Title
Nonlinear Operators II

Journal
Journal of Mathematical Physics, 38

Author
Schwartz, Charles

Publication Date
1997-07-01

Peer reviewed

Nonlinear operators. II

Charles Schwartz
Department of Physics, University of California, Berkeley, California 94720

(Received 20 December 1996; accepted for publication 15 April 1997)

This work extends the previous development of new mathematical machinery for nonlinear operators acting on a vector space. Starting from the usual concept of inner product, we find that Hermitian, anti-Hermitian, and unitary nonlinear operators can be defined without bringing in the ideas of a dual vector space or adjoint operators. After looking briefly at how these general ideas might be used in classical mechanics and to extend the linear Schrödinger equation of quantum theory, the topic of Lie groups and Lie algebras is studied. Many, but not all, of the familiar features of that topic are extended to nonlinear operators. New representations are found for a few simple cases of interest to physics, and some provocative implications for elementary particle theory are discussed. © *1997 American Institute of Physics.* [S0022-2488(97)04307-7]

I. INTRODUCTION

This paper continues a programmatic effort to see how far the conventional mathematics of quantum theory—which is based upon the application of linear operators in a Hilbert space—may be extended to include rather general nonlinear operators. Previous authors have investigated what happens when one adds nonlinear terms to the Schrödinger wave equation. The present study, by contrast, is not limited to any such particular equation, but rather reworks the more general mathematical structure of quantum theory: Physical states represented by vectors in an abstract Hilbert space and the operators that act upon these vectors, transforming them into other vectors.

A recent paper, titled ''NonLinear Operators and Their Propagators'' [1] and hereafter referred to as I, presented the beginnings of this program. Key to that work was the definition of the ''slash product'' A/B of two nonlinear operators and the development of an algebra and calculus appropriate for such operators. With the new mathematical tools many of the results familiar in the theory of linear operators could be extended to nonlinear operators: generalizing the exponential of an operator, time-dependent perturbation theory, the Baker–Campbell–Hausdorff theorem, and other results. The present paper presents still further progress.

These new analytical tools may be of practical use in some areas of classical physics as well. For example, in I it was shown how a powerful technique for the numerical computation of wave propagation, first developed for linear equations, could be extended to general nonlinear wave equations. However, the driving ambition of this work is an attempt to expand the frontiers of fundamental physics—the quantum theory. A particular focus here is to follow Wigner's group theoretical approach to the construction of elementary particle states and to see what new results of interest to physicists might be found by the consideration of nonlinear symmetry operators and their group representations. Thus most of the present paper works to rebuild the familiar mathematical infrastructure leading up to the theory of Lie groups and Lie algebras, extending it to accommodate nonlinear operators as well as the conventional linear ones.

After a review, in Sec. II, of the operator algebra and calculus previously developed, Sec. III goes into inner products, Hermitian, anti-Hermitian, and unitary operators, and we find that we do not need to speak of the adjoint of an operator nor of a dual vector space. Whereas so much of traditional quantum theory is based upon the assumption of superposition—mandating linear operators in a vector space—it is again surprising how much can still be achieved if one abandons that habit.

Sections IV and V give sample applications of these new techniques to classical mechanics and the Schrödinger wave equation, respectively. Section VI presents some special categories of nonlinear operators: amplitude invariant, phase invariant, and those that do not make use of the operator of complex conjugation.

In Sec. VII we show that the general mathematical structure of Lie groups and Lie algebras can be extended to nonlinear operators: using the generalized exponential function, the slash commutator, and the new form for similarity transformations. In Sec. VIII we look at the general question of finding representations, paralleling much of the familiar work on linear (matrix) representations. The problem of building direct product representations is looked at in Sec. IX; and some simple examples of nonlinear representations are presented in Sec. X. Further particular studies of Lie group representations—for SL$(2,R)$ and SU(2) in one and two dimensions—are given in Secs. XI and XII, where we find some intriguing new representations. The question of singularities, combined with a unique construction of a composite state is the topic of Sec. XIII, where we find a provocative result; and the possible application to the theory of elementary particle physics is discussed in Sec. XIV.

Appendices A and B present some additional results on power series, carried over from the previous paper: and Appendices C and D contain further new results concerning nonlinear operators.

II. REVIEW OF NONLINEAR OPERATOR ALGEBRA

Nonlinear operators $A,B,C,\ldots,$ act on vectors in a linear vector space to produce other vectors in that space.

$$A\psi=\phi. \tag{1}$$

Note the convention that operators act to the right.
The operators have the following algebra of addition and multiplication:

$$A+B=B+A, \tag{2}$$

$$(A+B)C=AC+BC, \tag{3}$$

$$(AB)C=A(BC), \tag{4}$$

and, as with linear operators, multiplication is not commutative. What distinguishes these from linear operators is that

$$Aa \text{ is not equal to } aA, \tag{5}$$

where $a,\ b,\ c,\cdots$ are ordinary numbers; and also that

$$A(B+C) \text{ is not equal to } AB+AC. \tag{6}$$

The central tool of analysis is the following definition:

$$A(1+\epsilon B)=A+\epsilon A/B+O(\epsilon^2), \tag{7}$$

where A/B is an operator called ''the slash product of A with B.'' The following properties were derived in I.
Linearity:

$$(A+B)/C=A/C+B/C, \tag{8}$$

$$A/(B+C)=A/B+A/C, \tag{9}$$

$$(aA)/B = a(A/B), \tag{10}$$

$$A/(bB) = b(A/B). \tag{11}$$

This last statement under Linearity, Eq. (11), is strictly true only if b is a real number—as the infinitesimal ϵ is taken as a real number. In certain cases, which will be detailed later on, this may also be true for complex numbers b.
The first identity:

$$(AB)/C = (A/C')B, \tag{12}$$

where

$$C' = (B/C)B^{-1} = C_B \tag{13}$$

is a new form of similarity transformation.
The second identity:

$$(A/B)/C - A/(B/C) = (A/C)/B - A/(C/B). \tag{14}$$

A key construct is the generalized exponential function of an operator.

$$E(A) = 1 + A + 1/2 \;\; A/A + 1/6 \;\; (A/A)/A + 1/24 \;\; ((A/A)/A)/A + \cdots \,. \tag{15}$$

The following properties were derived in I:

$$\frac{d}{dt} E(tA) = AE(tA) = E(tA)/A; \tag{16}$$

$$E(sA)E(tA) = E((s+t)A); \tag{17}$$

$$B_{E(A)} = (E(A)/B)E(-A) = \Sigma\, 1/n!\, S_n\,, \tag{18}$$

where

$$S_0 = B \quad \text{and} \quad S_n = [A/S_{n-1}] \tag{19}$$

and the ‘‘slash commutator’’ is defined as

$$[X/Y] = X/Y - Y/X. \tag{20}$$

In the special case of linear operators, just drop the / symbol (or replace it by a comma in the commutator) and all these formulas are familiar. The remarkable fact is that so many of the things commonly done with linear operators can be generalized in this manner to nonlinear operators, and these are not just abstract or formal generalizations but practical computable constructions.

III. INNER PRODUCTS, HERMITIAN AND UNITARY OPERATORS

In our vector space the general vector ψ will be represented by an ordered set of components (complex numbers), $\{\psi_1, \psi_2, \psi_3, \cdots\}$, usually written as the set $\{\psi_k\}$; or I may write $\psi|_k = \psi_k$, where the symbol $|_k$ means ‘‘take the kth component of the resulting vector standing to the left.’’ The inner product of two vectors will be written, as usual, as

$$\langle \phi | \psi \rangle = \langle \psi | \phi \rangle^* = \Sigma_k \phi_k^* \psi_k\,, \tag{21}$$

where * means complex conjugation. (Some other metric might be introduced but I shall stay with this simplest one here.)

The result of a general nonlinear operator A acting on the vector ψ will be another vector, represented as

$$A\psi=\{\mathscr{A}_k(\psi,\psi^*)\} \quad \text{or} \quad A\psi|_k=\mathscr{A}_k(\psi,\psi^*), \tag{22}$$

where the script symbols $\mathscr{A}$ are ordinary functions of their arguments, the set of complex variables ψ_k and their complex conjugates.

Now, consider the inner product,

$$\langle \phi|A\psi\rangle=\Sigma_k\phi_k{}^*\mathscr{A}_k(\psi,\psi^*). \tag{23}$$

and the complex conjugate of this equation,

$$\langle A\psi|\phi\rangle=\Sigma_k\phi_k\mathscr{A}_k(\psi,\psi^*)^*. \tag{24}$$

In the case of linear operators this leads to the definition of an adjoint (Hermitian conjugate) operator $A^\dagger$ which acts on the vector ϕ. But in the general study of nonlinear operators this idea of an adjoint operator seems to make no sense. Thus I do not speak of adjoint or dual vectors, nor of column and row vectors, which are concepts particular to linear operators and their matrix representation.

What comes as a surprise, however, is how much can still be achieved if we limit ourselves to studying only the ''expectation value''

$$\langle A\rangle=\langle \psi|A\psi\rangle \tag{25}$$

of the operator A in the ''state'' represented by the vector ψ.

Let us first examine the situation in the standard dynamical model, where $\psi=\psi(t)$ varies with time according to the equation

$$d_t\psi=\frac{d\psi}{dt}=A\psi \tag{26}$$

and the operator A is assumed time independent. What is the time derivative of the inner product?

$$d_t\langle\psi|\psi\rangle=\langle d_t\psi|\psi\rangle+\langle\psi|d_t\psi\rangle=\langle A\psi|\psi\rangle+\langle\psi|A\psi\rangle=\Sigma_k\psi_k{}^*\mathscr{A}_k(\psi,\psi^*)+\text{cc}. \tag{27}$$

All we can say from this equation is the following: If $\langle A\rangle$ is imaginary, then $\langle\psi|\psi\rangle$ will be time independent. But this is saying a lot.

Definition: A general nonlinear operator A will be called ''anti-Hermitian'' if its average value $\langle\psi|A\psi\rangle$ is imaginary for all vectors ψ; and it will be called ''Hermitian'' if $\langle\psi|A\psi\rangle$ is always real. (28)

If one takes the special case of linear operators, where A is represented by a matrix with matrix elements A_{jk}, these two definition lead to the familiar conditions

$$A_{jk}{}^*=-A_{kj} \quad \text{or} \quad A_{jk}{}^*=+A_{kj}, \tag{29}$$

respectively.

Definition: A general nonlinear operator U will be called ''unitary'' if

$$\langle U\psi|U\psi\rangle=\langle\psi|\psi\rangle \tag{30}$$

for all vectors ψ.

Again, for the special case of linear operators, this definition leads to the usual equation for a unitary matrix:

$$\Sigma_k U_{ki}{}^* U_{kj} = \delta_{ij}\,. \tag{31}$$

We have previously devoted much study to the generalized exponential function of a nonlinear operator $E(A)$. We know that the dynamical equation (26) is solved by

$$\psi(t) = E(tA)\,\psi(0); \tag{32}$$

and thus, if A is anti-Hermitian, we have already shown that

$$\langle \psi(t) | \psi(t) \rangle = \langle E(tA)\psi(0) | E(tA)\psi(0) \rangle = \langle \psi(0) | \psi(0) \rangle. \tag{33}$$

Thus $E(tA)$ (for real t) is unitary as long as A is anti-Hermitian.

Next, we want to calculate the time derivative of $\langle B \rangle = \langle \psi(t) | B\psi(t) \rangle$, where B is some general (t independent) operator. First, we recall that

$$d_t B\psi(t) = (B/A)\,\psi(t), \tag{34}$$

so we need to evaluate the slash product operating in our vector space.

Let us digress to do this in general. First we look at,

$$(1+\epsilon B)\psi|_k = \psi_k + \epsilon \mathscr{B}_k(\psi,\psi^*) \tag{35}$$

then

$$A(1+\epsilon B)\psi|_k = \mathscr{A}_k(\{\psi_j + \epsilon \mathscr{B}_j(\psi,\psi^*)\},\{\psi_j{}^* + \epsilon \mathscr{B}_j(\psi,\psi^*)^*\}) \tag{36}$$

and finally

$$A/B\psi|_k = \sum_j \; [\mathscr{B}_j(\psi,\psi^*)\partial_j + \mathscr{B}_j(\psi,\psi^*)^* \partial_j{}^*]\mathscr{A}_k(\psi,\psi^*), \tag{37}$$

where

$$\partial_j = \frac{\partial}{\partial \psi_j} \quad \text{and} \quad \partial_j{}^* = \frac{\partial}{\partial \psi_j{}^*}. \tag{38}$$

Thus we can write for general operators A and B

$$\langle \psi | A/B\psi \rangle = \sum_{j,k} \; \psi_k{}^*[\mathscr{B}_j \partial_j + \mathscr{B}_j{}^* \partial_j{}^*]\mathscr{A}_k\,, \tag{39}$$

where I have dropped the arguments (ψ,ψ^*) of $\mathscr{A}$ and $\mathscr{B}$ for easier reading. And this can be rewritten neatly as

$$\langle \psi | A/B\psi \rangle = \mathscr{D}(\mathscr{B})\langle A \rangle - \sum_k \; \mathscr{B}_k{}^* \mathscr{A}_k\,, \tag{40}$$

where the differential operator $\mathscr{D}(\mathscr{B})$ is defined as,

$$\mathscr{D}(\mathscr{B}) = \sum_j \; [\mathscr{B}_j \partial_j + \mathscr{B}_j{}^* \partial_j{}^*]; \tag{41}$$

and we note that $\mathscr{D}(\mathscr{B})$ is always real. We can now draw the following conclusions:

$$\text{If } A \text{ and } B \text{ are Hermitian operators, so is } A/B+B/A, \tag{42}$$

$$\text{If } A \text{ and } B \text{ are anti-Hermitian operators, so is } [A/B]=A/B-B/A. \tag{43}$$

And there is one more result, which looks awkward but will be useful: If A is anti-Hermitian and B is Hermitian, then

$$\text{Re}\langle[A/B]\rangle=-\mathscr{D}(\mathscr{A})\langle B\rangle. \tag{44}$$

Now we return to the calculation started above.

$$d_t\langle\psi(t)|B\psi(t)\rangle=\langle\psi(t)|B/A\psi(t)\rangle+\langle A\psi(t)|B\psi(t)\rangle=\mathscr{D}(\mathscr{A})\langle B\rangle. \tag{45}$$

Thus if B is a Hermitian operator (and we have A as anti-Hermitian)

$$d_t\langle B\rangle=-\text{Re}\langle[A/B]\rangle; \tag{46}$$

and if B is anti-Hermitian, then we can write (setting $B\to iB$),

$$d_t\langle B\rangle=i\ \text{Re}\langle[A/iB]\rangle. \tag{47}$$

In the case of linear operators, these formulas reduce to the well-known formula of quantum mechanics,

$$d_t\langle B\rangle=i\langle[H,B]\rangle, \tag{48}$$

where $H=iA$ is the Hamiltonian (in units $h=2\pi$).

Suppose we set $A=-iH$ for the general nonlinear equation of motion. H is a Hermitian operator and one wonders whether it is ''conserved.'' That is, Does $\langle H\rangle$ vary with time? From Eq. (46) we have,

$$d_t\langle H\rangle=\text{Re}\langle[iH/H]\rangle. \tag{49}$$

For linear operators this commutator is obviously zero; and for general nonlinear operators we also have $[H/H]=0$. But $[iH/H]$ is something else in general. Thus we must have a special constraint upon the structure of the nonlinear Hamiltonian in order to get ''conservation of energy'' in this form. The simplest statement is: If H, acting on any vector, does not involve the operation of complex conjugation, then $[iH/H]=0$; but with this restriction it may be difficult to insure that H is Hermitian.

Another condition to consider is the following:

$$He^{i\phi}=e^{i\phi}H(\text{real }\ \phi)\quad \text{or equivalently}\quad [H/i]=0. \tag{50}$$

Another phrase that describes this restriction is: the equation of motion obeys gauge invariance of the first kind. With this condition, we can seek stationary states of the equation of motion

$$\psi(t)=e^{-i\omega t}u\quad \text{and}\quad Hu=\omega u. \tag{51}$$

With such states, we obviously have $\langle\psi|\psi\rangle$ and $\langle\psi|H\psi\rangle$ independent of the time t. But this does not give us energy conservation as a general rule, i.e., for all $\psi(t)$, unless $[iH/H]=0$ is also satisfied. We shall return to this subject shortly. One should also note that the eigenvalue problem associated with such stationary states is entangled with the question of what normalization one should choose for the state vector.

In regular quantum theory, unitarity is usually invoked in connection with general scattering/measurement theory. An initial state is prepared; and it then evolves in time according to a wave function (state vector) $\psi(t)$ which may include some interactions of initially separated parts. Then, after the scattering process, we make measurements of the outcome by projecting this wave function onto various final states by means of detection apparatus. If any such final state is represented by a vector $|n\rangle$ in the vector space, then the probability that this particular final state will be detected is said to be calculated as

$$P_n=|\langle n|\psi(t)\rangle|^2 \tag{52}$$

and ''unitarity'' is the requirement that the sum of these probabilities over all possible final states $|n\rangle$ is equal to one. How does this work in our situation with nonlinear operators?

We make the usual assumption that the vectors $|n\rangle$ constitute a complete orthonormal basis in the vector space; and this leads to the result

$$\sum P_n=\langle\psi(t)|\psi(t)\rangle \tag{53}$$

regardless of whether we have linear or nonlinear operators in the equations of motion. But this is just the quantity we studied earlier; and we saw that, as long as the dynamical operator A is anti-Hermitian, this total probability is independent of the time t. Thus it can be evaluated back when the initial state was created, with the usual norm of 1.

IV. REAL VECTOR SPACE

If we restrict ourselves to a vector space and operators which involve only real numbers, not complex numbers, then the results of the previous section become even simpler. If the dynamical operator A obeys the condition

$$\langle\psi|A\psi\rangle=0 \quad \text{for all vectors} \quad \psi, \tag{54}$$

then $\langle\psi(t)|\psi(t)\rangle$ will be time independent.

Let's see how this works in a very familiar problem, Newton's law of motion in one dimension.

$$\frac{d^2x}{dt^2}=F(x). \tag{55}$$

Since this is second order in the time derivative, we introduce a new variable u and construct a two-component vector $\psi=\{x,u\}$ to satisfy the first order equation of motion $d\psi/dt=A\psi$. These are all time-dependent variables. We now construct the dynamical operator A so as to guarantee the condition $\langle\psi|A\psi\rangle=0$

$$A\psi=\{uW(x,u),-xW(x,u)\}. \tag{56}$$

In order to find the unknown function W, we look at the equations of motion,

$$\frac{dx}{dt}=uW, \tag{57}$$

$$\frac{du}{dt}=-xW, \tag{58}$$

take another time derivative of dx/dt and compare with the original Newton's law involving the force $F(x)$. The resulting equation is

$$\left[-x-1/2\ \ xu\ \frac{d}{du}+1/2\ \ u^2\ \frac{d}{dx}\right]W^2=F(x). \tag{59}$$

This looks unfamiliar, and messy. A solution of this equation (which I found after making a power series expansion in u^2) is

$$W(x,u)=1/u[V((x^2+u^2)^{1/2})-V(x)]^{1/2}, \tag{60}$$

where $F(x)=-dV/dx$. We also have a constant of the motion

$$\langle\psi|\psi\rangle=x^2+u^2. \tag{61}$$

I don't see that this adds anything useful to the study of Newton's equation of motion; but it illustrates our general approach.

V. NONLINEAR SCHRODINGER EQUATION

Here, we continue a discussion begun in I. Assume we have several components of a complex wavefunction which depend on one or more continuous variables x: $\psi_k=\psi_k(x,t)$; and we shall deduce equations of motion from an action

$$1/2\int dt\int dx\left\{\sum_k\ (i\psi_k{}^*d_t\psi_k-d_x\psi_k{}^*d_x\psi_k)-G(\rho,x)\right\}, \tag{62}$$

where G is a real (local) function of the densities $\rho_k=\psi_k^*\psi_k$ and may include as well some external force. Varying $\psi_k{}^*$ we get the equation of motion

$$d_t\psi_k=id_x^2\psi_k\quad\frac{-i\partial G}{\partial\rho_k}\ \psi_k=A\psi|_k \tag{63}$$

and we immediately see that our dynamical operator is anti-Hermitian: $\langle A\rangle$ is imaginary. This assures us conservation of probability. We also have a conserved current density just as in the usual Schrödinger equation. But what about conservation of energy? This was the question raised in a previous section.

If we look upon the action from the point of view of classical Lagrangian field theory, then we know how to derive the time-independent Hamiltonian from canonical variables: $H=\Sigma p_iq_i-L$. The result for the action given above is

$$H=\int dx\left\{\sum_k\ \left(d_x\psi_k{}^*d_x\psi_k+1/2\ \ \rho_k\ \frac{\partial G}{\partial\rho_k}\right)+1/2\ \ G\right\} \tag{64}$$

which is clearly different from

$$\langle iA\rangle=\int dx\sum_k\ \left(d_x\psi_k{}^*d_x\psi_k+\rho_k\ \frac{\partial G}{\partial\rho_k}\right) \tag{65}$$

unless G is linear in the ρ's.

Thus we have (re-)learned the lesson that for nonlinear systems the operator which gives the time development is not the same as the operator which gives the conserved quantity usually called energy; but in the usual linear quantum theory these two things are both called the Hamiltonian.

The above results can also be written in a more compact manner as follows. Let us write the action as

$$\int dt \quad 1/2\langle \psi | \frac{i\partial}{\partial t} - V|\psi\rangle, \tag{66}$$

where V is some general operator. Then we find that we can write

$$\langle iA\rangle = \langle V + 1/2(V/1 + iV/i)\rangle \tag{67}$$

and

$$H = \langle V\rangle + 1/4\langle (V/1 + iV/i)\rangle^*. \tag{68}$$

Note that for linear operators the expression $(V/1 + iV/i)$ vanishes.

A number of previous authors have explored nonlinear generalizations of the Schrödinger wave equation.[2,3] A persistent problem is how to achieve in a consistent mathematical way the physical separability of noninteracting systems. This is a deep concern, which we shall return to in studying group representations later in this paper.

VI. SPECIAL TYPES OF OPERATORS

Following the previous discussion, let us introduce a nomenclature for certain classes of nonlinear operators, as follows.

Type 1: Operators A which satisfy $Aa = aA$, for all positive real numbers a, or equivalently $[A/1] = 0$. These operators may be called ''amplitude invariant.'' (69)

Type 2: Operators A which satisfy $Ae^{i\phi} = e^{i\phi}A$ for all real numbers ϕ, or equivalently $[A/i] = 0$. These operators may be called ''phase invariant.'' (70)

Type 3: Operators A which do not involve the operation of complex conjugation when they act on any vector. (71)

If A is an operator of Type 3, then $A/iB = iA/B$, or more generally, $A/bB = bA/B$ for any complex number b and any operator B. Note that Type 2 status does not in general suffice for this result.

A nonlinear operator A which is of Type 1 and also of Type 2 has the property that $Aa\psi = aA\psi$ for any complex number a; and has been called ''homogeneous'' by Weinberg.[3]

An operator A of Type 3 also gives $(A/1 + iA/i) = [A/1] + i[A/i] = 0$; and this is also true if A is both Type 1 and Type 2. Recall the discussion at the end of the preceding section.

If an operator A is of Type 3 and also of Type 1 (or 2), then it must also be of Type 2 (or 1). But being of Type 1 and also of Type 2 does not in general imply being of Type 3.

If an operator A is of some given Type (1, 2, or 3), then aA, for any number a, is of that same type. If operators A and B are both of the same Type (1, 2, or 3), then $(A+B)$, AB, A/B are also of that same type. Thus if A is of some given type, then $E(A)$, and many other functions formed from A, will be of that same type.

A special nonlinear operator (already used in conventional quantum theory) is the operator K which takes the complex conjugate of all the components of the vector standing to its right.

$$K\psi = \psi^*; \tag{72}$$

and we note $KK = 1$. For any operator A we can define its complex conjugate as

$$A^*=KAK; \tag{73}$$

and, following Eq. (22)

$$A^*\psi|_k=\mathscr{A}_k(\psi^*,\psi)^*, \tag{74}$$

where the reader should note the two effects of complex conjugation. Furthermore,

$$(aA+bB)^*=a^*A^*+b^*B^*;\ \ (AB)^*=A^*B^*;\ \ (A/B)^*=A^*/B^*. \tag{75}$$

VII. LIE GROUPS AND ALGEBRAS

The usual Lie theory is based upon a set of linear operators X_i, called the generators of the group, which are closed under a commutator algebra

$$[X_i,X_j]=X_iX_j-X_jX_i=\Sigma_k f_{ij}{}^k X_k \tag{76}$$

and the numbers $f_{ij}{}^k$, called the structure constants, characterize the particular group. The group elements are exponentials of the generators and they obey the multiplication laws

$$\exp(aX_i)\exp(bX_j)=\exp(X'), \tag{77}$$

where X' is some linear combination of all the X's. Furthermore, the commutators obey the Jacobi identity (which says something about the structure constants) and all the equations are invariant in form under a similarity transformation,

$$X_i\rightarrow SX_iS^{-1}. \tag{78}$$

If one tries to extend this mathematics to nonlinear operators, none of the above machinery works in the form given. However, as shown in I, we can get the same mathematical structure to be consistent if we replace the ordinary commutator by the slash commutator in the Lie algebra,

$$[X_i/X_j]=X_i/X_j-X_j/X_i=\Sigma_k f_{ij}{}^k X_k, \tag{79}$$

and use the generalized exponential function $E(aX=\Sigma a_iX_i)$ for the group elements. The slash commutator acts linearly in its arguments if we restrict the coefficients a_i to be real, and we would also restrict the structure constants $f_{ij}{}^k$ to be real numbers. (We may be able to extend this to complex numbers with certain restrictions, such as being of Type 3, on the representations constructed for the operators X.)

If the generators X are anti-Hermitian operators, as this property was previously defined for nonlinear operators, then their slash commutators are also anti-Hermitian and the E function of such operators are unitary operators. This conforms fully to the familiar situation for linear operators.

The slash commutator obeys a Jacobi identity for any nonlinear operators, which the ordinary commutator fails to do for nonlinear operators. The Second Identity is essential in the proof of this.

Finally, there is the question of invariance under a similarity transformation. The form $S(X+Y)S^{-1}$ does not work nicely if S is nonlinear. We have the alternative form of transformation

$$X\rightarrow X_S=(S/X)S^{-1};$$

but, while it works nicely on linear expressions X, this operation does not work nicely on slash products

$$X_S/Y_S \text{ is not equal to } (X/Y)_S.$$

However, thanks to the First and Second Identities, one can show that this does work for the slash commutator

$$[X_S/Y_S]=[X/Y]_S. \tag{80}$$

Another interesting way to see the relation between ordinary commutators (used for elements of the group) and slash commutators (used for the elements of the algebra) is the following:

$$[(1+\epsilon X),(1+\delta Y)]=\epsilon\delta[X/Y]+\text{higher order in } \epsilon,\delta. \tag{81}$$

From the discussion of the previous section, we see that it is consistent to speak of any set of X's satisfying a Lie algebra to be of Type 1 or Type 2 or Type 3. And if a set of X's satisfy a Lie algebra with real structure constants, then the set of X^*'s also satisfy the same algebra.

VIII. REPRESENTATIONS

A representation of a Lie algebra is a set of explicit constructions of how the operators X_i act upon a general vector $\psi=\{\psi_1,\psi_2,\ldots,\psi_n\}$ in a space of dimension n, consistent with the given commutator rules. The objective is to see what new nonlinear representations one might find in addition to the known matrix representations for linear operators. This is a huge task, which I have barely started in this paper. However, some general features can be stated here. (Elsewhere, the word ''representation'' is often reserved for a linear, matrix, representation of the group or algebra and anything else is called a ''realization.'')

First is the matter of ''adding'' representations. Suppose we have one representation, call the operators X and they act on vectors ψ spanning a space of dimension n; and we also have another representation, with operators Y acting on the vectors ϕ of dimension m. Now we can form a larger vector space, of dimension $n+m$, as follows:

$$\Psi=\{\psi,\phi\} \tag{82}$$

and the operators act as,

$$X\Psi=\{X\psi,0\} \quad \text{and} \quad Y\Psi=\{0,Y\phi\}, \tag{83}$$

$$E(aX)\Psi=\{E(aX)\psi,\phi\} \quad \text{and} \quad E(bY)\Psi=\{\psi,E(bY)\phi\}. \tag{84}$$

Note that each generator effects a change upon its own vector space, and produces zero when acting upon the other vector space. The group element $E(X)$, on the other hand, acts as the unit operator upon ϕ.

From the above definitions, we can deduce that $(X+Y)$ is the correct generator in the composite space and

$$XY=YX=X/Y=Y/X=0, \tag{85}$$

$$E(X)E(Y)=E(Y)E(X)=E(X+Y)=E(X)+E(Y)-1. \tag{86}$$

This structure has exactly the ''block diagonal'' form familiar from the discussion of linear operators and matrices.

Building bigger spaces this way is the easy part. Next is the question of whether any given representation might be transformed into the ''block diagonal'' form that is transparently composite, as described above. If this is possible, then the representation is called reducible; if not, it is called irreducible.

Two representations $E=E(aX)$ and $E'=E(aX')$ of the same dimension are equivalent if there exists some invertible T such that $TE=E'T$ for every group element designated by the set of group parameters a. In terms of the generators, equivalence means $T/X = X'T$, as shown in I. Thus the representation by the set of generators X^* may or may not be equivalent to the representation generated by the set X.

I have nothing to say about Casimir operators because I have not been able to find the nonlinear generalization of the identity $[A,BC] = B[A,C] + [A,B]C$ which holds for all linear operators. There is also the question of how Schur's Lemma might be restated. As for taking the trace of an operator, see Appendix C.

IX. DIRECT PRODUCTS

With linear operators and matrices, there is also the construction of new representations as direct products of known ones. This is very important for physics, since it lets us describe composite systems in a consistent manner. For general nonlinear representations of a group or algebra, however, this at first seemed to me impossible. However, in an important special case—Type 2 (phase invariant) unitary representations—something may be achieved that mimics the linear situation.

Suppose we have the elements of a Lie algebra represented as follows:

$$X\psi|_k = \Sigma_{k'} R_{kk'}(\rho)\psi_{k'}, \tag{87}$$

where ρ stands for the set of variables $\rho_{pq} = \psi_p^* \psi_q$, which guarantees that this is of Type 2; and I have suppressed the fact that the matrix of functions R (call it an f-matrix for short) depends on this particular element X of the algebra. I will also require $R_{kk'}^* = -R_{k'k}$ so this will give us a unitary representation.

Now, we consider another representation of this same Lie algebra, with perhaps a different dimension. Call the operators Y, and let it be of this same form, but with some other f-matrix $S_{jj'}(\rho)$. The elements of the composite vector space will now be written as $\Psi = \{\Psi_{kj}\}$ and the action of the operators X will be taken as follows:

$$X\Psi|_{kj} = \Sigma_{k'} R_{kk'}(\rho^1)\Psi_{k'j} \quad (\text{``diagonal'' in } j), \tag{88}$$

where we now extend the definition of the arguments

$$\rho_{pq}^1 = \Sigma_j \Psi_{pj}{}^* \Psi_{qj}. \tag{89}$$

The action of the operators Y will be

$$Y\Psi|_{kj} = \Sigma_{j'} S_{jj'}(\rho^2)\Psi_{kj'} \quad (\text{``diagonal'' in } k) \tag{90}$$

and the arguments of S are

$$\rho_{pq}^2 = \Sigma_k \Psi_{kp}^* \Psi_{kq}. \tag{91}$$

Now we calculate the slash product X/Y acting on Ψ and find the result

$$X/Y\Psi|_{kj} = \Sigma_{j'k'} S_{jj'}(\rho^2) \quad R_{kk'}(\rho^1)\Psi_{k'j'}. \tag{92}$$

Key to this result is the fact that the arguments ρ^1 are all invariant under the action of $1+\epsilon Y$; this is a result of unitarity. When we calculate Y/X we get the identical result.

Thus with this construction of the composite vector space, we have $[X/Y] = 0$ and so the composite operators, $X+Y$, obey the original Lie algebra. This procedure mimics what is usually done in the case of linear representations by matrices; and this may be continued to products of additional subspaces.

Now we must see how this construction relates to the representation of the Lie algebra in each individual vector space. Let me start over again, as above, but now consider the operator Y and the f-matrix S as belonging to the same vector space as X and R. We must now see how the variables ρ^1_{pq} (we can drop the superscript now) are changed by the action of $(1+\epsilon Y)$. The result is

$$\rho_{pq}\to\rho_{pq}+\epsilon(\Sigma_{p'}S^*_{pp'}\rho_{\rho' q}+\Sigma_{q'}S_{qq'}\rho_{pq'}). \tag{93}$$

If we regard ρ as a matrix, along with R and S, and use the fact that S, like R, is anti-Hermitian, then we can write the result of this calculation succinctly as,

$$X/Y\Psi|_{kj}=\Sigma_{k'}\{RS+(\Sigma_{pq}[S^*,\rho]_{pq}\ \ \partial/\partial\rho_{pq})R\}_{kk'}\Psi_{k'j}. \tag{94}$$

where [,] is the usual commutator.

Thus we can proceed to write the equations of the Lie algebra, e.g., $[X/Y]=Z$, etc.; and see that these are a set of coupled, bilinear differential equations involving f-matrices R, S, etc., and the set of variables ρ_{pq}. In this form there is no reference to the elements of the composite vector space; it is just one way of formulating the problem of finding a single representation of the Lie algebra.

There is a warning here. In the one vector space the variables ρ_{pq}, as written after Eq. (87), are very redundant. For example, $\rho_{12}\rho_{21}=\rho_{11}\rho_{22}$, $\psi_1=\psi_2\rho_{11}/\rho_{12}$, etc.; but no such relations hold for the extended definition of Eq. (89). So, in building the representation f-matrices for the algebra in one space, we must avoid making use of these identities and then we can use these same f-matrices for the composite space construction. The one identity which does carry over is $\rho_{pq}{}^*=\rho_{qp}$.

While this construction of the direct product representation looks nice, it is unclear that it will be useful. There is still the physical necessity of being able to separate the variables associated with disjoint subsystems; and here they seem to be all intertwined in the ρ_{pq}. This subject will require further study. In Sec. XIII we take another tack on the construction of a direct product representation for a particular case.

X. SOME SIMPLE EXAMPLES

Consider a one-dimensional representation

$$A\psi=A(\psi), \tag{95}$$

where I have taken the representation to be of Type 3—a function of ψ but not involving ψ^*. Let us see what happens to this when we carry out a transformation

$$\psi\to\psi'=T\psi=T(\psi),\quad A\to A'=(T/A)T^{-1}. \tag{96}$$

We have

$$A'\psi'=T/A\psi=A(\psi)\,\frac{dT(\psi)}{d\psi}, \tag{97}$$

where I have assumed that T is also of Type 3. If I choose

$$T(x)=\exp\left(\lambda\int^x dy/A(y)\right), \tag{98}$$

then

$$A'\psi'=\lambda\psi' \tag{99}$$

and this is now in the form of a linear eigenvalue equation.

Suppose now that the operator A is of Type 2, in one dimension:

$$A\psi=\psi a(\rho); \quad \rho=\psi^*\psi; \tag{100}$$

and we look at a transformation T also of Type 2

$$\psi'=T\psi=\psi\tau(\rho). \tag{101}$$

We transform to $A'=(T/A)T^{-1}$ and consider

$$A'\psi'=\psi' a'(\rho'), \tag{102}$$

where

$$\rho'=\psi'^*\psi'=\rho|\tau(\rho)|^2. \tag{103}$$

By calculating $T/A\psi$, we come to the result:

$$a'(\rho')=a(\rho)+[a(\rho)+a(\rho)^*]\rho(d\tau(\rho)/d\rho)/\tau(\rho). \tag{104}$$

Thus given $a(\rho)$, our task is to find a function $\tau(\rho)$ that will produce any desired $a'(\rho')$—for example, $a'(\rho')=\lambda$, a constant, which gives us a linear representation of the transformed operator A'. This is done by

$$\tau(\rho)=\exp\left\{\int^{\rho}\left(\frac{du}{u}\right)[\lambda-a(u)]/[a(u)+a(u)^*]\right\}. \tag{105}$$

But if $a(\rho)$ is imaginary, then this solution is nonsense; and this is a very important special case: A is anti-Hermitian and $E(A)$ is unitary. Then we are faced with the equation

$$a'(\rho')=a(\rho) \tag{106}$$

along with Eq. (103) relating ρ' to ρ. Let us see the implications of this. The only way for $a'(\rho')$ to be a constant is if $a(\rho)$ is already that constant. Thus the linear representation is an equivalence class unto itself, for each value of the eigenvalue λ. If $a(\rho)$ is some nonconstant function, we can transform it into various other nonconstant functions by the choice of $\tau(\rho)$. Suppose we try to get the next simplest representation, linear plus cubic

$$a'(\rho')=\lambda(1+\beta\rho'). \tag{107}$$

We see that we need $\lambda=a(0)$ and

$$|\tau(\rho)|=\{[a(\rho)/a(0)-1]/(\beta\rho)\}^{1/2}. \tag{108}$$

The validity of this depends on $a(\rho)/a(0)$ being greater (or less) than 1 for all $\rho>0$ and the real constant β being correspondingly positive (or negative).

If we add the requirement that the transformation T be unitary, then $|\tau(\rho)|=1$ and $a'(\rho)=a(\rho)$ so each choice of $a(\rho)$ is inequivalent to any other choice. Such a restriction, however, is not generally called for.

Now, suppose we have two operators, A and B, that obey the very simple algebra,

$$[A/B]=0; \tag{109}$$

and let us look for one-dimensional representations. If we consider A and B to be of Type 3, then we get the requirement

$$B(\psi)\,\frac{dA(\psi)}{d\psi}=A(\psi)\,\frac{dB(\psi)}{d\psi}; \tag{110}$$

and the solution of this tells us that A is a constant times B, which is a trivial situation.

Alternatively, if we take both A and B to be of Type 2 and, furthermore, if we take them both to be anti-Hermitian, then we find (in a one dimensional vector space) that $[A/B]=0$ without any further restrictions on either operator. Thus the irreducible unitary representations of a nonlinear Abelian Lie group can be one dimensional with generators of Type 2.

A few examples of one- and two-dimensional nonlinear representations of three-element Lie algebras are worked out in the next two sections.

XI. THE GROUP SL(2,*R*)

This Lie algebra has three elements, A,B,C, which obey

$$[A/B]=B; \quad [A/C]=-C; \quad [B/C]=2A; \tag{111}$$

and I shall seek representations of Type 3, which is appropriate also for a real vector space. These will not be unitary representations. I start by assuming that we can take one of these operators, say A, to be linear and ''diagonal.''

In one dimension I find

$$A\psi=\lambda\psi; \quad B\psi=\lambda\psi^{1-1/\lambda}; \quad C\psi=-\lambda\psi^{1+1/\lambda}. \tag{112}$$

In order to avoid singular behavior as $\psi\rightarrow 0$, one should set λ equal to zero.

In two dimensions, write $\psi=\{s,t\}$ and note that the usual linear representation is the following:

$$A\psi=\{s/2,-t/2\}; \quad B\psi=\{t,0\}; \quad C\psi=\{0,s\}. \tag{113}$$

For the nonlinear representation I keep A the same as just written, but construct B and C as follows:

$$B\psi=\{tf,(t^2/s)g\}; \quad C\psi=\{(s^2/t)g,sf\}, \tag{114}$$

where f and g are functions of $u=st$. This satisfies the first two commutator equations for any functions f and g. I take the same two functions f and g in both B and C for the sake of symmetry. That is, defining the permutation operator P as

$$P\{s,t\}=\{t,s\},$$

the operators of this Lie algebra will be chosen so that $PAP=-A$, $PBP=C$ and $PCP=B$.

Finally, from the third commutator equation, $[B/C]=2A$, we get the differential equation

$$(f-g)^2+(f+g)u(d/du)(f-g)=1. \tag{115}$$

One set of solutions is simply,

$$f(u)-g(u)=1; \tag{116}$$

and an alternative set of solutions is given as,

$$(f+g)=[1-(f-g)^2]/[u(d/du)(f-g)]. \tag{117}$$

Either way, we have one arbitrary function in the solution.

Here are a few noteworthy special cases. If $g=0$ and $f=1$, we have the linear representation. If f and g are constants, independent of u, we have a Type 1 representation. If we take $f=-g=1/2$, we get a pretty answer

$$B\psi=\{1/2t,-1/2t^2/s\};C\psi=\{-1/2s^2/t,1/2s\}. \tag{118}$$

To avoid the apparent singularity at s or $t=0$, we can set

$$g=st\ h(st); \quad f=1+st\ h(st), \tag{119}$$

where h is any function which is finite at the origin. This form of the representation goes over to the linear one as the norm of ψ becomes small. Still another special case is

$$g=0;f=[1+\text{const.}/(st)^2]^{1/2}. \tag{120}$$

There is now the question whether all these representations are distinct or equivalent. The transformation T,

$$T\{s,t\}=\{s',t'\}=\{s\tau(u),t\tau(u)\}, \quad \text{where} \quad u=st, \tag{121}$$

is the most general one of Type 3 that leaves the operator A unchanged in form and furthermore respects the symmetry P described above. When we calculate how this transforms the operators B and C, we find that it gives us new functions $f\to f'$ and $g\to g'$ as follows:

$$f'(u')=f(u)+(f(u)+g(u))(u/\tau(u))\ \frac{d\tau(u)}{du}, \tag{122}$$

$$g'(u')=g(u)+(f(u)+g(u))(u/\tau(u))\ \frac{d\tau(u)}{du} \tag{123}$$

and we also have

$$u'=s't'=u\tau^2(u). \tag{124}$$

The first question is, can all the many representations described above be transformed into the linear one through an appropriate choice of the function $\tau(u)$? From Eqs. (122) and (123) we see that

$$f'(u')-g'(u')=f(u)-g(u). \tag{125}$$

Thus solutions of the first category, $f(u)-g(u)=1$ (or -1) are transformed into solution of that same category, and that includes the linear solution ($g=0$ and $f=1$ or -1.) However, it is not true that all solutions in this category are equivalent to the linear one. For example, the unique solution $f=-g=1/2$, which was noted as ''pretty'' above, is an equivalence class unto itself.

Solutions of the second category, where $f(u)-g(u)$ is different from 1 or -1, are also transformed into solutions of this same category, and these can never include the linear case. Many of these are equivalent to one another, but the identification of the various equivalence classes is more complicated.

XII. THE GROUP SU(2)

This Lie algebra has three elements, X,Y,Z, which obey

$$[X/Y]=-Z; \quad [Y/Z]=-X; \quad [Z/X]=-Y. \tag{126}$$

For a Type 3 representation this can be simply related to the SL$(2,R)$ Lie algebra studied above by introducing the imaginary i

$$Z=iA; \quad X=i/2(B+C); \quad Y=1/2(B-C). \tag{127}$$

but we know that this will not give us a unitary representation.

We will now seek representations that are of Type 2 and are unitary. We already know that the only result in one dimension is the trivial one, so we will construct a two-dimensional representation with $\psi=\{s,t\}$, where s and t are complex variables. The standard result in the linear case is

$$Z\psi=\{is/2,-it/2\}; \quad X\psi=\{it/2,is/2\}; \quad Y\psi=\{t/2,-s/2\} \tag{128}$$

which is $i/2$ times the Pauli spin matrices.

For the nonlinear representation we shall take Z linear, as just written, and work with the ''raising'' and ''lowering'' operators $X_+=X+iY$ and $X_-=X-iY$. Noting that since Z is linear, $[Z/iY]=i[Z/Y]$, we find from the two commutators with Z

$$X_+\psi=\{it\ f,it(s^*/t^*)g\}, \tag{129}$$

$$X_-\psi=\{is(t^*/s^*)g,is\ f\}, \tag{130}$$

where f and g are chosen to be real symmetric functions of the two arguments $\rho_1=s^*s$ and $\rho_2=t^*t$. This form insures that X, Y and Z are anti-Hermitian operators; and we have imposed the symmetry P from the previous section to get the same functions f and g in both constructions. Finally, from the commutator equation $[X/Y]=-Z$, now being very careful about i's in the slash products, we get the differential equation,

$$[(f-g)(1+\rho_2(\partial_2-\partial_1))-(\rho_2/\rho_1-1)g](f+g)=1, \tag{131}$$

where $\partial_1=\partial/\partial\rho_1$ and $\partial_2=\partial/\partial\rho_2$; and also a second equation gotten from this one by interchanging the subscripts 1 and 2 that appear explicitly here. The solution is

$$f=1/2(U+U^{-1}); \quad g=1/2(U-U^{-1}), \tag{132}$$

where

$$U=[1+h(\rho_1+\rho_2)/\rho_1\rho_2]^{-1/2}; \tag{133}$$

and h is an arbitrary real function of its argument.

We recover the linear solution if we choose $h=0$. If h is taken to be a constant times the square of its argument, then this solution is also of Type 1.

Note that taking the complex conjugate of these operators is the same as the mapping, $Z\to -Z$, $X\to -X$, $Y\to +Y$; both of these are symmetries of the Lie algebra.

Exploring the equivalence of these solutions, we find an equation for the mapping of the function $h(r)$ into $h'(r')$, with the following results. One equivalence class is $h(r)=0$, the linear case. Another set of equivalence classes is $h(r)=\alpha r^2$, for each value of α. All other functions $h(r)$ that are everywhere positive (or negative) form another equivalence class. There are still other equivalence classes, involving functions $h(r)$ that change sign.

XIII. SINGULARITIES—A COMPOSITE STATE

We have found some genuinely new nonlinear representations of these groups and obviously much further study is needed to see what else may arise from these mathematical constructions that is useful in physical theory.

Most of the new representations found in the previous two sections have singularities—occurring at the origin of one of the variables. One should not be too quick to discard them for this reason. Perhaps the domain on the vector space may be sensibly constrained to avoid the singularities; or perhaps some other construction will yield useful results.

Here, I want to report on one success: making a composite state out of the direct product of two vector spaces, each carrying the 2-dimensional nonlinear unitary representation of SU(2) reported above.

$$\psi=\{s,t\}; \quad Z\psi=\{i/2s,-i/2t\},$$

$$(X+iY)\psi=\{itf,it(s^*/t^*)g\},$$

$$f=f(s,t)=(U+1/U)/2; \quad g=g(s,t)=(U-1/U)/2,$$

and

$$U=[1+h(s^*s+t^*t)/(s^*st^*t)]^{-1/2}.$$

The composite vector, made from two of these representations, has four components; and I try to construct the ''singlet'' state, as follows:

$$\psi=\{\psi_1,\psi_2,\psi_3,\psi_4\}=\{p,q,-q,r\}. \tag{134}$$

The operators from one space will act on the element pairs (1,2) and (3,4) while the operators from the other space act on the pairs (1,3) and (2,4). In the linear situation, one would set $p=r=0$; but in our representations that appears to involve singularities in the functions $1/U$ given above. So let us proceed with caution.

First, when we act with $Z=Z^1+Z^2$ on this ψ, we get

$$Z\psi=\{ip,0,0,-ir\} \tag{135}$$

and so we will want to take the limit of p and $r\rightarrow 0$; but not too quickly. Now apply the raising operator in one subspace

$$(X^1+iY^1)\psi=\{-iqf(p,-q),irf(q,r),iq(p^*/q^*)g(p,-q),ir(q^*/r^*)g(q,r)\} \tag{136}$$

and in the other,

$$(X^2+iY^2)\psi=\{iqf(p,q),iq(p^*/q^*)g(p,q),irf(-q,r),-ir(q^*/r^*)g(-q,r)\}. \tag{137}$$

Now we add these two vectors and try to get the result {0,0,0,0}. The first and fourth elements cancel exactly, because the functions f and g depend only on the absolute magnitudes of their arguments. The second and third elements are the same

$$irf(q,r)+iq(p^*/q^*)g(p,q). \tag{138}$$

Now, as we let $r\rightarrow 0$,

$$|r|f(q,r)\rightarrow 1/2[h(q^*q)]^{1/2}/|q|+O(r^2); \tag{139}$$

and, as we let $p \to 0$,

$$|p|g(p,q) \to -1/2[h(q^*q)]^{1/2}/|q| + O(p^2). \tag{140}$$

Thus we get the desired result (zero) by taking the limits indicated with the following fixed phase relation:

$$\text{phase}(p) + \text{phase}(r) = 2 \quad \text{phase}(q). \tag{141}$$

Applying $X - iY$ leads to the same result.

What we have achieved is a ''spin zero'' state, behaving in the usual manner of a linear representation of SU(2), by a particular construction of the composite of two ''spin 1/2'' states each of which belongs to the novel (and singular) nonlinear representation found above. Note that this works for any choice of the function h.

This appears to be a nontrivial result. I cannot make a linear triplet (spin 1) state out of these two nonlinear representations.

XIV. DISCUSSION

This last result is provocative for elementary particle physics: Could this have something to do with building mesons and baryons as composites of two or three confined quarks? A ''free particle'' is one that can exist in a universe with many other free particles which can be ignored if they are far away. This is where the usual building of product wave functions or product representations in the linear quantum theory is crucial. The new nonlinear group representations introduced in this paper might be used to describe individual particles, each one alone in its own universe; but they are likely to lead to logical inconsistency when allowed to ''exist'' with other free particles. Thus these things, if they exist in our universe, cannot be realized as individual free particles. But perhaps certain composites built out of them can transform according to the usual linear transformations of the relevant symmetry group—such as SU(2), which is the little group for massive particles in four dimensional space–time.

If such composites of two or more such nonlinear representations are possible, then we might predict a vast number of ''elementary'' particles to be found in nature. However, it is a matter of dynamics which of them might be stable or metastable. The simple rule for the physical decomposition of such composites—apart from any other selection rules—is that no single one should be allowed out. This means that the stable composites would likely be comprised of either two or three of these basic nonlinear things. This compares to the common picture of quarks. The calculation of the previous section produced something that could describe spin zero mesons.

I have tried to make a composite of three of these things that would behave as a linear representation of a ''spin 1/2'' particle; but I failed.

Obviously, there is much more work to be done, investigating other groups and other nonlinear representations.

APPENDIX A: LOGARITHMS

We shall adopt a definition of the generalized logarithm function to be the inverse of the generalized exponential function

$$L(E(A)) = A \tag{A1}$$

from which follows $E(L(A)) = A$.

We know from earlier work that, if $[A/B] = 0$, then, $E(A)E(B) = E(A+B)$ and therefore

$$L(E(A)E(B)) = L(E(A)) + L(E(B)). \tag{A2}$$

Or, equivalently: if M and N are two commuting operators $(MN=NM)$,

$$L(MN)=L(M)+L(N). \tag{A3}$$

One can also show that

$$L(BCB^{-1})=(B/L(C))B^{-1}. \tag{A4}$$

We now want to find the series expansion for $L(1+tA)$. For ordinary variables or for linear operators this series is well known; but for nonlinear operators A it will be more complicated. First set

$$1+tA=E(B)=1+B+1/2B/B+1/6(B/B)/B+\cdots \tag{A5}$$

and then expand B as a power series in t:

$$B=tB_1+t^2B_2+t^3B_3+\cdots \tag{A6}$$

and solve term by term. The result is

$$L(1+tA)=tA-t^2/2A/A+t^3/4\{(1/3)(A/A)/A+A/(A/A)\}-\cdots . \tag{A7}$$

One should note that the derivative of this series with respect to t does not give a result simply related to $(1+tA)^{-1}$, as would be the case for ordinary functions or linear operators.

This may be an appropriate place to mention that when we write infinite series of operators, their convergence may be limited to a finite domain of some parameter. Look, for example, at the solution of the simple time dependent nonlinear equation shown in I (30). This obviously has a singularity at some finite value of the time t and therefore the formal infinite series we use for $E(tA)$ has a limited domain of convergence centered around $t=0$, which is $E=1$. Thus we expect the series for the logarithm, discussed above, to have in general a limited domain of convergence. (The question of analytic continuation needs further study.)

APPENDIX B: SOME ADDITIONAL POWER SERIES

We previously studied the expansion

$$V=V(t)=(1-tA)^{-1}=\Sigma t^n V_n(A), \tag{B1}$$

where $V_0=1$, $V_1=A$ and the other V_n are given by the recursion formulas:

$$V_n=1/(n-1)\Sigma V_m/V_{n-m} \quad \text{with} \quad m=1,\ n-1. \tag{B2}$$

Now consider $W(t)=BV(t)=\Sigma t^n W_n(B;A)$. Write this as $WV^{-1}=B$ and take the time derivative of this equation. Proceeding as in the earlier study of V, we find the recursion formulas

$$nW_n=\Sigma W_m/V_{n-m}, \quad \text{with} \quad m=0,\ n-1 \quad \text{and} \quad W_0=B. \tag{B3}$$

Thus with the expressions V_n known, one can generate the expressions W_n. In fact, one can see that each $W_n(B;A)$ can be simply gotten from $V_{n+1}(A)$ by replacing each first (leftmost) A with B.

[Note that a similar simple result was found for the expression $BE(tA)$; however, one should certainly not imagine that this simple result is a general property of power series with nonlinear operators.]

Note that B, above, can be anything at all; and so this result can be used to construct series expansions in a recursive manner for any integral power.

$$P_K(t)=(1-tA)^K=P_{K+1}V=\Sigma t^n W_n(P_{K+1}(t);A). \tag{B4}$$

With

$$P_K(t)=\Sigma t^n P_{K,n} \tag{B5}$$

this yields

$$P_{K,n}=\Sigma W_m(P_{K+1,n-m};A), \quad \text{where} \quad m=0,n. \tag{B6}$$

This formula gives us directly the expansion terms for $K=-2$, $K=-3$, etc. To get the terms for positive K, solve recursively from the right hand side of this last equation, in the order of increasing n, noting that $W_0(B;A)=B$.

APPENDIX C: TRACE

In the world of linear operators, taking the trace of a matrix is a common procedure. The best that I have been able to do in trying to extend this to nonlinear operators is the following. Write the action of a general nonlinear operator A on a vector ψ in the form introduced in Sec. IX

$$A\psi|_k=\Sigma_{k'}R_{kk'}\psi_{k'}, \tag{C1}$$

where R, called an f-matrix, is a matrix of functions whose arguments are all the components of ψ and ψ^*. Since the original description in Sec. III involved n functions (in an n-dimensional vector space) whereas Eq. (C1) involves n^2 functions, it is apparent that there is much redundancy in how the f-matrix R is expressed; but this is no hindrance here.

Now take another operator, call it B, and represent it as in Eq. (C1) by another f-matrix, call it S. Calculate the result of the slash product A/B, following the general procedure shown in Sec. III, and find the result

$$A/B\psi|_k=\Sigma_{k'}[\Sigma_{k''}R_{kk''}S_{k''k'}+d(S)R_{kk'}]\psi_{k'}, \tag{C2}$$

where the differential operator $d(S)$, which acts on the f-matrix R, is

$$d(S)=\Sigma_{jj'}[S_{jj'}\psi_{j'}\partial/\partial\psi_j+S_{jj'}{}^*\psi_{j'}{}^*\partial/\partial\psi_j{}^*]. \tag{C3}$$

This thing $d(S)$ is really the same as Eq. (41); but with lower case d I emphasize that it is a scalar, not a matrix, like R or S, nor a vector, like ψ.

We are now ready to give a definition of the trace of a general nonlinear operator. If an operator A is represented, as in Eq. (C1), by an f-matrix called R, and we write that relation as $A::>\{A\}=R$, then define

$$\mathrm{Tr}\{A\}=\Sigma_k R_{kk}. \tag{C4}$$

The result, seen from Eq. (C2) is this statement:

$$\text{If} \quad \mathrm{Tr}\{A\}=0, \quad \text{then} \quad \mathrm{Tr}\{A/B\}=\mathrm{Tr}(\{A\}\{B\}); \tag{C5}$$

and from this it follows that:

$$\text{If} \quad \mathrm{Tr}\{A\}=\mathrm{Tr}\{B\}=0, \quad \text{then} \quad \mathrm{Tr}\{[A/B]\}=0. \tag{C6}$$

This result (C6) is not as general as the familiar statement, for linear operators represented by ordinary matrices, that the trace of any commutator vanishes. But this result is just enough to be useful in the study of Lie algebras. It says that it is consistent for all the f matrix representations of a Lie algebra to have zero trace—just as in the linear case. (Whether all representations can be put into this form is another question.)

The condition $\mathrm{Tr}\{A\}=0$ is stronger than is necessary in order to make the second term of Eq. (C2) vanish when taking the trace of $\{A/B\}$. For linear operators it will always vanish. For nonlinear operators, if one says that $\mathrm{Tr}\{A\}$ is a constant, that is sufficient. And there are other special cases, for example: let $\mathrm{Tr}\{A\}$ depend on only the one variable $\langle\psi|\psi\rangle$, and furthermore let the operator B be anti-Hermitian.

APPENDIX D: THE DERIVATIVE OPERATOR

If the state vector depends upon a continuous variable x, as in $\psi=\{\psi_k(x)\}$, then we might want to use the derivative operator d/dx, which is itself a linear operator. If we have some nonlinear operator A which also acts on ψ, then in general we would not expect d/dx to commute with A. (The derivative of the square of some function is not the same as the square of the derivative of that function.) However, it can be readily shown that the slash commutator, $[(d/dx)/A]$, vanishes, provided that the operator A is itself independent of the coordinate x and acts upon the $\psi_k(x)$ locally in x. To see that this result is not trivial, note that the operator d^2/dx^2 does not have this property.

[1] C. Schwartz, J. Math. Phys. **38**, 484–500 (1997).
[2] I. Bialynicki-Birula and J. Mycielski, Ann. Phys. **100**, 62–93 (1976).
[3] S. Weinberg, Ann. Phys. **194**, 336–386 (1989); see further references given in both of these papers.