1971

A point and local position operator

Bernice Black Durand

Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Elementary Particles and Fields and String Theory Commons

Recommended Citation

https://lib.dr.iastate.edu/rtd/4876

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
DURAND, Bernice Black, 1942-
A POINT AND LOCAL POSITION OPERATOR.

Iowa State University, Ph.D., 1971
Physics, elementary particles

University Microfilms, A XEROX Company, Ann Arbor, Michigan
A point and local position operator

by

Bernice Black Durand

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major Subject: High Energy Physics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State University
Of Science and Technology
Ames, Iowa
1971
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>TITLE</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II</td>
<td>FORMALISM</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Hyperplanes</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Little Group Theory and Two Special</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Transformations</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>POSITION OPERATORS</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Formal Approach</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Historical Approach</td>
<td>32</td>
</tr>
<tr>
<td>IV</td>
<td>A POINT AND LOCAL POSITION OPERATOR</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>Pointness</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>Locality</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>A Point and Local Operator</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>The Hyperplane Operator</td>
<td>86</td>
</tr>
<tr>
<td>A</td>
<td>NOTATION</td>
<td>90</td>
</tr>
<tr>
<td>B</td>
<td>POINCARE IDENTITIES AND COMMUTATION RELATIONS</td>
<td>92</td>
</tr>
<tr>
<td>C</td>
<td>TETRAD IDENTITIES</td>
<td>94</td>
</tr>
<tr>
<td>D</td>
<td>THE SPIN OPERATOR S</td>
<td>95</td>
</tr>
<tr>
<td>E</td>
<td>PROPERTIES OF L(η'; η)</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td></td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>BIBLIOGRAPHY</td>
<td>104</td>
</tr>
<tr>
<td></td>
<td>ACKNOWLEDGMENTS</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

In classical physics the position of a particle can be measured to as great an accuracy as measuring instruments allow. The position of a system or property is usually considered to have all the properties of a coordinate, and to be related to velocity and momentum in the customary way,

\[
\frac{\text{d}x}{\text{d}t} = \frac{\text{d}p}{\text{d}t} = \frac{V}{m}. \quad (1.1)
\]

Position is also taken as canonically conjugate to momentum,

\[
[x_i, p_j] = \delta_{ij}. \quad (1.2)
\]

Nonrelativistically the center of mass is an example of a position which satisfies these properties plus others which will be enumerated in Chapter III.

The role and properties of position in relativistic quantum mechanics are less clear. There is no guarantee that a position is even an observable in the practical sense. One may not be able to do an experiment which measures it. If particles have some extent of the order of the Compton wavelength \( \lambda = \frac{h}{mc} \) (about \( 4 \times 10^{-11} \) cm. for the electron), a measurement within that distance must be made with photons or other particles of a frequency \( \omega \sim m \), in other words with energy equivalent to the mass of the measured particle. Such photons will obviously disturb the system they are measuring.
Because of these problems, position is not a very fashionable quantity in fundamental particle theory. Momentum and angular momentum can be measured, with the consequence that the coordinate or position representation has been dropped in most applications in favor of the momentum representation. It is nevertheless of interest to construct reasonable position operators for relativistic quantum mechanical systems. Such operators may provide a new way of describing particle interactions. The construction of these operators also displays the way in which the various properties of particles can be distributed in space.

The problems encountered in constructing relativistic position operators have led to as many operators as there are authors, although many of the operators can be shown to be equivalent. Lorentz covariance is one property which is subject to several interpretations. The research presented in Chapters II and IV is an attempt to put a certain kind of position operator into a formalism where its covariance properties are carefully displayed, thus allowing other properties to be satisfied without obscuring covariance.

The order of the chapters was chosen to build up to the results in Chapter IV in a natural way. Chapter II is a development of hyperplane formalism with no reference to position operators. This formalism is an extension and modification of Fleming's \(^1\) hyperplane formalism. The work in Chapter II is original.

Chapter III, on position operators, has two sections. The first is a formal statement of the requirements which can be imposed on position and the resultant operators which fulfill certain of the
conditions. This section will acquaint the reader with the general problem. There follows a "brief" chronological review of the literature which has been expanded with the thought of using it as the basis of a review paper. Much attention is paid to the early (pre-1950) works because a careful study of subsequent works has revealed that very few new ideas have been added since 1950. Different approaches and motivations have led to the same operators repeatedly.

Chapter III contains no original work, however the results of so many authors are included that it was necessary to do a considerable amount of interpretation.

Chapter IV contains much original work. It brings the hyperplane formalism and position operators together with the construction of a "point" and "local" operator. Pointness and locality are discussed in detail. The Dirac operator is analysed for these properties and then generalized to any spin. Hyperplane formalism is used to clarify covariance properties of the position and the results are discussed.

Notation will be found in Appendix A. Appendix B and Appendix C contain various identities. Appendix D is a derivation of a spin operator and Appendix E discusses a hyperplane transformation.
CHAPTER II

FORMALISM

Hyperplanes

One of the properties which can be postulated for a position operator is that it transform under an Inhomogeneous Lorentz Transformation (ILT) as a four-vector. Unfortunately, when certain combinations of other properties are postulated the resultant operator can appear non-covariant. This problem especially arises when the operator in question is defined as a three-vector time-dependent location of some spatial property. Then the "zeroth" component of a four-vector position would be the time $t$. Since $t$ is a c-number and $\mathbf{x}$ is a vector operator, an ILT will mix components to give a new position operator in which the space components are not all operators and the time component is not all c-number.

It is possible to eliminate many problems of apparent non-covariance by using a formalism which expresses the operators as functions of space-like surfaces (hyperplanes) rather than as functions of time. The two expressions are equivalent, since to say that the time coordinate of a point is $t = t_0$ is to say that the point lies on a spacelike hyperplane satisfying the equation $t = t_0$.

The hyperplane formalism of Fleming \cite{Fleming} at first glance seems to complicate the definitions of operators. However once it is mastered
it makes covariance easier to attain and dynamical position operators easier to visualize.

A hyperplane is a space-like three-dimensional surface in space-time. The time-like unit normal to a hyperplane defines its orientation.

\[
\begin{align*}
\eta^\mu &= \tau \\
x^\mu \eta_\mu &= \tau
\end{align*}
\]  

(2.1)

is drawn such that \( \eta^\mu \) is a time-like unit normal with \( \eta_\mu \eta^\mu = 1 \) and \( x^\mu \) is the four-vector from the origin to any point in the hyperplane. Implicit in this definition is the choice of the arbitrary origin and \( (\vec{x},t) \) axes to which the hyperplane parameters \( \eta^\mu \) and \( \tau \) are referred. Families of hyperplanes which share \( \eta \) but have varying \( \tau \) parametrize space-time.

A variation on this formalism is to imagine observers "sitting" on hyperplanes, so that the time-like unit normal to a hyperplane is some observer's time axis. Thus space-time is filled with possible coordinate systems, any of which may be used by an observer; and the hyperplane parameters can label operators or states observed by the different observers. The "preferred" coordinate system, which defines the origin for the parameters, is occupied by a metaobserver who observes observers.
The transformation from one hyperplane-system to another is simply the ILT. Thus the hyperplane formalism is a way of looking at the passive interpretation of ILT's. In order to study covariance under ILT's, it is necessary to distinguish between properties which depend on which observer is observing them and those which are observer-independent. Thus a new notation is introduced to label operators and tensors doubly: one label refers to the observer and one to the metaobserver. The metaobserver will be referred to by Greek letters \( \mu, \nu, \rho, \sigma, \ldots \) \((\mu = 0, 1, 2, 3)\) or Latin letters \( i, j, k, \ldots \) \((i = 1, 2, 3)\). The observers will be labeled by Greek letters \( \alpha, \beta, \gamma, \ldots \) or Latin letters \( a, b, c, \ldots \).

We define a tetrad, or a set of four vectors,

\[ \eta^\mu_\alpha. \]

Here \( \eta_\alpha \) means four unit vectors \((\hat{t}, \hat{x}_1, \hat{x}_2, \hat{x}_3)\) as seen by the observer \( \alpha \). They are his choice of coordinate axes. The \( \mu \) then refers to the \( \mu \)-th component, as seen by the metaobserver \( \mu \), of the \( \eta_\alpha \) vector. Thus

\[ \eta_0 = \eta^0 = (1, 0, 0, 0) \quad \eta_1 = (0, -1, 0, 0) \]

\[ \eta^1 = (0, 1, 0, 0). \quad (2.2) \]

in observer \( \alpha \)'s system. \( \eta^1_2 \) is the projection of observer \( \alpha \)'s \( y \)-axis unit vector along metaobserver \( \mu \)'s \( x \)-axis. A word of caution: \( \eta^\mu_\alpha \) is neither a vector nor a tensor; it is a set of four unit vectors. Tetrad identities are listed in Appendix C.
We have labeled the four axes of a hyperplane observer by $\eta^\mu_\alpha$. To fix his origin we need the further label $a^\mu$, the four vector from the metaobserver's origin to the observer's origin.

Thus $\{a^\mu, \eta^\mu_\alpha\}$ pinpoints a specific observer, as seen by the metaobserver. The diagram illustrates this notation.

One can see that $a^\mu$ and $\eta^\mu_\alpha$ are as good parameters for ILT's as $d^\mu$ and $\Lambda^\mu_\nu$, and it will be shown how to use them as such.

We introduce further state labels for the state $|m s \varepsilon \vec{p} \lambda\rangle$, namely $\eta^\mu_\alpha$ and $a^\mu$, which tell which observer is determining the eigenvalues used to label the state. Also we use the four-vector $p^\alpha$ instead of the three-vector $\vec{p}$. This injects some redundancy into the labeling but proves easier to work with. $p^\alpha$ is some $p$ chosen by the observer $\{a^\mu, \eta^\mu_\alpha\}$ who labels his axes with $\alpha$.

We can drop $m s \varepsilon$ because they look the same to all observers and we are not mixing irreducible representations. Thus a state now looks like

$$|\eta^\mu_\alpha a^\mu p^\alpha \lambda\rangle .$$

This is the state

$$|p^\alpha \lambda (m s \varepsilon)\rangle$$

as seen by observer $\{a^\mu, \eta^\mu_\alpha\}$.
The metaobserver would see the vector

\[ p^\mu = \eta^\mu_\alpha \, p^\alpha \]  \hspace{1cm} (2.3)

while the observer sees \( p^\alpha \). Another observer \( \{ a^\mu_\lambda, \eta^\mu_\alpha \} \) could select a state

\[ |\eta^\mu_\alpha \, a^\mu_\lambda \, p^\alpha \lambda> \]

and think he has chosen the same state as the first observer. However the metaobserver will see the \( p^\alpha \) chosen by the second observer as

\[ p'^\mu = \eta^\mu_\alpha \, p^\alpha \neq p^\mu \]  \hspace{1cm} (2.4)

The states are formally defined by

\[ p^\alpha |\eta^\mu_\alpha \, a^\mu_\lambda \, p^\alpha \lambda> = p^\alpha |\eta^\mu_\alpha \, a^\mu_\lambda \, p^\alpha \lambda> \]  \hspace{1cm} (2.5)

\[ p^\mu |\eta^\mu_\alpha \, a^\mu_\lambda \, p^\alpha \lambda> = \eta^\mu_\alpha \, p^\alpha |\eta^\mu_\alpha \, a^\mu_\lambda \, p^\alpha \lambda> \]  \hspace{1cm} (2.6)

\[ S^3 |\eta^\mu_\alpha \, a^\mu_\lambda \, p^\alpha \lambda> = \lambda |\eta^\mu_\alpha \, a^\mu_\lambda \, p^\alpha \lambda> \]  \hspace{1cm} (2.7)

\( S^3 \) is the \( \alpha = 3 \) component of the spin operator \( \vec{S} \) defined in Appendix D.

It will be convenient to shorten the notation from

\[ |\eta^\mu_\alpha \, a^\mu_\lambda \, p^\alpha \lambda> \]  to \[ |\eta \, a \, p^\alpha \lambda> \]. One must bear in mind that \( \eta \) and \( a \) do not add any group-theoretical information to a state; they merely specify who chose and labeled the state.
Little Group Theory and Two Special Transformations

When doing little group theory it is necessary to select some special momentum \( \hat{p}^\alpha \), usually defined as \((\epsilon m, \vec{0})\), and find the operators in its little group, i.e. which leave it alone. When observer \( \alpha \) chooses a

\[
\hat{p}^\alpha = \epsilon m (1, \vec{0})
\]  

(2.8)

the metaobserver \( \mu \) will see it as

\[
\hat{p}_\mu = \epsilon m \eta^0_\mu .
\]

(2.9)

We have now set up the necessary background to introduce two special Lorentz transformations defined to operate on states \( |\eta^\mu_\alpha a^\mu p^\alpha \rangle \). First we define these LT's on states for which \( a^\mu = 0 \), then generalize to the cases \( a^\mu \neq 0 \). Thus a state can be labeled \( |\eta^\mu_\alpha p^\alpha \rangle \).

When we mean for a transformation to act on vectors or operators we shall write it in matrix form, e.g. \( \Lambda^\mu_\nu \). When it is to act on a state it will be written without super or subscripts, e.g. \( \Lambda \). \( \Lambda^\mu_\nu \) is a parametrization of the unitary operation \( \Lambda \); the two are related by

\[
\Lambda = e^{-1/2 i \theta^{\mu\nu}} \omega_{\mu\nu}
\]

(2.10)

such that

\[
\Lambda^{-1} p^\mu_\Lambda = \Lambda^\mu_\nu p^\nu
\]

(2.11)

for any four-vector operator \( p^\mu \) and where \( \omega_{\mu\nu} \) and \( \Lambda_{\mu\nu} \) must be found for any given \( \Lambda \).
The first transformation is the familiar pure LT with no rotations which takes the state $|\eta^\mu_\alpha p^\alpha \lambda\rangle$ into the state $|\eta^\mu_\alpha \hat{p}^\alpha \lambda\rangle$, where $\hat{p}^\alpha$ is the arbitrarily chosen "special momentum" for little group theory. We call this transformation $\mathcal{L}(\eta;p)$ and define it by the equation

$$|\eta^\mu_\alpha p^\alpha \lambda\rangle = \mathcal{L}^{-1}(\eta;p) |\eta^\mu_\alpha \hat{p}^\alpha \lambda\rangle \quad (2.12)$$

Choosing $\hat{p}^\alpha = 0$ makes $\mathcal{L}(\eta;p)$ also the transformation to the rest system of particle, where $\hat{p}_{\text{rest}} = 0$.

To generalize this transformation to the state $|\eta^\mu_\alpha a^\mu p^\alpha \lambda\rangle$ we need some notation of ILT's. A general ILT is represented by $(d,A)$ with the rule

$$(d_2,A_2)(d_1,A_1) = (d_2 + A_2 d_1, A_2 A_1). \quad (2.13)$$

For any ILT $(d,A)$ it is true that

$$(d,A) = (d,1)(0,A). \quad (2.14)$$

Thus when we wish to study how an observer $\{a^\mu, \eta^\mu_\alpha\}$ and states $|\eta^\mu_\alpha a^\mu p^\alpha \lambda\rangle$ transform under an ILT we can deal with the problem in two steps: a translation $(d,1)$ and a homogeneous LT $(0,A)$.

Let parentheses ( ) mean that an observer is performing an operation and brackets { } mean that the metaobserver is performing it. Now we show how an ILT $(d,A)$ looks to observers $\{a,\eta\}$ and $\{a',\eta'\}$. (In referring to observers $\{a^\mu, \eta^\mu_\alpha\}$ it is convenient to abbreviate to $\{a,\eta\}$.)
The results

\[(d,1) = \{d,1\}\]  

and

\[(0,\Lambda) = \{a - \Lambda a,\Lambda\}\]  

(2.15)  
(2.16)

can be seen by inspection of the two diagrams below.

These can also be written

\[(d^\alpha,1) = \{d^\mu,1\}\]  

\[(0,\Lambda B^\alpha) = \{a^\mu - \Lambda^\mu \nu a^\nu,\Lambda^\mu \nu\}\]  

(2.17)  
(2.18)

to represent the matrix elements rather than the general unitary transformations.

Then

\[(d,\Lambda) = (d,1)(0,\Lambda)\]  

(2.19)

\[= \{d,1\}\{a - \Lambda a,\Lambda\} = \{d + a - \Lambda a,\Lambda\}\] .
Similarly
\[ \{d_\alpha \Lambda \} = (d + a - \Lambda a_\alpha \Lambda) . \tag{2.20} \]

However "a" in observer brackets ( ) is "-a" in metaobserver brackets \{ \}. Or,
\[ a^\alpha = -\eta^\alpha_\mu a^\mu. \tag{2.21} \]

We can now define \( \xi(\eta^\mu_\alpha, a^\mu; p) \) such that
\[ |\eta^\mu_\alpha a^\mu p^\alpha \lambda> = \xi^{-1}(\eta^\mu_\alpha, a^\mu; p) \eta^\mu_\alpha a^\mu \hat{p}^\alpha \lambda> \tag{2.22} \]

and
\[ (0, \xi^{-1}(\eta;p)) = (a - \xi^{-1}(\eta;p) a, \xi^{-1}(\eta;p)) \]
\[ = \xi^{-1}(\eta, a; p). \tag{2.23} \]

So when observer \{a^\mu, \eta^\mu_\alpha\} performs \( \xi^{-1}(\eta;p) \) the metaobserver sees \( \xi^{-1}(\eta, a; p) \).

Thus
\[ \xi^{-1}(\eta;p) = \xi^{-1}(\eta, 0; p). \tag{2.24} \]

The second transformation \( L(\eta';\eta) \) takes a state \[|\eta^\mu_\alpha p^\alpha \lambda>\] to the corresponding state \[|\eta'^\mu_\alpha \hat{p}^\alpha \lambda>\] of another observer. These states are defined by
\[ p^\alpha |\eta^\mu_\alpha p^\alpha \lambda> = p^\alpha |\eta'^\mu_\alpha p^\alpha \lambda> \tag{2.25} \]
\[ p'^\alpha |\eta'^\mu_\alpha p^\alpha \lambda> = p^\alpha |\eta'^\mu_\alpha p^\alpha \lambda> \tag{2.26} \]

and
\[ S^3 |\eta^\mu_\alpha p^\alpha \lambda> = \lambda |\eta^\mu_\alpha p^\alpha \lambda> \tag{2.27} \]
\[ S'^3 |\eta'^\mu_\alpha p^\alpha \lambda> = \lambda |\eta'^\mu_\alpha p^\alpha \lambda> \tag{2.28} \]
The question is how these states relate to each other.

Since

\[ p^\mu |p^\nu \lambda> = \eta_{\alpha}^\nu p^\alpha |p^\nu \lambda> \] (2.29)

and

\[ p^\mu |p_1^\nu \lambda> = \eta_{\alpha}^\nu p^\alpha |p_1^\nu \lambda> \]

\[ = \eta_{\gamma}^\nu \eta_{\alpha}^\nu p^\alpha |p_1^\nu \lambda> \] (2.30)

we can define

\[ \eta^\mu_{\gamma} \eta^\nu_{\gamma} \equiv L(n';n)^\mu_{\nu} \] (2.31)

so that

\[ p^\mu |p_1^\nu \lambda> = L(n';n)^\mu_{\nu} \eta_{\alpha}^\nu p^\alpha |p_1^\nu \lambda> \] (2.32)

Then

\[ L(n';n)^\mu_{\nu} = \eta^\mu_{\gamma} \eta^\nu_{\nu} \] (2.31)

is a Lorentz transformation.

It is always true that if

\[ p^\mu |p^\mu \lambda> = p^\mu |p^\mu \lambda> \] (2.33)

and \( \Lambda \) is a Lorentz transformation then

\[ p^\mu \Lambda |p^\mu \lambda> = \Lambda^\mu_{\nu} p^\nu \Lambda |p^\mu \lambda> \] (2.34)

Hence if

\[ |n'_{\alpha} p^\alpha \lambda> = L(n';n) |n_{\alpha} p^\alpha \lambda> \] (2.35)

then

\[ p^\mu |n'_{\alpha} p^\alpha \lambda> = L(n';n)^\mu_{\nu} p^\nu |n_{\alpha} p^\alpha \lambda> \] (2.36)
and the equations are consistent. This leads us to tentatively define

\[ |n_{\alpha}^\mu \hat{p}^\alpha \lambda> = L(n';n) |n_{\alpha}^\mu \hat{p}^\alpha \lambda> , \]  

(2.37)

which can be verified by testing the eigenvalues of \( p_{\beta}^\alpha \) and \( S^3 \) on \( L(n';n) |n_{\alpha}^\mu \hat{p}^\alpha \lambda> \). See Appendix E for this proof.

Some results with \( L \) follow.

\[ L^{-1} (n';n) p_{\alpha}^\beta L(n';n) = p_{\alpha} \]  

(2.38)

\[ L(n';n)_{\beta}^\alpha = n_{\mu}^\alpha n_{\beta}^\mu \]  

(2.39)

\[ L^{-1} (n';n') = L(n';n) \]  

(2.40)

\[ L(n;n') |n' p_{\alpha}^\lambda > = |n p_{\alpha}^\lambda > \text{ (up to a phase).} \]  

(2.41)

The proof of the equation (2.41) is useful for seeing the proof of the more general result.

\[ |n p_{\alpha}^\lambda > = L^{-1} (n;p) |n \hat{p}_{\alpha}^\lambda > \]

\[ = L^{-1} (n;p) L(n;n') \xi(n';p) |n' p_{\alpha}^\lambda > . \]  

(2.42)

We need to show that this equals \( L(n;n') |n' p_{\alpha}^\lambda > \). Throughout this proof it must be kept in mind that \( \xi_{\beta}^\alpha (n';p) = \xi_{\beta}^\alpha (n;p) \) and \( \xi^\mu_{\nu} (n;p) = n_{\alpha}^\mu \xi_{\beta}^\alpha (n';p) n_{\nu}^\beta \). Here \( n \) specifies the observer, as in equations (2.3) and (2.4). The notation can be confusing.
First multiply by $L^{-1}L$ from the right to get

$$L^{-1}(n;p) L(n;n') \cdot L(n';p) L^{-1}(n;n') L(n;n').$$

Now the middle three terms must be shown to equal $\varphi(n;p)$.

Let them act on an arbitrary vector $a^\nu$:

$$[L(n;n') \varphi(n';p) L^{-1}(n;n')]^\mu a^\nu.$$

$$L(n;n')^\mu \varphi(n';p)^\rho L^{-1}(n;n')^\sigma = n^\mu_\alpha n^\alpha_\rho \varphi(n';p)^\rho n^\sigma_\beta n^\beta_\nu. \tag{2.43}$$

But since $\varphi(n';p)$ is the transformation to the rest system, $\varphi(n';p)^\rho_\sigma$ can be written as functions of $n^\alpha_\nu p^\mu = p'^\alpha$. Then

$$n^\alpha_\rho \varphi(n';p)^\rho_\sigma n^\sigma_\beta$$
is a function of $p^\mu$ and

$$n^\mu_\alpha (n^\alpha_\rho \varphi(n';p)^\rho_\sigma n^\sigma_\beta) n^\beta_\nu$$
is a function of $n^\mu_\nu$, i.e. is $\varphi(n;p)^\mu_\nu$. Thus

$$[L(n;n') \varphi(n';p) L^{-1}(n;n')]^\mu a^\nu = \varphi(n;p)^\mu_\nu a^\nu \tag{2.44}$$
for any vector $a^\nu$. So we have now:

$$\varphi^{-1}(n;p) \varphi(n;p) L(n;n') |n^\alpha p^\lambda> = |n^\alpha p^\lambda> \tag{2.41}$$

$$= L(n;n') |n^\alpha p^\lambda>. \quad \text{QED.}$$

Now we define the transformation $L(n,a;n',a')$, which is the generalization of $L(n;n')$ and takes the state $|n^\mu_\alpha a^\mu_\alpha \hat{p}^\alpha >$ into the state $|n^\mu_\alpha a^\mu_\alpha \hat{p}^\alpha >$. See the following diagram.
By inspection we see that
\[ L(n,a;n',a') = \{ a - L(n;n') a', L(n;n') \} \]
\[ = \{ a, 1 \} \{ 0, L(n;n') \} \{ -a', 1 \} . \]  
(2.45)

To check the definition
\[ |n \ a \ p^\alpha \lambda> \equiv L(n,a;n',a') |n' \ a' \ p^\alpha \lambda> \]  
(2.46)
is straightforward, once we establish the formulation of the little group theory in this notation.

When we want to apply a general LT \( \Lambda \) to a state \( |n \ p^\alpha \lambda> \) (note this is the homogeneous case) we will use \( \mathcal{L}(n;p) \) as follows:
\[ \Lambda |n \ p^\alpha \lambda> = \mathcal{L}^{-1}(n;\Lambda p) \mathcal{L}(n;\Lambda p) \Lambda^{-1}(n;p) |n \ p^\alpha \lambda> \]
\[ = \mathcal{L}^{-1}(n;\Lambda p) R(n,\Lambda,p) |n \ p^\alpha \lambda> \]  
(2.47)
\[ = \sum_{\lambda'} D^S_{\lambda',\lambda}(R(n,\Lambda,p)) |n \ (\Lambda p)^\alpha \lambda'> . \]

Here
\[ R(n,\Lambda,p) = \mathcal{L}(n;\Lambda p) \Lambda^{-1}(n;p) \]  
(2.48)
is a rotation in the little group of \( \hat{p} \). \( D^S_{\lambda',\lambda} \) can be found by using Euler angle parameters relative to \( n^U_\alpha \). The \( \lambda \) is a summation index and does not refer the the \( n' \) observer.
One useful result is that

$$D_{\lambda',\lambda}^{S} (R(n',\Lambda,p)) = D_{\lambda',\lambda}^{S} (R(n,\Lambda,p)). \quad (2.49)$$

This is essentially because these are just coefficients which relate the state $|\hat{p}^\alpha \lambda\rangle$ to the states $|(\Lambda p)^\alpha \lambda'\rangle$ and both the primed and unprimed observers will calculate the same coefficients.

It is worth noting that "active" or "passive" transformations need not be mentioned because the observer-metaobserver approach eliminates this distinction by specifying who is making the transformation and on what.

Now to show that

$$|n \alpha p^\alpha \lambda\rangle = L(n,a;n',a') |n' \alpha p^\alpha \lambda\rangle. \quad (2.50)$$

$$= \varepsilon^{-1} (n,a;p) |n \alpha p^\alpha \lambda\rangle$$

$$= \varepsilon^{-1} (n,a;p) L(n,a;n',a') |n' \alpha p^\alpha \lambda\rangle$$

$$= \varepsilon^{-1} (n,a;p) L(n,a;n',a') \varepsilon(n',a';p) |n' \alpha p^\alpha \lambda\rangle$$

$$= \{a - \varepsilon^{-1}(n;p)a, \varepsilon^{-1}(n;p)\} L(n,a;n') \{a - L(n;a')a', L(n,a')\}$$

$$\{a' - L(n';p)a', \varepsilon(n';p)\} |n' \alpha p^\alpha \lambda\rangle$$

$$= \{a - \varepsilon^{-1}(n;p)\} L(n,a;n') \varepsilon(n';p)a',$$

$$\varepsilon^{-1}(n;p) L(n,a;n') \varepsilon(n';p) |n' \alpha p^\alpha \lambda\rangle.$$

We already showed that

$$\varepsilon^{-1}(n;p) L(n,a;n') \varepsilon(n';p) = L(n,a;n') \quad (2.42) \text{ and } (2.41)$$
\[ |n \ a \ p^\alpha \lambda > = \{a - L(n;n')a', L(n;n')\} |n' \ a' \ p^\alpha \lambda > \]
\[ = L(n,a;n',a') |n' \ a' \ p^\alpha \lambda > \]  \hspace{1cm} (2.50)
\[ = \{a,1\}{0,L(n;n')}\{-a',1\} |n' \ a' \ p^\alpha \lambda > \]
\[ = e^{-i \ p^\mu a_{\mu}} L(n;n') \ e^{i \ p^\mu a'_{\mu}} |n' \ a' \ p^\alpha \lambda >. \]  \hspace{1cm} (2.51)

Again we leave out the phase factor which each state carries, but the relative phase between states is seen from
\[ e^{i \ p^\mu a_{\mu}} |n \ a \ p^\alpha \lambda > = L(n;n') e^{i \ p^\mu a'_{\mu}} |n' \ a' \ p^\alpha \lambda >. \]  \hspace{1cm} (2.52)

Now for a general ILT acting on a general state. A general ILT can be broken into a translation and a pure LT. The translation \{d,l\} can be expressed by
\[ \{d,l\} |n \ a \ p^\alpha \lambda > = e^{-i \ p^\mu d_{\mu}} |n \ a \ p^\alpha \lambda >. \]  \hspace{1cm} (2.53)

In hyperplane variables this becomes
\[ \{d,l\} |n \ a \ p^\alpha \lambda > = (d,l) |n \ a \ p^\alpha \lambda > \]
\[ = e^{-i \ n^\mu_{\alpha} p^\alpha \ n^\gamma_{\mu} d_{\gamma}} |n \ a \ p^\alpha \lambda > \]
\[ = e^{-i \ p^\alpha d_{\alpha}} |n \ a \ p^\alpha \lambda > \]  \hspace{1cm} (2.54)
\[ = e^{-i \ p^\alpha d_{\alpha}} |n \ a \ p^\alpha \lambda > . \]

Here \( p^\alpha \), \( p^\alpha \), and \( d_{\alpha} \) are all as seen by observer \{a^\mu, n^\mu\}.

A useful identity is
\[ \{d,l\}{a' - L(n';n)a, L(n';n)} = \{a' - L(n';n)a, L(n';n)\}{L^{-1}(n';n)d,l} \]  \hspace{1cm} (2.55)
so that
\[ \langle d, l \mid n', a' \mid p^\alpha \rangle = \langle d, l \mid L(n', a' ; n, a) \mid n \rangle \mid p^\alpha \rangle \]
\[ = L(n', a' ; n, a) \{ L^{-1} (n' ; n) d, 1 \} \mid n \rangle \mid p^\alpha \rangle. \tag{2.56} \]

or
\[ e^{-i \int P^\mu d\mu} \langle n', a' \mid p^\alpha \rangle = L(n', a' ; n, a) e^{-i \int P^\mu (L^{-1} (n' ; n) d) d\mu} \mid n \rangle \mid p^\alpha \rangle. \tag{2.57} \]

Now let the observer \{a^\mu, n^\mu\} perform a homogeneous LT \( \Lambda_\beta^\alpha \). Then we define
\[ \Lambda(n)^\mu_\nu = n^\mu_\alpha \Lambda_\beta^\alpha n^\beta_\nu \tag{2.58} \]
and
\[ \Lambda(n, a) = \{ a - \Lambda(n)a, \Lambda(n) \}. \tag{2.59} \]

Next we define a special rotation \( \tilde{R} \) to be
\[ \tilde{R}(n) = \varepsilon (n; \Delta p) \Lambda(n) \varepsilon^{-1} (n; \Delta p). \tag{2.60} \]

Then the observer performing \( \Lambda \) on a state looks like
\[ \Lambda(n, a) \mid n \rangle \mid p^\alpha \rangle = \{ a - \Lambda(n)a, \Lambda(n) \} \mid n \rangle \mid p^\alpha \rangle \]
\[ = \{ a - \varepsilon^{-1} (n; \Delta p)a, \varepsilon^{-1} (n; \Delta p)\{ a - \tilde{R}(n)a, \tilde{R}(n) \}\mid n \rangle \mid p^\alpha \rangle \]
\[ = \varepsilon^{-1} (n, a; \Delta p) (0, \tilde{R}) \mid n \rangle \mid p^\alpha \rangle \]
\[ = \varepsilon^{-1} (n, a; \Delta p) \sum_{\lambda'} D^S_{\lambda', \lambda} (\tilde{R}) \mid n \rangle \mid p^\alpha \lambda' \rangle \]
\[ = \sum_{\lambda'} D^S_{\lambda', \lambda} (\tilde{R}) \mid n \rangle \mid (\Delta p)^\alpha \lambda' \rangle. \tag{2.61} \]
Now it can be shown that

\[ L(n, a; n', a') \Lambda(n', a') = \Lambda(n, a) L(n, a; n', a') \]  

(2.62)

so

\[ \Lambda(n, a) |n a p^\alpha \lambda> = \Lambda(n, a) L(n, a; n', a') |n' a' p^\alpha \lambda> \]

\[ = L(n, a; n', a') \Lambda(n', a') |n' a' p^\alpha \lambda> \]

\[ = \sum_{\lambda'} D^S_{\lambda'} \Lambda(\tilde{R'} |n a (\Lambda p)^\alpha \lambda'> \quad (2.63) \]

where \( \tilde{R} = \tilde{R}(n') \). We stated earlier that

\[ D^S_{\lambda'} \Lambda(\tilde{R}) = D^S_{\lambda'} \Lambda(\tilde{R}'), \]

(2.49)

when \( a = a' = 0 \). This still holds, as \( \tilde{R} \) has no dependence on \( a \) or \( a' \).

Combining translations and LT's we get the ILT

\[ (d, \Lambda) |n a p^\alpha \lambda> = (d, \Lambda)(0, \Lambda) |n a p^\alpha \lambda> \]

\[ = \{d, \Lambda\} \Lambda(n, a) |n a p^\alpha \lambda> \]

\[ = e^{-i(\Lambda p)^\alpha d_{\alpha}} \sum_{\lambda'} D^S_{\lambda'} \Lambda(\tilde{R}(\eta)) \]

\[ |n a (\Lambda p)^\alpha \lambda'>. \quad (2.64) \]
CHAPTER III

POSITION OPERATORS

Formal Approach

The idea of position in relativistic quantum mechanics is a carryover from classical mechanics and can be understood by comparing the classical structure with various possible generalizations to relativistic and quantum mechanical structures.

In classical nonrelativistic mechanics of spinless, non-interacting particles one can define the following algebra of ten generators of various transformations:

\[ H = \frac{p^2}{2m} \]

\[ \hat{P} = \hat{p} \]

\[ \hat{J} = \hat{q} \times \hat{p} \]

\[ \hat{K} = m\hat{q} \]  

(3.1)

Here \( \hat{p} \) and \( \hat{q} \) are canonically conjugate momentum and position obeying

\[ [q_i, q_j] = 0 \]

\[ [p_i, p_j] = 0 \]  

(3.2)

\[ [q_i, p_j] = \delta_{ij} \]
where \([ \ , \ ]\) is a Poisson bracket.

\(H\) generates time translations and is called the Hamiltonian. \(\dot{P}\) generates space translations and is called the momentum. \(\dot{J}\) generates space rotations and is called angular momentum. \(\dot{K}\) generates transformations to coordinate systems moving with uniform velocity and is called the boost operator.

These operators form the algebra of the Galilei group and they are a familiar and straightforward set.

\[
\begin{align*}
[P_i, P_j] &= 0 & [P_i, H] &= 0 \\
[J_i, H] &= 0 & [J_i, J_j] &= \epsilon_{ijk} J_k \\
[J_i, P_j] &= \epsilon_{ijk} P_k & [J_i, K_j] &= \epsilon_{ijk} K_k \\
[K_i, H] &= P_i \\
[K_i, K_j] &= 0 & [K_i, P_j] &= \delta_{ij} m
\end{align*}
\]

(3.3)

(3.4)

where \(m\) is mass.

For classical relativistic mechanics of noninteracting spinless particles, a corresponding algebra can be defined which generates the Inhomogeneous Lorentz Group (ILG).

\([\ , \ ]\) now becomes \(\frac{1}{i}\) (commutator). \(\hbar = c = 1\). All of equations (3.3) are retained but equations (3.4) become

\[
\begin{align*}
[K_i, K_j] &= -\epsilon_{ijk} J_k & [K_i, P_j] &= \delta_{ij} H.
\end{align*}
\]

(3.5)
The operators $H$ and $K$ become

$$H = (p^2 + m^2)^{1/2} \quad K = Hq$$  \hspace{1cm} (3.6)

while $\dot{J}$ and $\dot{P}$ stay the same. $\dot{q}$ and $\dot{p}$ also obey the same equations (3.2).

In nonrelativistic mechanics if $\dot{X}$ is the position we let $\dot{X} = \dot{q}$ and find that

$$[X_i, P_j] = \delta_{ij} \quad [J_i, X_j] = \varepsilon_{ijk} X_k \quad [X_i, K_j] = 0.$$  \hspace{1cm} (3.7)

The third of equations (3.7) follows from the fact that $K$ is a function of $\dot{q}$ and thus of $\dot{X}$.

In relativistic mechanics one might expect the position, as a function of time, to transform as follows under the various transformations:

A time translation of $\tau$, generated by $\dot{I}$:

$$X'(t) = X(t+\tau) = X(t) + \tau [X, H] + ...$$  \hspace{1cm} (3.8)

A space translation of $a$ in direction $i$, generated by $P_i$:

$$X'_j(t) = X_j(t) + a \delta_{ij} = X_j(t) + a[X_j(t), P_i] + ...$$  \hspace{1cm} (3.9)

which implies

$$[X_j(t), P_i] = \delta_{ij}.$$  \hspace{1cm} (3.10)

A space rotation of $\theta$ about the 3-axis:

$$X_1' = X_1 \cos \theta - X_2 \sin \theta \quad X_2' = X_1 \sin \theta + X_2 \cos \theta \quad X_3' = X_3$$  \hspace{1cm} (3.11)
These equations lead to

$$[J_i, X^j] = \varepsilon_{ijk} X^k.$$  \hfill (3.12)

A boost to a frame moving with velocity $\vec{v}$ in the $x$-direction:

$$x_1'(t') = \frac{1}{\sqrt{1-v^2}} (x_1(t) - vt)$$

$$x_{2,3}'(t') = x_{2,3}(t)$$  \hfill (3.13)

$$t' = \frac{1}{\sqrt{1-v^2}} (t - \frac{x_1}{v})$$

or, letting $s = \tanh^{-1} v$

$$x_1'(t') = x_1(t) \cosh s - t \sinh s$$

$$x_{2,3}'(t') = x_{2,3}(t)$$  \hfill (3.14)

$$t' = t \cosh s - x_1(t) \sinh s.$$  

These, after some manipulation, lead to

$$[X_i, K_j] = X_j[X_i, H].$$  \hfill (3.15)

Equation (3.15) is an unusual commutator equation, in that it is quadratic in $X$ on the right side. Equations (3.10), (3.12) and (3.15) are the expected commutation relations for $\vec{X}$ if $\vec{X}$ behaves the way coordinates behave.

A further generalization from nonrelativistic classical mechanics takes us to relativistic quantum mechanics of a spinless particle where a product $AB$ must be symmetrized to
1/2(AB + BA); [ , ] stands for 1/t times the commutator; and equations (3.1) become

\[ H = (p^2 + m^2)^{1/2} \quad \hat{p} = \hat{p} \quad \hat{J} = q \times \hat{p} \]

\[ \hat{K} = 1/2(\hat{q}\hat{p} + \hat{p}\hat{q}). \] (3.16)

\[ \hat{q} \text{ and } \hat{p} \text{ are now an irreducible set of Hermitian operators obeying (3.2).} \]

The requirements on X derived in equations (3.10), (3.12) and (3.15) become

\[ [X_i, P_j] = \delta_{ij} \quad [J_i, X_j] = \varepsilon_{ijk} X_k \]

\[ [X_i, K_j] = 1/2(X_j[X_i, H] + [X_i, H]X_j). \] (3.17)

Equations (3.3), (3.5) and (3.17) are expected to be satisfied by the operators \( \hat{H}, \hat{P}, \hat{J}, \hat{K}, \hat{X} \) in a relativistic quantum mechanical theory. Other nice properties for a position operator \( \hat{X} \) to have are that the velocity be represented by

\[ [X_i, H] = V_i \]

\[ V_i = \frac{p_i}{m} \text{ nonrelativistically} \quad V_i = \frac{p_i}{\hat{H}} \text{ relativistically} \] (3.18)

and that the components of \( X \) be simultaneously measurable

\[ [X_i, X_j] = 0. \] (3.19)

Condition (3.19) we will call locality.

It can be shown that

\[ [[X_i, H], H] = 0 \] (3.20)
which, with definition (3.18), means \([V_i, H] = 0\), or the velocity does not change with time. This says acceleration is zero and in the case where it is applied to two or more particles it rules out interactions. (3.20) is called the "No Interaction Theorem" and it is a difficulty stemming from the last of requirements (3.17), 
\[
[X_i, K_j] = \frac{1}{2}(X_j[X_i, H] + [X_i, H]X_j)
\]
which is usually taken to imply Lorentz covariance. The worldline of a particle traced out by the position \(\hat{\mathbf{x}}\) which obeys (3.17) is a straight line and thus very dull. Therefore the condition of "Lorentz covariance" bears re-examination. It was derived by assuming \(\hat{\mathbf{x}}(t)\) would transform the way coordinate points transform under Lorentz transformations. This means that if two different observers in coordinate systems moving with some relative velocity both look at a spacetime event, that is some point on the worldline, the point itself stays the same but their descriptions differ. In fact the condition (3.17(c)) is often called the "invariant worldline hypothesis". The diagram below illustrates the property of worldline invariance.
The Lorentz covariance has been imposed by making the position on the worldline behave as a point and therefore such position operators can be called point position operators. If one believes that particles are points or that all their dynamical properties behave as points then one wants point position operators.

It is possible to construct position operators \( \hat{X} \) as functions of \( H, \hat{P}, \hat{J}, \hat{K} \) which satisfy all of equations (3.17), (3.18), and (3.19) for spin = 0.

\[
\hat{X} = \hat{q}
\]  

(3.21)

is the solution and it is point, local, and gives velocity \( \hat{v} = \frac{\hat{p}}{\hat{H}} \).

When we introduce spin \( \hat{S} \) we must expand the set of Hermitian operators from which we construct \( H, \hat{P}, \hat{J}, \hat{K} \) to include \( \hat{S} \) as shown in equations (3.22)

\[
\begin{align*}
[q_i, q_j] &= 0 \\
[p_i, p_j] &= 0 \\
[q_i, S_j] &= 0 \\
[p_i, S_j] &= 0 \\
[S_i, S_j] &= \varepsilon_{ijk} S_k
\end{align*}
\]  

(3.22)

Then a canonical form for the generators is

\[
\begin{align*}
\hat{H} &= (\hat{p}^2 + m^2)^{1/2} \\
\hat{P} &= \hat{p} \\
\hat{J} &= \hat{q} \times \hat{p} + \hat{S} \\
\hat{K} &= \frac{1}{2}(\hat{H}\hat{q} + \hat{q}\hat{H}) + (\hat{H} + m)^{-1} \hat{p} \times \hat{S}.
\end{align*}
\]  

(3.23)

Restricting the solution to operators linear in \( \hat{S} \) one can find an \( \hat{X} \) satisfying (3.17), which is
\[ \dot{X} = \dot{q} - aH^{-1}(H+m)^{-1}(p\cdot S)p + aS - m^{-1}(H+m)^{-1} p \times S , \quad (3.24) \]

where \( a \) is any real number.

There is no value of \( a \) for which \([X_i, X_j]\) = 0. The solution \( \dot{X} = \dot{q} \) would satisfy all but (3.17 c).

There is another extension of the possible constructions of position operators, which is to add negative energy antiparticle states to the positive energy states which can be eigenstates of position. This necessitates including sign of energy and other operators in the set of Hermitian operators.

The vector \( \vec{p} \) corresponds to Dirac's original three matrices \( \rho_1, \rho_2, \rho_3 \) and to the Pauli matrices

\[
\begin{align*}
\rho_1 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{where } 1 \text{ is } \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
\rho_2 &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\
\rho_3 &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} .
\end{align*}
\]

\( \rho_3 \) in this case is the sign of energy operator. \( \rho_2 \) mixes the differently signed energy states. In more familiar terms

\[ \dot{\alpha} = 2\rho_1 \dot{S}, \quad \beta = \rho_3 . \quad (3.26) \]

The irreducible set of Hermitian operators \( \dot{q}, \dot{p}, \dot{S}, \dot{p} \) now obey (3.22), plus

\[ [q_i, \rho_j] = 0 = [p_i, \rho_j] = 0 = [S_i, \rho_j] \]

\[ [p_i, \rho_j] = 2\varepsilon_{ijk} \rho_k \quad (3.27) \]

\[ \{\rho_i, \rho_j\} = 2\delta_{ij} \text{ where } \{,\} \text{ is the anti-commutation bracket.} \]
The generators are now

\[ H = \rho_3 W \]
\[ \vec{p} = \vec{p} \]
\[ \vec{J} = \vec{q} \times \vec{p} + \vec{S} \]  

and

\[ \vec{K} = 1/2(Hq + qH) + \rho_3(p^3H + m)^{-1} \vec{p} \times \vec{S} = 1/2\rho_3(Hq + qH) + \rho_3(p^3H + m)^{-1} \vec{p} \times \vec{S}. \]

Now there is a solution \( \vec{X} \) to (3.17) and (3.19) which is valid for spin 1/2 and all other spins.

\[ \vec{X} = \vec{q} + \rho_2 \frac{1}{W} \vec{S} - \rho_2 \frac{1}{W^2(W + m)^{-1}} (\vec{p} \cdot \vec{S}) \vec{p} + W^{-1}(W + m)^{-1} \vec{p} \times \vec{S} + \rho_2 A \vec{p} \]  

\[ (3.29) \]

\[ A = 0 \text{ for } S = 1/2 \text{ and the forms of } \rho_1 \text{ and } \rho_2 \text{ change for different values of } S. \text{ In this case } \vec{X} \text{ has no sense, though } \vec{X} \text{ is both point and local, i.e. obeys (3.17(c)) and (3.19).} \]

The well-known position operators of Dirac, Foldy and Wouthuysen, and Newton and Wigner, Pryce, etc., can all be written in one of the above forms, as shown later in this chapter.

The Foldy-Wouthuysen transformation is the unitary transformation \( S \) which in operator form is

\[ S = -\frac{i}{2m} \vec{\alpha} \cdot \vec{p} \frac{m}{p} \tan^{-1} \left( \frac{p}{m} \right) \]  

\[ (3.30) \]

or

\[ e^{i\vec{S}} = \frac{\beta \left( \beta m + \vec{\alpha} \cdot \vec{p} + \beta H \right)}{(2H(H+m))^{1/2}} \]  

\[ (3.31) \]
where

\[ \beta = \text{sign of energy} \]
\[ \hat{\alpha} = \text{Dirac matrices} \]
\[ H = (p^2 + m^2)^{1/2} \]

S was invented to diagonalize the Dirac Hamiltonian

\[ H_{FW} = e^{iS} H_{Dirac} e^{-iS} = \beta H \quad (3.32) \]

When S is applied to the Dirac position operator \( \hat{x}_{Dirac} \) the result is

\[ \hat{x}_{Dirac}^{FW} = \hat{x}_{Dirac} - \frac{i\beta \hat{a}}{2i} + \frac{i\hat{S}(\hat{a} \cdot \hat{p}) \hat{b}}{2H^2(H+m)} - \frac{(\hat{a} \times \hat{b})H}{2H^2(H+m)} \quad (3.33) \]

Identify \( \frac{\beta}{2} \) with \( \tilde{S} \). Here \( \hat{x}_{Dirac} \) refers to the form \( \tilde{v}_p \).

The Dirac velocity is

\[ \hat{v}_{Dirac} = \hat{a} \]

from

\[ [\hat{x}_{Dirac}, H] = [\hat{x}, \beta m + \hat{a} \cdot \hat{p}] = \frac{\partial}{\partial p} (\beta m + \hat{a} \cdot \hat{p}) = \hat{a} \]. \quad (3.34) \]

We have let \( c = 1 \) or the velocity would be \( ca \) and have the expectation value of \( c \). A preferable velocity would be \( \hat{p}/H \). There is an operator \( \hat{x}_{mean} \) which is the inverse Foldy-Wouthuysen transformation of \( \hat{x}_{Dirac} \):

\[ \hat{x}_{mean} = e^{-iS} \hat{x}_{Dirac} e^{iS} \quad (3.35) \]

\( \hat{x}_{mean} \) is \( \pm \hat{p}/H \) depending on the sign of energy of the states being considered, so \( \hat{x}_{mean} \) is the velocity we would like, and
\[ \hat{x}_{\text{Dirac}} - \hat{x}_{\text{mean}} = \text{Zitterbewegung}. \] (3.36)

Thus \( \hat{x}_{\text{mean}} \) is the mean position, which moves with velocity \( \frac{\hat{p}}{\hat{m}} \), and \( \hat{x}_{\text{Dirac}} - \hat{x}_{\text{mean}} \) oscillates with amplitude of the order of a Compton wavelength about the mean position.

In the Foldy-Wouthuysen (FW) representation \( \hat{x}_{\text{Dirac}} \) is complicated (3.33). However, in form

\[ \hat{x}_{\text{mean}} \text{FW} = \hat{x}_{\text{Dirac}} = \hat{v}_p. \] (3.37)

It is \( \hat{x}_{\text{mean}} \) which reduces to the nonrelativistic position.

For spin 1/2 equation (3.29) can be shown to be identical to equation (3.33). \( A = 0, \; \omega = \Omega \), and the normalization must be changed. The FW transform of \( \hat{x}_{\text{mean}} \) is \( \hat{x}_{\text{Dirac}} \) in form and it goes to the nonrelativistic \( \hat{q} \) for \( S = 0 \).

It is interesting to note that since the canonical \( \hat{H} \) of equations (3.26) is the FW transform of the Dirac \( \hat{H} \), and the \( \hat{x} \) of equation (3.29) or (3.33) which is point and local is the FW transform of the canonical \( \hat{x}_{\text{Dirac}}(=\hat{q}) \), the Dirac Hamiltonian could be "derived" by insisting upon \( \hat{x} \) of equation (3.29), then finding the transformation which carries it into \( \hat{q} \), then applying that transformation to \( \hat{H} \) canonical.

For spins zero and one, with some manipulation, the transformation from \( \hat{x} \) to \( \hat{q} \) gives from the canonical \( \hat{H} \) the Klein-Gordon and Proca equations, respectively.

The above development, starting with equations (3.1), is largely due to Jordan and Mukunda. It has been presented because of its clarity in defining position operator properties.
Historical Approach

It is also informative to make a brief historical review of position operators, so as to introduce various lines of reasoning which lead to the above conditions and operators from different directions.

The unification of relativity and quantum mechanics, with the inclusion of spin, was achieved by Dirac in his theory of the electron. His own book still provides a very clear chain of reasoning for the formation of a relativistic quantum-mechanical dynamics.

Special relativity demands that space and time coordinates be treated in a symmetrical way, or that $x$ and $t$ (ct with $c = 1$) form a four-vector $x^\mu$. This four-vector then transforms covariantly under Inhomogeneous Lorentz Transformations (ILT's).

$$x'_\mu = A^\nu_\mu x^\nu + a_\mu. \quad (3.38)$$

Since physical states are set up as functions of the space-time point $(\vec{x}, t)$, the principle of superposition of states is relativistically sound.

It is the physical observables of quantum mechanics which might not fit nicely into relativity. The correspondence principle insists that physical observables be represented by Hermitian operators, which have real eigenvalues. But these observables may involve physical things at different spatial points at the same instant of time. In general it is not possible to have manifest covariance, i.e. symmetry in space and
time components, for a complete set of commuting observables.

Quantum mechanics also imposes the symmetrized product
\[
\frac{1}{2} (ab + ba)
\]
on operators \(a\) and \(b\) in place of simply \(ab\). This is because if \(a\) and \(b\) are real, \((ab + ba)\) and \(i(ab - ba)\) are real, but \(ab\) is not real unless \((ab - ba) = [a,b] = 0\).

Both the Schrödinger and Heisenberg pictures of quantum mechanics demand that the equations of motion be written in a Hamiltonian form. The Schrödinger picture puts time dependence in the states by defining states \(\psi\) at any instant of time \(t\) and position \(\vec{x}\), and then writing the Hamiltonian equation
\[
\hbar \frac{\partial}{\partial t} \psi (\vec{x},t) = H(t) \psi (\vec{x},t). \tag{3.39}
\]

The time development is contained in the total energy operator \(H(t)\), which is assumed to always be an observable. For an isolated system \(\hbar\) is a constant. This picture corresponds to a continuous motion being applied to the whole vector space of state vectors \(\psi(t)\).

The Heisenberg picture fixes the state vectors and allows the dynamical variables to change with time. Then the equations of motion apply to the dynamical variables rather than to the states. For any operator \(O\),
\[
\frac{dO}{dt} = [O,H]. \tag{3.40}
\]

The total energy \(H\) of Heisenberg theory is obtained by a simple transformation from \(\hbar\) of Schrödinger theory. Form (3.40) corresponds to classical Hamiltonian theory. Another carryover
from classical Hamiltonian mechanics is the Poisson bracket $[A,B]$ of two operators $A$ and $B$.

$$[A,B] = \sum_r \left( \frac{\partial A}{\partial q_r} \frac{\partial B}{\partial p_r} - \frac{\partial A}{\partial p_r} \frac{\partial B}{\partial q_r} \right) , \quad (3.41)$$

where $q$ and $p$ are canonical coordinates and momenta. Quantum mechanical commutators correspond to times Poisson Brackets. The algebra of Poisson Brackets is called a Lie Algebra.

The Heisenberg picture uses commutation relations among operators to derive dynamical results. If the commutation relations remain the same under coordinate transformations given by ILT's then the Lie Algebra is the algebra for the Inhomogeneous Lorentz group. Whenever a transformation is made on an operator $0$ taking it into $0' = e^{-iG}oe^{iG}$, one has

$$0' - 0 = [0,G] \quad (3.42)$$

where $G$ is the generator of the (infinitesimal) transformation.

To display relativistic symmetry, Dirac used the Schrödinger picture and derived his wave equation for the electron:

$$i\hbar \frac{\partial}{\partial t} \psi (\vec{x},t) = H \psi (\vec{x},t) \quad (3.39)$$

where

$$H = \rho_1 \vec{\sigma} \cdot \vec{p} + \rho_3 mc \quad (3.43)$$

and

$$\rho_1 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \rho_2 = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \\ i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix} \quad (3.44)$$
He required that

\[ p_\mu = i\hbar \frac{\partial}{\partial x^\mu} \tag{3.46} \]

and \( x^\mu \) and \( p^\mu \) both are four-vectors. A more familiar form of the equation uses

\[ \alpha = \rho_1 \sigma \]
\[ \beta = \rho_3 \cdot \]

In order to create this covariant equation it was necessary to assume that \( x^\mu \) and \( p^\mu \) were canonically conjugate coordinates and momenta. To get spin 1/2 from the equation it is necessary to assume that orbital angular momentum

\[ \mathbf{L} = \mathbf{x} \times \mathbf{p} \tag{3.48} \]

and that total angular momentum is conserved. One finds that
\[ \dot{L} + \frac{1}{2} \hbar \dot{\sigma} = 0 \]  

so \( \frac{1}{2} \hbar \dot{\sigma} \) is identified as the spin \( \vec{s} \). Then

\[ \dot{L} + \vec{s} = \dot{J} \]  

is conserved.

The assumptions about \( x^\mu \) are reasonable, namely

\[ [x^\mu, p^\nu] = i\hbar g^{\mu \nu}, \]

\( x^\mu \) transforms as a four-vector, and

\[ \dot{x} \times \dot{p} = \dot{L}. \]

However the time derivation of \( \dot{x} \), given by the Heisenberg picture as

\[ \frac{d\dot{x}}{dt} = \dot{x} = [\dot{x}, H] = c\dot{\alpha} \]

has led to some confusion. One expects

\[ \dot{x} = \dot{v} = \frac{\dot{p}}{E} = \frac{\dot{p}}{L} \]

for a classical particle. Yet the Dirac \( \dot{x} \) gives an eigenvalue of \( c \) for the velocity measured at any instant. Insight into this problem is provided by the following calculation.

\[ i\hbar \dot{\alpha} = \dot{\alpha} H - H\dot{\alpha} \]

\[ = 2\dot{\alpha} H - 2c\dot{\alpha}. \]

But

\[ i\hbar \ddot{\alpha} = 2\dot{\alpha} H, \]

since \( H \) and \( \dot{p} \) are both constants. Then

\[ \dot{\alpha} = \dot{\alpha}_0 e^{-\frac{2\dot{\alpha}}{\hbar} H t} \]

and

\[ \alpha = \frac{1}{2} i\hbar \dot{\alpha}_0 e^{-\frac{2\dot{\alpha}}{\hbar} H t} + \dot{p} L^{-1} \]
which gives
\[ \dot{x} = -\frac{1}{4} c^2 \hbar^2 \alpha e^{-\frac{1}{\hbar} \hbar^2 H t - 2 + c^2 \hbar^2 t + \dot{a}.} \] (3.57)

\( \dot{a} \) is a constant.

From (3.46) we see that the velocity \( \dot{ca} \) consists of two parts,
\[ \dot{ca} = \frac{1}{2} in \, \dot{\alpha} e^{-\frac{1}{\hbar} \hbar^2 H t - 1 + c^2 \hbar^2 t,} \] (3.58)
the constant part \( c^2 \hbar^2 t \) being connected with the momentum as classically and the oscillatory part \( \frac{1}{2} in \, \dot{\alpha} e^{-\frac{1}{\hbar} \hbar^2 H t - 1 \ \text{being}} \) of such high frequency \( \frac{2\hbar}{\hbar} \) as to be rendered unmeasurable in any practical situation. Time intervals of measurements are always much larger than \( \frac{\hbar}{2mc^2} \) so the "Zitterbewegung" cancels out. But the velocity \( \dot{ca} \) at any instant would give eigenvalues \( \pm c \).

From (3.57) we see that \( \dot{x} \) is similarly split into a constant, an oscillatory part giving rise to Zitterbewegung, and a part proportional to \( \hbar^2 t \) which gives the classical trajectory.

In further development of relativistic dynamics Dirac turned to the Heisenberg picture and set up the algebra of the ILG and considered different solutions to the equations given by commutation relations between \( P^\mu \) and \( M^\mu_\nu \).

\[ [P_\mu, P_\nu] = 0 \]
\[ [M^\mu_\nu, P_\rho] = -g_\mu_\rho P_\nu + g_\nu_\rho P_\mu \] (3.59)
\[ [M^\mu_\nu, M^\rho_\sigma] = -g_\mu_\rho M^\nu_\sigma + g_\nu_\rho M^\mu_\sigma - g_\mu_\sigma M^\nu_\rho + g_\nu_\sigma M^\mu_\rho. \]

Here he switches to a formalism where \( i \)'s and \( \hbar \)'s disappear. In trying to find dynamical operators which give Hamiltonian equations
of motion and Lorentz covariance he considers different forms for $p^\mu$ and $M^{\mu\nu}$ given by different subsidiary equations which describe the space-time point of view.

For example canonical coordinates and momenta with

$$[q_\mu, q_\nu] = 0 \quad \quad [p_\mu, p_\nu] = 0 \quad \quad [p_\mu, q_\nu] = g_{\mu\nu}$$

(3.60)

give

$$p_\mu = p_\mu \quad \quad M_{\mu\nu} = q_\mu p_\nu - q_\nu p_\mu$$

(3.61)

as the simplest solution to (3.59).

The practice of referring to the energy $p^0$ and momentum $p$ and position $q$ of a particle as being physical conditions at some instant of time $t$ leads to Dirac's "instant" form of dynamics. The subsidiary condition

$$q_0 = t = 0$$

(3.62)

defines a flat three-dimensional hyperplane in four-dimensional space.

Using (3.60) and (3.62) one gets

$$p_i = p_i \quad \quad M_{ij} = q_i p_j - q_j p_i$$

(3.63)

$$p_0 = (p_i p_i + m^2)^{1/2} \quad \quad M_{i0} = q_i (p_j p_j + m^2)^{1/2}.$$ 

Adding interactions $V$ and $V_i$ to $p_0$ and $M_{i0}$ respectively leads to quadratic conditions on the $V$'s which are hard to solve.

Using three-dimensional surfaces which do not refer to some
one instant is another possibility. To get a simple form, the subsidiary condition
\[ q^\mu q_\mu = \kappa^2 \quad q_0 > 0 \quad (3.64) \]
is taken. This follows from insisting that the surface be left invariant by rotations about the origin \( q_\mu = 0 \). Then the "point" form solutions to (3.59) are obtained as
\[ P_\mu = p_\mu + q_\mu \kappa^{-2} \left[ (p^\nu q_\nu)^2 - \kappa^2 (p^\sigma p_\sigma - m^2) \right]^{1/2} - p^\nu q_\nu \]
\[ M_{\mu\nu} = q_\mu p_\nu - q_\nu p_\mu. \quad (3.65) \]

In the instant form, \( P_i \) and \( M_{ij} \) were simple because they leave the instant \( q_0 = 0 \) invariant. In the point form \( M_{\mu\nu} \) is simple because it leaves the surface \( q^\mu q_\mu = \kappa^2, \quad q_0 > 0 \), which is one branch of an hyperboloid, invariant.

In the "front" form, a plane wave front given by
\[ q_0 - q_3 = 0 \quad (3.66) \]
is left invariant. By defining
\[ A_0 + A_3 = A_+ \quad A_0 - A_3 = A_- \quad (3.67) \]
and using \( +, -, 1, 2 \) as indices instead of \( 0, 1, 2, 3 \), one rewrites (3.66) as
\[ q_- = 0 \quad (3.68) \]
and finds solutions to (3.59)
\[ P_i = p_i \quad P_- = p_- \quad N_{12} = q_1 p_2 - q_2 p_1 \quad (3.69) \]
Here $P_i, P_-, M_{12}, M_{i-}$, and $M_{+-}$ are the generators which leave the front invariant. All of these solutions (3.63), (3.65), and (3.69) have certain advantages and disadvantages dynamically for working out actual problems with many particles and interactions.

The interaction terms must always be added to the generators which do not leave the instant or point or front invariant, and equations which must be obeyed by the interactions are hard to solve. Different interactions fit better into different forms. The instant form is the most familiar. The point form separates $P^\mu$ and $M^{\mu\nu}$ into a complicated four-vector and a simple six-vector which makes concise equations for interaction terms. The front form gives only three complicated generators and has no square roots. There is no conclusive argument in favor of any one of the forms. All of the forms use canonical positions $q^\mu$ and momenta $p^\mu$ as basic operators.

In 1949 Møller published lectures he had given in 1947 on position operators. He departed from the canonical $p$'s and $q$'s of equations (3.60) and considered generalizations to relativistic quantum theory of classical positions, such as center of mass. Between 1920 and 1947 very little else had been done with position operators. Born and Infeld and Pryce had
considered center of mass

\[ \hat{q} = \frac{1}{2} (\hbar^{-1} \hat{x} + \hat{x} \hbar^{-1}) \]  

(3.70)

such that

\[ \hat{L} = \hat{q} \times \hat{p} \]  

(3.71)

where \( \hat{x} \) is an integral over the spatial distribution of the particle. Later Pryce in 1948 published his derivations of several operators as possible relativistic quantum mechanical positions. Pryce, Möller, and Dirac have probably stimulated most of the work on position done since 1948. There are many papers in the 30's and 40's on the development of field theory and quantum theory which assume various position operators with no particular justification. Papapetrou of Greece wrote an inaccessible paper in 1939 which Möller acknowledges as similar to his.

Möller first considers classical relativistic generalizations of "center of gravity". Classically one has

\[ \hat{X} = \frac{1}{M} \sum_i m_i \hat{x}_i \]  

(3.72)

or

\[ \hat{X} = \frac{\int \mu(\hat{x},t) \hat{x} \, d^3x}{\int \mu(\hat{x},t) \, d^3x} \]  

(3.73)

as the definition of center of mass. Equation (3.72) refers to discrete masses and equation (3.73) refers to a continuous mass distribution. Furthermore \( \hat{X} \) gives the correct velocity:

\[ \frac{d\hat{X}}{dt} = \frac{\hat{p}}{M} \]  

(3.74)

The use of proper mass \( M^0 \) or relativistic mass \( M = \gamma M^0 \) is
unsatisfactory relativistically. \( M^0 \) gives an \( \chi \) with \( \chi \) not proportional to \( \dot{p} \) and \( M = \gamma M^0 \) gives a non-uniform motion unless there are no interactions at all.

However, using an energy density \( h \) from the symmetric energy-momentum tensor \( T_{\mu \nu} \), Møller defines

\[
\dot{\chi} = \frac{1}{M_0} \frac{h(x^0, t^0)}{c^2} \chi \, d^3x \tag{3.75}
\]

so that

\[
\dot{\chi} = \frac{c^2 \dot{p}}{H} \tag{3.76}
\]

He then varies his approach using proper time \( \tau \) instead of \( t \), discussing angular momentum, and adding spin to the picture. His final form for a "center of mass" is

\[
X_{\mu} = \frac{1}{\hbar} \int h(x^0, x^\mu) \, \chi_{\mu} \, d^3x - m_{\mu 0} P_{\mu}^{-1} = x_{\mu}(x^0). \tag{3.77}
\]

This \( X_{\mu} \) is a non-manifestly covariant four-vector. It is a function of the time \( (x^0) \) and in order to get simple expressions for Lorentz transformations, the LT's must leave the time constant. The \( m_{\mu 0} \) in (3.77) is the total angular momentum tensor and \( h \) is the energy density. \( X_{\mu} \) is a center of spin in the sense that the angular momentum is split into two parts, one of which is

\[
(x_{\mu} - X_{\mu}) P_{\nu} - (x_{\nu} - X_{\nu}) P_{\mu} \quad \text{where} \quad x_{\mu} \quad \text{are ordinary coordinates.}
\]

Møller's approach to finding a position operator which fits into quantum mechanics as well as relativity is to symmetrize the product \( h^{-1} \int h \chi_{\mu} d^3x \) in (3.77) and also the definition of \( m_{\mu \nu} \) as \( M_{\mu \nu} - (x_{\mu} p_{\nu} - x_{\nu} p_{\mu}) \) to give
His notation and somewhat informal (disorganized) approach make him hard to follow. He insists that $X_0$ be a c-number in any frame, which appears to contradict having $X_\mu$ be a four-vector. However, the asymmetry of space and time is not a defect, since one does not ask for the probability that a clock read a certain value, nor is a wave packet allowed to be so distributed in time as to disappear exponentially in the past or future. Time does not appear symmetrically to space even in simple relativity.

To make the position a function of proper time $\tau$ would be to make it a function of $\dot{v}$ which is a q-number and cannot be used as a parameter in quantum mechanics. Both $\dot{X}$ and $\tau$ are functions of the c-number $t$.

This $\dot{X}$ does not have commuting components.

$$[X_i, X_j] = \frac{1}{(M_0c)^2} (m_{ij} - cn_i P_j h^{-1} + cn_j P_i h^{-1})$$

where

$$\hat{n} = -\frac{c}{H} (\hat{m} \times \hat{p}),$$

$$\hat{m} = (m_{23}, m_{31}, m_{12}),$$

and

$$m_{\mu\nu} = M_{\mu\nu} + \frac{1}{(M_0c)^2} (M_{\mu\rho\rho\nu} - M_{\nu\rho\rho\mu}).$$

An $\dot{X}$ can be constructed which has commuting components. Møller calls this operator $\dot{X}$. It is defined so that
\[
\vec{M} = \frac{\vec{X}}{\vec{m}} + \vec{m} \times \vec{P}
\]

(3.83)

where

\[
\frac{\vec{X}}{\vec{m}} = \frac{m_0 c^2}{H} \frac{\vec{m}}{m} + \frac{1 - \frac{m_0 c^2}{H}}{p^2} \frac{(\vec{m} \cdot \vec{P}) \vec{P}}{p^2}
\]

(3.84)

The operator \(\vec{m}\) has the properties of an angular momentum, which \(\vec{X}\) does not quite have. Also \(\vec{m}\) is a constant of the motion.

If "center of gravity" refers to the rest frame center of mass and "center of mass" refers to (3.77) then \(\vec{X}\) lies between the two in any frame.

These definitions can be applied in any frame which has a well-defined energy-momentum tensor. As an example Möller applies it to Dirac's theory of electrons. He shows the energy-momentum tensor \(T_{\mu\nu}\) from which \(P_{\mu}\) and \(M_{\mu\nu}\) are derived.

An \(\vec{X}\) is defined as in (3.78), which obeys all the same commutation rules as (3.78), e.g. (3.79) and

\[
[X_{\mu}, P_{\nu}] = \delta_{\mu\nu} - \delta_{\mu\alpha} p_{\mu} p_{\alpha}^{-1}
\]

(3.85)

\[
[X_{\mu}, M_{\nu}\rho] = -\delta_{\mu\nu} X_{\rho} + \delta_{\mu\rho} X_{\nu} + \delta_{\nu\rho}: (X_{\rho} : p_{\mu} p_{\nu}^{-1}) - \delta_{\rho\nu} (X_{\nu} : p_{\mu} p_{\rho}^{-1})
\]

(3.86)

\[
[X_{\mu}, m_{\nu\rho}] = -\frac{1}{(m_0 c)^2} (m_{\mu\nu} p_{\rho} - m_{\mu\rho} p_{\nu} + p_{\mu \nu\rho} p_{\rho}^{-1} - p_{\mu} p_{\nu\rho} m_{\rho\nu} p_{\rho}^{-1})
\]

(3.87)

\[
[X_{\mu}, X_{\nu}] = \frac{1}{(m_0 c)^2} (m_{\mu\nu} - m_{\mu\rho} p_{\rho} p_{\nu}^{-1} + m_{\nu\rho} p_{\rho} p_{\nu}^{-1})
\]

(3.88)

\[
[X_{\mu}, M_{\nu\rho}] = 0.
\]

(3.89)
This $\vec{x}$ can be split into two terms, as Dirac did. The oscillatory term $\vec{x}$ is given by

$$\vec{x} = \frac{i\hbar c}{2} (\vec{\alpha} - H^{-1}\vec{c}\vec{p}) H^{-1} + c\vec{n} H^{-1}. \quad (3.90)$$

Again a position operator $\vec{X}$ with commuting components can be defined. If $\vec{x}$ is the purely coordinate vector then the regular center of mass is

$$\vec{X} = \vec{x} - \vec{\bar{x}} \quad (3.91)$$

and the commuting center of mass is

$$\vec{\bar{X}} = \vec{x} - \vec{\bar{x}} + \frac{c}{H + m_0 c^2} \vec{\bar{n}} = \vec{x} - \frac{i\hbar c}{2} (\vec{\alpha} - H^{-1}\vec{c}\vec{p}) H^{-1} - \frac{m_0 c^2}{H(H + m_0 c^2)} c\vec{n}. \quad (3.92)$$

Here $\vec{n}$ is

$$\vec{n} = -\frac{\hbar}{2m_0 c} \rho_3 (\vec{\sigma} \times \vec{p}) \quad (3.93)$$

and

$$\vec{\bar{n}} = \frac{\hbar}{2} \vec{\sigma} - \vec{n} \quad (3.94)$$

or

$$|\vec{\bar{m}}|^2 - |\vec{n}|^2 = \frac{3}{4} \vec{\sigma}^2 = |\vec{\bar{\sigma}}|^2. \quad (3.95)$$

A new spin $\vec{\sigma}$ can be defined, and then a new $\vec{\sigma}$ and $H$ in terms of these quantities. What is more important, though, is that all the $\vec{x}$'s increase linearly in time with the same velocity in any representation as long as they act only on pure positive or pure negative energy states. Only when a superposition of positive and negative energies is allowed does the $\vec{x}$ become an observable, but then the sign of the energy is not an observable. Møller concludes that it is nonsensical to work with superpositions of positive and negative energy states.
Pryce gives the first systematic and easy to understand generalization of position to relativistic quantum theory. He generalizes center of mass in six ways and considers the advantages of each. His generalizations are as follows.

(a) The coordinates of the mass-center are the weighted means of the coordinates of the several particles, the weights being the rest masses.

(b) Definition (a) is applied in the frame of reference where total momentum is zero, then transformed to any other frame by a Lorentz transformation.

(c) The weights are the dynamical masses (energies).

(d) Definition (c) is applied in the frame of reference where total momentum is zero and then transformed by LT to any other frame.

(e) The mean of (c) and (d), weighted with total energy and rest mass of the total system, respectively, is taken.

(f) Definition (a) is applied in the frame of reference where definition (a) is at rest, then transformed by LT to any other frame.

Definition (a) is not independent of the Galilean frame in which it is defined. In general, it is not at rest in the frame in which total momentum is zero. If the particles in the system interact, (a) does not move uniformly in a straight line.

Definition (b) is independent of the final frame of reference, but is not in general rectilinear nor at rest in the \( P_{\text{total}} = 0 \) frame.
Definition (c) is at rest in the $\vec{p}_{\text{total}} = 0$ frame and can be used for particles interacting through a field. However (c) is not independent of the frame of reference. Also the Poisson bracket of (c) with itself is not zero.

Definition (d) has been called the "invariant mass-center." It is independent of the final frame of reference and it can be used for interacting particles. However its Poisson bracket with itself does not vanish.

Definition (e) is constructed to have vanishing Poisson bracket with itself. Definition (a) also has zero Poisson bracket but cannot be used for interacting particles. Definition (e) is at rest in the $\vec{p}_{\text{total}} = 0$ frame but it is not independent of the frame in which it is defined.

Definition (f) has little merit.

Pryce concludes that (d) is the most satisfactory definition, being relativistically covariant and applicable to any system of particles interacting through a field. He discusses (c) and (e), admits that (a) has use in two-body problems but not much beyond that, and dismisses (b) and (f) as uninteresting.

Calling definition (c) $q^\mu$ and going through the usual energy-momentum tensor route, one gets

$$q^\mu = \frac{(t p^\mu + \mathcal{M}^{\mu 0})}{p^0} \quad (3.96)$$

$$\frac{dq^\mu}{dt} = \frac{p^\mu}{p^0} \quad (3.97)$$

Equation (3.96) gives the coordinates of $q^\mu$ defined at time $t$ in the frame to which the indices $\mu$ refer. If the mass-center
is defined at time \( s \) in another frame whose time axis is given by the unit vector \( \eta^\mu \) such that \( \eta^\mu \eta_\mu = 1 \) then

\[
q^\mu (\vec{\eta}, s) = \frac{sp^\mu + M^{\mu\nu}\eta_\nu}{p^\sigma\eta_\sigma}.
\] (3.98)

Here \( q^\mu (\vec{\eta}, s) \) is definition (c) applied in the \( \eta^\mu \) frame but looked at in the \( \mu \) frame. Thus Pryce (who called the unit vector \( n \) rather than \( \eta \)) first used "hyperplane" notation.

When one solves

\[
q^0 (\vec{\eta}, s) = t
\] (3.99)

for \( s \) and substitutes the result in (3.98), one finds

\[
q^\mu (\vec{\eta}) = \frac{tp^\mu}{p^0} + \frac{M^{\mu\nu}p_\nu - M^{\sigma\nu}p_\nu\eta_\nu}{p^0\eta_\sigma}
\] (3.100)

which reduces to (3.96) for \( \eta^\mu = (1, 0, 0, 0) \). If (c) were independent of the frame of reference then (3.100) would be independent of \( \eta^\mu \), which is not in general true.

Definition (d) substitutes

\[
\frac{p^\mu}{m} = \eta^\mu
\] (3.101)

in equation (3.100) to give

\[
\chi^\mu = \frac{tp^\mu}{p^0} + \frac{M^{\mu\nu}p_\nu}{m^2} + \frac{M^{\sigma\nu}p^\mu p_\nu}{m^2 p^0}
\] (3.102)

as the (d) center of mass. \( \chi^\mu \) is relativistically covariant.

Introducing

\[
N^i = M^{i0} + t p^i
\] (3.103)

We find

\[
\vec{q} = E^{-1} \vec{N}
\] (3.104)
and
\[ \dot{X} = m^{-2} \left( \dot{E} + \dot{M} \times \dot{P} - E^{-1} (\dot{P} \cdot \dot{N}) \dot{P} \right) \] (3.105)

with \( t \) no longer explicit. Now introducing
\[ \dot{S} = \dot{M} - \dot{q} \times \dot{P} \] (3.106)
\[ \Sigma = M - X \times P \] (3.107)

we find
\[ \dot{X} - \dot{q} = m^{-2} \dot{S} \times \dot{P} = E^{-2} \dot{P} \times \dot{\Sigma}. \] (3.108)

So \( \dot{X} \) and \( \dot{q} \) in general do not coincide, but travel in parallel straight lines at constant distance.

The Poisson brackets of \( \dot{P}, \dot{E}, \dot{M} \) and \( \dot{N} \) are the same as those for \( \dot{P}, \dot{E}, \dot{J} \) and \( \dot{K} \) in Appendix B, equations (B.4). Since \( \dot{q}, \dot{X}, \dot{S} \) and \( \dot{\Sigma} \) are defined in terms of \( \dot{P}, \dot{E}, \dot{M} \) and \( \dot{N} \), their Poisson brackets can be found and are

\[ \begin{align*}
[q_i, P_j] &= \delta_{ij} \\
[S_i, P_j] &= 0 \\
[q_i, M_j] &= \epsilon_{ijk} q_k \\
[S_i, M_j] &= \epsilon_{ijk} S_k \\
[q_i, q_j] &= -m^{-2} \epsilon_{ijk} S_k \\
[S_i, S_j] &= m^2 E^{-2} \epsilon_{ijk} \Sigma_k \\
[S_i, q_j] &= E^{-2} (P_j q_i - \delta_{ij} \dot{S} \cdot \dot{P}).
\end{align*} \] (3.109)

Equations (3.109) hold when \( \dot{q} \rightarrow \dot{X} \) and \( \dot{S} \rightarrow \dot{\Sigma} \). Also
\[ \begin{align*}
[X_i, N_j] &= \dot{X}_i X_j = E^{-1} p_i \dot{X}_j \\
\end{align*} \] (3.110)

which expresses the fact that \( \dot{X} \) defines an "invariant world line".
However
\[
[q_i, N_j] = E^{-1}(p_i q_j + \epsilon_{ijk} S_k)
\] (3.111)

and \( \hat{q} \) does not have an invariant world line.

Symmetrization of products in equations (3.96) through (3.111) gives correct quantum mechanical results.

Definition (e) is called \( \tilde{q} \) and is given by
\[
\tilde{q} = (m + E)^{-1} (E \tau + m \hat{x}) = \hat{q} + m^{-1} (m + E)^{-1} \hat{S} \times \hat{p} = \hat{x} - E^{-1} (m + E)^{-1} \hat{S} \times \hat{p}.
\] (3.112)

Analogously to \( \tilde{S} \) and \( \tilde{p} \) one has
\[
\tilde{S} = \hat{M} - \tilde{q} \times \hat{p}
\] (3.113)

which is like an elementary particle spin.

Whereas Pryce's definitions (a) through (f) were formulated for a collection of particles, they do carry over to a single particle, which also has an energy-momentum tensor. He states that it is impossible to find a position which is Lorentz covariant and has commuting coordinates.

For the Dirac electron Pryce derives
\[
\hat{q} = \hat{x} + \frac{\hbar}{2E^2} (\hat{p} \times \hat{\sigma} + \mi \beta \hat{\alpha}) \quad \hat{x} = \hat{x} + \frac{\hbar}{2m} \mi \beta (\hat{\alpha} - E^{-2}(\hat{\alpha} \cdot \hat{p}) \hat{p})
\] (3.114)

and
\[
\tilde{q} = \hat{x} + \frac{\hbar}{2} \left( \frac{1}{E} \mi \beta \hat{\alpha} + \frac{1}{E(E+m)} \hat{p} \times \hat{\sigma} - \frac{1}{E^2(E+m)} \mi \beta (\hat{\alpha} \cdot \hat{p}) \hat{p} \right),
\]

where \( \hat{x} \) is the coordinate vector. He also defines
\[
\tilde{S} = \frac{\hbar}{2E^2} (M^2 \hat{\sigma} - \mi \beta \hat{\alpha} \times \hat{p} + (\hat{p} \cdot \hat{\sigma}) \hat{p})
\] (3.115)
\[ \dot{\mathbf{x}} = -\frac{\hbar}{2m} (\mathbf{p} \times \mathbf{\alpha}) \]  

and

\[ \dot{\mathbf{x}} = \frac{\hbar}{2E} (m\mathbf{v} - i\hbar \mathbf{\alpha} \times \mathbf{p} + (E + m)^{-1} (\mathbf{p} \cdot \mathbf{\alpha}) \mathbf{p}). \]

He notes that \( \mathbf{\dot{x}}, \mathbf{\dot{p}}, \frac{\mathbf{E}}{E}, \) and \( \mathbf{\dot{S}} \) satisfy the same commutation relations as do \( \mathbf{\dot{x}}, \mathbf{\dot{p}}, \mathbf{\dot{B}}, \) and \( \frac{\hbar}{2} \mathbf{\dot{\sigma}} \). He finds the transformation between the two sets:

\[ U = (2E (E + m))^{-1/2} (H + E\beta) \]

so

\[ \mathbf{\dot{q}U} = \mathbf{U\dot{x}} \text{ etc.} \]

Since they act on mixed energy states, \( \mathbf{\dot{x}}, \mathbf{\dot{p}}, \mathbf{\dot{B}}, \) and \( \mathbf{\dot{\sigma}} \) are not really observables. However projection operators onto pure positive or negative energy states can be defined, in which case \( \mathbf{\dot{x}}, \mathbf{\dot{p}}, \frac{\mathbf{E}}{E}, \) and \( \mathbf{\dot{S}} \) are the observable projections of \( \mathbf{\dot{x}}, \mathbf{\dot{p}}, \mathbf{\dot{B}}, \) and \( \frac{\hbar}{2} \mathbf{\dot{\sigma}} \).

Noting that the observable projection of an operator which commutes with \( E \) is also the time average of that operator, Pryce concludes that \( \frac{\mathbf{p}}{H} \) is the velocity with the Zitterbewegung of \( \mathbf{\dot{\alpha}} \) averaged out and \( \mathbf{\dot{S}} \) is the average spin.

He then goes on to spins 0 and 1 and finds \( \mathbf{\dot{x}}, \mathbf{\dot{p}}, \mathbf{\dot{q}}, \) and \( \mathbf{\dot{Q}} \) for both cases. For all cases he finds \( \mathbf{\dot{x}} \), or definition (d), the most satisfactory. He interprets the uncertainty in measurement of its components, which is of order \( \frac{\hbar}{mc} \), as being a natural result of the measurement process. In order to measure to such a small distance as the Compton wavelength it is necessary to use enough energy to create a pair of particles. This argument is in some respects fallacious. It does not hold for spin zero, and its fallacies for
spinning particles will be discussed later.

So much attention has been paid to Dirac, Møller, and Pryce because in their work can be seen the seeds of essentially every position operator subsequently invented. Newton and Wigner\(^{10}\) did have a new approach but they came up with Pryce's \(\text{e}\) as their position operator.

Newton and Wigner\(^{10}\) are often cited as having written the definitive paper on position operators. They postulate four conditions for localized states of elementary systems. A localized state is one for which the three space coordinates \(\vec{x} = 0\) at \(t = 0\). An elementary system is a set of states which forms an irreducible representation space for the Inhomogeneous Lorentz Group. The postulates are:

a) the states form a linear set \(S_0\);

b) the set \(S_0\) is invariant under rotations about \(\vec{x} = 0\) and reflections of both spatial and time coordinates;

c) if a state \(\psi\) is localized at \(\vec{x} = 0, t = 0\) then a displacement makes \(\psi\) orthogonal to all states of \(S_0\); and

d) the infinitesimal operators of the ILG must be applicable to \(S_0\).

These postulates give for spin zero

\[ \psi = (2\pi)^{-3/2} p_0^{1/2} \]  \(3.119\)

as the only state localized at the origin. In coordinate space it becomes

\[ \psi(r) = \left(\frac{m}{r}\right)^{5/4} H_{5/4}(imr) \]  \(3.120\)
where \( m \) is mass, \( p_0 = (p^2 + m^2)^{1/2}, \) \( \text{H}(1) \) is the Hankel function of the first kind, and \( \hat{r} \) is \( \hat{x} \).

The state which is localized at \( \hat{x} \) at time \( t = 0 \) is found by applying the displacement operator \( \hat{p} \) to (3.119). It is

\[
\psi(\hat{x}) = (2\pi)^{-3/2} p_0^{1/2} e^{-i(\hat{p},\hat{x})}. \tag{3.121}
\]

This must be an eigenfunction of the position operator \( q^k \), with eigenvalue \( x^k \), which gives

\[
q^k = -i \left( \frac{\partial}{\partial p^k} + \frac{p^k}{2p_0} \right), \tag{3.122}
\]

in momentum space. In coordinate space

\[
q^k \phi(x) = x^k \phi(x) + \frac{1}{\delta m} \int \frac{e^{-m|x-y|}}{|x-y|} \frac{\partial \phi(y)}{\partial y_k} \, dy. \tag{3.123}
\]

This \( q^k \) is Pryce's (e). It has commuting components but no simple covariant meaning, since a state localized at the origin in one coordinate system is not localized in a moving system.

For spin not zero the localized wave functions are the \( 2s + 1 \) functions

\[
\psi_m = (2\pi)^{-3/2} 2^s p_0^{2s+1/2} (p_0 + m)^{-s} \times V_m(p_1p_2p_3; \xi_1 \ldots \xi_{2s}). \tag{3.124}
\]

where \( m \) runs from \(-s\) to \( s \) and \( V_m \) are given by

\[
V_m = E \gamma_m
\]

\[
E = E_1 E_2 \cdots E_{2s}
\]

\[
E_\alpha = \frac{1}{2} p_0^{-1} (\sum \gamma_k^\alpha p_k + m) \gamma_\alpha^0
\]
\[ y_\alpha \] are generalized \( y \) matrices
\[ \alpha = 1, 2, \ldots, 2s \]
\[ \xi \alpha \] are spin variables
\[ \nu_m = \frac{(s+m)!(s-m)!}{(2s)!} 1/2 \cdot 2^{-s} \]

The position operator is
\[ q^k = E \prod_{\alpha=1}^{2s} \left( 1 + y_\alpha \right) \frac{p_0^{2s+1/2}}{(p_0 + m)^s} \left( -i \frac{\delta}{\partial p_k} \right) \frac{p_0^{-1/2}}{(p_0 + m)^s} E \]

which for \( s = 1/2 \) is still Pryce's (e). The operators \( E \) are projection operators eliminating negative energy parts of any wave functions to which \( q^k \) is applied. The commutator of \( q^k \) with \( p^j \) is
\[ [q^k, p^j] = -i \delta_{kj} \]

and other commutators of \( q^k \) with Lorentz generators are the usual ones.

The localized states are not \( \delta \)-functions because they represent pure positive energy states. The operator \( q^k \) also keeps the positive energy nature of any state on which it acts. Newton and Wigner refute the argument that a position measurement within the Compton wavelength should lead to pair production. They claim it should lead to the position of the particle, not to a particle and some pairs. If the pair-creation argument is accepted then position is not observable in any traditional sense. The localized states and the operator \( q^k \) are reasonable for elementary systems. In scattering
problems where one customarily measures only cross sections the position of the particles during the interaction and the nature of the interaction in space might be measurable if one had the correct formulation.

Foldy and Wouthuysen investigated the transformation used by Pryce as equation (3.116) and stated it as equation (3.31). Their position operator (3.35) thus corresponds to Pryce's $\tilde{q}$ of equation (3.114) and (3.112), or definition (e). This is also Newton and Wigner's operator $q^k$ of equation (3.125).

Bakamjian and Thomas describe a system of particles by center of mass coordinates and internal coordinates. They give up having their particles trace out invariant world lines and derive

$$\dot{x} = \frac{1}{m} \dot{k} + \frac{i}{m(m+\Omega)} \dot{J} \times \dot{p} - \frac{\dot{p}(\dot{k}, \dot{p})}{m\hbar(m+\H)}$$

only they call $\dot{k}, \dot{J}, \dot{p}$ by the symbols $\dot{V}, \dot{\Omega}, \dot{R}$, respectively. They use representations similar to (3.28) for the Lorentz generators and a transformation similar to (3.31) to get their operator. Their notation is obscure and there seems to be nothing in their operator which wasn't in Pryce.

Case generalizes the Foldy-Wouthuysen transformation (3.31) to spin zero and one theories. The Hamiltonian in the absence of fields is given by

$$H = \rho_3 \left( mc^2 + \frac{\dot{p}^2}{2m} \right) + i\rho_1 \left( \frac{\dot{p}^2}{2m} - \frac{(S-P)^2}{m} \right),$$

where $\dot{p}$ and $\dot{S}$ are suitable spin zero or one operators. Letting the position be $\dot{x}$, which is multiplication by a coordinate, one
obtains for the velocity
\[ \mathbf{\dot{V}} = \frac{\mathbf{d}\mathbf{x}}{dt} = \frac{i}{\hbar} [H, \mathbf{x}] = \frac{\mathbf{p}}{m} (\rho_3 - i\rho_2) + i\rho_2 \frac{m^2}{2m} (\mathbf{S} \cdot \mathbf{p}) + (\mathbf{S} \cdot \mathbf{p}) \mathbf{S}. \]  
(3.129)

For spin zero this gives the result that the square of any velocity component is zero, while for spin one it is negative. Also for spin one the velocity components do not commute and \( \mathbf{S} \) is not a constant even in the absence of fields.

The transformation
\[ T = i\rho_1 \left( \frac{2}{2m} - \frac{(\mathbf{S} \cdot \mathbf{p})^2}{m} \right) + \rho_3 (mc^2 + \frac{p^2}{2m}) \]  
(3.130)
leads to Case's similarity transform
\[ T' = \rho_3 (T^2)^{1/2} \]  
(3.131)
which is his generalized Foldy-Wouthuysen transformation. Applying \( T' \) to the \( H \) of (3.128) gives
\[ H' = \rho_3 E_p \]  
(3.132)
where \( E_p \) is \( (m^2 + p^2)^{1/2} \). In the discussion of the Foldy-Wouthuysen results in equations (3.30) through (3.37) the symbol \( H \) was used instead of \( E_p \).

Application of \( T' \) to \( \mathbf{x} \) gives
\[ \mathbf{x}' = \mathbf{x} - \frac{\hbar \rho_2}{2} \left\{ -ic^2 \frac{\mathbf{p}}{E_p^2} - 2ic^2 \frac{\mathbf{p}(\mathbf{S} \cdot \mathbf{p})^2}{mE_p^2 (E_p + mc^2)} + \frac{i}{mE_p} (\mathbf{S} (\mathbf{S} \cdot \mathbf{p}) + (\mathbf{S} \cdot \mathbf{p}) \mathbf{S}) \mathbf{p} \right\} + \frac{\rho_2}{mE_p} \frac{(E_d - mc^2) \mathbf{p} \times \mathbf{S}}{E_p + mc^2} \]  
(3.133)

The operator \( \mathbf{x} \) whose transformation \( \mathbf{x}' \) is
as suggested by the spin 1/2 case, is (3.133) with $\rho_2$ replaced by $-\rho_2$. It is found that

$$\frac{d\mathbf{x}'}{dt} = \rho_3 c^2 \frac{\mathbf{p}}{E_p}$$

(3.135)

so $\mathbf{x}'$ is the integral spin mean position operator.

Case claims that his results can be extended to higher spin. Bunge derives a mean position operator without using the Foldy-Wouthuysen transformation. It is

$$\chi^\mu = x^\mu + \frac{\hbar}{2m_0c} i \gamma^\mu,$$

(3.136)

where $x^\mu$ is the Dirac operator (3.57). The space components $X_i$ perform a smooth motion and their eigenvalues are $x^i \pm \frac{\Lambda}{2}$ where $\Lambda = \frac{\hbar}{m_0c}$. The velocity is

$$v^\mu = \frac{dX^\mu}{dt} = \frac{1}{m_0} \gamma_0 p^\mu$$

(3.137)

so that the $\gamma^\mu$ term cancels the Zitterbewegung. Bunge chooses to interpret his results as describing an electron with mass smeared out in a region the size of the Compton wavelength and charge localized at a point with trembler motion.

Melvin in the appendix to a review article on elementary particles and symmetry principles, includes a table of the commutators of quantities which can be constructed out of generators of the inhomogeneous Lorentz group. He uses various of these quantities to construct a position operator with commuting components.
but without Lorentz covariance. This "centroidal position operator" is

\[ X^k = \frac{1}{2} (R^k R^{-1} + E^{-1} R^k) \]  \hspace{1cm} (3.138)

where \( R^k \) is an integral over the energy density. \( X^k \) has the canonical velocity and commutation relation with \( P_j \). Two operators, \( Q_+ \) and \( Q_- \), can be formed from \( X \) by projecting out positive and negative energy components. Then one has the Pryce (e) again.

Mathews and Sankaranarayanan 15,16,17 wrote several papers on observables of particles, three of which bear special relevance. For a Dirac particle they postulate a position \( \hat{x} \) which satisfies five conditions:

(i) \([X_i, P_j] = i \delta_{ij}\);

(ii) \([X_i, \frac{H}{E}] = 0 \) so that \( \hat{x} \) is defined separately on positive and negative energy states;

(iii) \( \hat{x} = \frac{\hat{p} H}{E} \);

(iv) \([X_i, X_j] = 0 \); and

(v) \( \hat{x} \) is a polar vector under space reflections and invariant with respect to time inversions.

They proceed to find the most general operator satisfying these conditions. Letting \( \hat{x} \) be the coordinate variable, they find

\[ \hat{x} = \hat{\hat{x}} + A \hat{\sigma} \times \hat{\hat{p}} + B \hat{\sigma} \times \hat{\hat{p}} \]  \hspace{1cm} (3.139)

where \( A \) and \( B \) are functions of \( \hat{p} \) and \( m \), homogeneous and of degree -2. This \( \hat{x} \) is in the Foldy-Wouthuysen representation and does not yet satisfy condition (iv). In the Dirac representation it is
\[ \dot{x} = \ddot{x} + \frac{1\beta}{2E} - \frac{i\beta(\ddot{\sigma} \cdot \ddot{p})}{2E^2(E + m)} + \frac{A}{E} (m(\ddot{\sigma} \times \ddot{p}) + i\beta \dot{p}^2) \]

(3.140)

For \( A = \frac{(E - 1)}{2p^2} \) and \( B = 0 \), \( \dot{x} \) becomes Pryce's (d) of equation (3.102) or (3.105). For \( A = \frac{(m - 1)}{2p^2} \) and \( B = 0 \), \( \ddot{x} \) becomes Pryce's (c) of equation (3.100). For \( A = -\frac{1}{2p^2} \) and \( B = 0 \), another operator is obtained. With \( A = B = 0 \), the familiar Pryce's (e) appears.

Now using condition (iv) to restrict \( \dot{x} \) to be "local", the following four cases are possible solutions.

\[
\begin{align*}
A = B &= 0 \\
A &= -\frac{1}{2p^2} \quad B = 0 \\
A &= -\frac{1}{2p^2} \\
A &= -\frac{1}{2p^2} \quad B &= +\frac{1}{2p^2}
\end{align*}
\]

(3.141)

Back in the Foldy-Wouthuysen picture these operators are

\[
\begin{align*}
\dot{x} &= \ddot{x} - \frac{1}{p^2} \dddot{\sigma} \times \dddot{p} \\
\dddot{x} &= \dddot{x} \\
\dddot{x} &= \dddot{x} - \frac{1}{2p^2} (\dddot{\sigma} \times \dddot{p})(1 + \beta) \\
\dddot{x} &= \dddot{x} - \frac{1}{2p^2} (\dddot{\sigma} \times \dddot{p})(1 - \beta).
\end{align*}
\]

(3.142)

Spin operators can be defined for each of these position operators, so that both spin and orbital angular momentum are constants of the motion.

They next find the unitary transformations taking the operator \( \dddot{x} \) into the others. The three transformations taking this Foldy-Wouthuysen (F.W.) operator (Pryce's (e)) into the other
operators can be called $U_1$, $U_2$, $U_3$. It is noted that if $U$ is the
F. W. transformation then $U_1 U$, $U_2 U$, and $U_3 U$ also take the Dirac
Hamiltonian into the form $BE$.

The transformations $U_1$, $U_2$, $U_3$ lead to localized states which
correspond to Newton and Wigner's states as the three operators of
(3.142) correspond to $\hat{X} = \hat{P}$ (Pryce's (e)). Positive and negative
energy eigenfunctions are obtained but they do not satisfy the regu­
larity condition of Newton and Wigner, which can be interpreted as
a condition that the states be uniquely defined in the rest system
of a particle. As long as a rest system exists, such terms as $\frac{\hat{S} \cdot \hat{P}}{\hat{P}}$
which appear in these states have no unique limit. For massless
particles, however, all four states and all four operators are allow­
able. Actually two of the operators are not independent so the
states reduce to two from four. One has positive parity and one has
negative parity.

Mathews and Sankaranarayanan next 17 extend their results to
spins 0 and 1. They drop the condition (iii) that $\hat{X} = \frac{\hat{P} H}{E}$ and
add the condition (vi) that $\hat{X}$ should be uniquely defined in the
zero-momentum limit if the particle has a rest system. The re­
quirements are not restricted to any spin but to find the operators
one needs Hamiltonians. They use the Hamiltonian

$$
H = \tau_3 \left( m + \frac{p^2}{2m} \right) + i\tau_2 \left( \frac{p^2}{2m} - \frac{(\hat{S} \cdot \hat{P})^2}{m} \right)
$$
(3.143)

where the $\tau$'s are Pauli matrices and $\hat{S}$ is 0 for spin zero and the
spin one representation of angular momentum for spin one. Since
the form is the same for both spin zero and spin one, they are
solved simultaneously even though the two cases have different
dimensionalities. Observables are pseudo-Hermitian which causes
problems of the sort noted by Case in equation (3.129). In
fact Case's transformation (3.131) is used to bring $H$ into the
form $H' = \tau_3 E$. Then position operators $\hat{x}$ are found similar to
(3.142).

$$\hat{x} = \hat{x} + (A + B\tau_3)(\hat{S} \times \hat{p})$$ (3.144)

where

$$\begin{align*}
A &= 0 & B &= 0 \\
A &= -2/p^2 & B &= 0 \\
A &= 1/p^2 & B &= 1/p^2 \\
A &= 1/p^2 & B &= -1/p^2 \\
\end{align*}$$ (3.145)

give four solutions. $A = 0$ and $B = 0$ leads to Case's operator
(3.133) in the Dirac representation. The other three cases are
new, in analogy to the spin $1/2$ situation. They can be ruled out
for massive particles by requirement (vi). For spin zero the
operator $A = 0, B = 0$ is the Newton-Wigner spin zero operator.

Wightman made a long and very rigorous study of the local-
izability of quantum mechanical systems. He points out that a
position operator which does not project out only positive energy
states is unphysical and that one which does not have commuting
components is unobservable. Until a satisfactory dynamical theory
is constructed, position operators can be determined only by rel-
ativistic kinematics and their actual observability will remain
questionable. Wightman extends Newton and Wigner's results from
localization at a point to localization in a region and makes the whole treatment mathematically rigorous.

Jordan and Mukunda \(^3\) did an excellent job of summarizing the Lorentz-covariant position operators for spinning particles as shown in the first part of this chapter.

Their (3.24) with \(a = 0\) is Pryce's (d). If one lets \(\hat{x} = \hat{q}\) instead of (3.24) one has Pryce's (e) and the Newton-Wigner position. This form is not covariant, while (3.24) is not local. Their (3.29) is the Foldy-Wouthuysen representation of the Dirac \(\hat{x}\).

Philips \(^{19}\) further extended the discussion of localized Lorentz invariant states. He proposed an alternative set of postulates to Newton and Wigner's, which includes Lorentz invariance, normalizability, and irreducibility instead of orthogonality. He solves for states only in the spin zero case. Since the Newton-Wigner solutions were not Lorentz invariant, some sacrifice is made by Philips of N-W conditions. Sacrificing orthogonality means he gets non-Hermitian position operators. His state is

\[
\psi(x) = \frac{m}{r} K_1(mr) \tag{3.146}
\]

which can be compared to the \(\tilde{n}-\tilde{w}\) state

\[
\psi_{nw}(x) = \left(\frac{m}{r}\right)^{5/4} K_{5/4}(mr). \tag{3.147}
\]

Here the argument is changed from (3.120). Philips finds no satisfactory position operator.

Sankaranarayanan and Good \(^{20}\) use the properties

(i) Hermiticity for all spin;
(ii) \([X_i, X_j] = 0\)
\([X_i, P_j] = i\delta_{ij}\);

(iii) charge-conjugation invariance;

(iv) \(\dot{\vec{x}} = \frac{\vec{p}}{E}\);

(v) polar vector under rotations and space reflections and
invariance under time reflections;

(vi) separate particle and antiparticle definitions; and

(vii) well-defined in rest system of the particle

to determine a position \(\vec{x}\). This is essentially the Newton-Wigner
\(\vec{x}\) but Sankaranarayanan and Good extend it to apply to mixed-energy
systems. It is also the Foldy-Wouthuysen mean position (3.35) for
spin 1/2, but it differs from the general Jordan-Mukunda (3.29) in
being designed to act on different wave functions.

Their three-vector \(\vec{x}\) is given by
\[
\vec{x} = \vec{x} + \left[ m(E + m) \right]^{-1} \vec{s} \times \vec{p} - \frac{is\vec{c}\vec{h}}{mE} + i\left[ mE^2(E + m) \right]^{-1} \vec{s} \cdot \vec{p} H\vec{p} + \frac{i\vec{p}}{2E^2},
\]

(3.148)

where \(\vec{x}\) is the coordinate \(\vec{x}\) and \(\vec{s}\) is spin. For spin zero this
becomes
\[
\vec{x} = \vec{x} + \frac{i\vec{p}}{2E^2},
\]

(3.149)

which is like the NW (3.122) except that the NW operator acts on
positive frequency states only. For spin 1/2 the Sankaranarayanan-
Good operator becomes
\[
\vec{x} = \vec{x} + \frac{i\vec{c}}{2E} - \left[ 2E^2(E + M) \right]^{-1} [i\vec{c} \cdot \vec{p} \vec{p} + \vec{s} \times \vec{p} E]
\]

(3.150)

which is the FW mean position operator, NW spin 1/2 operator, and
Pryce's (e). The formula for spin 1 is
\[
\hat{x} = \hat{x} + \frac{i\hat{p}}{2E^2} + \frac{i \partial}{m} E (E + m) (2E^2 - m^2)^{-1} \left[ (2E^2 + 2E_m + m^2) i \beta \hat{\alpha} \cdot \hat{p} \right] \\
-2iE(\hat{\alpha} \cdot \hat{p})^2 \hat{p} + [m(E + m)]^{-1} \hat{s} \times \hat{p} - (m(2E^2 - m^2))^{-1} \hat{x} \times 2(iE\hat{\alpha} \hat{\alpha} \cdot \hat{p} + \beta(\hat{\alpha} \times \hat{p}) \hat{\beta} \cdot \hat{p}).
\] (3.151)

They then construct a four-vector position \( \vec{y}_\mu \), the three-vector part of which is closely related to \( \hat{x} \). It is defined by
\[
\vec{y}_\mu = -\frac{1}{2m^2} [M_{\mu \rho}, P_{\rho}] + \] (3.152)
and it is used in Appendix D. \( M_{\mu \nu} \) and \( Y_\mu \) satisfy the commutation rules for the generators of the deSitter group and \( Y_\mu \) is a generalization of the component of position normal to the worldline of the particle.

Fleming 1 and 21 made a significant contribution to the study of position operators by extending the idea of Pryce expressed in equations (3.98) through (3.100) to create a whole formalism of hyperplanes. Chapter II gives some results of the hyperplane formalism, though not in Fleming's original notation. He used the term "local" to mean \([X_i, X_j] = 0\) and he changed the Lorentz covariance condition to be a "pointness" condition, as illustrated in the diagram on page 26 of Chapter III, and discussed in Chapter IV.

Lorentz covariance means frame independence. If an operator is defined so as to depend on the frame in which it is defined then it will not appear to be covariant. Fleming 1 points out
that positions which describe the location of some dynamical property might well be expected to be frame dependent. An example is the center of mass of the earth, which sits in a different rock inside the earth depending on who measures it. Spinning particles in general have mass distributions which depend on the frame from which they are observed. On the other hand an operator describing some space-time point can be defined invariantly. Fleming's operators and their interpretations will be dealt with in Chapter IV.

Berg considered position and spin operators which satisfy the usual canonical commutation relations (3.22) except for \([q_i, q_j] = 0\) and including \([q_i, E] = iV_i\). Also he requires that \(\hat{M} = \hat{\mathbf{q}} \times \hat{\mathbf{P}} + \hat{\mathbf{S}}\). This differs from Jordan and Mukunda in the last two requirements and also in lacking the last of conditions (3.17). His operator is the same as Jordan and Mukunda's for spin 1/2 and is Sankaranarayanan and Good's for other spin. For spin 1/2 all three become the FW operator.

Recently Johnson took an algebraic approach to constructing four-vector position operators \(X^U\) which form a Lie algebra when added to the ILG generators. He studies this algebra, constructs a Hamiltonian for all spins, and uses a proper-time framework. He shows the connection between his formalism and standard formalisms such as Dirac, Klein-Gordon, and Maxwell.

Miller was a student of Fleming whose work will be discussed in Chapter IV. Fleming elaborated his theory of position operators in a lecture series, the results of which will appear in Chapter IV. Currie, Jordan, and Sudarshan were responsible for one of the
proofs of the "invariant worldline hypothesis" referred to early in this chapter, as well as the "no interaction theorem". McDonald \(^{27,28}\) has done a great deal of work in the hyperplane formalism, including a construction of Newton and Wigner's \(^{10}\) localized states and local operator in hyperplane variables \(^{27}\) and a construction of a relative position operator for two-particle systems. \(^{26}\)

While some authors have been omitted from this review, to my knowledge none who made a really original contribution to the subject has been overlooked.
A POINT AND LOCAL POSITION OPERATOR

Pointness

The requirement that a relativistic operator be Lorentz co-
variant means that it will transform as a scalar, four-vector, or
tensor under the Inhomogeneous Lorentz Group. For a space-
and time-independent four-vector operator $A^\mu$ this is stated as

$$[A^\mu, M^{\nu\rho}] = i(g^{\nu\rho}A^\rho - g^{\rho\nu}A^\nu) \quad (4.1)$$

where $M^{\nu\rho}$ is the generator of rotations and pure Lorentz trans­
formations. Setting $M^{\rho\sigma} = K^\rho_\sigma$ we get

$$[A^i, K^j] = i\delta^i_j A^0 \quad (4.2)$$

$$[A^0, K^i] = i A^i$$

where $K$ generates pure Lorentz transformations, or boosts. Any
four-vector which is not a function of space and time will obey
(4.1) and (4.2), while a scalar will commute with $K$ and $M^{\nu\rho}$ and
a tensor will have a more complicated commutator.

If, however, $A^\mu$ is allowed to be a function of the coordinates
$x$ and $t$ then $A^\mu(\vec{x}, t)$ will have a more complicated commutator with
$M^{\nu\rho}$. For example, letting $A^\mu(\vec{x}, t)$ be a vector field we get

$$[A^\mu(\vec{x}, t), M^{\nu\rho}] = i(g^{\nu\rho}A^\rho(\vec{x}, t) - g^{\rho\nu}A^\nu(\vec{x}, t)) + i(x^\nu \frac{\partial}{\partial x^\rho} - x^\rho \frac{\partial}{\partial x^\nu})A^\mu(\vec{x}, t). \quad (4.3)$$
In three-vector terms

\[ [A^i_K^j] = i\delta^{ij}A^0 + i(t \frac{\partial}{\partial x_j} - x^j \frac{\partial}{\partial t})A^i(x,t). \]

\[ [A^0_K^j] = iA^j + i(t \frac{\partial}{\partial x_j} - x^j \frac{\partial}{\partial t})A^0(x,t). \]  

(4.4)

Any vector function of \( x \) and \( t \) will have to have its commutation relations derived by considering how it is defined. In particular the position operator will have a complicated commutator with \( K \), to which we now turn.

A classical point particle traces out a worldline in space-time which can be described by a four-vector position operator \( X^\mu(\tau) \) where \( \tau \) is proper time. The position measured by any observer will be the three-vector \( \mathbf{x}(t) \) where \( t \) is that observer's time parameter. The four vector \( X^\mu(\tau) \) is

\[ X^\mu(\tau) = (t(\tau), \mathbf{x}(\tau)). \]  

(4.5)

Given \( \tau \), one can solve for \( t \) and then \( \mathbf{x}(t) \), though the usual route is to solve for \( \tau \) as a function of \( t \).

Let us examine how this point particle position transforms under an infinitesimal Lorentz transformation. Two observers, primed and unprimed, observe the worldline \( X^\mu(\tau) \). Both measure the position at time \( t_0 \), so

\[ X^0(\tau') = X^0(\tau) = t_0. \]  

(4.6)

Thus they measure two different points on the worldline, at parameter values \( \tau \) and \( \tau' \). The usual covariance holds at any point. If \( \Lambda^\mu_\nu \) is the Lorentz transformation, then
\[ X^\mu(\tau') = \Lambda^\mu_\nu X'^\nu(\tau') \]
\[ = X'^\mu(\tau') + \varepsilon^\mu_\nu X'^\nu(\tau') \quad (4.7) \]
when \( \Lambda^\mu_\nu \) is an infinitesimal transformation of the coordinates.

Also, looking at the variation along the worldline, we find

\[ X^\mu(\tau') = X^\mu(\tau) + \frac{dX^\mu}{d\tau} d\tau. \quad (4.8) \]

The diagram illustrates these equations.

Combining (4.7) and (4.6) we find

\[ X'^\mu(\tau') - X^\mu(\tau) = \frac{dx^\mu}{d\tau} d\tau - \varepsilon^\mu_\nu x^\nu \quad (4.9) \]

to first order. Use of (4.6) for \( \mu = 0 \) leads to

\[ d\tau = \frac{\varepsilon^0_\nu x^\nu}{dt/d\tau} \quad \text{for} \quad x^0 = t \quad (4.10) \]

so that

\[ X'^\mu(\tau') - X^\mu(\tau) = \varepsilon^0_\nu \frac{dx^\mu}{dt} x^\nu - \varepsilon^\mu_\nu x^\nu. \quad (4.11) \]

Note that (4.7) is for fixed \( \tau \) while (4.11) is for fixed \( t \) but varying \( \tau \) and the whole derivation is for a moving point.

The \( \mu = 0 \) term of (4.11) is zero by construction. For infinitesimal transformations \( \varepsilon^0_\nu = 0 \) so for \( \mu = i \)
\[ x^i(t') - x^i(t) = \varepsilon^0_j \frac{d}{dt} x^j - \varepsilon^i_j x^j - \varepsilon^i_0 x^0. \tag{4.12} \]

The \( \varepsilon^i_j \) refers to rotations, so for pure boosts
\[ x^i(t') - x^i(t) = \varepsilon^0_j \frac{d}{dt} x^j - \varepsilon^i_0 t. \tag{4.13} \]

Translating this into commutator language we have the result
\[ [x^i(t), \mathcal{K}^j] = i\delta^i_j t - i \frac{d}{dt} x^j, \tag{4.14} \]

which resembles (4.4) for fields with only \( t \)-dependence. This resemblance is due to the similarity of the operator \( \dot{x} \) and the coordinate \( x \).

A worldline traced out by a point particle is called an invariant worldline by Pryce, Fleming, and Currie, Jordan, and Sudarshan. This is because two observers will agree on the uniqueness of the worldline, and the points they measure at fixed \( \tau \) will coincide. Equation (4.7) states the covariance of the worldline.

But it is the pointness of the particle which allows it to have an invariant worldline in this case, and pointness leads to equation (4.14), which gives new information about \( \dot{x}(t) \). Fleming first used the term "point" in this sense.

The commutator (4.14) can be rewritten
\[
[x^i(t), \mathcal{K}^j] = i\delta^i_j t - i \dot{x} x^j \\
= i\delta^i_j t + [x^i, \mathcal{H}] x^j \\
= i\delta^i_j t + \frac{1}{2}(x^j [x^i, \mathcal{H}] + [x^i, \mathcal{H}] x^j) \tag{4.15}
\]
where the second line uses the Poisson bracket relation
\[ \dot{X} = -i \{X, \dot{H}\} \]  
(4.16)
and the third line symmetrizes the commutators for the quantum mechanical case.

The above derivation has been for the time-dependent case. To remove the time dependence we use
\[ X^i(t) = e^{iHt} X^i e^{-iHt} \]  
(4.17)
and calculate
\[ [X^i, K^j] = e^{-iHt} [X^i(t), K^j] e^{iHt} - i\delta^{ij} t. \]  
(4.18)
It should be noted that all of the $X's$ in (4.15) were time-dependent.

Thus the final result for a three-vector position of a point particle, with no explicit time-dependence, is
\[ [X^i, K^j] = \frac{1}{2} \{X^j [X^i, H] + [X^i, H] X^j\} \]  
(4.19)
which agrees with previous authors cited in Chapter III.

The step (4.17) to (4.18) made use of
\[ [X^i(t), P^j] = i\delta^{ij} \]  
(4.20)
so that the canonical commutation relation for $\dot{X}$ with $\dot{P}$ is assumed, as well as the fact that $[\dot{X}, H] = i\dot{X}$ is a legitimate velocity.

Whenever a three-vector position $\dot{X}$ obeys the pointness relation (4.19), $\dot{X}(t)$ and $t$ can be combined into a four-vector
\[ X^\mu(\tau) = (t(\tau), \dot{X}(\tau)) \]  
(4.5)
which is a function of a scalar parameter $\tau$. This four-vector $X^\mu(\tau)$ is covariant for fixed $\tau$ in the sense that it obeys the
standard relation (4.7). If \( \dot{x} \) does not obey (4.19), then \( \dot{x} \) is not a point position operator, but it may still be covariant.

Not all the position operators one might consider are pointlike. The "position" of a particle could refer to the average position of any number of dynamical properties, such as mass, charge, or spin. In classical physics the center of mass is usually taken to be the logical definition of the position of a particle or system of particles, and the classical center of mass does behave as a point. But the center of total energy classically will shift from point to point depending on who measures it. For example, it can be shown for a spherically symmetric rotating classical particle the center of energy is

\[
\dot{x} = \dot{x}_{\text{non-rotating}} + \frac{1}{2mc^2} \mathbf{v} \times \mathbf{s}
\]  

(4.21)

where \( \mathbf{s} = I \cdot \omega \). If two observers are moving with velocity \( \mathbf{v}' \) relative to each other, their values of \( \dot{x} \) will not shift by \( \mathbf{v}'t \), as they would for \( \dot{x}_{\text{non-rotating}} \). There will be an extra term \( \frac{1}{2mc^2} \mathbf{v}' \times \mathbf{s} \) which shifts the position. Thus the position does not trace out an invariant worldline.

As has been seen in Chapter III there are many ways to generalize center of mass, and in many cases the quantum mechanical relativistic "center of" something will be dependent on the frame in which it is measured. Thus a position of a dynamical property in general is not a point operator obeying (4.19). It was long assumed that such an operator was not covariant, that is did not obey (4.7), which is the usual statement of covariance. But even though the four-vector made
from $\hat{x}(t)$ and $t$ does not look covariant, it must be examined carefully before concluding noncovariance. And even though it may not be covariant, the commutators of $\hat{x}$ with other operators are interesting.

The center of mass as defined by Möller in equation (3.77) or Pryce's definitions (a), (c), and (e) or Newton and Wigner's (3.123) are all examples of frame-dependent position operators which do not obey (4.7), and have therefore been called noncovariant. However, using Fleming's formalism such an operator can be made to appear covariant by replacing $t$ and other timelike quantities by contractions with the $\eta^\mu$ of equation (2.1). In one's own frame (called the "instantaneous" frame by Fleming)

$$\eta^\mu = (1,0,0,0), \quad (4.22)$$

so that $\eta^\mu A_\mu$, where $A_\mu$ is any four-vector, gives $A_0$, the timelike component of $A_\mu$.

$$\eta^\mu A_\mu = A_0 \quad \text{in the instantaneous frame.} \quad (4.23)$$

When the center of mass is calculated by an observer it is done at one instant of time in that observer's frame:

$$X_i(t) = \frac{\int d^3 x \ \Theta_{00}(\bar{x},ct) \ x_i}{\int d^3 x \ \Theta_{00}(\bar{x},ct)} \quad (4.24)$$

$$X_0(t) = ct,$$

where $\Theta_{00}(\bar{x},ct)$ is the energy per unit volume at the point $\bar{x}$ and instant $t$. How $X_i(t)$ transforms is determined by how that averaging procedure looks to a second observer in motion relative to the first.
The second observer will believe that the averaging was done over a time interval, not at a single instant, and that what was measured was a combination of energy and momentum. Labeling the second observer with primes, one finds that what he believes was measured is \( \Lambda_0^{-1} \lambda_0^{-1} \nu_\lambda \Theta^\prime_0(x') = \Theta_{00}(x) \) over the points \( \Lambda_0^{-1} \nu \nu_\nu = ct = x_0 \) in the unprimed observer's frame. Here \( \Lambda_\mu^{-1} \nu \) is the Lorentz transformation from the primed to the unprimed frame.

Using (4.3) we can rewrite the above equations. The set of space-time points over which \( \Theta_{\mu\nu} \) is measured form a spacelike hyperplane which is labeled by (4.22) in the unprimed frame and by some different \( \eta' \mu \) in the primed frame,

\[
\eta'_\mu = \Lambda_\mu^\nu \eta_\nu = \Lambda_\mu^0.
\]  

(4.25)

Note Fleming uses \( \tau \) to mean the perpendicular distance from an observer's origin to a hyperplane. If we include translations in our transformations, \( \tau \) transforms as

\[
\tau' = \tau + \eta'\nu a_\mu.
\]  

(4.26)

The two observers represent \( \Theta_{00}(x) \) by

\[
\Theta_{00}(x) = \eta^{\nu} \Theta_{\nu\nu}(x) \eta^\nu
= \eta^{\nu} \Theta_{\nu'\nu}(x') \eta'^\rho,
\]  

(4.27)

which is a scalar quantity, being the double inner product of \( \Theta \) with \( \eta \). It is a simple step to then rewrite the symbols of integrations as

\[
\int d^3x = \int d^4x \delta(x_0 - ct) = \int d^4x \delta(nx - \tau)
\]  

(4.28)

and finally define
\[ X_\mu(n,\tau) = \int \frac{d^4x \delta(\eta x - \tau) \eta^\rho \delta^\lambda(x) \eta \lambda x_\mu}{d^4x \delta(\eta x - \tau) \eta^\rho \delta^\lambda(x) \eta \lambda} \]  

(4.29)

as the center of mass. \( X_\mu \) is a four vector but its dependence on \( n \) and \( \tau \) (see (2.1)) really constrains it to be a three-vector function of \( t \) on any hyperplane. The right hand side of (4.29) is manifestly covariant, being constructed of Lorentz scalars and vectors, so

\[ X'_\mu(n',\tau') = \Lambda_\mu^\nu X_\nu(n,\tau) + a_\mu. \]  

(4.30)

Here \( a_\mu \) is included for generality but can be left out of the covariance argument.

Equation (4.30) appears to be a covariance equation for \( X_\mu(n,\tau) \). The diagram below illustrates the problem.

If the unprimed observer measures the center of mass over a constant-time slice of the worldtube of the particle, then (4.30) states that the primed observer can get the same number as long as he relabels all quantities and does the measurement over a \( t' \)-not-constant slice. If the primed observer measures over a \( t' = \) constant slice, that is,
carries out the same operations as the unprimed observer, he will in general get a different point for $\tilde{X}'(t')$, even if that slice goes through $\tilde{X}(t)$ as measured by the unprimed observer. That is, the worldline of $\tilde{X}$, the center of mass, is different for different observers.

If covariance is to mean that the operations of physics look the same in all frames, then (4.30) is not a statement of covariance, and the operator (4.24) still is not covariant. (4.30) refers to the purely kinematical situation of two observers looking at one hyperplane. Two hyperplanes as viewed by one observer is a dynamical situation, while two hyperplanes as viewed by two observers is a mixed dynamic and kinematic situation. Pointness is an example of a statement about the latter kind of situation. Most of the position operators which are "covariantly" defined in the usual sense will obey the pointness commutation relation (4.15). However, we cannot rigorously preclude the existence of a time-dependent covariant position operator obeying some commutation relation different from (4.15)
Locality

The term "local" is used to describe another property of position operators. If a particle is to have a "local" position then that particle must in some way be localized in space. Thus one way to find a local position operator is to find an operator whose eigenstates are the most localized possible.

Newton and Wigner \(^{10}\) followed this route. The particular requirements which they imposed on their localized states are given preceding equation (3.119). The orthogonality condition \(c\) requires that states localized about different points be non-overlapping,

\[
(\psi(\hat{\alpha}),\psi(0)) = \frac{1}{(2\pi)^3} \int \frac{d^3p}{p_0} \psi^*(\hat{\beta}) \psi(\hat{\beta})
\]

\[
= \delta^3(\hat{\alpha}) \quad (4.31)
\]

where \(\psi^*(\hat{\beta}) = e^{-i\hat{p} \cdot \hat{\beta}} \psi(\hat{\beta})\). This condition is satisfied if the momentum space wave function \(\psi(\hat{p})\) has the form

\[
\psi(\hat{p}) = (2\pi)^{-3/2} p_0^{1/2}. \quad (4.32)
\]

Note that \(p_0^{1/2}\) is a quantity which does not have nice properties under Lorentz transformations.

The Fourier transform of \(\psi(\hat{p})\) gives the coordinate-space wavefunction \(\psi(x)\),

\[
\psi_x(x) = (\text{constant})(\frac{m}{r})^{5/4} H_{5/4}^{(1)}(imr) \quad (4.33)
\]

where \(H_{5/4}^{(1)}\) is the Hankel function of the first kind and \(r\) measures the distance from the point \(x\) about which \(\psi\) is localized. Since \(\psi_x(x)\) is not zero for \(r \neq 0\), \(\psi_x(x)\) is not a delta function even
though $\psi^*\psi$ is completely localized. However $\psi$ falls off exponentially with increasing $r$, and is large only for distances from the point of localization which are smaller than the Compton wavelength of the particle. Only positive energy solutions were used to derive $\psi_x(r)$. These localized states are not covariant, and have the expected properties only on hyperplanes with $t = \text{constant}$.

The function $\psi_x(p)$

$$\psi(p) = (2\pi)^{-3/2} e^{-i\mathbf{p} \cdot \mathbf{x}} p_0^{1/2}$$

is an eigenfunction of the operator

$$\hat{x} = i\hat{\mathbf{p}} - \frac{1}{2} \frac{i\mathbf{p}}{p^2 + m^2}$$

with eigenvalue $\hat{x}$. $\hat{x}$ is an Hermitian operator, whereas $i\hat{\mathbf{p}}$ alone is not. Furthermore, the components of $\hat{x}$ commute,

$$[x^i, x^j] = 0$$

while $\hat{x}$ and $\hat{p}$ have the usual commutation relations

$$[x^i, p^j] = i\delta^{ij}.$$  

The time derivative $\dot{x}$ is given by

$$\dot{x} = i[x, p^0] = \frac{\mathbf{p}}{p^0}$$

as might be expected. Thus the Newton-Wigner $\dot{x}$ is apparently a good generalization of position.

The property of locality has been expressed in recent work in terms of the commutator $[x^i, x^j]$. Local position operators are required to satisfy the same commutation relation.
\[ [x^i, x^j] = 0 \]  

(4.36)

as was obtained for the operator (4.35) constructed from the localized eigenstates. The physical interpretation of (4.36) is that a measurement of one component of position does not interfere with the measurement of another component as may be seen from the proof of uncertainty relations. For any operators \( \alpha \) and \( \beta \)

\[ \Delta \alpha \Delta \beta \geq \frac{1}{2} |\langle [\alpha, \beta]\rangle| \]  

(4.38)

Thus, if \( [\alpha, \beta] = 0 \), \( \alpha \) and \( \beta \) can both be measured to arbitrary accuracy. On the other hand, (4.20) is the usual canonical commutation relation for \( \vec{x} \) and \( \vec{p} \), and the position operator \( \vec{x} \) and the momentum \( \vec{p} \) will satisfy the standard uncertainty relation.

Not all position operators are local, as will be seen in the next section. The nonlocality of \( \vec{x} \) will usually appear in the form

\[ [x^i, x^j] = (\text{factors}) S^{ij} \]  

(4.39)

where \( S^{ij} \) is some suitably defined spin tensor. This suggests that locality is a nontrivial property for a system with spin. This result is not surprising; since even classically if a system has an angular velocity along the z-axis, the x and y axes will be constantly carried into one another.
A Point and Local Operator

Newton and Wigner proved and it has subsequently been shown in several ways (see Fleming for example) that for a system of positive energy states no position operator can be defined which is both point and local. Before Fleming the condition was stated "not covariant and local" but the pointness commutation relation (or invariant worldline condition) equation (4.19)

\[ [X^i, X^j] = \frac{1}{2} \{ [X^j, X^i_H] + [X^i, H] X^j \} \tag{4.19} \]

was taken as the definition of covariance. Some examples of familiar operators follow which are not both point and local.

Fleming constructs a point operator with noncommuting components (nonlocal) which can be interpreted as the operator which describes the point in space-time occupied by the rest-frame center of mass. This operator \( \hat{X}^P(t) \) has the self-commutator

\[ [X^P_i(t), X^P_j(t)] = \frac{i\hbar}{(mc)^2} \varepsilon_{ijk} (S^k_\parallel + \frac{M}{p_0^*} S^k_\perp) \tag{4.40} \]

where \( S^\parallel \) and \( S^\perp \) are the components of spin parallel and perpendicular to the momentum. Here spin is defined by the equations

\[ S_i = \frac{1}{2} \varepsilon_{ijk} S_{jk} \tag{4.41} \]

\[ S_{\mu\nu} = M_{\mu\nu} - L_{\mu\nu} \tag{4.42} \]

and

\[ L_{\mu\nu} = X_\mu : P_\nu - X_\nu : P_\mu \tag{4.43} \]

The symbol : stands for a symmetrized product. \( X_\mu \) is whichever position operator one chooses. In the case of (4.40), \( X_\mu \) is the four-vector corresponding to \( \hat{X}^P(t) \). \( L_{\mu\nu} \) is thus an "orbital
angular momentum" and $S_{\mu\nu}$ is a "spin". $M_{\mu\nu}$ is total angular momentum.

The regular center of mass is neither point nor local. Its form is

$$x_i^{c.m.}(t) = \frac{ct}{p_o} p_i + M_{10} p_o^{-1}$$

(4.44)

which can be derived from (4.24), and

$$[x_i^{c.m.}(t), x_j^{c.m.}(t)] = -\frac{i\hbar}{p_o^2} \varepsilon_{ijk} (S_k'' + \frac{mc}{p_o} S_k')$$

(4.45)

where $\hat{S}$ is the same $\hat{S}$ as in (4.40).

A local operator can be constructed, as Pryce did in his case (e), by taking a weighted mean of (4.40) and (4.44), but the local operator is not point, because $x_i^{c.m.}$ is not point. It is in fact Newton and Wigner's operator and it can be interpreted as a "center of spin". Defining spin using (4.41), (4.42), and (4.43) and the local operator

$$x_i^L(t) = \{mc x_i^P(t) + p^0 x_i^{c.m.}(t)\} (mc + p_o)^{-1}$$

(4.46)

one finds that

$$[S_{ij}^L, x_k^L(t)] = 0.$$  

(4.47)

If $\hat{S}$ is a legitimate spin and $\hat{X}$ is a center of spin, then any rotation generated by $\hat{S}$ should leave $\hat{X}$ alone, which is the statement (4.47). $S_{ij}^L$ can be shown to behave as an angular momentum and

$$\hat{S}^L = \hat{n}^2 s(s + 1),$$

so it is a reasonable spin operator. Thus the interpretation that $\hat{S}^L$ is spin and $\hat{X}^L$ is the center of spin is logical.
It should be noted that there are trivial examples of point and local operators, such as a center of mass measured only in the rest frame, or the restrictions to spin zero of nearly all the operators discussed so far. However since pointness and locality are incompatible for an operator on positive energy states of particles with spin, a natural way to construct a point and local operator is to drop the positive energy restriction.

It is convenient at this point to list all the properties one would like a position operator $\hat{x}$ to have:

a) $[x^i, p^j] = i\delta_{ij}$

b) $[x^i, j^j] = -i \varepsilon^{ijk} x^k$

c) $[x^i, k^j] = \frac{1}{2} (x^j [x^i, H] + [x^i, H] x^j)$

d) $[x^i, H] = i \dot{x}^i = i \frac{p^i}{p^0}$ \hspace{1cm} (H = p^0)

e) $[x^i, x^j] = 0$.

It is properties c) and e) which are incompatible if only positive energy states are used.

The Dirac operator $\hat{x}$ which acts on mixed energy states and which is discussed in (3.46), (3.48), (3.57), and the accompanying text, was constructed by Dirac to behave as a canonical coordinate $\hat{q}$. He postulated Lorentz covariance as for a coordinate, which implies pointness (condition c)). Canonical coordinates are self-commuting, so $\hat{x}_{\text{Dirac}}$ is local (condition e)). The failure of condition d) has been discussed in the section on Dirac in Chapter III. Zitterbewegung is an unmeasurable quantity in any practical
situation, so essentially all of the conditions desirable for a position operator hold.

The theory of quantum electrodynamics, which originates with the Dirac equation, is the most successful theory we have yet constructed in physics. It therefore makes sense to generalize the Dirac position operator to hold for any spin, in the hope that such a position operator would be useful for extended field theories. Perhaps, for example, the usual ideas of interactions at a "point", or "locally" defined quantities could be modified to incorporate a point and local position operator.

The Foldy-Wouthuysen transformation is a unitary transformation to a representation in which \( H \) is diagonal and the wave functions are easier to manipulate. Since commutation relations stay the same under such a transformation, the Foldy-Wouthuysen transform of the Dirac \( \gamma \) is generalized rather than the Dirac \( \gamma \) itself. Equation (3.33) gives this operator

\[
\gamma_{\text{Dirac FW}} = \gamma_{\text{Dirac}} - \frac{i\sigma \cdot \mathbf{p}}{2H} + \frac{i\sigma \cdot \mathbf{p}}{2H^2(H+m)} - \frac{\sigma \times \mathbf{p}}{2H(H+m)}
\]  

(3.33)

Here \( H = (p^2 + m^2)^{1/2} \).

For generalizing to higher spin it is convenient to use \( \rho_1, \rho_2, \rho_3 \) and \( \mathbf{\tilde{S}} \) instead of \( \sigma, \beta, \) and \( \mathbf{\tilde{\sigma}} \).

\[
\mathbf{\tilde{S}} = \frac{1}{2} \mathbf{\tilde{\sigma}} \quad \quad \quad \mathbf{\tilde{\alpha}} = 2\rho_1 \mathbf{\tilde{S}} \quad \quad \quad \beta = \rho_3
\]

and the \( \rho \)'s are defined in (3.25) - (3.27). The \( \rho \) matrices and \( \mathbf{\tilde{S}} \) are easier to generalize to any spin and \( \mathbf{\tilde{S}} \) is a more natural and visualizable quantity than \( \mathbf{\tilde{\sigma}} \). Then (3.33) becomes
\[ \hat{\chi}_{\text{Dirac FW}} = \hat{\chi}_{\text{Dirac}} + \frac{\rho_{2S} \hat{S}}{H} - \frac{\rho_{2S} \hat{S} \cdot \hat{p} \hat{p}}{H^2(H+m)} - \frac{\hat{S} \times \hat{p}}{H(H+m)}. \] (4.48)

The \( \hat{\chi}_{\text{Dirac}} \) on the right side of this equation is just a canonical coordinate, so it is renamed \( \hat{Q} \). The \( \hat{\chi}_{\text{Dirac FW}} \) is to be the point and local operator so it is renamed \( \hat{\chi}_{\text{PL}} \). Also \( H \) is a confusing notation for \( (p^2 + m^2)^{1/2} \), as the operator \( H \) is

\[ H = \beta(p^2 + m^2)^{1/2}, \]

so \( (p^2 + m^2)^{1/2} \) is renamed \( E \), and

\[ H = \beta E. \] (4.49)

\( \hat{P} \) is capitalized to indicate its operator character. Then

\[ \hat{\chi}_{\text{PL}} = \hat{Q} + \frac{\rho_{2S} \hat{S}}{E} - \frac{\rho_{2S} \hat{S} \cdot \hat{P} \hat{P}}{E^2(E + m)} - \frac{\hat{S} \times \hat{P}}{E(E+m)} \] (4.50)

The form of (4.50) is a natural form to expect. All of the terms are polar vectors, which position must be. In order to specify the distortions perpendicular to the momentum of a spinning system it is necessary to define two directions. \( \hat{S} \times \hat{P} \) is one of them and \( \hat{P} \times (\hat{S} \times \hat{P}) = \hat{S} \hat{P}^2 - \hat{P} (\hat{S} \hat{P}) \) is the other. All three types of terms appear in (4.50), although they are not weighted quite as expected.

One of the flaws of (4.50) is that for any spin a complete determination of the state of the system requires tensor combinations of \( \hat{S} \) through rank 2S. For spin \( 1/2 \) the combination \( S^i S^j \) gives back \( S^k \) and terms linear in spin are all that are necessary. Thus (4.50) does contain all possible spin terms for spin \( 1/2 \).
An interesting example of the necessity for higher rank spin terms is the determination of the spin density matrix of the deuteron. This requires measurements of components of a second order spin tensor (e.g. one must measure the spin alignment as well as the polarization of the deuteron in a scattering experiment).

For spin one, a possible term to include in the position operator would be something proportional to

\[ \hat{\mathbf{s}} \times (\hat{\mathbf{s}} \times \hat{\mathbf{p}}) = \hat{\mathbf{s}}(\hat{\mathbf{s}} \cdot \hat{\mathbf{p}}) - \hat{\mathbf{p}}(\hat{\mathbf{s}} \cdot \hat{\mathbf{s}}) \]

\[ = (\hat{\mathbf{s}} \cdot \hat{\mathbf{s}} - 1 \hat{\mathbf{s}} \cdot \hat{\mathbf{s}}) \cdot \hat{\mathbf{p}} \quad (4.51) \]

which should be symmetrized for quantum mechanics to

\[ \frac{1}{2} \hat{\mathbf{s}}(\hat{\mathbf{s}} \cdot \hat{\mathbf{p}}) + \frac{1}{2} (\hat{\mathbf{s}} \cdot \hat{\mathbf{p}}) \hat{\mathbf{s}} - \hat{\mathbf{p}}(\hat{\mathbf{s}} \cdot \hat{\mathbf{s}}) \cdot \hat{\mathbf{p}}. \quad (4.52) \]

The pointness and locality of these terms is certainly not manifest, and undoubtedly higher orders of \( \hat{\mathbf{p}} \) would have to be included if it were possible at all to make (4.52) part of a point and local operator.

\( \hat{x}_{\text{FL}} \) of equation (4.50) agrees with Jordan and Mukunda's (3.29) except that theirs includes for spin > \( \frac{1}{2} \) a term \( \rho_2 \hat{\mathbf{a}} \hat{\mathbf{p}} \) which is an axial vector.
The Hyperplane Operator

The hyperplane formalism of Chapter II differs from Fleming's in its structure in a very basic way. As has been seen in the discussion of pointness, Fleming uses a spacelike hyperplane to describe an event or process which can be viewed by different observers. His \( \eta^\mu \) and \( \tau \) label the hyperplane in any observer's frame. \( \eta^\mu \) and \( \tau \) transform as shown in equations (4.25) and (4.26). Thus "covariance" means that an operator measuring one event on one hyperplane can be formulated so that two observers will get the same result. They agree on the hyperplane but their descriptions of \( \eta^\mu \) and \( \tau \) differ as shown in (4.25) and (4.26).

In Chapter II a hyperplane is attached to each observer as well as to each particle. \( \eta_\alpha^\mu \) and \( a_\alpha^\mu \) label the rest-frame hyperplane of each observer as seen from some metaobserver's system. Covariance in this formalism means that operators have the same form for all observers.

Observer \( \eta \) is described by the metaobserver as \( \{ a_\alpha^\mu , \eta_\alpha^\mu \} \) and observer \( \eta' \) is described as \( \{ a'_\alpha^\mu , \eta'_\alpha^\mu \} \). The four-vector \( a'-a \) gives the displacement of \( \eta' \)'s origin from \( \eta \)'s origin. The \( \eta \)'s label the two observers' choices of axes, and the transformation \( L(\eta,a;\eta',a') \) of equations (2.45) and (2.31) takes the primed observer into the unprimed.

With this notation it is possible to study the differences in the descriptions of an operator by two observers. If the form of the operator changes under \( L(\eta,a;\eta',a') \) then the operator is frame dependent. It is usually convenient to let \( a = a' \) so that the
origins coincide. Thus for example the spin operator defined in Appendix D is frame-dependent. For an observer \( \eta^\alpha_\mu \) this operator is

\[
S^a = \eta^a_\mu w^\mu - \frac{(\eta^0_\nu \eta^\nu_\rho)(\eta^3_\rho \rho^0)}{(\eta^0_\sigma \rho^\sigma) + m}
\]  

(4.53)

or

\[
\vec{S} = \vec{w} - \frac{w^0_\rho}{E + m}
\]

(4.54)

in his own frame.

The momentum \( \vec{p} \) for an observer is given by

\[
p^a = \eta^a_\mu p^\mu.
\]

(4.55)

The term \( \vec{Q} \) is the differential operator \( \vec{\nabla}_p \) and it transforms as

\[
Q^a = \eta^a_\mu Q^\mu.
\]

(4.56)

The energy \( E \) is

\[
E = p^0 = \eta^0_\mu p^\mu,
\]

(4.57)

while \( m = m \) in any frame. The symbol \( \epsilon^{abc} \) used in cross products is an invariant, as are \( \delta_{ij} \) and \( g^{\mu\nu} \).

The \( \rho \) matrices transform component-by-component from one frame to another. If \( \rho_2 \) has the form given in (3.44) for the metaobserver then \( \rho_2(\eta) \) as seen by observer \( \eta^\alpha_\mu \) will have components

\[
\rho_2(\eta)_{\alpha\beta} = -i (\eta^1_\alpha \eta^3_\beta + \eta^2_\alpha \eta^4_\beta - \eta^3_\alpha \eta^1_\beta - \eta^4_\alpha \eta^2_\beta).
\]

(4.58)

This form transforms from one observer to another as

\[
\rho_2(\eta) = L^{-1}(\eta;\eta') \rho_2(\eta')L(\eta;\eta')
\]

(4.59)

as would be expected. It should also be noted that \( \rho_2 \) acting on a state of definite sign of energy changes the sign of the energy.
according to
\[ \rho_2 \begin{pmatrix} \hat{p} \lambda \, m \, s \, \varepsilon \end{pmatrix} = \varepsilon = \begin{pmatrix} \hat{p} \lambda \, m \, s \, \varepsilon \end{pmatrix} = \begin{pmatrix} \hat{p} \lambda \, m \, s \, \varepsilon = \varepsilon \end{pmatrix}. \] (4.60)

All of the operators of equation (4.50) for \( \hat{x}_{PL} \) behave "nicely" under a change of frame except \( \hat{S} \). The \( \hat{S} \) of equations (4.53) and (4.54) was used to define states in Chapter II, and it is constructed from generators of the ILG in a reasonable way. It can be written in terms of metaobserver four-vectors, but it is not the three-vector part of a four-vector in any frame except the rest frame of the particle. If it is used for \( \hat{S} \) in equation (4.50) then \( \hat{x}_{PL} \) will not have manifest covariance.

However if an arbitrary vector \( \hat{S} \) is postulated which obeys all of the commutation relations it should, namely
\[
[S^i, P^j] = 0 \quad [S^i, U^j] = -i \varepsilon^{ijk} S^k \quad [S^i, K^j] = -i \varepsilon^{ijk} K^k \quad [S^i, S^j] = i \varepsilon^{ijk} S^k,
\] (4.61)

then a form for \( \hat{K} \) can be found which solves (4.61). This is the "canonical" \( \hat{K} \) of Chapter III, equation (3.23) or (3.28). In this representation
\[ \hat{J} = \hat{L} + \hat{S} \] (4.62)
and
\[ \hat{L} = \hat{q} \times \hat{p} \] (4.63)
which is simpler than the \( L^{uv} \) of Appendix D. This \( \hat{S} \) will transform as \( \hat{J} \) transforms. It can be made covariant by treating it as in Appendix D as the three-by-three part of the antisymmetric tensor \( S^{uv} \) with \( Q^u \) replacing \( Y^u \) in the definition of \( L^{uv} \).
\[ s^i = \frac{1}{2} \epsilon^{ijk} s^{jk} \]
\[ = \frac{1}{2} \epsilon^{ijk} (M^{jk} - L^{jk}). \] (4.64)

The choice of a representation for \( \hat{s} \) is obviously arbitrary, as any \( L^{\mu \nu} \) can be defined using any position operator \( X^\mu \) with \( P^\nu \). This "canonical" \( \hat{s} \) in hyperplane terms is

\[ S^a = \frac{1}{2} \epsilon^{abc} S^{bc} \]
\[ = \frac{1}{2} \epsilon^{abc} \eta^b_\alpha \eta^c_\mu S^{\lambda \mu}. \] (4.65)

It is now possible to write \( \hat{X}^a_{PL} \) in a form which is preserved under the transformation \( L(\eta';\eta) \).

\[ X^a_{PL}(\eta) = Q^a + \frac{\rho_2(\eta) S^a}{p^0} - \frac{\rho_2(\eta) S^{bp} p^a}{p^0 (p^0 + m)} - \frac{\epsilon^{abc} b p^c}{p^0 (p^0 + m)}. \] (4.66)

Miller \cite{24} and McDonald \cite{27,28} have used the Fleming formalism to discuss hyperplane position operators. The formalism of Chapter II is useful for understanding covariance, but cumbersome in practice.
APPENDIX A

NOTATION

Covariant quantities are represented by subscripted indices and contravariant quantities by superscripted indices. The contravariant four-vector $x^\mu$ has positive components

$$x^\mu = (x^0, x^1, x^2, x^3) \quad (A.1)$$

while the covariant four-vector $x_\mu$ has the space components changed in sign:

$$x_\mu = (x_0, x_1, x_2, x_3) = (x^0, -x^1, -x^2, -x^3) \quad (A.2)$$

$$= g_{\mu\nu} x^\nu .$$

The metric tensor $g_{\mu\nu}$ in equation (A.2) is

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (A.3)$$

and it is used to raise and lower indices.

The summation convention is used and sums are over upper and lower indices together.

$$x^2 = x^\mu x_\mu = x^0 x_0 - \dot{x} \cdot \dot{x} . \quad (A.4)$$
The letters q and x are both used to represent position and appear both in upper and lower case. Likewise s and σ are both used for spin, J and M are both used for the angular momentum defined in equations (B.4), and K and N are both used for the boost operator defined in (B.4). The notation has been kept as consistent as possible, but when dealing with 25 different position operators it is easy to overlook a minor inconsistency, such as the placement of i's.

Greek letters \( \mu, \nu, \rho, \ldots \) used as indices take on the values 0, 1, 2, 3. Roman letters \( i, j, k, \ldots \) take on the values 1, 2, 3. When referring to two different observers as in Chapter II, letters from the beginning of the alphabet are used for one observer and letters from the middle or end of the alphabet are used for the other. Thus \( \alpha, \beta, \gamma, \ldots \) and \( a, b, c, \ldots \) refer to one while \( \mu, \nu, \rho, \ldots \) and \( i, j, k, \ldots \) refer to the other.

In general upper case letters refer to operators while lower case letters refer to eigenvalues. In Chapter III it was impossible to be consistent with this notation.

Also in general \( h = c = 1 \) although \( h \)'s and \( c \)'s have been left in definitions when authors included them.

When working in three dimensions
\[
\epsilon^{ij} = -\epsilon_{ij}.
\] (A.5)
POINCARÉ IDENTITIES AND COMMUTATION RELATIONS

\[ M^{\mu\nu} = -\eta^{\mu\nu} \]
\[ M^{\mu\nu} M_{\mu\nu} = -2(j^2 + k^2) \]
\[ \epsilon^{ijk} j^k = M^{\mu i} = k^i \]  \hspace{1cm} (E.1)

\[ [H^{\mu\nu}, H^{\rho\sigma}] = i(g^{\nu\rho} H^{\mu\sigma} + g^{\mu\sigma} H^{\nu\rho} - g^{\mu\rho} H^{\nu\sigma} - g^{\nu\sigma} H^{\mu\rho}) \]

\[ p^0 = E \quad p_\mu p^\mu = E^2 - p^2 = m^2 \]  \hspace{1cm} (E.2)

\[ [p^\mu, M^{\nu\rho}] = i(g^{\mu\nu} p^\rho - g^{\mu\rho} p^\nu) \quad [p^\mu, p^\nu] = 0 \]

\[ w^\mu = \frac{1}{2m} \epsilon^{\mu\nu\rho\sigma} p_\nu \gamma_\rho \gamma_\sigma \]
\[ \eta^\mu = \frac{i}{8m} [p^\mu, \epsilon^{\nu\rho\sigma} \gamma_\nu \gamma_\rho \gamma_\sigma] \]
\[ \eta^\mu = (\frac{1}{m} p \cdot J, \frac{E}{m} J - \frac{1}{m} p \times k) \]
\[ \eta^\mu \eta_\mu = -s(s+1) \]

\[ [w^\lambda, w^\mu] = i \eta^\lambda \eta^\mu - i \frac{1}{m^2} \{ \rho^\lambda(p^\sigma \eta_\mu) - \rho^\mu(p^\sigma \eta_\lambda) \} \quad (E.3) \]
\[ p^\mu \eta_\mu = 0 \]
\[ [p^\mu, w^\nu] = 0 \]
\[ [M^{\mu\nu}, w^\lambda] = i(g^{\lambda\nu} w^\mu - g^{\lambda\mu} w^\nu) \quad [p^i, w^j] = 0 \]
\[ J^i = \frac{1}{2} \epsilon^{ijk} j^k \]
\[ K^i = n^i \]
\[ [E, J^i] = 0 \]
\[ [E, K^i] = i p^i \]
\[ [p^i, J^j] = [J^i, p^j] = -i \epsilon^{ijk} p^k \quad (B.4) \]
\[ [p^i, K^j] = [K^i, p^j] = i \delta^{ij} E \]
\[ [J^i, K^j] = [K^i, J^j] = -i \epsilon^{ijk} k^k \]
\[ [J^i, J^j] = i \epsilon^{ijk} j^k \]
\[ [K^i, K^j] = -i \epsilon^{ijk} j^k \]
APPENDIX C

TETRAD IDENTITIES

\[ \eta^\mu_\alpha \eta^\beta_\mu = g_{\alpha \beta} \]
\[ \eta^\mu_\alpha \eta^\beta_\mu = g_{\alpha \beta} \]  \hspace{1cm} (C.1)

\[ \eta^\mu_\alpha \eta^\beta_\mu = g_{\alpha \beta} \]
\[ g^{\alpha \beta} \eta^\mu_\alpha = \eta^\mu_\beta \]

For any vector such as \( p^\mu \),

\[ p^\alpha = \eta^\alpha_\mu p^\mu \]
\[ p^\mu = \eta^\mu_\alpha p^\alpha \]  \hspace{1cm} (C.2)

while for any tensor such as \( A^\mu_\nu \)

\[ A^\alpha_\beta = \eta^\alpha_\mu A^\mu_\nu \eta^\nu_\beta \]  \hspace{1cm} (C.3)
\[ A^\mu_\nu = \eta^\mu_\alpha A^\alpha_\beta \eta^\beta_\nu \]
APPENDIX D

THE SPIN OPERATOR \( \mathbf{S} \)

The operator

\[
W^\mu = - \frac{1}{2m} \epsilon^{\mu \nu \rho \sigma} \rho_\nu V_\rho S^\sigma
\]

is called the polarization or Pauli-Lubanski vector. Its properties are found in Appendix B. Its commutator with itself,

\[
[V^\lambda, W^\mu] = i \gamma^\lambda_{\mu} - \frac{i}{m^2} \{p^\lambda(p^\sigma S^\mu) - p^\mu(p^\sigma S^\lambda)\},
\]

can be simplified by use of the Sankaranarayanan-Good position operator

\[
\gamma^\mu = \frac{1}{2m^2} (p_\rho S^{\mu \rho} + m^{\mu \rho} p_\rho).
\]

Define

\[
L^\lambda_{\mu} \equiv -(p^\lambda \gamma^\mu - p^\mu \gamma^\lambda)
\]

so

\[
[V^\lambda, W^\mu] = i (\gamma^\lambda_{\mu} - L^\lambda_{\mu}) = i S^\lambda_{\mu} \text{ by definition.}
\]

The properties of \( S^\lambda_{\mu} \) are

\[
\begin{align*}
S^{00} &= 0 \\
S^{0i} &= \{- \frac{E}{m^2} (\vec{p} \times \vec{j}) + \frac{1}{m^2} \vec{p} \times (\vec{p} \times \vec{k})\}^i \\
S^{ij} &= \epsilon^{ijk} \{\vec{j} - \frac{1}{m^2} \vec{p} \times (E \vec{k} + \vec{p} \times \vec{j})\}^k
\end{align*}
\]
It is well known that if $W^\mu$ is describing the polarization of a particle with four-momentum $P^\mu$, then in the rest frame of the particle where $P^\text{rest} = 0$ and $P^0\text{rest} = m$ we have

$$W^\mu\text{rest} = (0, \vec{J}\text{rest}) = (0, \vec{S}\text{rest}) \quad (D.7)$$

since $\vec{J} = \vec{S}$ in the rest frame. The fact that the three-vector part of $W^\mu$ becomes $\vec{S}$ in the rest system suggests the following operations.

The LT $\Lambda^\mu_\nu$ defined by

$$\Lambda^0_\nu = \frac{1}{m} p^\nu$$

$$\Lambda^i_0 = -\frac{1}{m} p^i$$

$$\Lambda^i_j = \delta^i_j - \frac{1}{m(E+m)} p^i p_j$$

is the LT used to transform to the system moving with momentum $\vec{P}$, i.e. the rest system. In the rest system

$$p^i\text{rest} = \Lambda^i_\nu p^\nu = 0$$

and

$$p^0\text{rest} = \Lambda^0_\nu p^\nu = m \quad (D.9)$$

When we transform $W^\mu$ and $S^{\mu\nu}$ to the rest system using LT $\Lambda^\mu_\nu$ as above we get

$$W^\mu\text{rest} = (0, \frac{E}{m} \vec{J} - \frac{\vec{P} \times \vec{K}}{m^2} - \frac{\vec{P} (\vec{P} \cdot \vec{J})}{m(E+m)})$$

$$S^{0i}\text{rest} = 0 \quad (D.10)$$

$$S^{ij}\text{rest} = \varepsilon^{ijk} \left( \frac{E}{m} \vec{J} - \frac{\vec{P} \times \vec{K}}{m^2} - \frac{\vec{P} (\vec{P} \cdot \vec{J})}{m(E+m)} \right)^k.$$
So if we let $s_{ij} = \epsilon_{ijk} s^k$, then in the rest system

$$s_{\text{rest}}^{ij} = \epsilon_{ijk} s^k_{\text{rest}} = \epsilon_{ijk} w^k_{\text{rest}}.$$  \hfill (D.11)

That is, $\hat{s}_{\text{rest}} = \hat{w}_{\text{rest}}$, where now $\hat{s}$ is well defined as

$$s^i = \frac{1}{2} \epsilon_{ijk} s^{jk}. \hfill (D.12)$$

Proceeding with commutation relations we already have

$$[w^\lambda, w^\mu] = i s^\lambda_{\mu}$$

so

$$[w^i, w^j] = i s^{ij} = \frac{i}{m} \epsilon_{ijk} \left( p^0 w^k - p^k w^0 \right). \hfill (D.13)$$

and then

$$[\hat{w}^i_{\text{rest}}, \hat{w}^j_{\text{rest}}] = i s^{ij}_{\text{rest}} = i \epsilon_{ijk} w^k_{\text{rest}}. \hfill (D.14)$$

So $\hat{w}_{\text{rest}}$ does behave like an angular momentum.

Now we note that

$$\hat{w}^i_{\text{rest}} = w^i - \frac{\mu^0 p^i}{E+m}. \hfill (D.15)$$

We define

$$s^i \equiv \hat{w}^i_{\text{rest}} = w^i - \frac{\mu^0 p^i}{E+m} \hfill (D.16)$$

to hold in any frame. Then remembering that $[p^i, w^j] = 0$ we have

$$[s^i, s^j] = i \epsilon_{ijk} s^k \hfill (D.17)$$

and we have a useful

$$\hat{s} = \hat{w} - \frac{\mu^0 p}{E+m} \hfill (D.18)$$

in any frame.
To put $S$ into hyperplane terms let the observer $\alpha$ define

$$S^a = \eta^a - \frac{\eta^0 p^a}{p^0 + m} = \eta^a \eta^\mu - \frac{(\eta^0 w^\nu)(\eta^a p^\nu)}{(\eta^0 p^\sigma) + m}$$

(D.19)

the latter being put into metaobserver terms. This is the usual way to form a hyperplane operator: replace "time-like" components of vectors such as $p^0$ by the contraction $\eta^0 p^\mu$ and three-vectors such as $p^a$ by $\eta^a p^\mu$.

This form can be checked by letting $\alpha \rightarrow \mu$, that is $a \rightarrow i$, or the observer becomes the metaobserver. Then

$$S^i = \eta^i \eta^\mu - \frac{(\eta^0 w^\nu)(\eta^i p^\nu)}{(\eta^0 p^\sigma) + m} = \eta^i - \frac{\eta^0 p^i}{p^0 + m}$$

(D.20)

by straightforward $\eta$-identities. We cannot define an $S^0$, nor claim that $S^i$ is a three-vector. $S^i$ has no "covariant" properties. Thus $\lambda$ in the state labels is the eigenvalue of

$$S^3 = \eta^3 \eta^\mu - \frac{(\eta^0 w^\nu)(\eta^3 p^\nu)}{(\eta^0 p^\sigma) + m}$$

(D.21)

or, to the observer himself,

$$S^3 = \omega^3 - \frac{\eta^0 p^3}{\lambda + m}$$

(D.22)
APPENDIX E

PROPERTIES OF \( L(n'; n) \)

Proof that \( |n'_{\alpha} \hat{p}^{\alpha} \lambda> \equiv L(n'; n) |n_{\alpha} \hat{p}^{\alpha} \lambda> \) is a good definition by checking the eigenvalues of \( p'^{\alpha} \) and \( s'^{3} \) on \( L(n'; n) |n_{\alpha} \hat{p}^{\alpha} \lambda> \).

\[ \begin{align*}
p'^{\alpha} L(n'; n) |n_{\alpha} \hat{p}^{\alpha} \lambda> &= \\
&= n'^{\alpha}_{\mu} p^{\mu} L(n'; n) |n_{\alpha} \hat{p}^{\alpha} \lambda> \\
&= L(n'; n) n'^{\alpha}_{\mu} L(n'; n)_{\nu}^{\mu} p^{\nu} |n_{\alpha} \hat{p}^{\alpha} \lambda> \\
&= L(n'; n) n'^{\alpha}_{\mu} (n'^{\nu}_{\gamma} n_{\delta}^{\gamma}) n_{\delta}^{\nu} p^{\delta} |n \hat{p}^{\alpha} \lambda> \quad \text{(E.1)} \\
&= L(n'; n) g^{\alpha}_{\gamma} g^{\gamma}_{\delta} p^{\delta} |n \hat{p}^{\alpha} \lambda> \\
&= L(n'; n) \hat{p}^{\alpha} |n \hat{p}^{\alpha} \lambda> \\
&= \hat{p}^{\alpha} L(n'; n) |n \hat{p}^{\alpha} \lambda>
\end{align*} \]

\[ \begin{align*}
s'^{3} L(n'; n) |n_{\alpha} \hat{p}^{\alpha} \lambda> &= \\
&= S^{3}(n') L(n'; n) |n_{\alpha} \hat{p}^{\alpha} \lambda> \quad \text{(E.2)} \\
&= L(n'; n) L^{-1}(n'; n) S^{3}(n') L(n'; n) |n_{\alpha} \hat{p}^{\alpha} \lambda> \\
&= L(n'; n) S^{3}(n) |n_{\alpha} \hat{p}^{\alpha} \lambda> \\
&= \lambda L(n'; n) |n_{\alpha} \hat{p}^{\alpha} \lambda> .
\end{align*} \]
The justification of
\[ L^{-1}(n'; n) S^3(n'; n) L(n'; n) = S^3(n) \]
going as follows. First
\[
\eta^\alpha \mu L^{-1}(n; n') p^\mu L(n; n') = \eta^\alpha \mu L(n; n')^\mu \nu p^\nu
\]
\[ = \eta^\alpha \mu (L(n; n') p)^\mu \]
also
\[ = \eta^\alpha \mu \eta^\mu \beta n'^\nu \beta p^\nu. \tag{E.3} \]
\[ = \eta^\alpha \nu p^\nu \]
\[ = \eta^\beta \nu \eta^\mu \beta \eta^\alpha \nu p^\nu \]
\[ = L^{-1}(n; n')^\mu \nu \eta^\alpha \nu p^\nu \]
\[ = (L^{-1}(n; n') \eta)^\alpha \nu p^\nu. \]

In other words
\[ n(LP) = (L^{-1} n)p. \tag{E.4} \]

This is true of course for any four-vector operator, not just for \( P \). Then any function \( f \) of the product of \( \eta^\alpha \mu \) and any four-vector \( A^\mu \) also has the property
\[ f(\eta^\alpha \mu (LA)^\mu) = f((L^{-1} n)^\alpha \mu A^\mu). \tag{E.5} \]

Specifically \( S(\eta^\alpha \mu, \nu^\mu, p^\mu) \) obeys this so that
\[ L^{-1}(n'; n) S^3(n', P, \hbar) L(n'; n) = S^3(n', L P, L \hbar) \]
\[ = S^3(L^{-1} n', P, \hbar) \]
\[ = S^3(L^{-1}(n'; n)n', P, \hbar) \]
\[ = S^3(L(n;n')n', P, \hbar) \]
\[ = S^3(n, P, \hbar). \quad (E.G) \]

Thus the eigenvalues of \( p' \alpha \) and \( S'^3 \) on \( L(n'; n) |n \alpha \lambda> \)
are the same as on \( |n' \alpha \lambda> \), and the definition holds, at least up to a phase.
BIBLIOGRAPHY

ACKNOWLEDGMENTS

I would like to thank Professor Derek Pursey for suggesting the work on both hyperplanes and position operators, and for helping substantially in carrying out that work. I am also grateful to Henry Black for supporting me throughout most of my education, Loyal Durand for encouraging me to finish this thesis, and Laurel Hermanson for the excellent typing.