$$g^{(m,m)}(x_1,\cdots,x_m;x_1,\cdots,x_m)\times g^{(m,m)}(x_1',\cdots,x_m';x_1',\cdots,x_m')\geq|g^{(m,m)}(x_1,\cdots,x_m;x_1',\cdots,x_m')|^2,\quad \text{(A4)}$$

which are similar to the inequalities (3.12) and (3.14) of Ref. 5. For $N=3$, the inequality (A3) expresses that, in addition, the determinant

$$\begin{vmatrix} g^{(m,m)}(x^{(1)};x^{(1)}) & g^{(m,m)}(x^{(1)};x^{(2)}) & g^{(m,m)}(x^{(1)};x^{(3)}) \\ g^{(m,m)}(x^{(2)};x^{(1)}) & g^{(m,m)}(x^{(2)};x^{(2)}) & g^{(m,m)}(x^{(2)};x^{(3)}) \\ g^{(m,m)}(x^{(3)};x^{(1)}) & g^{(m,m)}(x^{(3)};x^{(2)}) & g^{(m,m)}(x^{(3)};x^{(3)}) \end{vmatrix}\geq 0,\quad \text{(A5)}$$

where $x^{(j)}$ now stands for the set of variables $x_1^{(j)}$, $x_2^{(j)},\cdots,x_m^{(j)}$. In the case of complete $(2m)$th-order coherence, we have $|g^{(m,m)}|=1$, and hence we can write

$$g^{(m,m)}(x^{(i)};x^{(j)})=\{g^{(m,m)}(x^{(j)};x^{(i)})\}^*,\quad =\exp[-i\psi(x^{(i)},x^{(j)})],\quad \text{(A6)}$$

where ψ is real. It can then be seen, on evaluating the determinant, that (A5) can only be satisfied as an equality, and we then obtain the relation

$$\psi(x^{(3)},x^{(2)})=\psi(x^{(3)},x^{(1)})-\psi(x^{(2)},x^{(1)}).\quad \text{(A7)}$$

Setting $x^{(1)}=0$ (i.e., $x_1^{(1)}=x_2^{(1)}=\cdots=x_m^{(1)}=0$, which is always permissible by suitable choice of the origin), we finally obtain

$$\psi(x^{(3)},x^{(2)})=f(x^{(3)})-f(x^{(2)}),\quad \text{(A8)}$$

where $f(x^{(j)})=\psi(x^{(j)},0)$ is a function of $x^{(j)}$ only. Hence, from (A6) we find that

$$g^{(m,m)}(x^{(i)};x^{(j)})=\exp\{-i[f(x^{(i)})-f(x^{(j)})]\},\quad \text{(A9)}$$

which, when written in full, gives the required relation (4.9): viz.,

$$g^{(m,m)}(x_1,\cdots,x_m;x_1',\cdots,x_m')=\exp\{-i[f(x_1,\cdots,x_m)-f(x_1',\ldots,x_m')]\}.\quad \text{(A10)}$$

It may be noted that the results obtained in this appendix are also true of classical fields and can be obtained in a strictly similar manner.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 9 SEPTEMBER 1967

Energy-Momentum Conservation Implies Translation Invariance: Some Didactic Remarks

DAVID N. WILLIAMS*

Service de Physique Théorique, Centre d'Etudes Nucléaires de Saclay, BP n° 2,91, Gif-sur-Yvette, France

(Received 10 October 1966)

We discuss from a rigorous viewpoint two more-or-less familiar cases where energy-momentum conservation implies invariance under space–time translations. First, if a closed linear operator on a Hilbert space has a domain that is invariant under spectral projections belonging to the four-momentum operators, and if it "conserves energy-momentum," it necessarily commutes with the appropriate representation of the translations. (Bounded operators, such as the S matrix, are a special case.) At least for separable spaces, the domain restriction characterizes the closed operators for which the theorem is true. Second, if a bounded bilinear form between momentum states of m and n particles in a Fock space (or more generally, a bounded multilinear form) conserves energy momentum, the corresponding tempered distribution has a conservation delta function at points where the mass shell is a C_∞ manifold; but no derivatives of delta functions can occur. In this connection, we are led to a result that seems to be new: the cluster parameters ("connected amplitudes") of a family of bounded bilinear forms, labeled by (m, n), are also bounded bilinear forms. The two systems, of course, mutually conserve energy momentum.

I. INTRODUCTION

THAT translation invariance implies momentum conservation is a familiar example of the relation between continuous symmetries and conservation laws, which is classically and elegantly expressed by Noether's theorem.[1] Conversely, to every constant of the motion corresponds the infinitesimal generator of an invariance group of the Hamiltonian or

* Present address: The University of Michigan, Ann Arbor, Michigan.

[1] E. Noether, Nachr. Akad. Wiss. Goettingen, II. Math. Physik. Kl., 235 (1918).

Lagrangian[2]; and for quantum theory in general, we have the following (presumably) well-known formal argument. Let $|p'\rangle$ be an "eigenstate" of the total four-momentum operators P_μ, with eigenvalues p'_μ. Let A be a linear operator that conserves energy momentum, i.e.,

$$\langle p''|A|p'\rangle = 0 \quad \text{for} \quad p'' \neq p'.$$

In other words, $A|p'\rangle$ is an eigenstate of P with the same eigenvalue p', so it follows that

$$[A, P_\mu] = 0.$$

Hence A is invariant under translation by any space–time four-vector b:

$$[A, e^{iP\cdot b}] = 0.$$

We see in Sec. II that it is a simple exercise to make this argument rigorous under the conditions stated in the abstract, including the case where A is a bounded operator, such as the S matrix.[3]

Actually, the question to what extent the converse of the energy-momentum conservation theorem is true has some relevance for elementary particle physics, where experimental statements are commonly statements about momentum space, involving only macroscopic space–time localization. In this situation the conservation law is verified more directly than the invariance principle.

Some theorists have argued that this matter of practice should be given the status of a matter of principle.[4] Either they deny the operational significance of microscopic space–time[5] or for some more conservative reason they propose to base the theory of strong interactions on momentum space and to treat space–time as a derived concept.[6] Of course, if one advocates this view, he is not thereby prevented from postulating translation invariance, since microscopic displacements could conceivably have a sense, even if microscopic space–time does not; but it seems more in the spirit of things for those who take the S matrix as the fundamental observable quantity to postulate instead the conservation law.[7]

Whether for reasons of practice or principle, we think there is at least a pedagogical value in spelling out some of the contexts in which energy-momentum conservation implies translation invariance, with the most direct applications being to S-matrix theory. Because the results are to some extent known, and because the proofs as well have very likely occurred to those who have wondered about the question with enough mathematical curiosity, we make no particular claim of originality for our rather straightforward discussion. On the other hand, a rigorous treatment does lead us indirectly to a potentially useful piece of information about the S matrix which is new, as far as we know. Namely, the connected S-matrix elements in momentum space (cluster amplitudes) are not only tempered distributions but kernels of bounded operators.

In Sec. II, we use the spectral theory to formulate the property of energy-momentum conservation for operators on a Hilbert space, and for a certain class of operators we transform the formal argument already given into a proof of the theorem on translation invariance. We discuss to what extent the conditions imposed characterize the operators for which the theorem is true.

In Secs. III and IV, we reformulate and prove the theorem by a different method, for bounded multilinear forms on Cartesian products $(\mathcal{H}_{m_1}, \cdots, \mathcal{H}_{m_r}, \mathcal{H}_{n_1}, \cdots, \mathcal{H}_{n_s})$, where $\mathcal{H}_m$ is the m-particle subspace of a Fock space. By "multilinear" we mean antilinear on each space $\mathcal{H}_{m_i}$ and linear on each $\mathcal{H}_{n_j}$. Such forms may correspond to operators between the spaces $\mathcal{H}_m$ and $\mathcal{H}_n$, where $m = \sum m_i$ and $n = \sum n_j$, but in general they do not. Whether such a general situation has a practical application, we do not know, but the generality costs nothing extra. The second proof deals directly with transition amplitudes in momentum space (tempered distributions), and the idea is to show that energy-momentum conservation is expressed only by delta functions in the transition amplitudes, and not by derivatives of delta functions. This leads at once to translation invariance. We are careful not to write delta functions at points where the mass shell is not a differentiable manifold, because they are not well defined at such points.

In Sec. V we mention that the result extends to the cluster parameters for momentum space amplitudes.

[2] A. Messiah, *Mécanique Quantique* (Dunod Cie., Paris, 1960), Vol. II, Chap. XV.

[3] Although we should not be surprised to learn that the argument in question is known, we have not succeeded in finding it in the literature. For the case of the S matrix, H. P. Stapp [Phys. Rev. **125**, 2139 (1962)] mentions without proof that translation invariance and energy-momentum conservation are equivalent. For bounded operators, the exercise is indeed not only simple but trivial, given the standard results of the spectral theory.

[4] E.g., G. F. Chew, Sci. Progr. (G.B.), **51**, 529 (1963); H. P. Stapp in Ref. 3; E. Lubkin, Nuovo Cimento **32**, 171 (1964).

[5] This strikes us as a radical view because we are not able to imagine all possible theories by means of which the concept could acquire an operational meaning. We do not intend by that a value judgement on the plausibility of theories motivated by such a view.

[6] There have been several interesting attempts in this direction, based on the S matrix. Among them we mention M. L. Goldberger and K. M. Watson, Phys. Rev. **127**, 2284 (1962); M. Froissart, M. L. Goldberger, and K. M. Watson, *ibid.* **131**, 2820 (1963); H. P. Stapp, *ibid.* **139**, B257 (1965); A. Peres, Ann. Phys. (N.Y.) **37**, 179 (1966). The last of these contains a more complete list of references.

[7] This is, for example, the attitude of Stapp. See Ref. 3.

Although this is a trivial fact, we again follow a "didactic" route in an attempt to clarify in what sense it is true. We apply some elementary theorems on Hilbert–Schmidt operators to find that the cluster amplitudes corresponding to a family of bounded *bilinear* forms are themselves kernels of bounded operators between the m- and n- particle Hilbert spaces, which conserve energy momentum if the original amplitudes do.

Finally, in an appendix, we prove that, on a separable Hilbert space, a closed operator commutes with all spectral projections if and only if it commutes with the translations. (The "only if" part is valid for nonseparable spaces as well.) This result is probably known to mathematicians, since it is only a slight generalization of the theorem for bounded operators, but neither the theorem nor its proof seems to be readily accessible to nonspecialists (such as the author).

II. FORMULATION, THEOREM, AND PROOF

What do we mean when we say that an operator conserves energy-momentum? We give ourselves a Hilbert space $\mathcal{H}$ and commuting self-adjoint energy-momentum operators P_μ, $\mu = 0, 1, 2, 3$, defined on a common dense submanifold of $\mathcal{H}$. That a linear operator on $\mathcal{H}$ conserves energy-momentum means at least that its matrix elements do not connect subspaces of $\mathcal{H}$ belonging to disjoint subsets of the spectrum of P_μ.

In other words, let

$$P_\mu = \int_{\mathrm{R}^4} p_\mu \, dE(p)$$

be the simultaneous spectral decomposition of P_μ, where $dE(p)$ is the spectral measure, with support on the spectrum of P_μ.[8] For any Borel set $\Delta \subset \mathrm{R}^4$, consider the projection operator[9]

$$E(\Delta) = \int_\Delta dE(p).$$

The subspace of $\mathcal{H}$ belonging to the part of the spectrum of P_μ contained in Δ is $E(\Delta)\mathcal{H} \equiv \mathcal{H}(\Delta)$. Let A be a linear operator on $\mathcal{H}$ with domain $D(A)$, which we may assume to be dense or not, as we please. Then we say that A conserves energy–momentum if, whatever be the Borel set Δ or $f \in D(A)$, the condition

$$E(\Delta)f = 0$$

implies that

$$E(\Delta)Af = 0.$$

Thus, if $f \in \mathcal{H}(\Delta') \cap D(A)$ and $g \in \mathcal{H}(\Delta)$, with Δ' and Δ disjoint, we have the minimum requirement just mentioned:

$$\langle g, Af \rangle = 0.$$

This equation is equivalent to the definition; for if $f \in D(A)$ and $E(\Delta)f = 0$, it follows that $f \in \mathcal{H}(\mathrm{R}^4 - \Delta)$, where $\mathrm{R}^4 - \Delta$ is the complement of Δ. Then for any $g \in \mathcal{H}$

$$\langle g, E(\Delta)Af \rangle = \langle E(\Delta)g, Af \rangle = 0;$$

hence

$$E(\Delta)Af = 0.$$

We aim to study under what conditions the fact that A conserves energy momentum implies that it commutes with all spectral projections $E(\Delta)$, and hence with all translations

$$T(b) = \int e^{ip\cdot b} \, dE(p), \quad b \in \mathrm{R}^4.$$

To make sense out of such a statement, we have to know something about the domains of the operators that occur. Following Riesz and Sz.-Nagy,[10] we define the domain of a product A_1A_2 to be the set of all vectors $f \in D(A_2)$ such that $A_2 f \in D(A_1)$. We write $A_1 \subseteq A_2$ if A_2 is an extension of A_1; i.e., $D(A_2) \supset D(A_1)$ and $A_2 f = A_1 f$ for $f \in D(A_1)$. We say that a bounded operator B defined on all of $\mathcal{H}$ commutes with A if $BA \subseteq AB$. We say that A is closed if, whenever both $f_n \in D(A)$ and Af_n are Cauchy sequences in the norm of $\mathcal{H}$, it follows that $\lim f_n = f \in D(A)$ and $\lim Af_n = Af$.

What we actually prove is the following theorem, which perhaps does not characterize the operators for which energy-momentum conservation and translation invariance are equivalent, but which probably comes close enough for practical purposes.

Theorem A: Let A be a closed linear operator on a separable Hilbert space $\mathcal{H}$. Then the following statements are equivalent:

(i) A conserves energy momentum, and $D(A)$ is invariant under spectral projections; i.e., $E(\Delta)D(A) \subset D(A)$ for all Δ;

(ii) $E(\Delta)A \subseteq AE(\Delta)$ for all Δ;

(iii) $T(b)A = AT(b)$ for all b.

If $\mathcal{H}$ is nonseparable, then we still have (i) $\Leftrightarrow$ (ii) $\Rightarrow$ (iii).

[8] Very readable summaries on the "SNAG" theorem are given by R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics, and All That* (W. A. Benjamin, Inc., New York, 1964), pp. 91–93; R. Jost, *The General Theory of Quantized Fields* (American Mathematical Society, Providence, Rhode Island, 1965), pp. 16–17.

[9] Recall that the Borel sets of R^n are the smallest family of sets that contains all denumerable unions, intersections, and complements of open sets.

[10] F. Riesz and B. Sz.-Nagy, *Functional Analysis* (Frederick Ungar Publishing Company, New York, 1955), Chap. VIII, Secs. 114–116.

The only nontrivial part of the proof is the relation between the statements (ii) and (iii). Because this result belongs properly to the functional calculus of self-adjoint operators, we take it for granted here and reserve the proof for the Appendix. Certainly its formal equivalent is a part of the folklore of quantum mechanics.

We complete the proof of Theorem A by showing the equivalence of (i) and (ii), without assuming that $\mathcal{H}$ is separable (or even that A is closed). To prove that (i) implies (ii), note that, for $f \in D(A)$ and any Δ,

$$E(\Delta)AE(\mathsf{R}^4 - \Delta)f = 0,$$

from energy-momentum conservation. Because

$$E(\Delta) + E(\mathsf{R}^4 - \Delta) = 1,$$

we have

$$E(\Delta)Af = E(\Delta)AE(\Delta)f,$$
$$AE(\Delta)f = E(\Delta)AE(\Delta)f.$$

Hence $E(\Delta)A \subseteq AE(\Delta)$.

It only remains to show that (ii) implies (i). But that is trivial. First, $E(\Delta)D(A) \subset D(A)$, from the definition of the expression (ii). That A conserves energy momentum follows at once from (ii) and the definition of energy-momentum conservation. Thus, the theorem is proved.

We have not made any restrictions on the spectrum of P_μ. For closed operators and separable spaces, Theorem A says that it is not possible to relax the condition of the invariance of $D(A)$ under spectral projections. Whether the domain requirement is automatically implied in the case of closed operators by energy-momentum conservation as formulated here, we do not know; nor are we inclined to worry about it. The condition that A be closed, or at least have a closure fulfilling the other conditions, seems essential for the proof in the Appendix of the relation between (ii) and (iii); but we do not know whether it can be relaxed. We also do not know whether the statement (iii) $\Rightarrow$ (ii) is true for nonseparable spaces.

At any rate, the conditions of the theorem seem sufficiently general for most practical applications in physics.

III. ALTERNATIVE FORMULATION IN FOCK SPACE

Of course, nothing more has to be said in order to apply the theorem to a Fock space. But in that case, we have constructed another proof, for a certain class of operators and forms, which we think instructive. In the first proof, the nontrivial part was contained in the spectral theory. In the second, the basic mathematical tools are the nuclear theorem for tempered distributions,[11] plus a theorem of Schwartz on the structure of a distribution with support on a submanifold of some R^n.

For simplicity we put ourselves in the relativistic Fock space $\mathcal{F}$ corresponding to spinless particles with a single mass $M > 0$. The generalization of the discussion to Fock spaces with denumerable numbers of different types of particles with various spins and nonzero masses is trivial. Thus,

$$\mathcal{F} = \bigoplus_{n=0}^{\infty} \mathcal{H}_n,$$

where for $n \geq 1$,

$$\mathcal{H}_n = \operatorname{Sym}\left[L_2\left(\frac{d^3p_1}{\omega_1}\cdots\frac{d^3p_n}{\omega_n}, \mathsf{R}^{3n}\right)\right],$$
$$\omega_i \equiv \omega(\mathbf{p}_i) = (M^2 + \mathbf{p}_i^2)^{\frac{1}{2}} > 0,$$

is the symmetrized Hilbert space of momentum-space wavefunctions of n free particles.

Because each $\mathcal{H}_n$, for $n \geq 1$, is identified with an L_2 space of functions, with a measure that "dominates" Lebesgue measure (and is dominated by it: the zero sets are the same), we can give a meaning to the "support" of a vector $f \in \mathcal{H}_n$. Namely, let h be any element in the equivalence class of almost everywhere equal functions that corresponds to f; we write $h \in f$. Let supp h be the support of h in the usual sense, i.e., the complement of the largest open set of R^{3n} on which h vanishes. Then define

$$\operatorname{supp} f = \bigcap_{h \in f} \operatorname{supp} h.$$

If f can be represented by a continuous function $h \in f$, we have (exercise for the reader)

$$\operatorname{supp} f = \operatorname{supp} h.$$

First we consider bounded operators on $\mathcal{F}$. To each bounded linear operator B, and to each ordered pair of spaces $(\mathcal{H}_m, \mathcal{H}_n)$, we associate the bounded bilinear form

$$B_{mn}(f, g) = \langle f, Bg\rangle,$$

where $f \in \mathcal{H}_m$ and $g \in \mathcal{H}_n$. We say that B_{mn} conserves energy momentum if

$$B_{mn}(f, g) = 0$$

for all f and g having supports that nowhere satisfy the energy-momentum conservation equations. More explicitly, let $\mathbf{P} = (\mathbf{p}_1, \cdots, \mathbf{p}_m)$ and $\mathbf{Q} = (\mathbf{q}_1, \cdots, \mathbf{q}_n)$. Then B_{mn} vanishes if

$$(\mathbf{P}, \mathbf{Q}) \in \operatorname{supp} f \times \operatorname{supp} g$$

[11] L. Schwartz, *Théorie des Distributions* (Hermann et Cie., Paris, 1959), Vol. II, Chap. VII.

implies that, for at least one μ,

$$t^\mu(\mathbf{P}, \mathbf{Q}) \equiv \sum_{i=1}^{m} p_i^\mu - \sum_{j=1}^{n} q_j^\mu \neq 0,$$

where all four-vectors are on the positive sheet of the mass hyperboloid; e.g., $p_i^0 = \omega(\mathbf{p}_i)$. We say that B conserves energy-momentum if each B_{mn} does.[12]

It is not difficult to see that this definition is equivalent to the one given before (in the cases where it applies). It is perhaps worth remarking that, for the S matrix, the above statement of energy-momentum conservation for the transition amplitudes is equivalent to the analogous requirement on the observable transition probabilities,[13] as the reader can immediately see for himself.

In the next section, we prove again that if B conserves energy momentum, it commutes with the standard unitary representation of space–time translations defined on $\mathcal{F}$. We do it by considering the tempered distributions $B_{mn}(\mathbf{P}, \mathbf{Q})$, defined by restricting $B_{mn}(f, g)$ to pairs of functions in the appropriately symmetrized Schwartz spaces[11] $(\mathcal{S}_m, \mathcal{S}_n)$ of test functions which are C_∞ and decrease at infinity with all derivatives faster than any inverse polynomial. That we get a tempered distribution on the entire subspace of test functions in $\mathcal{S}(\mathrm{R}^{3(m+n)})$ that are symmetric in the first m and last n three-vectors follows from[14]:

(i) the fact that B_{mn} is a bounded bilinear form (after accounting for the antilinearity of the first factor)

$$|B_{mn}(f, g)| \leq C \, \|f\| \, \|g\|,$$

where $\|f\|$ indicates the scalar product norm in $\mathcal{F}$;

(ii) the fact that the topology of $\mathcal{S}_m$ is finer than that induced from the strong topology of $\mathcal{H}_m$;

(iii) the "théorème nucléaire" of Schwartz.[15]

If B conserves energy momentum, the tempered distributions B_{mn} have their supports on the sets where

$$t^\mu(\mathbf{P}, \mathbf{Q}) = 0.$$

[12] It will generally be obvious how to take into account the case m or $n = 0$, corresponding to the vacuum with zero energy momentum, so we most often do not mention it explicitly.

[13] I am indebted to D. Iagolnitzer for drawing my attention to this point, as well as to the fact that translation invariance of the probabilities does not imply translation invariance of the amplitudes (although Poincaré invariance does).

[14] This simple proof occurred to several other people before it occurred to the author. In the literature, the question is raised indirectly by E. H. Wichmann and J. H. Crichton [Phys. Rev. **132**, 2788 (1963)], who give a lucid discussion of the cluster decomposition property which assumes that the S-matrix amplitudes are tempered distributions. K. Hepp [Helv. Phys. Acta **37**, 659 (1964)] states it as a fact, without giving the proof. We have found references to the proof in J. R. Taylor, Phys. Rev. **142**, 1236 (1966), and D. Iagolnitzer, "S-Matrix Theory and Phenomenological Space–Time Description," Saclay preprint (to be published).

[15] L. Gårding and J. L. Lions, Nuovo Cimento, Suppl. **14**, 9 (1959).

Our second method of proving the translation invariance of B is to show that, on a sufficiently large space of test functions, B_{mn} factorizes into a product of a delta function for energy-momentum conservation times a "tempered distribution" on the manifold defined by the conservation law. As mentioned in the Introduction, the essential point is to show that derivatives of delta functions cannot occur, because they conflict with the boundedness of B_{mn}, considered as a bilinear form.

Before passing to the theorem and proof, note that, as far as the discussion so far has been concerned, we have never used the fact that the bound C is the same for each bilinear form B_{mn}; we could just as well have a family of positive constants C_{mn} which could be unbounded for large (m, n), corresponding to a class of unbounded operators on $\mathcal{F}$. (The "number of particles" operator is a simple example.) Actually, we never need to know that B_{mn} is a bounded *bilinear* form; we can do just as well with the weaker statement that it is a bounded *multilinear* form which satisfies

$$|B_{mn}(f_1 \cdots f_m, g_1 \cdots g_n)| \leq C_{mn} \prod_{i,j} \|f_i\| \, \|g_j\|,$$

with $f_i, g_j \in \mathcal{H}_1$. Such forms can correspond to a larger class of unbounded operators[16]; or, on the other hand, they might not correspond to operators at all, not even between $\mathcal{H}_n$ and $\mathcal{H}_m$.

In fact, the whole discussion goes through for bounded multilinear forms of the type

$$B_{m_1 \cdots m_r, n_1 \cdots n_s}(f_1, \cdots, f_r, g_1, \cdots, g_s),$$

where $f_i \in \mathcal{H}_{m_i}$ and $g_j \in \mathcal{H}_{n_j}$. Energy-momentum conservation is defined in the obvious way, and we still have the reduction to tempered distributions. Although we have in mind no particular situation where such generality might be useful, there is no reason not to state our result for such cases. *A priori*, as we see in Sec. V, we would have said that the cluster amplitudes are an example of bounded multilinear forms on Cartesian products of $\mathcal{H}_1$, if the Hilbert–Schmidt theorems did not tell us that they are really bounded bilinear forms.

IV. THEOREM FOR MULTILINEAR FORMS

The theorem below is stated for forms. If the forms come from closed operators on $\mathcal{F}$, it extends immediately to the operators, by linearity and

[16] The domain specified is translation invariant, but not invariant under spectral projections. It follows from Theorem A, with Theorem B in Sec. IV, that if the operators conserve energy momentum, they are not closed on this domain.

continuity, modulo questions of domain. Certainly there is no problem for bounded operators.

Theorem B: Let $T(b)$ be the unitary representation of space–time translations on $\mathcal{F}$, defined on each $\mathcal{H}_m$ by

$$[T(b)f](\mathbf{P}) = \exp\left(-i\sum_{j=1}^{m} p_j \cdot b\right) f(\mathbf{P}).$$

If $B_{m_1 \cdots m_r, n_1 \cdots n_s}$ is a bounded multilinear form on $(\mathcal{H}_{m_1}, \cdots, \mathcal{H}_{n_s})$ which conserves energy momentum, then

$$B_{m_1 \cdots n_s}[T(b)f_1, \cdots, T(b)g_s] = B_{m_1 \cdots n_s}(f_1, \cdots, g_s).$$

To save writing, the proof is given in detail only for bounded bilinear forms B_{mn}. Very little modification is needed to extend it to multilinear forms, and it will hardly tax the reader to provide it himself.

Consider B_{mn} as a bounded linear transformation $B_{mn}\colon \mathcal{H}_n \to \mathcal{H}_m$. Because

$$T(b)B_{mn}T(b)^{-1} - B_{mn}$$

is also a bounded (i.e., continuous) linear transformation of $\mathcal{H}_n$ into $\mathcal{H}_m$, it suffices to prove

$$\langle f, T(b)B_{mn}T(b)^{-1}g\rangle = \langle f, B_{mn}g\rangle,$$

where f and g are arbitrary elements of two sets of vectors, each of which spans (by means of finite linear combinations) a dense submanifold of $\mathcal{H}_m$, $\mathcal{H}_n$ respectively. In particular, we always choose $(f, g) \in (\mathcal{S}_m, \mathcal{S}_n)$, with m and n running over the positive integers.

In order to avoid a possible difficulty about defining delta functions and their rth derivatives of the form

$$\prod_{\mu=0}^{3} \delta^{(r_\mu)}[t^\mu(\mathbf{P}, \mathbf{Q})],$$

(where r is a "four-vector" with nonnegative integers as components) at zeros of t where the Jacobian matrix has rank less than four, we make one further restriction on the support of one of the elements, say g, of (f, g). Namely, if $n \geq 2$, we demand that there shall be no $\mathbf{Q} \in \operatorname{supp} g$ for which all corresponding mass hyperboloid four-vectors q_i are collinear; at least two of the four-vectors are to be linearly independent. Hepp[17] has observed that even the smaller subspace of functions in $\mathcal{S}_n$ with supports having *no* two of the corresponding four-vectors collinear ("disjoint velocities") is dense in $\mathcal{H}_n$. The reader may easily verify that, with this restriction, the Jacobian matrix evaluated for $t = 0$ indeed has rank four.

Now we consider the tempered distribution $B_{mn}(\mathbf{P}, \mathbf{Q})$, restricted to the open set Ω of points $(\mathbf{P}, \mathbf{Q})$ where $\mathbf{Q}$ satisfies the condition just mentioned. The support of B_{mn} in Ω is a C_∞ manifold, which we denote $\operatorname{supp}_\Omega B_{mn}$, of dimension 3 if $m = n = 1$, and of dimension $3(m + n) - 4$ if $m, n \geq 2$. In the latter case, $\operatorname{supp}_\Omega B_{mn}$ is the set of simultaneous zeros of the C_∞ functions $t^\mu(\mathbf{P}, \mathbf{Q})$, which forms a C_∞ manifold by the implicit function theorem.[18] It is not difficult to see that $\operatorname{supp}_\Omega B_{mn}$ can even be covered by a finite number of coordinate neighborhoods.

Some theorems of Schwartz[19] tell us that, on Ω, B_{mn} can be written as a finite sum:

$$B_{mn}(\mathbf{P}, \mathbf{Q}) = \sum_r \prod_{\mu=0}^{3} \delta^{(r_\mu)}[t^\mu(\mathbf{P}, \mathbf{Q})] R_{mn}^{(r)}(\mathbf{P}, \mathbf{Q}),$$

where R_{mn}^r is a "tempered distribution" on $\operatorname{supp}_\Omega B_{mn}$. Of course, if $m = n = 1$, the product in the expression above runs only over $\mu = 1, 2, 3$, and the whole discussion simplifies because the manifold is just R^3.

Finally, we smear with test functions $(f, g) \in (\mathcal{S}_m, \mathcal{S}_n)$ satisfying $\operatorname{supp} f \times \operatorname{supp} g \subset \Omega$. This set of pairs of test functions is invariant under any translation $[T(b)f, T(b)g]$. Since $T(b)$ is unitary, and since B_{mn} is a bounded bilinear form, we have

$$|B_{mn}[T(b)f, T(b)g]| \leq C\, \|f\|\, \|g\|,$$

with the right-hand side independent of b.

To see the behavior of the left-hand side, we substitute the decomposition of $B_{mn}(\mathbf{P}, \mathbf{Q})$. Integrating by parts, we get

$$B_{mn}[T(b)f, T(b)g] = \int \prod_{i,j} d^3p_i\, d^3q_j \sum_r \delta[t(\mathbf{P}, \mathbf{Q})] R_{mn}^{(r)}(\mathbf{P}, \mathbf{Q}) \times \prod_{\mu=0}^{3} (-1)^{r_\mu} \frac{\partial^{r_\mu}}{(\partial t^\mu)^{r_\mu}} \left[e^{it\cdot b} \frac{\bar{f}(\mathbf{P}) g(\mathbf{Q})}{\prod_{i,j} \omega(\mathbf{p}_i)\omega(\mathbf{q}_j)} \right].$$

Suppose that some derivative of a delta function occurs; that is, there is a term with $r \neq 0$ such that $R_{mn}^{(r)} \not\equiv 0$. Consider the terms of highest homogeneous order in r. Carrying out the differentiation gives a

[17] K. Hepp, Commun. Math. Phys. **1**, 7 (1965).

[18] L. Auslander and R. E. MacKenzie, *Introduction to Differentiable Manifolds* (McGraw-Hill Book Company, Inc., New York, 1963), Chap. II. Our excuse for sketching the proof of these well-known facts about the mass shell is that it takes only a few words, and hopefully makes things clearer. For more details, see K. Hepp, Helv. Phys. Acta **36**, 355 (1963); and **37**, 55 (1964); H. P. Stapp, "Studies in the Foundations of S-Matrix Theory," University of California, Lawrence Radiation Laboratory Report, UCRL 10843.

[19] L. Schwartz, Ref. 11, Vol. I, pp. 100–103, applied to Ex. 2, p. 114, and using the temperedness of B_{mn}.

polynomial in b. We can always find a pair (f, g) in our set (because it is dense) such that the coefficient of at least one $(b^0)^{r_0}(b^1)^{r_1}\cdots(b^3)^{r_3}$ in a term of highest homogeneous degree is nonzero. As a function of b, this term cannot be canceled identically by other terms coming from r of the same or lower order. Thus the left-hand side of our inequality contains a polynomial of nonzero degree, which cannot be bounded as a function of b, conflicting with the right-hand side. We conclude that there are no derivatives of delta functions.

But then the remaining delta function implies that

$$B_{mn}[T(b)f, T(b)g] = B_{mn}(f, g),$$

which is what we set out to prove.

V. CLUSTER AMPLITUDES

To avoid a possible point of confusion, we follow Wichmann and Crichton[20] in emphasizing that a large class of amplitudes, labeled in this case by (m, n), has a cluster parametrization, which is given by a purely combinatorial algorithm, having very little to do with the mathematical nature of the amplitudes involved. The *cluster decomposition property* of the S matrix, for example, is logically independent from the cluster parametrization. The relation between the two is rather one of convenience; the cluster property has an especially simple and useful expression in terms of cluster parameters. That, of course, is why cluster parameters are interesting, but we do not assume here that the cluster property holds, nor, for the moment, that we have energy-momentum conservation. We seek only to determine the general structure of the cluster amplitudes for a family of bounded bilinear forms, in the interest of having as much *relevant* information as possible when we apply the theorem on translation invariance.

To help in defining the cluster amplitudes, we introduce some notation. To each bounded bilinear form B_{mn} we associate a kernel defined by

$$B_{mn}(f, g) = \int d\mathbf{P}\, d\mathbf{Q} B_{mn}(\mathbf{P}, \mathbf{Q})\bar{f}(\mathbf{P})g(\mathbf{Q}),$$

where $d\mathbf{P}$ and $d\mathbf{Q}$ are the invariant measure elements for $\mathcal{H}_m$ and $\mathcal{H}_n$. When f and g are in $\mathcal{S}_m$ and $\mathcal{S}_n$, the kernel $B_{mn}(\mathbf{P}, \mathbf{Q})$ is the same as the tempered distribution already considered; but it is also defined as a respectable mathematical object for $(f, g) \in (\mathcal{H}_m, \mathcal{H}_n)$. By the Riesz representation theorem, we may associate L_2 (equivalence classes of) functions $B_{mn}(\mathbf{P}, g) \in \mathcal{H}_m$ and $B_{mn}(f, \mathbf{Q}) \in \mathcal{H}_n$ to any $g \in \mathcal{H}_n$ and $f \in \mathcal{H}_m$; and we have

$$B_{mn}(f, g) = \int d\mathbf{P} \bar{f}(\mathbf{P}) B_{mn}(\mathbf{P}, g),$$
$$= \int d\mathbf{Q} B_{mn}(f, \mathbf{Q}) g(\mathbf{Q}).$$

In other words, we may "integrate" in either order.

Cluster amplitudes B^c_{mn} for a family of such kernels may be defined recursively on (m, n) as follows[21]:

(i) if m or n is zero,

$$B_{mn} = B^c_{mn};$$

(ii) if m and n are nonzero,

$$B_{mn}(\mathbf{P}, \mathbf{Q}) = \sum_I \prod_i B^c_{m_i n_i}(\mathbf{P}_{I_i}, \mathbf{Q}_{I_i}),$$

where I labels the partitions of the variables $(\mathbf{P}, \mathbf{Q})$ into disjoint sets labeled I_i, each of which contains nonzero numbers m_i and n_i of $\mathbf{p}$'s and $\mathbf{q}$'s. Within each partition, the natural order is preserved. Solving, we may write, for m and n nonzero,

$$B^c_{mn}(\mathbf{P}, \mathbf{Q}) = \sum_I \eta(I) \prod_i B_{m_i n_i}(\mathbf{P}_{I_i}, \mathbf{Q}_{I_i}),$$

where $\eta(I)$ is a numerical factor that does not concern us.

As it stands, B^c_{mn} is well defined as a tempered distribution, which contains an over-all delta function for energy-momentum conservation if the B_{mn} do, and as a bounded multilinear form for finite sums of products of one-particle wavefunctions. In addition, we can prove:

Theorem C: The cluster amplitudes B^c_{mn} for a family of bounded bilinear forms are also bounded bilinear forms. In particular, for m and n nonzero, $B^c_{mn}(\mathbf{P}, \mathbf{Q})$ is the kernel of a bounded linear transformation from $\mathcal{H}_n$ into $\mathcal{H}_m$.

For the proof, we may assume that m and n are nonzero; otherwise the result is trivial. Let us consider what meaning we may assign to $B^c_{mn}(\mathbf{P}, g)$ for $g \in \mathcal{H}_n$. The plan of the proof is to show that:

(i) this expression is well defined as an element of $\mathcal{H}_m$, and $B^c_{mn}(f, g)$, defined in this way for all $(f, g) \in (\mathcal{H}_m, \mathcal{H}_n)$, is a bilinear extension of the form already defined if f or g is a sum of products of one-particle wavefunctions;

(ii) the domain of the adjoint of this linear transformation is all of $\mathcal{H}_m$, so that we know from an

[20] See Ref. 14.

[21] By analogy with the definition of truncated Wightman functions due to R. Haag, Phys. Rev. **112**, 669 (1958). Any other consistent choice of momentum-dependent, but measurable phases in this definition would be harmless for our purpose.

extension of the Hellinger–Toeplitz theorem[22] that B^c_{mn} is bounded.

From the definition of the cluster amplitudes, it is enough to look at typical terms of the form

$$K^I(\mathbf{P}, g) = \int d\mathbf{Q} \prod_i B_{m_i n_i}(\mathbf{P}_{I_i}, \mathbf{Q}_{I_i}) g(\mathbf{Q}).$$

To define such a term, we first partition $\mathbf{Q}$ into two disjoint parts, $\mathbf{Q}_{I_1}$ and the remaining $n - n_1$ three-vector variables, $(\mathbf{Q}_{I_2}, \cdots, \mathbf{Q}_{I_l})$. Because it is an L_2 function, we may consider

$$g[\mathbf{Q}_{I_1}, (\mathbf{Q}_{I_2}, \cdots)] \equiv g(\mathbf{Q})$$

as the kernel of a Hilbert–Schmidt (H–S) operator from $\mathcal{K}_{n-n_1}$ into $\mathcal{K}_{n_1}$. Standard theorems on H–S operators[23] tell us that the product of the bounded linear operator $B_{m_1 n_1}$ and the H–S "operator" g is an H–S operator, and that

$$K_{m_1, n-n_1}(\mathbf{P}_{I_1}, \mathbf{Q}_{I_2}, \cdots) \equiv \int d\mathbf{Q}_{I_1} B_{m_1 n_1}(\mathbf{P}_{I_1}, \mathbf{Q}_{I_1}) g[\mathbf{Q}_{I_1}, (\mathbf{Q}_{I_2}, \cdots)]$$

is the kernel of the resultant H–S mapping from $\mathcal{K}_{n-n_1}$ into $\mathcal{K}_{m_1}$. That means precisely that $K_{m_1, n-n_1}$ is in the L_2 space of functions of $3(m_1 + n - n_1)$ variables (always with respect to the invariant measure).

Thus, we may repeat the process, partitioning the variables $(\mathbf{P}_{I_1}, \mathbf{Q}_{I_2}, \cdots, \mathbf{Q}_{I_l})$ into two parts, $\mathbf{Q}_{I_2}$ and the rest, $(\mathbf{P}_{I_1}, \mathbf{Q}_{I_3}, \cdots, \mathbf{Q}_{I_l})$. Then we find that

$$K_{m_1, m_2, n-n_1-n_2}(\mathbf{P}_{I_1}, \mathbf{P}_{I_2}, \mathbf{Q}_{I_3}, \cdots) \equiv \int d\mathbf{Q}_{I_2} B_{m_2 n_2}(\mathbf{P}_{I_2}, \mathbf{Q}_{I_2}) K_{m_1, n-n_1}(\mathbf{P}_{I_1}, \mathbf{Q}_{I_2}, \cdots)$$

is an H–S kernel from the space corresponding to $(\mathbf{P}_{I_1}, \mathbf{Q}_{I_3}, \cdots, \mathbf{Q}_{I_l})$ into the space corresponding to $\mathbf{P}_{I_2}$, and hence L_2 in the space corresponding to all the variables.

Continuing in this way, we find that $K^I(\mathbf{P}, g)$ is L_2 in the nonsymmetrized space corresponding to $\mathcal{K}_m$. It is clear that we have defined in this way a linear map K^I: $\mathcal{K}_n \to \mathcal{K}_m$ that is an extension of the multilinear form defined trivially for wavefunctions of the type $\prod_i g_i(\mathbf{Q}_{I_i})$.

Consider the adjoint of K^I. By definition, a vector $f \in \mathcal{K}_m$ is in the domain of the adjoint if there exists a vector $h \in \mathcal{K}_n$ such that

$$\langle f, K^I g \rangle = \langle h, g \rangle$$

for all $g \in \mathcal{K}_n$. In our case, we find that such a vector exists for every f, so that the adjoint is everywhere defined. The proof is to show that we can calculate the scalar product on the left-hand side of the above equation by integrating first on the $d\mathbf{P}_{I_i}$ successively in some order, then on $d\mathbf{Q}$. By the same argument as before, the $\mathbf{P}$ integration defines for us a vector h; our only problem is to see that we get the same scalar product.

First[24] we consider the scalar product $\langle f, K^I g \rangle$ as an iterated integral on $\mathbf{P}$ and $\mathbf{Q}$, computed in the order (beginning at the right)

$$\int d\mathbf{P}_{I_1} \cdots \int d\mathbf{P}_{I_l} \int d\mathbf{Q}_{I_l} \cdots \int d\mathbf{Q}_{I_1}.$$

The $\mathbf{Q}$ integrations are defined as already described, and we have used Fubini's theorem to write the $\mathbf{P}$ integration in iterated form. Next we note that after doing the integrations on $d\mathbf{Q}_{I_{l-1}} \cdots d\mathbf{Q}_{I_1}$, we have to integrate the kernel of the bounded operator $B_{m_l n_l}$ with a function that is L_2 in $\mathbf{Q}_{I_l}$ and then with a function that is L_2 in $\mathbf{P}_{I_l}$, for fixed values of the remaining variables. We have already observed that, from the definition of the kernel, we can interchange the order of these two integrations. Thus, we may integrate first on $d\mathbf{P}_{I_l} d\mathbf{Q}_{I_{l-1}} \cdots d\mathbf{Q}_{I_1}$; and by our previous argument, the remaining integrand is L_1 in $(\mathbf{P}_{I_1}, \cdots, \mathbf{P}_{I_{l-1}}, \mathbf{Q}_{I_l})$, being a product of two L_2 functions. By Fubini's theorem, we now see that we get the same scalar product if we do the $d\mathbf{Q}_{I_l}$ integration last, integrating in the order

$$\int d\mathbf{Q}_{I_l} \int d\mathbf{P}_{I_1} \cdots \int d\mathbf{P}_{I_l} \int d\mathbf{Q}_{I_{l-1}} \cdots \int d\mathbf{Q}_{I_1}.$$

At this stage it is not difficult to verify that the $d\mathbf{P}_{I_l}$ integration can be interchanged successively with each preceding $d\mathbf{Q}_{I_i}$ integration, because the $\mathbf{P}$ and $\mathbf{Q}$ integrations are decoupled for $i \neq l$. Thus we arrive at the sequence of integrations

$$\int d\mathbf{Q}_{I_l} \int d\mathbf{P}_{I_1} \cdots \int d\mathbf{P}_{I_{l-1}} \int d\mathbf{Q}_{I_{l-1}} \cdots \int d\mathbf{Q}_{I_1} \int d\mathbf{P}_{I_l}.$$

Reasoning by finite descent, we repeat the whole process; and at last we find that the scalar product can be calculated by integrating in the order

$$\int d\mathbf{Q} \int d\mathbf{P}_{I_1} \cdots \int d\mathbf{P}_{I_l},$$

where we have used Fubini's theorem for the last time to replace the iterated $\mathbf{Q}$ integrations by a single multiple integration. Therefore, the adjoint of K^I has

[22] F. Riesz and Sz.-Nagy, Ref. 10, pp. 305–306.

[23] N. Dunford and J. T. Schwartz, *Linear Operators* (Interscience Publishers, Inc., New York, 1958), Part II, Chap. XI, Sec. 6.

[24] The reader who treats the following argument as a recipe for pencil and paper will find it straightforward.

all of $\mathcal{H}_m$ for its domain, and we conclude that K^I is bounded.

The original linear transformation was defined on the dense submanifold of $\mathcal{H}_n$ spanned by wavefunctions of the product form. Thus the extension K^I is unique because it is continuous and, in particular, it does not depend on the order in which we choose to do the original $\mathbf{Q}$ integrations. We are justified in claiming that K^I is well defined for each I, and that the theorem is proved.

Note that by the same argument the converse of Theorem C is also true. If the B^c_{mn} are bounded bilinear forms, so are the B_{mn}.

Now apply Theorem B. It is clear that if the B_{mn} conserve energy momentum, so do the B^c_{mn}. In that case, the cluster amplitudes are translation invariant. We could reach the same conclusion directly from the translation invariance of B_{mn}.

VI. CONCLUSION

We have verified that energy-momentum conservation implies translation invariance in a fairly general class of theories related to Hilbert space, and in particular for the S matrix. We have also shown that the cluster amplitudes for a family of bounded bilinear forms can be discussed in the same framework, as bounded bilinear forms.

As indicated in the title, our hope in this discussion has been not so much to achieve the virtue of originality as that of clarity. If we have not succeeded in even this modest aspiration, we hope that the reader will agree that it is no reflection on the utility or the simplicity of the mathematical tools that we have chosen.

ACKNOWLEDGMENTS

It is a pleasure to thank L. Streit, M. Jacob, and D. Iagolnitzer for conversations which led me to attack this problem, and R. Stora for his patient reading of the manuscript and his constructive suggestions. I am also grateful to C. Bloch for the hospitality of the Centre d'Études Nucléaires de Saclay.

APPENDIX

Here we prove that $E(\Delta)A \subseteq AE(\Delta)$ for all Borel sets if and only if $T(b)A \subseteq AT(b)$ for all translations.[25] The proof that the commuting of the spectral projections with A implies the commuting of the translations with A is rather easy, given some basic results of measure theory and the fact that A is closed. It is not necessary in this case to assume that $\mathcal{H}$ is separable.

The proof of the converse for separable $\mathcal{H}$ is a little more delicate. Modulo a straightforward reduction, our discussion imitates an argument of Sz.-Nagy,[26] used in the proof of Stone's theorem to show that the spectral projections commute with all bounded operators that commute with all elements of the corresponding continuous, one-parameter, unitary group.

Our basic method of proving the two statements is to show that each operator in one of the two sets, labeled by Borel sets or by four-vectors, can be approximated strongly by finite linear combinations of operators in the other set, and to use the fact[27] that if B_n is a strongly convergent sequence of bounded operators with bounded limit B, and if A is a closed operator such that $B_nA \subseteq AB_n$ for all n, then $BA \subseteq AB$.

By means of the functional calculus for bounded functions of commuting self-adjoint operators (such as P_μ), the approximation of the operators in one class by those of the other can be reduced to that of the approximation of the corresponding functions. Namely, let $h(p)$ be a bounded function on R^4, measurable with respect to the spectral measure, i.e., with respect to all the measures $\langle f, dE(p)g\rangle$; and let $h_n(p)$ be a uniformly bounded sequence of such functions, which converges to $h(p)$ almost everywhere with respect to the spectral measure. Then the corresponding bounded operators

$$h_n = \int h_n(p)\, dE(p)$$

converge strongly to[28]

$$h = \int h(p)\, dE(p).$$

In our case we have to consider two classes of such functions, composed on the one hand of finite linear combinations of characteristic functions of Borel sets,

$$\xi_\Delta(p) = \begin{cases} 1 & \text{if} \quad p \in \Delta, \\ 0 & \text{if} \quad p \notin \Delta, \end{cases}$$

and on the other hand of finite linear combinations of exponentials, $\exp(ib\cdot p)$, i.e., of trigonometric polynomials. These functions are certainly bounded. The characteristic functions are measurable with respect to the spectral measure, because on locally compact Hausdorff spaces such as R^n the Borel sets are measurable with respect to any measure; and continuous functions, such as exponentials, are measurable with respect to any measure on such

[25] Note that $T(b)A \subseteq AT(b)$ for all translations implies by definition that $T(b)D(A) \subset D(A)$, and hence from the group property that $T(b)A = AT(b)$.

[26] F. Riesz and Sz.-Nagy, Ref. 10, p. 383.

[27] Ref. 10, p. 302.

[28] Ref. 10, Sec. 126.

spaces.[29] Our problem is to show that sufficiently many functions in each class can be approximated in the sense described above by functions in the other class.

To emphasize the point at which the separability of $\mathcal{H}$ enters, we divide the work among three lemmas, which together add up to the required results. We do not assume that $\mathcal{H}$ is separable unless we say so explicitly. But we always assume that A is closed.

Lemma 1: If $E(\Delta)A \subseteq AE(\Delta)$ for all Borel sets, then $T(b)A \subseteq AT(b)$ for all b.

Proof: We refer to a basic theorem of measure theory, according to which any measurable function is the limit of an *everywhere*-convergent sequence of simple functions.[30] A simple function is a finite linear combination of characteristic functions of pairwise-disjoint, measurable sets. The sequence can be chosen to be uniformly bounded if the limit function is bounded.[31] Since a continuous function on R^n is, in particular, Borel-measurable, the result follows from our previous remarks.

Lemma 2: If $T(b)A \subseteq AT(b)$ for all b, then $E(\Delta)A \subseteq AE(\Delta)$ for all compact Δ.

Proof: We have to express ξ_Δ for any compact Δ as the limit of a uniformly bounded, everywhere-convergent sequence of trigonometric polynomials. This can be achieved by the argument of Sz.-Nagy mentioned before.[26] First we take a decreasing sequence $\{U_n\}$ of bounded open neighborhoods of Δ, such that $\bigcap_{n=1}^{\infty} U_n = \Delta$. Applying Urysohn's lemma,[32] we choose a continuous, nonnegative, real function f_n which is unity on Δ, has support in $\bar{U}_n$ (the closure of U_n), and is bounded by unity. Next, we choose an increasing sequence of compact cubes $\square_n < \bar{U}_n$, such that $\bigcup_{n=1}^{\infty} \square_n = \mathsf{R}^4$; and we let g_n be the continuous periodic function defined by f_n in $\square_n$. The uniformly bounded sequence $\{g_n\}$ converges everywhere to ξ_Δ.

Finally, we apply Weierstrass's approximation theorem[33] to approximate g_n uniformly to within $1/n$ by a trigonometric polynomial t_n of the same period. The sequence $\{t_n\}$ is uniformly bounded and converges everywhere to ξ_Δ.

[29] M. A. Naimark, *Normed Rings* (P. Noordhoff Ltd., Groningen, The Netherlands, 1964), Appendix III.
[30] P. R. Halmos, *Measure Theory* (D. Van Nostrand Company, Inc., Princeton, New Jersey, 1950), p. 86.
[31] Ref. 30, Ex. 2, p. 86.
[32] Ref. 29, p. 28.
[33] Ref. 29, p. 33.

Lemma 3: Let $\mathcal{H}$ be separable. If $E(\Delta)A \subseteq AE(\Delta)$ holds for all compact Δ, it holds for all Borel sets.

Proof: Every Borel set is "summable" with respect to the spectral measure; i.e., $\langle f, E(\Delta)f\rangle$ is finite for all $f \in \mathcal{H}$. According to a basic result of measure theory,[34] if a set is summable with respect to some measure, there is a denumerable family of compact sets $\Delta_n \subset \Delta$ (which can even be chosen pairwise-disjoint) such that the set $R = \Delta - \bigcup_{n=1}^{\infty} \Delta_n$ is a set of zero measure (the difference of two sets is the set of points in the first, not in the second). We want to find a similar family with the property that the remainder R, which is a Borel set in our case, has spectral measure zero, i.e., such that $\langle f, E(R)f\rangle = 0$ for all f.

This equation is true for all vectors in $\mathcal{H}$ if and only if it is true for a dense set in $\mathcal{H}$, because $E(R)$ is a projection, hence bounded, hence continuous. Since $\mathcal{H}$ is separable, we can choose a denumerable dense set of vectors f_i.

Corresponding to each f_i, we choose a decomposition of Δ as above, such that

$$R_i \equiv \Delta - \bigcup_{n=1}^{\infty} \Delta_n^{(i)}$$

has measure zero for the corresponding measure. It follows that

$$R \equiv \bigcap_{i=1}^{\infty} R_i = \Delta - \bigcup_{i=1}^{\infty} \bigcup_{n=1}^{\infty} \Delta_n^{(i)}$$

is a Borel set which satisfies

$$\langle f_i, E(R) f_i\rangle = 0$$

for all f_i, since any subset of a set of zero measure has zero measure, and $R \subset R_i$.

A denumerable union of a denumerable union is still a denumerable union, so by taking all the compact sets in each decomposition of Δ and relabeling them, we get a denumerable family of compact sets $\Delta_n \subset \Delta$ such that

$$E(\Delta) = E(\Delta - R) = E\left(\bigcup_{n=1}^{\infty} \Delta_n\right).$$

Now we have only to note that the characteristic functions of the increasing sequence of compact sets $C_N = \bigcup_{n=1}^{N} \Delta_n$ are uniformly bounded and converge pointwise to $\xi_{\Delta-R}$.

Thus, $E(\Delta) = \lim E(C_N)$, and the lemma follows from the property of closed operators that has been the theme of our discussion.

[34] Ref. 29, p. 129.