclusions disappear if in 6B) strict locality is replaced by macroscopic locality. They disappear even if in 5B) one allows fields with an infinite number of components. These questions will be one of the principal concerns of this series of lectures. In particular we shall focus on the problem of statistics.

The framework sketched under B is usually called the Wightman frame not only because Wightman gave the most precise mathematical characterization of (most of) the assumptions, but also because he pointed out one of the most important methods in analyzing the consequences, namely the study of the vacuum expectation values of products of fields [10].

C. Local Quantum Theory (Field Theory without Fields)

The point of view described under B has some unsatisfactory features. First, there is the association of fields with particle types which, though not necessarily tied to this framework, is nevertheless heuristically in the background. It is quite clear that if one insists on such an association and wants to go beyond a purely phenomenological description then one is forced to make a sharp distinction between elementary and composite particles. No satisfactory criterion (of more than approximative value) for such a distinction has been given. The opposite point of view, paraphrased by Chew as "the democracy of particles" has grown since the early thirties. It was in particular the realization that - in spite of the existence of β-decay - the neutron should not be considered as a composite of a proton and an electron which supported the idea that the era of atomistic thinking in physics was over, that the division of structures into elementary building blocks could be regarded only as an approximative model, very successful in nonrelativistic situations but with no place in high energy physics. The most emphatic and consistent advocate of this attitude was Heisenberg who worked hard to develop a theory involving a minimal number of basic fields. These fields should then be associated with the fundamental quantum num-

bers or conservation laws rather than with observed particle types. Although in recent years the possibility of a quark model of elementary particles has given the atomistic idea a new lease on life this does not necessarily contradict the trend towards "democracy." If the quarks turn out to be real and heavy and the models successful then one has an unexpected extension of the regime in which nonrelativistic approximations are meaningful but still one is led to a theory with few fields associated with fundamental quantum numbers rather than particles.

We may look at the relation between fields and particles from another angle: the collision theory of particles in the framework of local field theory. We adopt for the moment the field theoretic framework sketched above leaving the detailed physical interpretation of the fields open and keeping only the generic statements

(i) the operator $\phi^{i}(f)$ shall represent a physical operation on the system which can be performed in the space-time region given by the support of the weight function f,

(ii) the Hilbert space $\mathfrak{H}$ shall contain (the state vectors of) all the different single particle states; the fields shall be complete in this Hilbert space, i.e. there shall be no operator in $\mathfrak{H}$ apart from multiples of the identity

operator which commutes with the $\phi^i(f)$ for all i and f.

One finds then that the identification of those vectors in $\mathfrak{H}$ which correspond to arbitrary configurations of incoming or outgoing particles is already implied and hence the S-matrix elements for all possible processes in such a model are uniquely determined. [11], [12] See also the exposition in [13]. In other words it is neither necessary nor possible to add any further independent assumption concerning the physical significance of the fields. Their only role is to fix for each space-time region $\mathcal{O}$ the set of operators $\mathfrak{F}(\mathcal{O})$ which correspond to physical operations performable in the region. For reasons which will be discussed a little later we shall call $\mathfrak{F}(\mathcal{O})$ the "field algebra of the region $\mathcal{O}$." Once the correspondence

$$\mathcal{O} \rightarrow \mathfrak{F}(\mathcal{O}) \qquad (1.15)$$

is established the theory is defined and its physical content fixed. We may consider (1.15) as the intrinsic definition of the theory and the description by a set of fields as a special way to parametrize (1.15). Loosely speaking, if we start from the frame B then $\mathfrak{F}(\mathcal{O})$ will consist of all functions of the operators $\phi^i(f)$ for all weight functions f which have their support in $\mathcal{O}$.

Let me try to explain what I mean by the word "intrinsic" in this context by an analogy. In geometry we would consider as intrinsic objects the space and its points. Now we may introduce coordinates, thus assigning to each point an n-tuple of numbers. These coordinates are not intrinsic because for the same geometry the introduction of a coordinate system can be done in many different ways; any choice involves some convention. **In a similar way, suppose we are given the collection (net) of field algebras $\mathcal{F}(\mathcal{O})$ for** the various regions $\mathcal{O}$. **Then the theory is fixed. We may** try to find a system of local fields $\phi^i(f)$ **which generates** the net of field algebras $\mathcal{F}(\mathcal{O})$. If such a system ϕ^i can be found it may be regarded as a coordinate system for the net $\mathcal{F}$. As in the above analogy the system ϕ^i is not uniquely determined by the net $\mathcal{F}$. There are many other field systems $\phi^{k'}, \phi^{l''}, \ldots$ which lead to the same net and hence to the same physics.

This non-uniqueness of the choice of a set of fields within one physical theory has been stressed by H. J. Borchers [14]. I shall call therefore the collection of all local fields associated with the same net $\mathcal{F}$ a "Borchers class." It is true that the definition of such a class of "relatively local" fields used by Borchers is slightly different from the one I use here but this is a difference in technicalities

and not in spirit. So let us not worry about it at this stage. To fix the ideas I want to illustrate the nature of a Borchers class in an example well known to all of you, the theory of a free scalar field A. In that case we may define other "local functions" of A by the Wick-ordered powers:

$$A_n(x) = :A(x)^n:$$

We can also obtain local vector fields, e.g.

$$B_\mu(x) = :\partial_\mu A(x)\, \partial^\nu A(x)\, \partial_\nu A(x):$$

or tensor fields

$$B_{\mu\nu}(x) = :\partial_\mu A(x)\, \partial_\nu A(x):$$

etc.

From any one of these fields one generates either the same local net $\mathcal{F}$ as from the original field A or a subnet of it.

Let us now look a little closer at the nature of $\mathcal{F}(\mathcal{O})$ and the necessary structural relationships within a net $\mathcal{F}$. We have called $\mathcal{F}(\mathcal{O})$ the "field algebra" of $\mathcal{O}$, implying thereby that $\mathcal{F}(\mathcal{O})$ is an algebra. In fact it is a $*$-algebra, i.e. an algebra permitting an adjoint operation. This means that if F_1 and F_2 belong to $\mathcal{F}(\mathcal{O})$ and if α_1, α_2 are any complex numbers, then the operators

$$F_1 F_2 \quad , \quad \alpha_1 F_1 + \alpha_2 F_2 \quad , \text{ and } \quad F_1^*$$

again belong to $\mathcal{F}(\mathcal{O})$. If we have in mind the specific construction of $\mathcal{F}(\mathcal{O})$ by functions of a set of basic fields then this statement is obvious. If, on the other hand we want to start from the physical significance of $\mathcal{F}(\mathcal{O})$ as the set of all "operations performable in the region $\mathcal{O}$," then the statement that $\mathcal{F}(\mathcal{O})$ must be a *-algebra is not quite so evident. It is tied to peculiarities of the general quantum theoretic description buried in our assumption 1). I am not too happy about the status of our understanding of this formalism from operational principles but it would carry us too far astray if we tried to discuss such questions here.

Some further more technical specifications about the nature of $\mathcal{F}(\mathcal{O})$ have to be added. For many purposes it is convenient to consider only bounded operators. This is no restriction from the physical point of view. Suppose q is an unbounded self adjoint operator and ϱ a function from the real numbers to the real numbers which maps the whole real line on a finite interval so that $|\varrho(\xi)| \leq a$ for all ξ. Then the operator $\varrho(q)$ is bounded and has a norm $\|\varrho(q)\| \leq a$. If the mapping $\xi \to \varrho(\xi)$ has a unique inverse then the measurement of the observable q and that of the observable $\varrho(q)$ are the same physical operations. Only the scale of the measured values has been regauged from ξ to $\varrho(\xi)$.

Thus we want to assert that each $\mathcal{F}(\mathcal{O})$ shall be a *-algebra of bounded operators. The set of all bounded operators in a Hilbert space $\mathcal{H}$ is denoted by $\mathcal{B}(\mathcal{H})$. There are several important topologies in the set $\mathcal{B}(\mathcal{H})$ which have been studied extensively in the mathematical literature. It turns out that those topologies which appear natural in the mathematical context also have a rather direct physical significance in our discussion. Let me say some words at this stage about the relevant mathematical concepts (topologies, von Neumann rings, C*-algebras, etc.) and the most important theorems relating to them. I shall collect this material in an appendix so that the reader who is familiar with these mathematical matters can ignore the digression.

For the moment we need only note that $\mathcal{F}(\mathcal{O})$ will be taken to be a von Neumann algebra and that this specification is a matter of convenience and choice rather than a restrictive assumption concerning the physics.

The adaptation of 5B and 6B to this language is trivial. Instead of 5B we have the simple transformation law

$$U(a,\Lambda)\,\mathcal{F}(\mathcal{O})\,U^{-1}(a,\Lambda) = \mathcal{F}(\Lambda\mathcal{O}+a) \qquad (1.16)$$

meaning that the field algebra of one region is mapped by the Poincaré operators $U(a,\Lambda)$ exactly onto the field algebra of the image region. In the case of 6B we still

have the somewhat artificial sounding assumption: Each $F \in \mathcal{F}(\mathcal{O})$ can be uniquely decomposed into

$$F = F_B + F_F \qquad (1.17)$$

with F_B and F_F from $\mathcal{F}(\mathcal{O})$. (The indices B and F stand for "Bose part" and "Fermi part"). If $F^{(1)}$ and $F^{(2)}$ belong to algebras of two regions which are space-like separated, then

$$[F_B^{(1)}, F_B^{(2)}] = [F_B^{(1)}, F_F^{(2)}] = 0 \qquad (1.18)$$

$$[F_F^{(1)}, F_F^{(2)}]_+ = 0 \qquad (1.19)$$

Again, if F is an observable then $F_F = 0$ and, of course, the product of two Bose type or two Fermi type operators is of Bose type, the product of a Bose-type with a Fermi-type operator is of Fermi type.

The unsatisfactory feature of this formulation of the locality principle does not only lie in its complicated structure. Let us denote by $R(\mathcal{O})$ the subalgebra of $\mathcal{F}(\mathcal{O})$ generated by the observables. Then for $\mathcal{O}_1$ spacelike to $\mathcal{O}_2$ and $A_1 \in R(\mathcal{O}_1)$, $A_2 \in R(\mathcal{O}_2)$ we must have

$$[A_1, A_2] = 0 \qquad (1.20)$$

The commutativity for observables is the well known condition for compatibility of their measurement and the com-

patibility is demanded by Einstein's causality principle. Let me emphasize again that the assumption (1.20) is not made in view of our knowledge about Bose-statistics but is an expression of the causality principle. This suggests a number of questions:

1) Why should we consider a "field algebra" $\mathcal{F}$ which is essentially larger than the "observable algebra" R?

2) While it is clear that for nonobservable elements of $\mathcal{F}$ we need not have space-like commutativity in order to satisfy the causality principle (as illustrated by the example of Fermi-type operators) it is to be expected that the causality principle puts some restrictions on the commutation relations of the "local operations" which are not observables. The relations (1.18), (1.19) specify one possible commutation structure. But is this the only possibility? Does causality alone imply that there exists only the Bose-Fermi alternative?

We may answer the first question for the moment in a preliminary, so to speak phenomenological, way. Recall that we insisted on incorporating in $\mathcal{H}$ the state vectors of all types of particles. From collision theory and the causality requirement (1.20) for observables it follows that the observable algebra cannot connect the vacuum state with a state of a single Fermi particle. Hence if we restrict our attention to the observable algebra and if Fermi particles occur then $\mathcal{H}$ must be decomposable into subspaces $\mathcal{H}_\sigma$

so that R transforms each $\mathcal{H}_\sigma$ into itself. In other words, we must have "superselection rules" [15];

the matrix elements of any observable between vectors from different subspaces $\mathcal{H}_\sigma$ vanish. This brings us again close to the second question; we should understand why superselection rules appear and whether we can say anything about their structure.

To investigate these questions one can start from two different angles. The first approach would be to realize that the concept of an "operation" on the system is more general than that of an "observable" and that correspondingly we get less stringent requirements for causally disjoint "operations" than for causally disjoint "observables." Since we shall not follow this approach I shall not try to make the notion of a "physical operation" more precise at this stage but refer to [16] in which this concept is partly used. The second approach, and this is the one we shall take, starts from the observation that all the physical information of the theory must already be contained in the net of observable algebras R and in fact even in the restriction of R to one of the invariant subspaces $\mathcal{H}_\sigma$, say to the space $\mathcal{H}_0$ which results by the application of R on the vacuum state vector. In $\mathcal{H}_0$ we have an irreducible but still faithful representation of the observable algebra. A justification of these claims will be

given a little later. First I want now to formulate the framework as it results if we focus only on the observable algebra and its irreducible representation in the "vacuum sector" $\mathcal{H}_0$. The assumptions 1), 2), 3) concerning quantum physics, Poincaré invariance and P_μ-spectrum are not changed. We may, however, anticipate that in the spectrum of P_μ we do not necessarily find the mass-hyperboloids of all relevant particles (because some of these states may lie in other superselection sectors).

4C. To each double cone K in space-time one has a von Neumann ring $R(K)$, the algebra of observables localized in the region K.

Note: For reasons becoming apparent later we do not at this stage consider arbitrary regions of space-time but only the simplest set of Poincaré covariant, finitely extended regions, namely double cones.

Instead of 5B we have

5C $$U(a,\Lambda)\, R(K)\, U^{-1}(a,\Lambda) = R(\Lambda K + a)\,. \qquad (1.21)$$

In order to formulate the primitive causality we need consider also the algebras of regions which are not double cones. Let us denote by $\bigvee_i R_i$ the von Neumann ring generated by all the rings R_i, so that one has

$$\vee_i R_i = \{\cap R_i'\}' \qquad (1.22)$$

If $\mathcal{O}$ is any space-time region then the following definition suggests itself.

Definition.

$$R(\mathcal{O}) = \vee_i R(K_i) \quad \text{over all } K_i \subset \mathcal{O}\,. \qquad (1.23)$$

Note that this introduces no essential additional restriction on the relation of the $R(K)$ because, if $\mathcal{O}$ is itself a double cone then on the right hand sight K also appears so that in this case the only consequence is the monotony requirement

$$R(K_1) \subset R(K_2) \text{ if } K_1 \subset K_2 \qquad (1.24)$$

We can now formulate a (strengthened) version of the primitive causality assumption.

7C. Let $\{K_i\}$ be an arbitrary covering of the base of a double cone K,
then $\vee_i R(K_i) \supset R(K)\,.$ (1.25)

Note that this assumption implies first that (at least for the simple types of regions considered here) the algebra of the region is generated by the algebras of an arbitrary set of covering subregions. This property, called "additivity" is

suggested by the field theoretic background. For a field ϕ we can decompose $\phi(f)$ into a sum of $\phi(f_i)$ where the f_i have their respective support in the subregions corresponding to an arbitrarily chosen covering of the support of f. This decomposition of the smeared out fields corresponds to the additivity property of the rings. Secondly, (1.25) demands the hyperbolic propagation character of the equations of motion. Loosely speaking the Cauchy data on the base of a double cone determine the quantities everywhere in the double cone.

Another requirement, mentioned before, is

8C. Irreducibility

$$\bigvee_{all\,K} R(K) = \mathfrak{B}(\mathfrak{H}), \text{ or } \bigcap \{R(K)'\} = \{\lambda 1\}. \tag{1.26}$$

Finally there is the Einstein causality which we shall incorporate in a somewhat stronger assumption:

6C. Duality (for double cones)

$$R(K)' = R(K'), \tag{1.27}$$

where K' is **the causal complement of** K, i.e. the set of points which are space-like to K

Clearly this assumption is a strengthened version of the locality postulate since the latter could also have been written

$$R(K') \subset R(K)'. \tag{1.28}$$

Beyond the locality requirement the assumption of duality implies that the rings $R(K)$ are maximal. If we have a net of von Neumann rings $R(K)$ satisfying the requirements 1), 2), 3), 4C, 5C, and 1.28 we may ask whether we can find a richer net $\bar{R}(K) \supset R(K)$ still satisfying the mentioned requirements. Now if Q belongs to $\bar{R}(K)$ it has to commute at least with $R(K')$. Hence, if R satisfies (1.27) then $\bar{R}(K) = R(K)$, i.e. then R is already maximal. One may ask whether conversely maximality also implies duality. Suppose we want to enrich the net R by adjoining one more element Q to $R(K)$ for one special K Because of the covariance we have to adjoin the element $U(a,\Lambda)\, Q\, U^{-1}(a,\Lambda)$ to $R(\Lambda K + a)$. There will be a subset S_K of Poincaré transformations for which $\Lambda K + a$ is totally spacelike to K Therefore Q can be adjoined to $R(K)$ if and only if the following two requirements are met

(i) $$Q \in \{R(K')\}'$$

(ii) $$\left[Q, U(a,\Lambda) Q U^{-1}(a,\Lambda)\right] = 0 \quad \text{for } (a,\Lambda) \in S_K.$$

In the case of fields (framework B) Borchers [14] has derived the interesting result that if $\phi^i, (i = 1, \cdots, N)$, is a complete set of local fields and if ϕ^0 is another field (not assumed local) satisfying

$$\left[\phi^0(x), \phi^i(y)\right] = 0 \quad \text{for } (x-y) \text{ space-like}$$

then ϕ^0 itself is local. This so called "transitivity of locality" suggests that perhaps the requirement (ii) on Q may follow from (i) under rather general circumstances. If this could be shown then we could always enrich the net R by adjoining elements from $R(K')'$ until the resulting net satisfies duality and is then maximal. For the moment, however, it is not known under what circumstances (ii) can be inferred from (i) and hence it is not clear whether the duality assumption is an extra restriction (beyond locality and maximality). For the net of von Neumann rings arising in free field theories Araki has shown that duality holds [17] .

D. Algebraic Approach

In the early fifties a mathematical fact of seemingly great importance to quantum field theory was noticed and emphasized independently by several authors. Studying the canonical commutation relations for an infinite system of degrees

of freedom $(q_k, p_k;\ k = 1, \cdots \infty)$ it was found that there are many inequivalent irreducible representations of the q_k, p_k by operators in a (separable) Hilbert space. The number of such inequivalent representations is not denumerable and a complete classification in any constructive sense seems impossible.*) This is in contrast to the situation of a finite number of degrees of freedom where there is essentially one unique equivalence class of irreducible representations. It was also realized that the many inequivalent representations could not be dismissed as pathological. In fact, if a Quantum Field Theory could be defined at all by field equations and kinematical commutation relations then the selection of the representation space of the kinematical relations so that it fits with the assumptions 2) and 3) is a problem determined by the dynamics. In his talk at the Lille Conference 1956 [18] I. E. Segal confronted a rather critical and disbelieving audience with

---

* For the canonical anticommutation relations this phenomenon had already been pointed out by J. von Neumann, Comp. Math. 6, 1 (1938) but for some reason its relevance for quantum field theory had not been realized till much later.

the claim that the representation problem was irrelevant, that one did not need operators in a Hilbert space but only an abstract C*-algebra. In a very interesting earlier paper [19] Segal had studied the mathematical structure of quantum mechanics and pointed out there that many questions of physical interest (e.g. the determination of spectral values) could be answered without reference to a Hilbert space if one chooses the algebra of observables to be a C* algebra. The skepticism of the physicists about the possibility of a purely algebraic approach to field theory was due to the lack of a convincing idea as to how a typical scattering experiment could be discussed in this frame. Also the physical significance of the myriads of inequivalent representation was not understood. But seven years later we realized that Segal's claim had been essentially correct.

I give a brief description of the algebraic version of general quantum physics, i.e., the mathematical and conceptual structure which constitutes assumption 1) in this language:

1 D) The central mathematical object is a C* algebra $\mathcal{A}$. A state ω is mathematically described by a positive linear form over $\mathcal{A}$. Every complex valued function on $\mathcal{A}$ satisfying the two properties (with $A, B \in \mathcal{A}$ and α, β complex numbers)

$$\omega(\alpha A+\beta B)=\alpha\omega(A)+\beta\omega(B) \qquad \text{(linearity)}$$

$$\omega(A^*A)\geq 0 \qquad \text{(positivity)}$$

is a state.

We may add the normalization convention

$$\omega(1)=1$$

One can immediately distinguish pure states and mixtures (see appendix). Every element $C\in\mathcal{O}$ induces a linear transformation of states $\omega\rightarrow\omega_C$ defined by

$$\omega_C(A)=\omega(C^*AC)\ . \qquad (1.29)$$

Such a transformation maps the pure states into pure states. The norm is changed by the factor $\frac{\omega(C^*C)}{\omega(1)}$.

The correspondence of these mathematical objects to physics is the following. A "state" represents a statistical ensemble of physical systems. The norm $\|\omega\|=\omega(1)$ may be regarded as a measure for the total number of systems in the ensemble (in arbitrary units, hence usually put equal to 1). The mixing of states is linearly represented here. For instance $\omega'=\lambda_1\omega_1+\lambda_2\omega_2$ (with $\lambda_1,\lambda_2>0$) is the mixture of ω_1 and ω_2 with weights λ_1 , λ_2 . (In the Hilbert space version this corresponds to the mixing of density matrices, not the superposition of wave functions!).

The physical significance of the elements of the algebra is a double one. On the one hand, if $A \in \mathcal{A}$ is self adjoint, then A may be regarded as an "observable" in the standard sense of this term. If A is regarded in this role then $\omega(A)$ is the expectation value of the observable A in the state ω. Secondly, any element of $\mathcal{A}$ with norm less or equal to one (whether self adjoint or not) represents an "operation." By this we mean the change of state produced if an apparatus acts during a finite amount of time on the systems constituting the ensemble ω. The "operation" may include a selection process by which a certain fraction of the systems in the original ensemble is rejected by the apparatus (prototype of an "operation" is an arrangement of Nicol prisms and quarter wave plates). The fraction of the original systems which is transmitted will be called the transition probability for the state through the apparatus. If $C \in \mathcal{A}$ is regarded as an operation then the change of state is given by (1.29) and the transition probability by $\frac{\omega(C^*C)}{\omega(1)}$.

Of course the measurement of an "observable" also implies an "operation" in the above sense if the measured systems are available for subsequent further observations. Such an operation transforms, however, in general pure states into mixtures and does not coincide with (1.29) except in the special case when C is a projection (i.e. $C = C^* = C^2$). In

general, we have the double role of the algebraic elements and it is useful to keep both roles in mind.

These remarks may suffice to indicate how an experiment may be described in the algebraic frame, without recourse to Hilbert space. A typical experiment may be schematized by a source which prepares the initial state and an analyzing apparatus involving a selection process. The result of the experiment is then the transition probability of the initial state through the apparatus. The description of the source and of the initial state will be done by a combination of two methods both corresponding to actual experimental practice. The first is filtering, the second monitoring. In the first, one uses an "operation" with as small a range as possible on an entirely unknown original state. In the second, one obtains information measuring the transition probabilities through a certain number of monitoring apparatus.

We may now compare with the Hilbert space formulation. There one uses ordinarily an irreducible representation of the observable algebra $\mathcal{O}$ by operators in a Hilbert space $\mathcal{H}$. Let us denote the operator representing the algebraic element A by $\pi(A)$. Picking up any vector $\Psi \in \mathcal{H}$ we obtain a state on $\mathcal{O}$ by

$$\omega_\Psi(A) = \langle \Psi, \pi(A)\Psi \rangle . \qquad (1.30)$$

If we let Ψ run through $\mathfrak{H}$ we obtain precisely all those states which result from one of them by transformations of the form (1.29) i.e. by application of some operation from $\mathcal{A}$. We call one such family of states a "superselection sector" because, if Ψ_1 and Ψ_2 are state vectors belonging to two unitarily inequivalent irreducible representations π_1 resp. π_2 they are so to speak incomparable. A linear superposition of such vectors is meaningless. It may, of course, be formally defined in the representation $\pi_1 \oplus \pi_2$ but then it only corresponds to the mixing of the states, not to a coherent superposition. Thus, existence of unitarily inequivalent irreducible representations of $\mathcal{A}$ is synonymous with existence of superselection sectors. It also means that there are "pure operations" (i.e. linear transformations of states mapping pure states into pure states) which are not induced by elements of the observable algebra (not of the form (1.29)).

In the example of the algebra generated by an infinite system of canonical quantities we remarked that there is a tremendously large multitude of inequivalent irreducible representations and, correspondingly,an overwhelmingly rich supply of states. This phenomenon is typical for the nets of local algebras encountered in field theory. One may ask whether really all of these representations should be considered or

whether some additional physical restriction has to be imposed to limit the number of superselection sectors which are regarded as physically relevant. The answer to this question depends somewhat on the point of view and on a reasonable balance between considerations of principle and those of practical nature. From the point of view of principle one can say that actually it is impossible to know precisely which state is prepared by a given source. The available information is always such that it determines not a state but a so called "weak neighborhood" in state space. This is a set of states whose common feature is the validity of some finite set of inequalities

$$| \omega(A_i) - a_i | < \epsilon_i \; ; \; i = 1,2,\cdots,N; \; \epsilon_i > 0. \quad (1.31)$$

For instance in monitoring a state one is only able to make a finite number of measurements (say $A_1,\cdots,A_N$) with a limited accuracy $(\epsilon_1,\cdots,\epsilon_N)$. Any choice of measurements A_i, mean values a_i and error limits ϵ_i defines one weak neighborhood in state space.

Theorem [20].

Let π and π_0 be two irreducible representations and ω_0 some vector state of π_0. If π is faithful then every weak neighborhood of ω_0 contains also a vector state of π.

This means that it is impossible to prepare an ensemble by a realistic source in such a way that we know in which superselection sector the corresponding state lies. The distinction based on unitary inequivalence of representations is much too fine to be physically measurable. We might say that all faithful representations are physically equivalent. We might select any one of them (or use no representation at all) in order to discuss a specific experimental set-up.

On the other hand it is usually very convenient to simplify the description by an idealization which restricts the set of states considered (adding some information about the occurring states which is neither needed in principle nor available in reality). We shall do this also subsequently and thereby reduce the number of relevant superselection sectors to a manageable size and the distinction between them to physically important quantum numbers. This discussion will be exemplified below.

Thus far we have described how the assumption 1), the principles of Quantum Physics are expressed in the algebraic approach. The expression of the other assumptions is to a large part a straightforward adaptation of the formulations under C.

For instance, instead of the von Neumann rings $R(K)$, we consider now C*-algebras $\mathcal{A}(K)$ with the obvious monotony property that $K_1 \supset K_2$ implies $\mathcal{A}(K_1) \supset \mathcal{A}(K_2)$. The total algebra of observables is defined as

$$\mathcal{A} = \overline{\bigcup_K \mathcal{A}(K)}\,, \qquad (1.32)$$

where the bar means the completion in the norm topology.

One feature of physical importance, tied to the definition (1.32), should be stressed. Every element of $\mathcal{A}$ can be approximated by an element from some finite region K uniformly with respect to all states. Hence $\mathcal{A}$ contains only elements which correspond to still essentially local quantities, observables which are "quasilocal." Truly "global" quantities, like (bounded functions of) the total energy are excluded as unmeasurable, and this is satisfactory because their measurement would require an infinitely extended apparatus. In this respect the difference between the norm topology and the strong or weak operator topologies is crucial. Take the example of a Poincaré operator $U(a,\Lambda)$. It is a global operation because its effect does not become weaker in faraway regions of space. The norm of the difference between $U(a,\Lambda)$ and any element from any $\mathcal{A}(K)$ is always greater than 1 because, no matter how large K is we can always find states which are essentially different from the vacuum with respect to measurements outside of K. On such states the effect

of $U(a,\Lambda)$ is very different from that of any $A \in \mathcal{O}(K)$. If, on the other hand, we first pick a state vector Ψ in the Hilbert space $\mathcal{H}_0$ then we can find a region K large enough so that with respect to observations in the causal complement K' this state and the vacuum state are almost identical. Then one can find an $A \in \mathcal{O}(K)$ so that $\|(U(a,\Lambda)-A)\Psi\| < \epsilon$. This means that $U(a,\Lambda)$ can be approximated by local quantities in the strong operator topology in the vacuum representation.

We have seen that $U(a,\Lambda)$ cannot belong to $\mathcal{O}$. Hence, in a strictly algebraic formulation the Poincaré invariance cannot be expressed in terms of unitary operators. Instead we have to each Poincaré transformation (a,Λ) a corresponding automorphism $\alpha_{a,\Lambda}$ of the algebra $\mathcal{O}$ (see appendix for an explanation of the term "automorphism"). Again we must have the product relation

$$\alpha_{a_1\Lambda_1}\,\alpha_{a_2\Lambda_2} = \alpha_{a_1+\Lambda_1 a_2,\,\Lambda_1\Lambda_2}\,,$$

and, in fact, now we have this relation really applying to the Poincaré group itself, not to the covering group because the arbitrariness of phase of a state vector Ψ does not enter here. The elements of the algebra and the states considered as positive forms over the algebra are free from this ambiguity. The transformation law reads then

$$\alpha_{a,\Lambda}\left(\mathcal{O}(K)\right) = \mathcal{O}(\Lambda K + a)\,. \qquad (1.33)$$

From (1.32) we see one reason for the occurrence of many inequivalent irreducible representations of $\mathcal{A}$. A state ω over $\mathcal{A}$ defines a state over each subalgebra $\mathcal{A}(K)$. This is called the restriction of $\mathcal{A}$ to $\mathcal{A}(K)$ or, for brevity, in more physical terms the "partial state" of the region K. It will be denoted by $\omega|_K$. Now consider a sequence of mutually space-like regions K_n, moving to infinity as $n \to \infty$. We may prescribe arbitrarily a partial state $\omega|_{K_n}$ for each of these regions and there exists always a total state ω which is the simultaneous extension of this collection of partial states to the algebra $\mathcal{A}$. (Compare [21]). All vector states occurring in the Hilbert space of one irreducible representation of $\mathcal{A}$ have a common asymptotic behavior for their partial states $\omega|_{K_n}$ as $n \to \infty$. In other words for any two such normalized states we have $\| \omega^{(1)}|_{K_n} - \omega^{(2)}|_{K_n} \| \to 0$. Omitting fine points the proof uses the following facts: **The** two state vectors in question, **being associated with the same** irreducible representation π, **are related by**

$$\Psi^{(2)} = \pi(C)\,\Psi^{(1)} \quad \text{with} \quad C \in \mathcal{A}.$$

Since C is quasilocal it cannot change the partial state far away unless the state $\Psi^{(1)}$ already has correlations between its partial states in regions which are infinitely far apart. This is however not possible for a pure state.

We may formalize this in the following way. Let K_ℓ be a sequence of increasing double cones which exhaust space-time:

$$K_{\ell+1} \supset K_\ell \quad ; \quad \cup K_\ell = \text{all space-time} . \tag{1.34}$$

Take any sequence of elements $A_\ell \in \mathcal{A}(K_\ell')$ with $\|A_\ell\| \leq 1$. Then for every fixed $C \in \mathcal{A}$ we have due to locality

$$\| [C, A_\ell] \| \to 0 .$$

Since A_ℓ tends to commute with all elements of the algebra, the representatives $\pi(A_\ell)$ in an irreducible representation approach multiples of the identity operator (by a slight generalization of Schur's lemma). One has therefore the "cluster property"

$$(\Psi, \pi(A_\ell C)\Psi) \to (\Psi, \pi(A_\ell)\Psi)(\Psi, \pi(C)\Psi)$$

for normalized Ψ, and the first factor is, in fact, even independent of the direction of the unit vector Ψ in the representation space. This includes the statement that all vector states in an irreducible representation have the same asymptotic tail of partial states.

We see therefore how to construct a great variety of inequivalent representations of $\mathcal{A}$. We just have to choose at random sequences of partial states for space-like separated regions moving to infinity and then extend each such sequence

to a total state on $\mathcal{A}$. If the asymptotic tails of two such sequences do not coincide we obtain states belonging to inequivalent representations. Simple examples of such states are those which describe a non-vanishing density of particles extending to infinity in space. For the purpose of elementary particle physics it is, however, both legitimate and convenient to idealize the "cosmology" by the claim that all states of interest to us coincide with the vacuum state asymptotically for observations in far away regions.

For the remainder of these lectures we shall limit our attention to states ω for which the difference of the partial states $(\omega-\omega_0)|_{K_\ell'}$ in the causal complement K_ℓ' of the double cone K_ℓ goes to zero in norm as $\ell \to \infty$ for a sequence of double cones K_ℓ of the type (1.34). Thus, for any "state of interest" ω and any positive ϵ there is a double cone K_ϵ so that

$$\| (\omega-\omega_0)|_{K_\epsilon'} \| < \epsilon . \qquad (1.35)$$

This restriction on the "states of interest" is a very strong one. Let me point out that it is, in fact, too strong to be reasonable in Quantum Electrodynamics because there Gauss's law asserts that an electric charge located in a finite region can be determined by means of field strength measure-

ments on the surface of an arbitrarily large sphere. Hence, no matter how large we choose K there is always some element of $\mathcal{A}(K')$ for which a state with a localized charge and the vacuum give markedly different expectation values. We should really require the asymptotic coincidence of states not for their restrictions to the complements K_ℓ' of a sequence of increasing finite regions K_ℓ, but the partial states of finitely extended regions moving out to infinity. Still, excluding the much more difficult case of long range forces from our consideration, the restriction (1.35) appears reasonable.

Another condition on the states we wish to consider, a condition which is not unrelated to (1.35) but not quite a consequence of it is the following: all states considered shall lead (via the GNS-construction) to representations in which the Poincaré automorphism group $\alpha_{a,\Lambda}$ can be (continuously) implemented by unitary operators $U(a,\Lambda)$ and furthermore the resulting spectrum shall be contained in the forward cone (no negative energies).

With these limitations on the "states of interest" and some well supported properties of the vacuum representation*) one finds that all the representatives we want to consider are "strongly locally equivalent." This means that for any two

---

* For instance that $\pi_0(\mathcal{A}(K))''$ is a factor of Type III.

such representations their restrictions to a subalgebra $\mathcal{A}(K')$(or to $\mathcal{A}(K)$ for that matter) are unitarily equivalent. The set of partial states of a region $\mathcal{O}$ occurring in any of these representations are identical as long as $\mathcal{O}$ has a non-void space-like complement $\mathcal{O}'$. The distinction between the surviving superselection sectors appears only when we consider the total algebra $\mathcal{A}$. In particular we have now not only the asymptotic coincidence of states as demanded by (1.35) but in each superselection sector of interest and for each region K we can find states which coincide exactly with ω_0 on $\mathcal{A}(K')$. Such states are called "exactly localized" in K. If such a state lies in a different superselection sector than the vacuum, say it lies in the sector σ, then we may visualize the situation by the physical picture that the state has a "charge quantum number σ" but that this charge is strictly localized within K. It can be strictly localized in any region, no matter how small, but it has to sit somewhere. Again we see that this is not realistic if the charge is necessarily accompanied by an extended, observable field as in Electrodynamics. Our limitations probably exclude the most interesting case of "gauge invariances of the second kind" and refer only to charge quantum numbers which are not the sources of fields, to cases in which only the analogue of gauge transformations of the first kind exist.

Let me briefly indicate the essential point in the argument which leads from our limitations on the states of interest to the strong local equivalence of the corresponding representations. If $\pi^{(1)}$ and $\pi^{(2)}$ are two disjoint representations of a C^*-algebra $\mathcal{C}$ (meaning that no subrepresentation of $\pi^{(1)}$ is unitarily equivalent to any subrepresentation of $\pi^{(2)}$) and if $\omega^{(1)}$ and $\omega^{(2)}$ are normalized vector states in the respective representations, then $\| \omega^{(1)} - \omega^{(2)} \| = 2$. Thus (1.35) demands that no representation of interest, when restricted to a sufficiently far out region K_ℓ', can be disjoined from the vacuum representation of $\mathcal{O}(K_\ell')$. For factors of Type III non-disjointness implies unitary equivalence. The assumed translational covariance allows us to extend this unitary equivalence to the algebras of other regions.

With these limitations on the "states of interest" we can quickly summarize the assumptions on which the remainder of these lectures will be based. We can now without loss of generality identify the local algebras $\mathcal{O}(K)$ with the Von Neumann algebras $R(K)$ in the vacuum representation (since all "states of interest" are normal states on this net of Von Neumann algebras; see Appendix for definition of the term "normal state"). We can then take over all the structural assumptions concerning the net $R(K)$ as described under C. The question to be asked

is: do there exist unitarily inequivalent irreducible representations π_σ of the total algebra $\mathcal{A} = \overline{\cup \mathcal{A}(K)}$ which are strongly locally equivalent, i.e. for which $\pi_\sigma(\mathcal{A}(K'))$ is unitarily equivalent to $\pi_0(\mathcal{A}(K'))$ for every K. If so, these representations will be the relevant superselection sectors for us. We wish to classify them.*

II. Structure of Superselection Rules; Charge Quantum Numbers; Statistics

The question to be studied here was described at the end of the last section. Consider a representation π_σ in a Hilbert space $\mathcal{H}_\sigma$ and pick some double cone K. Since the representations π_σ and π_0 when restricted to $\mathcal{A}(K')$ are unitarily equivalent we can find a unitary mapping from $\mathcal{H}_0$ to $\mathcal{H}_\sigma$, denoted by V, such that

$$V\pi_0(A)\Psi = \pi_\sigma(A)V\Psi \quad \text{for all } A\in\mathcal{A}(K') \text{ and } \Psi\in\mathcal{H}_0. \tag{2.1}$$

We may omit the symbol V if we identify the two spaces $\mathcal{H}_0$ and $\mathcal{H}_\sigma$ (identifying the vector Ψ with its image $V\Psi$). This identification also makes $\pi_\sigma(A)$ into an operator from $\mathcal{B}(\mathcal{H}_0)$ which can be expressed in terms of the $\pi_0(\mathcal{A})$. We shall also omit the symbol π_0 because we take the operator algebras $\pi_0(\mathcal{A}(K))$ in the vacuum sector as identical by definition to the C^*-algebras $\mathcal{A}(K)$. Then (2.1) reads

* This analysis was initiated by Borchers [22]; the following section is based on [23,24,25].

$$\pi_\sigma(A) = A \qquad \text{for all} \quad A \in \mathcal{A}(K'). \tag{2.2}$$

On the other hand, if $B \in \mathcal{A}(K_1)$ and $A \in \mathcal{A}(K_1')$, then as soon as $K_1 \supset K$ we have

$$[\pi_\sigma(A), \pi_\sigma(B)] = [A, \pi_\sigma(B)] = 0 .$$

Hence by duality

$$\pi_\sigma(\mathcal{A}(K_1)) \subset \mathcal{A}(K_1) \quad \text{for } K_1 \supset K . \tag{2.3}$$

This means first of all that

$$\pi_\sigma(\mathcal{A}) \subset \mathcal{A} . \tag{2.4}$$

Also, π_σ is a faithful representation and therefore preserves the algebraic structure and the norm. The transition from the vacuum sector to the sector σ can therefore be described by a "localized endomorphism" ρ_σ. In detail this means that the representative $\pi_\sigma(A)$ may be considered as the image $\rho_\sigma(A)$ of a norm preserving mapping of the algebra <u>into</u> itself and such that ρ_σ acts like the identity mapping on $\mathcal{A}(K')$ and only reshuffles the elements of the algebra of the finite region K. If the mapping ρ_σ is <u>onto</u> the whole of $\mathcal{A}$ we call it an <u>automorphism</u>. In general, however, the image of ρ may be smaller than $\mathcal{A}$. In either case the endomorphism is called <u>localized</u> in K because it does

nothing to the elements belonging to the causal complement region K'.

Let me summarize this discussion and the resulting notation. Due to the unitary equivalence of all representations $\pi_\sigma(\mathcal{A}(K'))$ we can identify all representation spaces with $\mathcal{H}_0$ and as a consequence all operators $\pi_\sigma(A)$ with operators from $\pi_0(\mathcal{A})$. Since we consider now π_0 as the defining representation of $\mathcal{A}(K)$ we omit the symbol π_0. So any $\pi_\sigma(A)$ will be identified with some element of $\mathcal{A}$, namely with $\rho_\sigma(A)$, the image of A by the endomorphism ρ_σ. Of course the identification of $\mathcal{H}_\sigma$ with $\mathcal{H}_0$ and the corresponding identification of $\pi_\sigma(A)$ with $\rho_\sigma(A)$ is not canonical (=natural) and highly non-unique. It depends on the arbitrary choice of a reference region K and even beyond that on the choice of the V in (2.1) which is determined only up to a unitary operator from $\mathcal{A}(K)$. Still, any localized endomorphism ρ will lead from a representation $\pi^{(1)}$ to a strongly locally equivalent representation, denoted by

$$\pi^{(2)} = \pi^{(1)} \circ \rho \qquad \text{(in detail } \pi^{(2)}(A) = \pi^{(1)}(\rho(A))\text{)}$$

and any representation of interest can be obtained from the defining representation $\pi_0(\mathcal{A}) = \mathcal{A}$ by an endomorphism localized in an arbitrarily chosen region K:

$$\pi = \pi_0 \circ \rho$$

with $\varrho(A) = A$ for $A \in \mathcal{O}(K')$.

There is one class of localized automorphisms which does not lead to a change of sectors: the "inner" automorphisms. Let $\mathcal{A}(K)$ be the set of all unitary elements of $\mathcal{O}(K)$, then

$$\sigma_U(A) = UAU^{-1} \quad \text{with } U \in \mathcal{A}(K);\ A \in \mathcal{O} \qquad (2.5)$$

defines an inner automorphism localized in K. The unitary U is determined by σ_U up to a phase factor. One has the

Lemma 2.1. The representations $\pi_0 \circ \varrho_1$ and $\pi_0 \circ \varrho_2$ are unitarily equivalent if and only if $\varrho_2 = \sigma_U \varrho_1$ where σ_U is an inner automorphism.

Proof: Omitting as explained before the symbol π_0, unitary equivalence means that there exists a unitary $V \in \mathcal{B}(\mathcal{H}_0)$ such that

$$\varrho_2(A) = V \varrho_1(A) V^{-1}.$$

We only have to show that V in fact belongs to $\mathcal{O}$. Since ϱ_2 and ϱ_1 are localized in some finite regions we can choose a K large enough so that it encloses both localization regions. Then

$$\varrho_2(A) = \varrho_1(A) = A \qquad \text{for } A \in \mathcal{O}(K').$$

Hence $V \in \{\mathcal{O}(K')\}' = \mathcal{O}(K)$, (by duality)

Next we note

Lemma 2.2 If ϱ_1 and ϱ_2 are endomorphisms localized respectively in the space-like separated double cones K_1 and K_2 then

$$\varrho_1\varrho_2 = \varrho_2\varrho_1 \qquad (2.6)$$

Proof: Consider the 7 regions drawn in Figure 2 (for simplicity we draw only the base of the double cones)

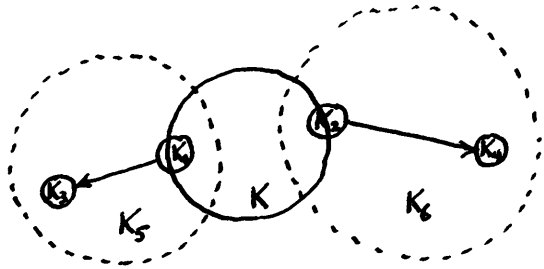

Figure 2.

K_1 and K_2 are the two given, space-like situated regions. K is arbitrary. We want to test the action of $\varrho_1\varrho_2$ on $\mathcal{O}(K)$. We shift K_2 by translation to K_4 and K_1 by opposite translation to K_3 till K_4 and K_3 are space-like to K . In addition the cone K_5 enclosing K_1 and K_3 will be space-like to K_6 , the cone enclosing K_2 and K_4 . Let ϱ_3 and ϱ_4

denote the endomorphisms localized respectively in K_3, K_4 which result by shifting ρ_1, ρ_2 by the relevant translations. Due to our covariance assumption the representation $\pi_0 \circ \rho_1$ is equivalent to $\pi_0 \circ \rho_3$ and $\pi_0 \circ \rho_2$ to $\pi_0 \circ \rho_4$. Hence by Lemma 2.1 we have

$$\rho_3 = \sigma_{U_5} \rho_1 \quad ; \quad \rho_4 = \sigma_{U_6} \rho_2 \; ;$$

where σ_{U_5}, σ_{U_6} are inner automorphisms implemented by unitary operators U_5, U_6 which (by duality) belong to $\mathcal{O}(K_5)$ respectively $\mathcal{O}(K_6)$. Since K_3 and K_4 are space-like to K we have

$$\rho_3 \rho_4 (A) = A = \rho_4 \rho_3 (A) \qquad \text{for } A \in \mathcal{O}(K).$$

Thus

$$\sigma_{U_5} \rho_1 \sigma_{U_6} \rho_2 (A) = \sigma_{U_6} \rho_2 \sigma_{U_5} \rho_1 (A) .$$

The right hand side is rewritten as

$$\sigma_{U_6} \rho_2 \left(U_5 \rho_1 (A) U_5^{-1} \right) = U_6 \rho_2 (U_5) \rho_2 \rho_1 (A) \rho_2 (U_5^{-1}) U_6^{-1}$$

$$= U_6 U_5 \rho_2 \rho_1 (A) U_5^{-1} U_6^{-1}$$

since $\rho_2(U_5) = U_5$ due to the support properties.

By the same manipulations the left hand side becomes

$$U_5 U_6 \rho_1\rho_2(A)\, U_6^{-1} U_5^{-1} \quad .$$

But U_5 and U_6 commute too. Hence we have (2.6).

We shall now specialize the discussion temporarily to the case of those sectors which can be reached from the vacuum sector by localized automorphisms. This allows a simpler analysis than the more general endomorphic case. We shall call such sectors therefore "simple sectors." We find then first of all a classification of simple sectors into Bose-type and Fermi type according to the following lemma:

Lemma 2.3 Let γ_1 and γ_2 be two automorphisms, leading to the same sector and based on space-like separated double cones K_1, K_2. By lemma (2.1) they are related by an inner automorphism, i.e. $\gamma_2 = \sigma_U \gamma_1$. One has then

$$\gamma_1(U) = \pm U \qquad (2.7)$$

where the sign depends only on the sector, not on the choice of γ_1, γ_2.

Sectors for which the + sign holds are called Bose-type; those for which the - sign holds Fermi type.

Proof: Since γ_1 and γ_2 commute, so do γ_1 and σ_U. But

$$\gamma_1 \sigma_U(A) = \gamma_1(UAU^{-1}) = \gamma_1(U)\,\gamma_1(A)\,\gamma_1(U^{-1}) = \sigma_{\gamma_1(U)}\,\gamma_1(A).$$

Thus $\gamma_1 \sigma_U = \sigma_{\gamma_1(U)} \gamma_1$. By commutativity

$$\sigma_U = \sigma_{\gamma_1(U)} \quad \text{or} \quad \gamma_1(U) = \epsilon_{\gamma_1,\gamma_2} U \qquad (2.8)$$

where ϵ is a numerical phase factor depending possibly on γ_1 and γ_2 . Interchanging the role of γ_1 and γ_2 and correspondingly replacing U by U^{-1}, we have

$$\gamma_2(U^{-1}) = \epsilon_{\gamma_2,\gamma_1} U^{-1} \; . \qquad (2.9)$$

Multiplying (2.8) and (2.9) we get

$$\gamma_1(U)\gamma_2(U^{-1}) = \gamma_1\left(U \gamma_1^{-1}\gamma_2(U^{-1})\right) = \epsilon_{\gamma_1,\gamma_2}\epsilon_{\gamma_2,\gamma_1} . \qquad (2.10)$$

By the definition of σ_U and commutativity we have

$$\gamma_1^{-1}\gamma_2 = \sigma_U .$$

Thus the left hand side of (2.10) is 1 and we have

$$\epsilon_{\gamma_1,\gamma_2} = \epsilon_{\gamma_2,\gamma_1}^{-1} . \qquad (2.11)$$

Let us take now a third region K_3 , space-like to both K_1 and K_2 , and choose an automorphism γ_3 localized in K_3 and leading to the same sector. Then

$$\gamma_3 = \sigma_W \gamma_2 = \sigma_V \gamma_1 \; ; \quad \gamma_2 = \sigma_U \gamma_1$$

with $W \in \mathcal{O}(K_2 \cup K_3)$; $V \in \mathcal{O}(K_1 \cup K_3)$; $U \in \mathcal{O}(K_1 \cup K_2)$.

Also $\sigma_W \sigma_U = \sigma_V$ and hence we can choose $V = WU$.

According to the earlier discussion we have

$$\gamma_1(V) = \epsilon_{\gamma_1,\gamma_3} V$$

But

$$\gamma_1(V) = \gamma_1(WU) = \gamma_1(W)\gamma_1(U) = W\gamma_1(U) = \epsilon_{\gamma_1,\gamma_2} WU = \epsilon_{\gamma_1,\gamma_2} V$$

due to the support properties. Hence

$$\epsilon_{\gamma_1,\gamma_3} = \epsilon_{\gamma_1,\gamma_2}$$

Repeating this process we see that $\epsilon_{\gamma_1,\gamma_2}$ is independent of its second argument and therefore by (2.11) also of its first argument. It is a fixed number depending only on the sector. Due to (2.11) its value can only be ± 1.

To see that the sign appearing in (2.7) has something to do with Bose - or Fermi - statistics we sketch the relation of this description with the conventional formulation by means of a field algebra. First it is evident that the set Γ of localized automorphisms forms a group, since $\gamma_1\gamma_2$ and γ^{-1} can be performed within Γ .*

The localized inner automorphisms form an invariant subgroup N since

$$\gamma\sigma_U\gamma^{-1} = \sigma_{\gamma(U)} \ .$$

---

* Actually we should restrict attention to those automorphisms which lead to sectors in which the Poincaré group is implementable. This subset of Γ has been called Γ_c in Ref. [24]. But it is shown in [24] that Γ_c is a group too and possibly we have $\Gamma_c = \Gamma$ so that we do not bother here to make the distinction.

The quotient group $\hat{\mathcal{G}} = \Gamma/\mathcal{J}$ (2.12)

corresponds by lemma (2.1) as a set to the collection of all simple sectors. We see that this set has a group structure and moreover this group is Abelian because if γ_1 and γ_2 are in Γ we can always find a γ_2' in the same equivalence class as γ_2 which has its support space-like to that of γ_1 (just shifting it by a sufficiently large translation). Then γ_1 and γ_2' commute due to lemma 2.2. Using an additive notation for the multiplication in the Abelian group $\hat{\mathcal{G}}$ we see that the sectors can be labeled by generalized charge quantum numbers σ such that along with σ_1 and σ_2 also $\sigma_1 + \sigma_2$ and $-\sigma_1$ occur. Of course it is not implied that all such linear combinations belong actually to different sectors; e.g. we might have a relation like $-\sigma = \sigma$ in which case the chain would really consist only of two different sectors (a case usually described by a multiplicative quantum number). Still we have the typical chains: if the simple sector σ contains a single particle state, then the sector $n\sigma$ (with $n > 1$) contains the states with n such particles and $-\sigma$ contains the antiparticle.

Let us now turn to the construction of a field algebra for the set of simple sectors. I shall do this here only for the case of a single additive charge; i.e. the case where $\hat{\mathcal{G}}$ is the (additive) group of all positive and negative integers. For the case of a general $\hat{\mathcal{G}}$ I refer to [24]. The sectors are now labeled by an integer n running from $-\infty$ to ∞. We wish to consider all sectors simultaneously and take therefore the direct sum of the representations π_n. This "universal" representation π acts in the Hilbert space

$$\mathcal{H} = \sum_{\oplus} \mathcal{H}_n \quad . \tag{2.13}$$

In $\mathcal{H}$ we wish to implement the automorphisms from Γ by unitary operators ψ so that

$$\pi\left(\gamma_\psi(A)\right) = \psi\, \pi(A)\, \psi^{-1} . \tag{2.14}$$

If Ψ is a general vector in $\mathcal{H}$ we denote its projection on $\mathcal{H}_n$ by $\Psi(n)$. As before we may consider all $\mathcal{H}_n$ to be copies of $\mathcal{H}_0$ and regard $\Psi(n)$ as a vector of $\mathcal{H}_0$; then $\Psi \in \mathcal{H}$ is described by the set of vectors $\{\Psi(n)\}$ of $\mathcal{H}_0$. To describe the representation $\pi(\mathcal{A})$ we pick an arbitrary reference automorphism γ_1 localized say in the region K_0 transferring unit charge so that we can reach all sectors by applying powers γ_1^n on the vacuum representation. Then π is taken to be

$$\left(\pi(A)\Psi\right)(n) = \pi_n(A)\Psi(n) = \gamma_1^n(A)\Psi(n). \quad (2.15)$$

To understand the notation keep in mind that $\pi_n(A) = \gamma_1^n(A)$ is regarded as an operator in $\mathcal{H}_0$ and each $\Psi(n)$ as a vector in $\mathcal{H}_0$. We implement the automorphism γ_1 by a unitary operator V_1 acting in $\mathcal{H}$ and defined by

$$\left(V_1\Psi\right)(n) = \Psi(n+1) . \quad (2.16)$$

Correspondingly γ_1^n is implemented by the unitary V_1^n. One checks that this definition indeed gives

$$\pi\left(\gamma_1^m(A)\right) = V_1^m \pi(A) V_1^{-m}$$

because according to (2.16) and (2.15) we get

$$\left(V_1^m \pi(A) V_1^{-m}\Psi\right)(n) = \left(\pi(A)V_1^{-m}\Psi\right)(n+m) =$$

$$= \gamma_1^{n+m}(A)\Psi(n) = \left(\pi\left(\gamma_1^m(A)\right)\Psi\right)(n) .$$

An arbitrary $\gamma \in \Gamma$ **is in some equivalence class modulo** $\mathcal{J}$, **which class contains also one of the powers** γ_1^m. **Hence any** γ **can be written as**

$$\gamma = \sigma_U \gamma_1^m \quad (2.17)$$

and implemented by

$$\Psi = \pi(U)V_1^m . \quad (2.18)$$

The choice of the unitary Ψ **implementing** γ **is again fixed up to the arbitrary phase in the choice of** U **(in the determination of** U **from** σ_U **).**

Consider now the commutation relations between two such operators ψ' and ψ'' which are based on space-like separated regions and lead to the same sector (the corresponding automorphisms γ', γ'' being equivalent modulo $\mathcal{J}$ to the same power of γ_1, say to γ_1^m. Then we have $\gamma'' = \sigma_W \gamma'$ where W now can be chosen to be $W = \psi''\psi'^{-1}$. Lemma (2.3) gives

$$\gamma'(W) = \epsilon_m W \quad ; \quad \epsilon_m = \pm 1 .$$

Thus

$$\psi'\psi''\psi'^{-1}\psi'^{-1} = \epsilon_m \psi''\psi'^{-1}$$

or

$$\psi'\psi'' = \epsilon_m \psi''\psi' \quad ; \qquad (2.19)$$

also, one immediately sees that

$$\epsilon_m = \epsilon_1^m . \qquad (2.20)$$

Thus the commutativity of space-like based automorphisms leads to the alternative between commutativity or anticommutativity for the corresponding implementing operators. If we use these operators in $\mathcal{H}$ to generate states $\psi'\Omega$, $\psi''\Omega$ then the state $\psi'\psi''\Omega$ may be interpreted as the "product state": we have the partial state equal to $\psi'\Omega$ in the one region and the partial state equal to $\psi''\Omega$ in the other region and Ω in the space-like complements of both regions. The sign ϵ_m determines whether linear combinations of such product state vectors behave like vectors

in the symmetrized or antisymmetrized direct product space of the starting sectors. In other words it determines whether localized states in the sector $\psi'\Omega$ are of the Bose type or of Fermi type.

We still have to define the field algebra $\mathcal{F}(K)$ associated with the region K, the gauge group $\mathcal{G}$ and its representation $\mathcal{U}(\mathcal{G})$ in $\mathcal{H}$ and show

Lemma 2.4 The gauge invariant part of the field algebra of a region K is precisely the observable algebra of the region; i.e.

$$\pi\left(\mathcal{A}(K)\right) = \mathcal{F}(K) \cap \mathcal{U}(\mathcal{G})' \qquad (2.21)$$

The field algebra $\mathcal{F}(K)$ of the region K is of course defined as the von Neumann ring generated by all the ψ implementing automorphisms γ which are localized in K. The gauge group, abstractly, is defined as the dual group of $\hat{\mathcal{G}}$. Its elements are the characters of $\hat{\mathcal{G}}$; i.e. functions from $\hat{\mathcal{G}}$ to the complex numbers which furnish a one-dimensional representation of $\hat{\mathcal{G}}$. In our example, where $\hat{\mathcal{G}}$ is the additive group of integers, $\mathcal{G}$ is the group of the unit circle in the complex plane. Writing an element $g \in \mathcal{G}$ as $e^{i\delta}$ we have the character $g(n) = e^{in\delta}$. The unitary representative $\mathcal{U}(g)$ in the Hilbert space $\mathcal{H}$ is obviously given by

$$\left(\mathcal{U}(g)\,\Psi\right)(n) = g(n)\,\Psi(n)\,. \qquad (2.22)$$

For the proof of lemma (2.4) and the discussion of the general case of simple sectors (when the "charge group" $\hat{\mathcal{G}}$ is generated by several, possibly not independent, elements) see ref. [24]. If there are several charges then the construction of the field algebra is not unique, the commutation properties of operators transferring different types of charge are not intrinsically fixed. One finds, however, that one can always achieve the "normal" commutation relations described in (1.17), (1.18), (1.19).

Non simple sectors. Endomorphic case.

We have seen that for simple sectors one always has the Bose-Fermi alternative and an Abelian gauge group. A more complicated structure results if $\mathcal{A}$ admits localized endomorphisms (for which the image $\varrho(\mathcal{A})$ is strictly smaller than $\mathcal{A}$). We denote the set of localized endomorphisms (including the automorphisms Γ) by Δ (respectively $\Delta(K)$ if the localization region is specified to be K). To avoid misunderstandings I should perhaps repeat that the term "endomorphism" is used here for a one-to-one mapping of $\mathcal{A}$ into itself, conserving all the C^*-algebraic structure. Perhaps a better term might be "isomorphic injection."

We begin with some remarks on intertwining operators. Given ρ , $\rho' \in \Delta$ then the operator $S \in \mathcal{B}(\mathcal{H}_0)$ is called an intertwining operator from ρ to ρ' if

$$\rho(A)\, S = S\, \rho'(A) \quad \text{for all } A \in \mathcal{A}. \tag{2.23}$$

Such intertwining operators exist if the representations $\rho(\mathcal{A})$ and $\rho'(\mathcal{A})$ are not disjoint. Actually in that case S belongs to $\mathcal{A}(K)$ where K is the union of the support regions of ρ and ρ' Let us write

$$\tilde{S} = (\rho\,;\, S\,;\, \rho') \tag{2.24}$$

If we have two such triples $\tilde{S}_1 = (\rho_1; S_1; \rho_1')$ and $\tilde{S}_2 = (\rho_2\,;\, S_2\,;\, \rho_2')$ we can immediately construct an intertwining operator from $\rho_1 \rho_2$ to $\rho_1' \rho_2'$. One checks that the operator $\rho_1(S_2)\, S_1 = S_1\, \rho_1'(S_2)$ performs this function. Thus one has a cross product of the triples, defined by

$$\tilde{S}_1 \times \tilde{S}_2 = \left(\rho_1 \rho_2 \,;\, \rho_1(S_2) S_1 \,;\, \rho_1' \rho_2' \right). \tag{2.25}$$

This cross product is associative but in general not commutative

$$\tilde{S}_1 \times (\tilde{S}_2 \times \tilde{S}_3) = (\tilde{S}_1 \times \tilde{S}_2) \times \tilde{S}_3 = \tilde{S}_1 \times \tilde{S}_2 \times \tilde{S}_3. \tag{2.26}$$

If all the ρ_i' have their supports mutually spacelike and if the same holds for the ρ_i then the cross product (2.26) is commutative. The proof is analogous to that of lemma 2.2.

Consider now n endomorphisms $\varrho_1, \varrho_2, \cdots, \varrho_n$ and a permutation P of the numbers $1, 2, \cdots, n$. The products $\varrho_1 \varrho_2 \cdots \varrho_n$ and $\varrho_{P(1)} \varrho_{P(2)} \cdots \varrho_{P(n)}$ lead to the same sector and hence there is a unitary intertwining operator $\epsilon_P(\varrho_1, \cdots, \varrho_n)$ from $\varrho_1 \cdots \varrho_n$ to $\varrho_{P(1)} \cdots \varrho_{P(n)}$:

$$\varrho_1 \cdots \varrho_n(A)\, \epsilon_P(\varrho_1, \cdots, \varrho_n) = \epsilon_P(\varrho_1, \cdots, \varrho_n)\, \varrho_{P(1)} \cdots \varrho_{P(n)}(A). \tag{2.27}$$

The relation (2.27) does not yet define $\epsilon_P(\varrho_1, \cdots, \varrho_n)$ uniquely but we can find a unique and natural determination of the ϵ_P by comparing them for different sets of arguments. Let $\varrho_1', \cdots, \varrho_n'$ be any other set of n endomorphisms with ϱ_i' not disjoint from ϱ_i so that we have intertwining operators $\tilde{R}_i = (\varrho_i; R_i; \varrho_i')$. Then $\tilde{R}_1 \times \cdots \times \tilde{R}_n$ intertwines from $\varrho_1 \cdots \varrho_n$ to $\varrho_1' \cdots \varrho_n'$; $\epsilon_P(\varrho_1', \cdots, \varrho_n')$ from $\varrho_1' \cdots \varrho_n'$ to $\varrho'_{P(1)} \cdots \varrho'_{P(n)}$. Denoting for brevity by $R_1 \times \cdots \times R_n$ the operator which is the middle piece of the triple $\tilde{R}_1 \times \cdots \times \tilde{R}_n$ the operators $\epsilon_P(\varrho_1, \cdots, \varrho_n) R_{P(1)} \times \cdots \times R_{P(n)}$ and $R_1 \times \cdots \times R_{n-1} \times R_n \epsilon_P(\varrho_1', \cdots, \varrho_n')$ both intertwine between $\varrho_1 \cdots \varrho_n$ and $\varrho'_{P(1)} \cdots \varrho'_{P(n)}$. One finds

Lemma 2.5 It is possible to choose the operators $\epsilon_P(\varrho_1, \cdots, \varrho_n)$ in such a way that

(i) $\epsilon_P(\varrho_1, \cdots, \varrho_n) = 1$ whenever all the supports of the

ρ_i are mutually space-like.

(ii) For any other set ρ_i' and any choice of intertwining operators R_i from ρ_i to ρ_i' we have

$$R_1 \times \cdots \times R_n \, \epsilon_\rho(\rho_1', \cdots, \rho_n') = \epsilon_\rho(\rho_1, \cdots, \rho_n) R_{\rho(1)} \times \cdots \times R_{\rho(n)}. \tag{2.28}$$

The unitary operators $\epsilon_\rho(\rho_1, \cdots, \rho_n)$ are uniquely determined by these two requirements and they satisfy (2.27).

The proof of this lemma proceeds by a straightforward computation. An immediate consequence of the lemma (in particular of the uniqueness of ϵ_ρ) is the multiplication law

$$\epsilon_\rho(\rho_1, \cdots, \rho_n)\, \epsilon_\sigma(\rho_{\rho(1)}, \cdots, \rho_{\rho(n)}) = \epsilon_{\rho\sigma}(\rho_1, \cdots, \rho_n). \tag{2.29}$$

If we put all ρ_i equal, a more convenient notation is

$$\epsilon_\rho(\rho, \rho, \cdots, \rho) \equiv \epsilon_\rho^{(n)}(P). \tag{2.30}$$

One then has by (2.29)

$$\epsilon_\rho^{(n)}(P)\, \epsilon_\rho^{(n)}(Q) = \epsilon_\rho^{(n)}(PQ), \tag{2.31}$$

i.e. the $\epsilon_\rho^{(n)}$ form a unitary representation of the permutation group of n elements which is (up to unitary equivalence) characteristic of the sector ρ^n . Note that $\epsilon_\rho^{(n)}(P)$ intertwines ρ^n with itself. Hence one might be inclined to think that $\epsilon_\rho^{(n)}(P)$ should be trivial. This is however not so because the $\epsilon_\rho(\rho_1, \cdots, \rho_n)$ are defined to be 1 not for equal ρ_i but for space-like separation of the supports of the ρ_i and it turns out that the representation

$\epsilon_\rho^{(n)}$ is one-dimensional ($\epsilon_\rho^{(n)}(P) = \pm 1$) if and only if ρ is an automorphism. It is an instructive exercise to check that in this case, putting $n = 2$ and choosing P as the transposition $1 \leftrightarrow 2$ $\epsilon_\gamma(P)$ reduces to the sign factor in (2.7)

The representation $\epsilon_\rho^{(n)}$ may be characterized in terms of the Young tableaux it contains and we may compare the set of Young tableaux occurring for fixed ρ and varying n . This analysis is most satisfactory if the sector ρ has an adjoint sector ρ^* (related to the existence of anti-particles). Then one finds that associated with ρ there is a number λ , depending only on the sector generated by ρ, not on ρ itself. $\lambda^{-1} = p_\rho$ is a positive or negative integer. For positive p_ρ all possible Young tableaux appear in the sectors ρ^n for which the number of rows does not exceed p_ρ . For negative p_ρ one has all Young tableaux whose number of columns is limited by $|p_\rho|$. The first case is familiar as the para-Bose case of order p_ρ the second as the para-Fermi case of order $|p_\rho|$.

Thus we conclude that a particle which has an anti-particle and belongs to a non-simple sector is a paraboson or parafermion of some fixed order.

III. Parastatistics.

Field theoretic models for the description of strange statistics (i.e. statistics which are neither of the Bose- or Fermi-type) have been given by H. S. Green [26] and further discussed by numerous authors. See e.g. the clear survey by Greenberg [27]. Here we shall analyse in terms of the concepts developed in section I the physical content of the simplest example of such a model, the case of a parafermi field of order 2. I shall sketch only the essential line of argument referring for details of proofs to [28]. The discussion can be carried through using only the field quantities at one time, say $t = 0$. Lorentz invariance plays no role. The model itself may be described as follows. Take two Fermi fields $\psi^{(1)}$ and $\psi^{(2)}$ commuting with each other. The commutation relations are

$$\left[\psi^{(i)*}_{\underset{\sim}{}}(\underset{\sim}{x}), \psi^{(i)}(\underset{\sim}{y})\right]_+ = \delta^3(\underset{\sim}{x}-\underset{\sim}{y}); \quad \left[\psi^{(i)}(\underset{\sim}{x}), \psi^{(i)}(\underset{\sim}{y})\right]_+ = 0; \quad i=1,2 \tag{3.1}$$

$$\left[\psi^{(i)}(\underset{\sim}{x}), \psi^{(j)}(\underset{\sim}{y})\right]_- = \left[\psi^{(i)*}_{\underset{\sim}{}}(\underset{\sim}{x}), \psi^{(j)}(\underset{\sim}{y})\right]_- = 0 \quad \text{for } i \neq j. \tag{3.2}$$

Then define the "parafield" ψ as the sum of these two:

$$\psi(\underset{\sim}{x}) = \psi^{(1)}(\underset{\sim}{x}) + \psi^{(2)}(\underset{\sim}{x}) \tag{3.3}$$

and demand that only such quantities which are expressible in terms of this parafield ψ should occur in the theory.

Before we can analyse the physical content we have to know what the observable algebras corresponding to various space regions are. There are several possibilities but the choice is limited by the following two restrictions:

i) The observable algebra $\mathcal{A}(V)$ of the space region V shall be a subalgebra of the parafield algebra $\mathcal{F}_p(V)$ of the same region. The latter is the *-algebra generated by the $\psi(f)$ for all test functions f with support in V .

ii) If V_1 and V_2 are disjoint, then $\mathcal{A}(V_1)$ and $\mathcal{A}(V_2)$ shall commute (locality).

Of course the net of algebras $\mathcal{A}(V)$ shall be covariant with respect to translations and the total algebra $\mathcal{A}$ is defined as

$$\mathcal{A} = \overline{\cup \mathcal{A}(V)} .$$

The largest algebra satisfying these requirements is the algebra $\mathcal{A}_0$, which consists precisely of all <u>even</u> elements of $\mathcal{F}_p$ (elements invariant under the substitution $\psi \to -\psi$). So $\mathcal{A}$ must be contained in (or possibly be equal to) $\mathcal{A}_0$. Now one remarks that $\mathcal{A}_0$ can also be characterized in terms of the algebra generated by two Fermi fields, $\phi^{(1)}$, $\phi^{(2)}$ with normal commutation relations:

$$\left[\phi^{(i)}(\underset{\sim}{x}), \phi^{(j)}(y)\right]_+ = 0 , \quad \left[\phi^{(i)}(\underset{\sim}{x}), \phi^{(j)*}(y)\right]_+ = \delta_{ij}\,\delta^3(\underset{\sim}{x}-y) . \quad (3.4)$$

To see this we first embed $\mathcal{F}_P$ in a large algebra $\tilde{\mathcal{F}}$ generated by the "Green components" $\psi^{(1)}$, $\psi^{(2)}$ and one additional element K_2 (**standing** for 'Klein transformation"). The relations of the $\psi^{(i)}$ are given in (3.1), (3.2). The relations involving K_2 are

$$K_2^{-1} = K_2 = K_2^{*} \; ; \tag{3.5}$$

$$\left[K_2 , \psi^{(1)}(\underset{\sim}{x})\right]_- = 0 \; ; \quad \left[K_2 , \psi^{(2)}(\underset{\sim}{x})\right]_+ = 0 .$$

It is readily seen that with the definition

$$\psi^{(1)}(\underset{\sim}{x}) = \phi^{(1)}(\underset{\sim}{x}) K_2 \; ; \quad \psi^{(1)*}(\underset{\sim}{x}) = \phi^{(1)*}(\underset{\sim}{x}) K_2 \; ; \tag{3.6}$$

$$\psi^{(2)}(\underset{\sim}{x}) \; ; \; \phi^{(2)}(\underset{\sim}{x}) K_2 \; ; \quad \psi^{(2)*}(\underset{\sim}{x}) \; ; \; \phi^{(2)*}(\underset{\sim}{x}) K_2$$

we obtain the normal commutation relations (3.4). The general element of $\tilde{\mathcal{F}}$ is of the form $F + F' K_2$ where F and F' are expressible in terms of the $\psi^{(i)}$ For the elements of $\mathcal{O}_0$ F' is zero and F even. Rewriting such an element in terms of the $\phi^{(i)}$ and K_2 one sees that K_2 drops out because in each monomial the factors K_2 can be shifted to the right using (3.5) and an even power of K_2 is the identity.

Let us denote the algebra generated by the two normal Fermi fields $\phi^{(i)}(\underset{\sim}{x})$ (with $\underset{\sim}{x}$ ranging in V) by $\mathcal{F}(V)$ and consider substitutions of the form

$$\phi^{(i)}(\underset{\sim}{x}) \longrightarrow \sum_{\kappa} g_{\kappa i}\, \phi^{(\kappa)}(\underset{\sim}{x}) \,. \qquad (3.7)$$

Such a substitution generates an automorphism of $\mathcal{F}$ denoted by α_g as long as the 2 x 2 matrix g is unitary. One finds

Lemma 3.1 $\mathcal{A}_0(V)$ is the subalgebra of $\mathcal{F}(V)$ consisting of precisely those elements which are invariant under the automorphisms α_g when g runs through SO(2) (real, orthogonal matrices with determinant +1).

In other words: $\mathcal{A}_0$ is the "gauge invariant" part of $\mathcal{F}$ when we take SO(2) as the gauge group (acting on $\mathcal{F}$ according to (3.7)).

Since the observable algebra $\mathcal{A}$ must be contained in $\mathcal{A}_0$ other possibilities for $\mathcal{A}$ result if we take a larger gauge group $\mathcal{G}$ and again define $\mathcal{A}$ as the gauge invariant part with respect to this group. We shall just consider one such example, the case where the gauge group is the largest automorphism group of the form (3.7), namely the group U(2). The resulting observable algebra will be denoted by $\mathcal{A}_2$.

Summing up: instead of expressing the observables in terms of the parafield we can express them also as functions

of two normal Fermi fields. The characterization of the observable algebra within this Fermi field algebra $\mathcal{F}$ is best done by specifying the "gauge group" $\mathcal{G}$ under which the observables are invariant. In our case (starting from a para fermi field of order 2) the minimal gauge group is SO(2), leading to the maximal observable algebra $\mathcal{A}_0$.

Given $\mathcal{F}$, $\mathcal{G}$ and $\mathcal{A}$ we have the following structure [23]: The different superselection sectors (families of states which are of interest) are in one-to-one correspondence with the "spectrum" of the gauge group.* If $\mathcal{G}$ is Abelian we have only Bose- or Fermi statistics. Let us consider from this point of view the two examples mentioned above and compare the conclusions with the parafield description.

Example 1. $\mathcal{G} = SO(2)$; $\mathcal{A} = \mathcal{A}_0$.

The gauge group is Abelian. Its spectrum consists of the integers $n = 0, \pm 1, \pm 2, \cdots$. We thus have just one ordinary charge quantum number, distinguishing the superselection sectors. States of charge $n = +1$ are obtained from the vacuum state vector Ω by applying $\phi^{(1)}(f) + i\,\phi^{(2)}(f)$. Note that $\phi^{(1)*}(f) + i\,\phi^{(2)*}(f)$ leads to the same sector. The sector $n = -1$ is reached from Ω by $\phi^{(1)} - i\,\phi^{(2)}$ or, equally well, by $\phi^{(1)*} - i\,\phi^{(2)*}$.

* The "spectrum of $\mathcal{G}$" consists of the equivalence classes of irreducible representations of $\mathcal{G}$.

The particles with $n = \pm 1$ are ordinary Fermions. How does that fit with the parafield description? We have, according to (3.6)

$$\psi(f) = (\phi^{(1)}(f) + i\phi^{(2)}(f))K_2 \qquad (3.8)$$

The comparison between the two descriptions is simplest if $K_2\Omega = \Omega$. Then, according to (3.8), $\psi(f)\Omega$ is a state with charge $n = +1$. Suppose g and f are test functions with far separated supports. Then

$$\psi(g)\psi(f)\Omega = (\phi^{(1)}(g) + i\phi^{(2)}(g))K_2(\phi^{(1)}(f) + i\phi^{(2)}(f))K_2\Omega =$$
$$= (\phi^{(1)}(g) + i\phi^{(2)}(g))(\phi^{(1)}(f) - i\phi^{(2)}(f))\Omega$$

is a state with $n = 0$. Any polynomial of ψ applied to Ω will produce only states with $n = +1$ and $n = 0$ because in a product ψ will alternatingly raise or lower the charge, depending on its position. The fact that there are both symmetric and antisymmetric wave functions allowed for the states generated from the vacuum by two parafield operators has (in this example) nothing to do with parastatistics but results from the fact that the two operators produce different particles (the first a negatively charged, the second a positively charged one). The effect of ψ depends on the position it has within a product.

Example 2. $\mathcal{G} = U(2)$; $\mathcal{H} = \mathcal{H}_2$.

We may first note that $\mathcal{H}_2$ is generated by the bilocal densities

$$\varrho(\underset{\sim}{x},\underset{\sim}{y}) = \phi^{(1)*}(\underset{\sim}{x})\,\phi^{(1)}(\underset{\sim}{y}) + \phi^{(2)*}(\underset{\sim}{x})\,\phi^{(2)}(\underset{\sim}{y}) \qquad (3.9)$$

The irreducible representations of U(2) may be labeled by two quantum numbers ($\mathcal{B}$, $\mathcal{I}$) with the relation

$$\mathcal{B} + 2\mathcal{I} \quad = \text{even.} \qquad (3.10)$$

$B = 0, \pm 1, \pm 2$ may be conveniently interpreted as baryon number, $\mathcal{I} = 0, \frac{1}{2}, 1, \cdots$ as isospin,

thinking of the theory of nuclei with strict charge independence. To each allowed pair ($\mathcal{B}$, $\mathcal{I}$) we have a sector. In a sector with isospin I each state appears with a multiplicity ($2\mathcal{I}+1$). Any vector in the ($2\mathcal{I}+1$)-dimensional subspace spanned by an isospin multiplet gives exactly the same expectation values over the observable algebra and corresponds therefore to the same physical state as any other vector in this subspace. With the conventional choice of the three components of isospin (the Pauli matrices σ_1 and σ_3 real, σ_2 purely imaginary) one finds that the subgroup SO(2) is placed within U(2) in such a way that the charge quantum number n is related to the second component of the isospin

$$\mathcal{I}_2 = 2n \,. \qquad (3.11)$$

Applying the parafield algebra on the vacuum we get therefore only states with $\mathcal{I}_2 = 0$ or $\mathcal{I}_2 = \frac{1}{2}$. The former appear if we have an even number of Ψ-factors, which leads to

an even baryon number; the latter appear for odd B. According to (3.10) even B implies integer I and then we can find in each multiplet a vector with $I_2 = 0$. For odd B, half integer I, we have in each multiplet a state with $I_2 = \frac{1}{2}$. Hence the restriction to $n = 0, 1$ does not limit the selection of states. Applying the parafield algebra to Ω we obtain all relevant states over $\mathcal{A}$, only the multiplicity is changed as compared to the representation space of the field algebra $\mathcal{F}$. We obtain with $\mathcal{F}_p$ each state only once instead of the $(2I+1)$-dimensional multiplets. In this second example the parafield model gives a complete description. The parastatistics of particles corresponds then to the fact that there is one hidden parameter (the charge I_2) which is not observable. We see that in this case the parastatistics may be reduced to ordinary (Bose-Fermi) statistics if one introduces this hidden degree of freedom as an additional distinctive quantum number. It appears that all reasonable parafield models can be reduced in this way to the Bose-Fermi case [28] although a general theorem to this effect based on the structure analysis described in section II has not yet been obtained.

Mathematical Appendix

The set of all bounded operators acting in a Hilbert space $\mathcal{H}$ is denoted by $\mathcal{B}(\mathcal{H})$. There are several topologies in this set $\mathcal{B}(\mathcal{H})$ which are important in our context. A topology means that we define what is a "neighborhood" of an element in the set. Actually it is sufficient here to define the neighborhoods of the origin. The three most important topologies in $\mathcal{B}(\mathcal{H})$ arising in the context of physics are:

a) Uniform topology. Since an operator $A \in \mathcal{B}(\mathcal{H})$ has a finite norm, $\|A\|$, we may define a neighborhood of the origin as the set of all $A \in \mathcal{B}(\mathcal{H})$ which have a norm less than ϵ. Any $\epsilon > 0$ gives us one such neighborhood. Convergence of a sequence $A_n \in \mathcal{B}(\mathcal{H})$ in this topology means that the $\|A_n - A_m\| < \epsilon_N$ for all $n, m > N$ and $\epsilon_N \to 0$. Such a sequence is called "uniformly convergent" or a Cauchy sequence in the norm topology.

b) Strong topology. We pick an arbitrary vector $\Psi \in \mathcal{H}$ and an arbitrary number $\epsilon > 0$. The corresponding neighborhood of the origin, denoted by $\mathcal{N}(\Psi, \epsilon)$ consists of all $A \in \mathcal{B}(\mathcal{H})$ satisfying

$$\|A\Psi\| < \epsilon$$

where now the left hand side is the length of the vector $A\Psi$ (not the norm of the operator A). Convergence of a sequence A_n in this topology ("strong convergence") means that for every vector Ψ the sequence of image vectors $A_n\Psi$ satisfies

$$\| A_n\Psi - A_m\Psi \| \leq \epsilon_N \text{ for } n,m > N; \epsilon_N \to 0.$$

c) Weak topology. We pick an arbitrary pair of vectors $\Phi, \Psi \in \mathcal{H}$ and $\epsilon > 0$ and define a corresponding (weak) neighborhood $\mathcal{N}(\Phi,\Psi;\epsilon)$ as the set of all $A \in \mathcal{B}(\mathcal{H})$ satisfying $|(\Phi, A\Psi)| < \epsilon$. Weak convergence of a sequence A_n means the convergence of the sequences of matrix elements $(\Phi, A_n\Psi)$ for all pairs Φ, Ψ.

Example of a strongly convergent sequence which does not converge uniformly. Take a complete orthonormal basis Ψ_n in $\mathcal{H}$ and let E_n by the projector on the subspace spanned by the first n basis vectors. One easily sees that the strong limit of the sequence E_n as $n \to \infty$ is the unit operator (completeness relation). Yet for arbitrarily large n we still have $\| E_n - E_{n+1} \| = 1$, i.e. there is no uniform convergence.

Example of a weakly convergent but not strongly convergent sequence. Take Ψ_n as above and define A_n by

$A_n \Psi_m = \Psi_{n+m}$. This sequence A_n converges weakly to zero as $n \to \infty$ but obviously has no strong limit.

If A and B belong to $\mathcal{B}(\mathcal{H})$ and α, β are complex numbers then

$$\begin{aligned} &\alpha A + \beta B \\ &A \cdot B \\ &A^* \text{ (adjoint of } A) \end{aligned} \qquad \text{(A. 1)}$$

also belong to $\mathcal{B}(\mathcal{H})$. A subset of $\mathcal{B}(\mathcal{H})$ which is closed under the three operations (A. 1) is called a *-algebra of bounded operators. If in addition it contains the unit operator and is closed in the strong topology it is called a "von Neumann ring." If it is closed in the uniform topology it is called a (concrete) C^*-algebra. Since the three topologies listed under a), b) and c) are decreasing in strength we have in general that the weak closure of a set is larger than the strong closure and this again larger than the uniform closure. A weakly closed set is always strongly closed, a strongly closed one is always uniformly closed. Hence any von Neumann ring is also a C^*-algebra but the converse is not true. It turns out, on the other hand, that for a *-algebra of bounded operators the weak and the strong closures coincide. Thus a von Neumann ring is also always weakly closed.

If S is any subset of $\mathfrak{B}(\mathcal{H})$ one defines the "commutant" S' as the set of all operators from $\mathfrak{B}(\mathcal{H})$ which commute with every element of S. Let S^* denote the set of all the adjoint operators of the members of S. One has

a) The commutant of any subset of $\mathfrak{B}(\mathcal{H})$ which is closed under the *-operation is a von Neumann ring, i.e. $\{S \cup S^*\}'$ is a von Neumann ring.

b) If R is a von Neumann ring, then

$$R = R'' \quad (R'' \text{ denotes the commutant of the commutant}).$$

c) $\{S \cup S^*\}''$ is the smallest von Neumann ring containing S.

Let us now consider abstract*-algebras. Such an algebra $\mathfrak{A}$ consists of elements A, B, ... for which the three operations (A.1) are defined, satisfying the usual laws. As in the case of an abstract group the elements now are not regarded as operators in some space but just as objects which can be connected by the operations (A.1). It turns out that under rather general conditions the algebraic structure determines a natural norm $\|A\|$ for the elements of the algebra satisfying

$$\|A\| = \|A^*\| = \|AA^*\|^{1/2}; \quad \|\alpha A\| = |\alpha| \, \|A\| \tag{A.2}$$

and the inequalities

$$\|A+B\| \leq \|A\| + \|B\| \qquad \|AB\| \leq \|A\| \, \|B\| . \tag{A.3}$$

If these conditions prevail, then the *-algebra may be equipped with this natural norm topology and completed in it. It is then called an abstract C^*-algebra.

Given an abstract C^*-algebra $\mathcal{A}$ one considers the "positive linear forms" over it. Denote this set by $\mathcal{A}^{*(+)}$. An element ω from $\mathcal{A}^{*(+)}$ assigns to each $A \in \mathcal{A}$ a complex number $\omega(A)$ subject to the two conditions

(i) linearity $\omega(\alpha A + \beta B) = \alpha\omega(A) + \beta\omega(B)$, (A.4)

(ii) positivity $\omega(A^*A) \geq 0$ (A.5)

The norm of such a form is defined as

$$\|\omega\| = \sup_{\|A\|=1} |\omega(A)| \tag{A.6}$$

One has

$$\|\omega\| = \omega(1) . \tag{A.7}$$

The set of positive linear forms is a convex cone i.e.

$$\omega = \lambda_1\omega_1 + \lambda_2\omega_2 \tag{A.8}$$

belongs to $\mathcal{A}^{*(+)}$ if ω_1 and ω_2 belong and λ_1, λ_2 are positive numbers. A positive linear form ω is called "extremal" if no non-trivial decomposition of the form (A.8) is possible.

A representation of $\mathcal{A}$ (by operators in a Hilbert space $\mathcal{H}$) assigns to each $A \in \mathcal{A}$ an operator $\pi(A)$ from

$\mathcal{B}(\mathcal{H})$ in such a way that the algebraic relations (A. 1) are conserved. In other words it is an isomorphic mapping from the abstract C^*-algebra to a concrete C^*-algebra of operators $\pi(\mathcal{A})$. The representation is called faithful if $\pi(A)=0$ implies $A = 0$. In this case the norm of the operator $\pi(A)$ is the same as the C^*-norm of the abstract element A. We shall only consider faithful representations in these lectures. Given a representation one has immediately a family of positive linear forms over $\mathcal{A}$, which are associated with this representation. Pick any vector Ψ in the representation space, then

$$\omega_\Psi(A) = (\Psi, \pi(A)\Psi) \tag{A.9}$$

is such a form. Also any positive trace class operator ρ in $\mathcal{B}(\mathcal{H})$ defines a positive linear form over the algebra by

$$\omega_\rho(A) = \mathrm{Tr}(\rho\pi(A)) \tag{A.10}$$

Let us denote by $\mathcal{F}_\pi$ the family of forms associated with the representation π according to (A. 10) and by $\mathcal{F}_\pi^V$ the subset of "vectorial" forms (A. 9). The former are the convex combinations of the latter.

The connection between positive linear forms and representations can also be followed in the opposite direction. Given an $\omega \in \mathcal{A}^{*(+)}$ one can construct a representation π_ω so that we have a vector Ψ in the representation space giving

back the form ω by (A. 9). This vector Ψ is moreover a "cyclic vector", i.e. the vectors $\pi_\omega(\mathcal{A})\Psi$ are dense in $\mathcal{H}$. This is achieved by the GNS-construction (the letters standing for Gelfand, Naimark, Segal). It proceeds as follows: First note that the algebra is itself a linear space and that an $\omega \in \mathcal{A}^{*(+)}$ defines a semidefinite scalar product between the elements of $\mathcal{A}$ by

$$(B, A) = \omega(B^* A) .$$

To obtain a positive definite scalar product one has to divide the algebra into equivalence classes modulo the set J which latter consists of all elements $Z \in \mathcal{A}$ for which $\omega(Z^* Z) = 0$. Let us denote the class of $A \in \mathcal{A}$ by $\hat{A}$. This contains all the elements of $\mathcal{A}$ which differ from A by an element of the set J. The set of these classes is a linear space with the positive definite scalar product

$$(\hat{B}, \hat{A}) = \omega(B^* A) , \tag{A. 11}$$

where one checks that the right hand side is independent of the choice of the elements B, A in the respective classes $\hat{B}$, $\hat{A}$. Thus $\mathcal{A}/J$ is a "pre Hilbert space," i.e., it may be considered as a dense set of vectors in a Hilbert space $\mathcal{H}$. We obtain a representation of $\mathcal{A}$ by operators in $\mathcal{H}$ defining the operator $\pi(A)$ representing A by

$$\pi(A)\hat{B} = \widehat{AB} \tag{A. 12}$$

The right hand side involves a choice of B from the class $\hat{B}$ but one checks again that the class of AB does not change when B varies within one class (the set J is a left ideal of $\mathcal{O}$). The cyclic vector Ψ giving back the expectation value ω corresponds to the class of the unit element of $\mathcal{O}$ We have, by (A. 11) and (A. 12)

$$(\Psi, \pi(A)\Psi) = (\hat{1}, \pi(A)\hat{1}) = \omega(1^* A 1) = \omega(A).$$

A representation π is called irreducible if there is no invariant subspace in the representation space $\mathcal{H}$. A criterion for irreducibility is Schur's lemma: π is irreducible if and only if $(\pi(\mathcal{O}))'$ consists only of multiples of the identity. An equivalent criterion is $(\pi(\mathcal{O}))'' = \mathcal{B}(\mathcal{H})$. The GNS-construction leads to an irreducible representation if and only if the form ω from which the construction starts is extremal.

If π is reducible we may consider the restriction of π to one of the invariant subspaces. This is called a subrepresentation of π. One calls two representations disjoint if they contain no subrepresentations which are unitarily equivalent. A representation is called "primary" or a "factor" if

$$(\pi(\mathcal{O}))'' \cap (\pi(\mathcal{O}))' = \{\lambda 1\},$$

i. e. if the von Neumann algebra generated by the representers $\pi(\mathcal{O})$ has no nontrivial center. A form ω is called primary if the GNS-construction starting from it leads to a primary representation.

We frequently use the notion of an automorphism or an endomorphism of a C^*-algebra. In each case we mean a one-to-one mapping carrying $A \in \mathcal{A}$ to $\rho(A) \in \mathcal{A}$ and such that the algebraic structure and the norm are conserved. If the image set $\rho(\mathcal{A})$ is equal to $\mathcal{A}$, ρ is called an automorphism. If $\rho(\mathcal{A})$ is smaller than $\mathcal{A}$ then we call ρ an endomorphism.

References

[1] N. Bohr and L. Rosenfeld, Dan. Mat. Fys. Medd. XII, No. 8 (1933).

[2] E. P. Wigner, Group Theory and its application to the Quantum Mechanics of Atomic Spectra, Academic Press, New York, 1959.

[3] R. F. Streater and A. S. Wightman, "PCT, Spin Statistics and all that," Benjamin, New York, 1964.

[4] B. Schroer, private communication.

[5] H. Epstein, Nuovo Cim. 27, 886 (1963).

[6] R. A. Brandt, Am. Phys. 44, 221 (1967).

[7] W. Zimmermann, Brandeis Lectures 1970 and earlier work quoted therein.

[8] R. T. Powers, Commun. Math. Phys. 4, 145 (1967).

[9] K. B. Sinha, Univ. of Rochester Thesis, 1969.

[10] A. S. Wightman, Phys. Rev. 101, 860 (1956).

[11] R. Haag, Phys. Rev. 112, 669 (1959).

[12] D. Ruelle, Helv. Phys. Acta 35, 147 (1962).

[13] R. Jost, The General Theory of Quantized Fields, American Math. Soc., Providence, 1965.

[14] H. J. Borchers, Nuovo Cim. 15, 784 (1960).

[15] G. C. Wick, E. P. Wigner and A. S. Wightman, Phys. Rev. 88, 101 (1952).

[16] R. Haag and D. Kastler, Journal Math. Phys. 5, 848 (1964).

[17] H. Araki, J. Math. Phys. 5, 1 (1964).

[18] I. E. Segal in "Les Problemes Mathematiques de la Theorie Quantique des Champs," C.N.R.S. 1959.

[19] I. E. Segal, Ann. Math. 48, 930 (1947).

[20] J. M. G. Fell, Trans. Am. Math. Soc. 94, 365 (1960).

[21] H. Roos, Commun. Math. Phys. 16, 238 (1970).

[22] H. J. Borchers, Commun. Math. Phys. 1, 291 (1965).

[23] S. Doplicher, R. Haag, J. E. Roberts, Commun. Math. Phys. 13, 1, (1969).

[24] S. Doplicher, R. Haag, J. E. Roberts, Commun. Math. Phys. 15, 173 (1969).

[25] S. Doplicher, R. Haag, J. E. Roberts, "Locality and Statistics," forthcoming manuscript.

[26] H. S. Green, Phys. Rev. 90, 270 (1953).

[27] O. W. Greenberg, "Parafield Theory" in "Proceedings of the Conference on the Mathematical Theory of Elementary Particles," edited by R. Goodman and I. E. Segal. MIT Press, 1966.

[28] N. Drühl, R. Haag and J. E. Roberts, Comm. Math Phys. 18, 204 (1970).