Foldy-Wouthuysen transformations in an indefinite-metric space. II. Theorems for practical calculations †

R. A. Krajcik* and Michael Martin Nieto

Theoretical Division, Los Alamos Scientific Laboratory, University of California, Los Alamos, New Mexico 87545

(Received 15 September 1975)

We present a number of theorems and lemmas which are useful in indefinite-metric spaces for explicitly calculating Foldy-Wouthuysen (FW) transformations, both as power-series expansions in c^{-1} and as exact transformations. We mention two applications.

I. INTRODUCTION AND NOTATION

In the preceding paper[1] we proved the theorem that the necessary and sufficient conditions that a "pseudounitary" (Foldy-Wouthuysen)[2,3] transformation exists which will diagonalize a nondiagonal "pseudo-Hermitian" matrix $\mathcal{O}$ on a (nonsingular) indefinite-metric space are that all the eigenvalues of $\mathcal{O}$ be real and all the eigenvectors of $\mathcal{O}$ have nonzero norm. We also discussed physical applications of this theorem as well as demonstrating that the pseudounitary transformation of a pseudo-Hermitian matrix also yields a pseudo-Hermitian matrix.

On a more practical level, if one is given the existence of an FW transformation from the above theorem, it is then of use to have techniques available with which to calculate the FW transformation either order by order in c, or preferably exactly. For the Dirac case, one can do both[4] these things by writing the transformation U as

$$U=e^{iS}, \tag{1.1}$$

$$S=\sum_n c^{-n}S_n, \tag{1.2}$$

where the S_n are independent of c, calculating U order by order in c, and then finding that the closed form of the sum is the exact FW transformation.

It is the purpose of this paper to derive theorems and lemmas for the analogous calculations in indefinite-metric spaces. In the following section we will state four such theorems, commenting on the physical significance and applications of them. Lemmas will also be derived in this section, but the proofs of the theorems themselves will be reserved for the Appendix.

Before proceeding, we wish to discuss our terminology and notation. In Ref. 1, we stated preference for the more descriptive terminology "metric-unitary" and "metric-Hermitian" over the commonly used "pseudo." In this paper we will use our terminology meaning a transformation U is "metric-unitary" if

$$U^\dagger MU=M, \tag{1.3}$$

where M is the metric, and an operator $\mathcal{O}$ is "metric-Hermitian" if

$$(M\mathcal{O})^\dagger=M\mathcal{O}. \tag{1.4}$$

As noted in Ref. 1, one can take M diagonal, meaning

$$M^\dagger=M. \tag{1.5}$$

We now introduce the following useful notation.

Definitions. The "adjoint commutator" of A with B is the operation

$$[{}^\dagger A, B]\equiv A^\dagger B-BA. \tag{1.6}$$

Thus, if B is the metric M, then from (1.4) and (1.5) an operator S is metric-Hermitian if its adjoint commutator with M is zero:

$$[{}^\dagger S,M]\equiv S^\dagger M-MS=0. \tag{1.7}$$

We also define the symbolic notation that in an equation quantities like $([{}^\dagger A,)^n$ and $(\,])^n$ mean the objects in the round brackets are explicitly written out n times. Thus, for example,

$$\begin{aligned}([{}^\dagger A,)^2B(\,])^2&\equiv[{}^\dagger A,[{}^\dagger A,B]]\\&=[{}^\dagger A,(A^\dagger B-BA)]\\&=A^\dagger A^\dagger B-2A^\dagger BA+BAA.\end{aligned} \tag{1.8}$$

II. THEOREMS AND LEMMAS

Theorem I. Let

$$U=e^{iS}, \tag{2.1}$$

where S is not necessarily self-adjoint. Then

$$\begin{aligned}U^\dagger\mathcal{O}U&=e^{-iS^\dagger}\mathcal{O}e^{iS}\\&=\sum_{n=0}^{\infty}([{}^\dagger S,)^n\,\mathcal{O}(\,])^n\,\frac{(-i)^n}{n!}.\end{aligned} \tag{2.2}$$

Lemma I. When $S=S^\dagger$, Eq. (2.2) becomes

$$U^\dagger \Theta U=\sum_{n=0}^{\infty}([S,)^n\Theta(\,])^n\frac{(-i)^n}{n!}, \qquad (2.3)$$

which in the positive-definite-metric Hilbert space case is also equal to $U^{-1}\Theta U$.

In the Appendix theorem I is used to prove theorem II. Theorem III will be the converse of theorem II, but is treated separately owing to the lemmas and the associated discussions.

Theorem II. Let the operator Θ be invariant under the transformation

$$U^\dagger \Theta U=\Theta, \quad U=e^{iS}, \qquad (2.4)$$

where S can be expanded as an analytic power series in a real parameter λ, and S goes to zero as $\lambda\to 0$:

$$S=\sum_{k=1}\lambda^k S_k\,. \qquad (2.5)$$

Then the adjoint commutators of the S_k and hence of S with Θ are zero. That is,

$$[\,^\dagger S_k,\Theta]\equiv S_k^\dagger\Theta-\Theta S_k=0, \qquad (2.6)$$

$$[\,^\dagger S,\Theta]\equiv S^\dagger\Theta-\Theta S=0. \qquad (2.7)$$

An important lemma follows as a special case of theorem II. Suppose U in Eq. (2.4) is a Foldy-Wouthuysen transformation, and Θ in Eq. (2.4) is the metric operator. Further, suppose that the expansion parameter of (2.5) is the inverse of the velocity of light, c^{-1}. Then the first line of Eq. (2.4) is the definition of metric-unitary and Eqs. (2.6) and (2.7) are the definitions of metric-Hermitian. Thus we have lemma II.

Lemma II. Let U be a metric-unitary Foldy-Wouthuysen transformation which can be expanded as

$$U=e^{iS}, \quad S=\sum_{k=1}c^{-k}S_k\,. \qquad (2.8)$$

Then S is metric-Hermitian, and so too are the S_k.

Lemma II is of much practical value. For example, in our series[5] on the first-order Bhabha fields of arbitrary spin, we will soon discuss[6] the Foldy-Wouthuysen transformation of the Poincaré generators. Lemma II will allow the S_n of the transformation to be determined uniquely, up to a phase.

Now going to the converse of Theorem II we have theorem III.

Theorem III. Define U and S as in Eqs. (2.4) and (2.5),

$$U=e^{iS}, \quad S=\sum_{k=1}\lambda^k S_k\,. \qquad (2.9)$$

Suppose the adjoint commutators of the S_k and hence of S with an operator Θ are zero:

$$[\,^\dagger S_k,\Theta]=[\,^\dagger S,\Theta]=0. \qquad (2.10)$$

Then the transformation $U^\dagger\Theta U$ leaves Θ invariant:

$$U^\dagger\Theta U=e^{-iS^\dagger}\Theta e^{iS}=\Theta. \qquad (2.11)$$

Again letting M be the metric operator in theorem III yields the important converse of lemma II.

Lemma III. Let S be metric-Hermitian; then

$$U=e^{iS} \qquad (2.12)$$

is metric-unitary.

Finally we quote a theorem which is of use in obtaining exact FW transformations.

Theorem IV. Let V and F be given by

$$V=e^{iF}, \quad F^\dagger=-F. \qquad (2.13)$$

Then

$$\Theta'(F)\equiv V^\dagger\Theta V=e^{iF}\Theta e^{iF} \qquad (2.14)$$

can be given by

$$\Theta'(F)=\sum_{n=0}^{\infty}\frac{i^n}{n!}(\{F,)^n\Theta(\})^n\,. \qquad (2.15)$$

The manner in which theorem IV helps in obtaining exact FW transformations is seen by recalling the contents of Ref. 1. The metric-unitary transformation which diagonalizes the Hamiltonian is given by a matrix whose columns are the "class A" eigenvectors. Thus, the transformation matrix can be found by starting with the rest system eigenvectors and performing a Lorentz transformation on them.

To cite a specific example we are interested in, using the notation of Refs. 5 and 6 for the Bhabha half-integer-spin system and defining

$$F\equiv-\theta[\alpha_4,\vec{\alpha}\cdot\hat{v}]\,, \qquad (2.16)$$

$$\tanh\theta=v/c, \qquad (2.17)$$

$$\Theta\equiv M=\eta_4\alpha_4, \qquad (2.18)$$

we will show elsewhere[6] that

$$e^{iF}Me^{iF}=M\cosh\theta-i(\vec{\alpha}\cdot\hat{v})\sinh\theta. \qquad (2.19)$$

The result Eq. (2.19) is crucial in obtaining the Foldy-Wouthuysen matrix.

APPENDIX: PROOFS OF THEOREMS

Theorem I. Let $U=e^{iS}$, where S is not necessarily equal to $S^\dagger$. Then

$$U^\dagger \mathcal{O} U = e^{-iS^\dagger}\mathcal{O}e^{iS} \tag{A1}$$

$$\begin{aligned}&= \left(1 - iS^\dagger + \frac{(-i)^2}{2!}(S^\dagger)^2+\cdots\right)\mathcal{O}\left(1+iS+\frac{i^2}{2!}S^2+\cdots\right)\\ &= \mathcal{O}+i\mathcal{O}S - iS^\dagger\mathcal{O}+\frac{(-i)^2}{2!}(S^\dagger)^2\mathcal{O}+\frac{(-i)^2}{2!}\mathcal{O}S^2-(-i)^2S^\dagger\mathcal{O}S+\cdots\end{aligned} \tag{A2}$$

$$= -i[{}^\dagger S,\mathcal{O}]+\frac{(-i)^2}{2!}[{}^\dagger S,[{}^\dagger S,\mathcal{O}]]+\cdots . \tag{A3}$$

We will now show that Eq. (A3) can be written to all orders as

$$e^{-iS^\dagger}\mathcal{O}e^{iS}=\sum_{n=0}\frac{(-i)^n}{n!}([\,{}^\dagger S,)^n\mathcal{O}(\,])^n . \tag{A4}$$

To see this result consider[4]

$$\mathcal{O}'(a)\equiv e^{-iaS^\dagger}\mathcal{O}e^{iaS}=\sum_{n=0}a^nA_n . \tag{A5}$$

The object is to find the coefficients A_n. This can be done by alternately setting $a=0$ and differentiating with respect to a. The first term, A_0, is an identity from setting $a=0$ in (A5):

$$A_0=\mathcal{O}=([\,{}^\dagger S,)^0\mathcal{O}(\,])^0. \tag{A6}$$

Now differentiating (A5),

$$\begin{aligned}\frac{d}{da}(\mathcal{O}'(a)) &\equiv \frac{d}{da}(e^{-iaS^\dagger}\mathcal{O}e^{iaS})\\ &= -iS^\dagger\mathcal{O}'(a)+i\mathcal{O}'(a)S\\ &= -i[\,{}^\dagger S,\mathcal{O}'(a)]\\ &= \sum_{n=1} n a^{n-1}A_n .\end{aligned} \tag{A7}$$

Letting $a\to 0$ in (A7) gives

$$-i[\,{}^\dagger S,\mathcal{O}]=A_1. \tag{A8}$$

Differentiating (A7) with respect to a yields

$$\begin{aligned}\frac{d}{da}(-i[\,{}^\dagger S,\mathcal{O}'(a)]) &= (-i)^2[\,{}^\dagger S,[\,{}^\dagger S,\mathcal{O}'(a)]]\\ &= \sum_{n=2}n(n-1)a^{n-2}A_n .\end{aligned} \tag{A9}$$

Letting $a\to 0$ gives

$$A_2=\frac{(-i)^2}{2!}[\,{}^\dagger S,[\,{}^\dagger S,\mathcal{O}]] . \tag{A10}$$

Similarly, every derivative with respect to a of (A5) will produce another $-i[\,{}^\dagger S,\mathcal{O}'(a)]$ inside the nest of factors $[\,{}^\dagger S,[\,{}^\dagger S,\ldots$ on the left-hand side, and another term to the factorial on the right-hand side. Hence, letting $a\to 0$,

$$A_n=(-i)^2([\,{}^\dagger S,)^n\mathcal{O}(\,])^n . \tag{A11}$$

Now, setting $a=1$ in Eq. (A5) completes the proof.

Theorem II. Consider the problem where $\mathcal{O}$ is left invariant by the transformation U:

$$U^\dagger\mathcal{O}U=\mathcal{O}, \tag{A12}$$

$$U\equiv e^{iS}. \tag{A13}$$

Suppose further that S can be expanded as a power series in λ which goes to zero as $\lambda\to 0$ [i.e., S is analytic in λ, λ is a real c number, $S(\lambda=0)=0$, but $(dS/d\lambda)$ $(\lambda=0)$ is not necessarily zero]. Then

$$S\equiv\sum_{k=1}\lambda^kS_k . \tag{A14}$$

However, since $\mathcal{O}$ is not a function of λ, we have from theorem I

$$\sum_{n=0}(-i)^n([\,{}^\dagger S,)^n\mathcal{O}(\,])^n=\mathcal{O}. \tag{A15}$$

Using (A6) to subtract the first term, this means

$$\sum_{n=1}\frac{(-i)^n}{n!}([\,{}^\dagger S,)^n\mathcal{O}(\,])^n=0. \tag{A16}$$

But since the λ^n are independent, each coefficient of λ^n must vanish, meaning

$$\sum_{k=1}^{n}\frac{(-i)^k}{k!}\prod_{i=1}^{k}([\,{}^\dagger S_{l(ik)},)\mathcal{O}(\,])^k=0, \tag{A17}$$

where the $l(ik)$ are integers $\geqslant 1$, in the product we are again using our symbolic notation, and

$$\sum_{i=1}^{k}l(ik)=n. \tag{A18}$$

Now consider the structure of these general coefficients of λ^n:

$$0=\sum_{k=1}^{n}\frac{(-i)^k}{k!}\prod_{i=1}^{k}([\,{}^{\dagger}S_{l(ik)},\,)\mathcal{O}(\;])^k$$
$$=-i[\,{}^{\dagger}S_n,\mathcal{O}]+\frac{(-i)^2}{2!}[\,{}^{\dagger}S_{n-1},[\,{}^{\dagger}S_1,\mathcal{O}\,]\,]$$
$$+\frac{(-i)^2}{2!}[\,{}^{\dagger}S_1,[\,{}^{\dagger}S_{n-1},\mathcal{O}]\,]+\cdots$$
$$+\frac{(-i)^n}{n!}[\,{}^{\dagger}S_1,[\,{}^{\dagger}S_1,[\,{}^{\dagger}S_1,\ldots[\,{}^{\dagger}S_1,\mathcal{O}\,]\cdots]]]. \tag{A19}$$

First let $n=1$. Then we have

$$[\,{}^{\dagger}S_1,\mathcal{O}]=0. \tag{A20}$$

Let $n=2$. Then because of the structure of the coefficient for λ^2,

$$0=-i[\,{}^{\dagger}S_2,\mathcal{O}]+\frac{(-i)^2}{2!}[\,{}^{\dagger}S_1,[\,{}^{\dagger}S_1,\mathcal{O}]\,]. \tag{A21}$$

Thus, from the result for $n=1$, Eq. (A20), we have

$$[\,{}^{\dagger}S_2,\mathcal{O}]=0. \tag{A22}$$

For $n=3$,

$$0=\sum_{k=1}^{3}\frac{(-1)^k}{k!}\prod_{i=1}^{k}([{}^{\dagger}S_{l(ik)},)\mathcal{O}(\;])^k$$
$$=-i[{}^{\dagger}S_{l(11)},\mathcal{O}]+\frac{(-i)^2}{2!}[{}^{\dagger}S_{l(12)},[{}^{\dagger}S_{l(22)},\mathcal{O}]\,]$$
$$+\frac{(-i)^3}{3!}[{}^{\dagger}S_{l(13)},[{}^{\dagger}S_{l(23)},[{}^{\dagger}S_{l(32)},\mathcal{O}]\,]\,]. \tag{A23}$$

Hence,

$$l(11)=3,\quad l(12)+l(22)=3, \tag{A24}$$

meaning

$$l(12)=1,\quad l(22)=2 \text{ or } l(12)=2,\quad l(22)=1, \tag{A25}$$

$$l(13)=l(23)=l(33)=1. \tag{A26}$$

Thus,

$$0=-i[{}^{\dagger}S_3,\mathcal{O}]+(-i)^2[{}^{\dagger}S_1,[{}^{\dagger}S_2,\mathcal{O}]\,]$$
$$+(-i)^2[{}^{\dagger}S_2,[{}^{\dagger}S_1,\mathcal{O}]\,]$$
$$+(-i)^3[{}^{\dagger}S_1,[{}^{\dagger}S_1,[{}^{\dagger}S_1,\mathcal{O}]\,]\,]. \tag{A27}$$

Equations (A20), (A22), and (A27) mean

$$[{}^{\dagger}S_3,\mathcal{O}]=0. \tag{A28}$$

By induction, if $[{}^{\dagger}S_1,\mathcal{O}]=[{}^{\dagger}S_2,\mathcal{O}]=\cdots[{}^{\dagger}S_{n-1},\mathcal{O}]=0$, then from Eq. (A19) for the coefficient of λ^n we have

$$[{}^{\dagger}S_n,\mathcal{O}]=0. \tag{A29}$$

Since n is arbitrary, we have the result that

$$e^{-iS^{\dagger}}\mathcal{O}e^{iS}=\mathcal{O} \tag{A30}$$

implies

$$[{}^{\dagger}S_n,\mathcal{O}]=0,\quad n\geq 1,\text{ or } [{}^{\dagger}S,\mathcal{O}]=0. \tag{A31}$$

That is,

$$S_n^{\dagger}\mathcal{O}=\mathcal{O}S_n,\quad n\geq 1,\quad S^{\dagger}\mathcal{O}=\mathcal{O}S, \tag{A32}$$

where the restrictions on S are given in Eq. (A14):

$$S=\sum_{k=1}\lambda^k S_k. \tag{A33}$$

Theorem III. Conversely, if $[{}^{\dagger}S_n,\mathcal{O}]=0$, then $[{}^{\dagger}S,\mathcal{O}]=0$, and we have

$$\sum_{n=1}\frac{(-i)^n}{n!}([{}^{\dagger}S,\,)^n\,\mathcal{O}(\;])^n=0, \tag{A34}$$

so that

$$\sum(-1)^n([{}^{\dagger}S,\,)^n\mathcal{O}(\;])^n=\mathcal{O}, \tag{A35}$$

or

$$U^{\dagger}\mathcal{O}U=\mathcal{O}. \tag{A36}$$

Theorem IV. The proof of theorem IV is similar to that of theorem I.[4] Consider

$$\mathcal{O}'(bF)\equiv e^{ibF}\mathcal{O}e^{ibF}\equiv\sum_{n=0}^{\infty}b^n B_n. \tag{A37}$$

To find the coefficients B_n, alternately set $b=0$ and differentiate with respect to b. Setting $b=0$ yields

$$B_0=\mathcal{O}. \tag{A38}$$

Differentiating (A37) with respect to b gives

$$\frac{d}{db}\mathcal{O}'(bF)=i\{F,\mathcal{O}'(bF)\}=\sum_{n=1}nb^{n-1}B_n. \tag{A39}$$

Setting $b=0$ in (A39) gives

$$B_1=i\{F,\mathcal{O}\}. \tag{A40}$$

Similarly one obtains

$$B_2=\frac{i^2}{2!}\{F,\{F,\mathcal{O}\}\} \tag{A41}$$

and in general

$$B_n=\frac{i^n}{n!}(\{F,\,)^n\mathcal{O}(\;\})^n. \tag{A42}$$

Now set $b=1$ in (A37), yielding

$$\mathcal{O}'(F)=e^{iF}\mathcal{O}e^{iF}=\sum_{n=0}^{\infty}\frac{i^n}{n!}(\{F,\,)^n\mathcal{O}(\;\})^n. \tag{A43}$$

†Work supported by the United States Energy Research and Development Administration.
*Present Address: Institute of Geophysics and Planetary Physics; Scripps Institute of Oceanography; University of California, San Diego, La Jolla, Calif. 92093.
[1]R. A. Krajcik and M. M. Nieto, preceding paper, Phys. Rev. D 13, 2245 (1976).
[2]L. L. Foldy and S. A. Wouthuysen, Phys. Rev. 78, 29 (1950); S. Tani, Prog. Theor. Phys. 6, 267 (1951).
[3]Also see the earlier papers by M. H. L. Pryce, Proc. R. Soc. London A150, 166 (1935), and R. Becker, Nachr. Akad. Wiss. Göttingen, Math. Physik Kl. II, No. 1, 39 (1945).
[4]J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill, New York, 1964), p. 49, footnote 1.
[5]R. A. Krajcik and M. M. Nieto, Phys. Rev. D 10, 4049 (1974); 11, 1442 (1975); 11, 1459 (1975); 13, 924 (1976).
[6]R. A. Krajcik and M. M Nieto (in preparation), paper VI of the series of Ref. 5, on the indefinite-metric and Foldy-Wouthysen transformations of Bhabha fields.