

Massive vector mesons and gauge theory

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2000 J. Phys. A: Math. Gen. 33 4317

(<http://iopscience.iop.org/0305-4470/33/23/309>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 171.66.16.118

The article was downloaded on 02/06/2010 at 08:11

Please note that [terms and conditions apply](#).

Massive vector mesons and gauge theory

Michael Dütsch[†] and Bert Schroer^{‡§}

[†] II Institut für Theoretische Physik, Universität Hamburg, Luruper Chaussee 149,
D-22761 Hamburg, Germany

[‡] Institut für Theoretische Physik der FU, Berlin, Germany

E-mail: duetsch@mail.desy.de and schroer@cbpfsu1.cat.cbpf.br

Received 11 February 2000

Abstract. We show that the requirements of renormalizability and physical consistency imposed on perturbative interactions of massive vector mesons fix the theory essentially uniquely. In particular, physical consistency demands the presence of at least one additional physical degree of freedom which was not part of the originally required physical particle content. In its simplest realization (probably the only one) these are scalar fields as envisaged by Higgs but in the present formulation without the ‘symmetry-breaking Higgs condensate’. The final result agrees precisely with the usual quantization of a classical gauge theory by means of the Higgs mechanism. However, the emphasis is shifted: instead of invoking the gauge principle (and the Higgs mechanism) on the local quantum field theory, the principles of local quantum physics restricted by the perturbative renormalizability demand ‘explains’ (via Bohr’s correspondence) the classical gauge principle as a selection principle among the many *a priori* (semi)classical possibilities of coupling vector fields among each other. Our method proves an old conjecture of Cornwall, Levin and Tiktopoulos stating that the renormalization and consistency requirements of spin-1 particles lead to the gauge theory structure (i.e. a kind of inverse of ’t Hooft’s famous renormalizability proof in quantized gauge theories) which was based on the on-shell unitarity of the S -matrix. Since all known methods of renormalized perturbation theory are off-shell, our proof is different and a bit more involved than the original arguments. We also speculate on a possible future ghost-free formulation which avoids ‘field coordinates’ altogether and is expected to reconcile the on-shell S -matrix point of view with the off-shell field theory structure.

1. Introduction

In the development of understanding of the renormalization aspects for spin-1 interactions, the classical concepts of gauge and fibre bundles have played a crucial role. Without the closely related Higgs mechanism it would be hard to imagine, how with the state of quantum field theory (QFT) at the end of the 1960s, the incorporation of the electroweak interaction into the framework of renormalizable field theory could have been achieved.

The main motivation for this paper arises from the remark that to the best of our knowledge the Higgs mechanism via Higgs condensates within the setting of gauge theories does not have a well understood intrinsic quantum physical content. Hence we choose an alternative procedure by which we will construct the same physical results for interacting massive vector mesons in a quite different way which does not rely on the above concepts. In our approach based on the well known real-time causal formulation of perturbative QFT, the renormalizability in the presence of self-interacting vector mesons is the basic input

§ Present address: CBPF, Rua Dr Xavier Sigaud, 150, 22290-180 Rio de Janeiro, RJ, Brazil.

requirement and the uniqueness and the gauge appearance of the polynomial interaction densities are consequences. In other words, we would phrase 't Hooft's famous statement that gauge structure implies renormalizability the other way around: maintaining perturbative renormalizability in the presence of vector mesons explains the gauge structure. Since local quantum theory is more fundamental than classical theory, this brings interacting vector mesons into harmony with Bohr's correspondence principle which is properly understood: it is the quantum theory which tells the classical one which possibility among many couplings involving vector fields and lower-spin fields it has to follow, namely the gauge-invariant one.

We would not have gained much, and a cynic might claim that we have replaced one mystery (the gauge recipe) with another one (the renormalization prescription), but fortunately, we have some slightly more tangible results to offer. Our method brings into the open the long-looming suspicion that the appearance of additional physical degrees of freedom (the alias Higgs particle but now without vacuum condensates) is a necessity, following from perturbative consistency up to second order (no claim outside of perturbation theory is made!)†. In addition it suggests strongly that the physics of zero-mass theory should be approached from massive vector mesons, the latter being conceptually (but not necessarily analytically) simpler. So as often happens in physics, the new aspect does not so much lie in the physical results as such, but rather in the novel way in which they are obtained and in the interpretation associated with this derivation.

To avoid misunderstandings we emphasize that our approach supports the principle of local gauge invariance because we obtain the same results (Lie-algebraic structure, the necessity of the scalar 'Higgs' field) by starting from quantum physical requirements instead of the conventional, semiclassical, differential geometric ideas. However, it is also important to emphasize that crucial theoretical observations have been made in the first 5 years of the existence of gauge theories (e.g. 't Hooft and Veltmann [44,45], Becchi *et al* [3]). After almost 25 years there seems to be a consensus (based on the present theoretical aspects of the standard model) that gauge theory, despite its impressive early successes, did not (yet) reach the level of maturity of other great theoretical developments in this century. Our present attempt at the physics of vector mesons outside (but not in contradiction to) gauge theory is very much motivated by this state of affairs. We consider this attempt as somewhat preliminary. Although it starts from the very physical Wigner particle concept, it shares with the gauge approach the necessity of extension by unphysical degrees of freedom (ghosts), which in this case originates from renormalizability. This can be formulated as a cohomological extension within the Wigner one-particle spaces. The advantage of the present approach is that the auxiliary role of the ghosts as 'renormalization catalysers'‡ becomes more manifest. The reader can find some speculative remarks on where to look for a ghost-free formulation in the final sections.

Since some of our concepts are not new, it may be helpful to remind the reader of some historical precedents. Ever since theories in which vector mesons or higher-spin particles became physically relevant in the late 1950s, there were two points of views to deal with such a problem: either to start from the Wigner particle picture and stay close

† Unfortunately, this is not a structural theorem, such as, for example, the theorem connecting spontaneous symmetry breaking with the appearance of 'Goldstone modes' or the theorem connecting the appearance of 'massive photons' with charge screening; rather it still remains in the setting of renormalized perturbation theory [43].

‡ The ghosts are objects outside local quantum physics which are neither present in the original physical degrees of freedom (which one wants to couple) nor in the final physical cohomological subspace. However, without their intermediate presence one does not know how to set up the renormalization machinery. In the perturbative framework we shall need fermionic and bosonic (scalar) ghost fields. The first correspond to the 'Faddeev-Popov ghosts' and the latter to the 'Higgs ghosts' (which are also called 'would-be Goldstone bosons') or 'Stückelberg fields'.

to particles and scattering theory, or to quantize classical gauge field theory (canonically or by functional integrals) and to make contact with (infra)particles at a later stage. In fact Sakurai, who introduced the term ‘Yang–Mills theory’ [29], and who wanted to use quantized classical field theory for the description of massive vector mesons in strong interactions, encountered difficulties in reconciling the two points of view. In most of his contributions he therefore took a phenomenological non-geometric point of view within a dispersion theoretical setting. This was particularly advisable since, as a result of some criticism of Pauli, the use of the Yang–Mills model for the description of massive vector mesons was cast into question.

The discovery of the electroweak theory and the renormalization of Yang–Mills theories [44] led to a drastic change of that picture. The first approach, which tried to make the gauge principle for vector mesons more palatable from a particle physics point of view, found some protagonists in the early days of gauge theories (Lewellyn-Smith [26], Bell [4], Cornwall *et al* [8]). It was argued that ‘tree unitarity’ of the S -matrix implies that one can choose the fields such that the model describes a spontaneously broken gauge theory. Cornwall, Levin and Tiktopoulos even gave a heuristic argument that renormalizability requires tree unitarity. On the other hand, the gauge or Yang–Mills point of view enjoyed general popularity and became the predominant one, irrespective of whether it could be derived from a more particle-dominated on-shell approach or not. Here we show that the tree unitarity viewpoint may be substituted by renormalizability within a suitable cohomologically extended Wigner particle setting. The presence of additional physical degrees of freedom (Higgs particles) follows from perturbative consistency, even though there is no concept of gauge symmetry breaking.

The main reason in favour of the gauge point of view was not only its success in describing problems of quantum matter coupled to external electromagnetic fields as well as its aesthetical mathematical appeal, but rather that it opened a path to a renormalizable theory of massive vector mesons, as was demonstrated by ’t Hooft and Veltman [44, 45]. This created, at least for some physicists, the impression that in addition to causality and spectral properties the general framework of QFT needs an additional principle for interacting massive vector mesons, namely the gauge principle. Here we propose a more intrinsic alternative framework which produces the same physical (gauge-invariant) results solely on the principles of local quantum physics (LQP, for an introduction and overview see [21]) and perturbative renormalizability. It uses a simplified free form of the BRS formalism [3] for a cohomological extension of the Wigner representation theory as a mathematically less formal operator substitute for the Faddeev–Popov formalism [19] in functional representations. This special cohomological extension (which is only available for *massive* vector mesons!) lowers the operator dimension of free massive vector fields from their physical value of two to the formal (classical) value of one without losing their point-like nature. Such a lowering is motivated by renormalizability within the framework of causal perturbation and the resulting theory is for all practical purposes equivalent to the gauge-invariant Lagrangian approach[†]. Our contention is that the renormalizability of interacting massive vector mesons alone uniquely fixes the theory, including its additional Higgs-like matter content. In other words, the gauge and Higgs aspect from which one starts the Lagrangian quantization in the standard formulation is reduced to the concept of ‘renormalizability’ with which we feel more comfortable because it occurs on the more fundamental level of local quantum physics where there is still a future chance for

[†] Although most of the free fields which correspond to the (m, s) Wigner representation cannot be obtained in the setting of free Lagrangians \mathcal{L}_0 , one can always rewrite a causal interaction W in terms of Lagrangian field coordinatizations. This is necessary for path integral formulations, but not for the causal formulation of perturbations.

a profound understanding. The remaining intrinsic feature of the Higgs mechanism is the Schwinger charge screening[†] caused by the Higgs-like matter degrees of freedom.

In more recent times a similar point of view has been taken up by Aste, Scharf and one of the authors (MD) [2]. We show, however, that their ‘free perturbative operator gauge invariance’, which is a pure quantum formulation of gauge invariance, is not needed as an *a priori* perturbative requirement for pure massive theories: our physical consistency is a weaker requirement but fixes the theory to the same extent. Some aspects of our point of view appeared in the work of Grigore [20]. It has also been mentioned in the setting of Schwinger’s work and of general quantum field theory by one of the authors (BS) [33].

Recently, there has been an approach to understand the local observable $*$ -algebras of non-Abelian gauge theories without addressing the particle content (which in the LQP framework is anyhow part of the separate more difficult representation theory associated with states on the algebra) [10]. This approach was specially aimed at the zero-mass theories with infrared problems, because the method does not require the existence of an adiabatic limit. In that case one cannot use the scattering theory of the physical particles and the BRS operators cannot be written as bilinear operators in free fields but they receive interacting contributions in every order. Therefore, one has to face the more difficult problem of a changing position of the physical cohomology space inside the extended space depending on the perturbative order. The zero-mass limit in our approach is conceptually and analytically complicated because it is expected to lead to charge liberation (the opposite of the Schwinger–Swieca charge screening) and the decoupling of the physical consistency (Higgs) degree of freedom. Those physical matter fields which will be charged in that limit cannot maintain their point-like localization, rather one expects that their localization cannot be better than semi-infinite space-like (Mandelstam string-like). In fact the tight relation between these various phenomena generates the hope that by controlling those off-shell infrared problems one may actually gain an insight into this (even in perturbation theory) notoriously difficult localization structure.

For reasons of brevity we exemplify our procedure in a particular class, namely self-interacting models of massive vector mesons. In section 2 we review the causal approach for massive vector mesons and its simplification as a result of the existence of a natural Fock reference space supplied by scattering theory. Our presentation uses a simplified quadratic BRS formalism. Physical consistency means that the S -matrix constructed on Fock space (which is the extended space, i.e. it includes ghost states) induces a well defined (unitary) operator on the physical ‘sub’-space.

As a justification for the introduction of ghost fields we then describe in section 3 the apparent clash between renormalizability and the operator dimension $\dim A = 2$ of the free vector meson operators in the usual causal setting and its resolution via cohomological ghost extension, which can be pre-empted on the level of the Wigner one-particle spaces.

In section 4 we formulate physical consistency in an alternative way which does not rely on scattering theory. The existence of the lowest-dimensional interpolating physical (i.e. BRS invariant) fields (those which we want to describe in our model as observable particles) not only fixes the form of these fields, but it also determines the interaction density (including the necessity of the alias Higgs particle) in agreement with section 2. In particular, we find

[†] Schwinger, in a little noticed paper and some more extensive published lecture notes [38], thought about a massive phase in QED through the mechanism of charge screening but without (Higgs) vacuum condensates. In order to make his non-perturbative ideas of massive vector mesons (the Maxwellian interaction is renormalizable) more palatable, he invented the two-dimensional Schwinger model. (In fact, in the Lowenstein–Swieca treatment of this model there is a chiral condensate (coming from the θ -degeneracy), but after the dust has settled, the physical content is described in terms of a massive free field only). Schwinger’s ideas were later elevated by Swieca into a theorem about vector mass generation via charge screening [43].

that the physical interpolating fields are composites in the elementary fields of the extended theory.

So the LSZ-type power series in terms of the *physical incoming Fock space operators* are modified (as we comment in section 5). This leads to a loss of the specific iterative law for the perturbative representation of the coefficient functions in terms of retarded products (for the fields). Only the general LSZ-like identification of coefficient functions of local fields involving multiple commutators of the local field with the incoming free field (generalized formfactors) remains valid [37]. The reason for this complication, which prevents the interchange of computations with the descent to the physical fields, is that the Wick basis used for writing the latter in terms of linear combinations of composites (including ghosts) is not a natural basis for the physical fields (i.e. the fields which commute with the BRS charge Q). Section 5 also contains some still speculative remarks on where one has to look for, if one wants to have a ghost-free formulation. In view of the fact that the previous sections have made clear that ghosts behave in some sense analogous to catalysers[†] in chemistry, this is not an academic problem but really goes to the root of understanding of renormalizability for higher spin where the standard causal approach breaks down. We are led to believe that such an approach must bypass the Bogoliubov–Shirkov transition operator $S(g)$ and be on-shell, i.e. directly deal with the on-shell S -matrix and multi-particle formfactors of physical fields.

We emphasize again that we consider the present formalism as transitory towards a completely ghost-free formulation of interactions involving higher-spin particles. Although it does not achieve this aim, it creates a critical Bohr–Heisenberg attitude towards the many unobservable and formal aspects of the standard gauge formalism.

2. Consistent perturbative construction of the S -matrix for massive vector fields

The aim of this section is to construct the Stückelberg–Bogoliubov–Shirkov transition functional $S(g)$, which is the generating functional for the time-ordered products of Wick polynomials. As for most functional quantities this object is not directly observable, but it gives rise to fundamental physical observables such as the S -matrix, formfactors and correlation functions of observable fields. The notation should not be misread as the S -matrix by which we always mean the scattering operator computed with the LSZ or Haag–Ruelle scattering theory. Our model is that of self-interacting massive vector mesons. Our procedure is related to that of Scharf, Aste and the first author [2, 30], but similarly to a previous discussion by the other author [34] and to Grigore [20] as well as older articles such as Lewellyn-Smith [26], we simply rely on physical consistency within the framework of local quantum physics and do not require such technical tools as ‘operator gauge invariance’ although they tend to simplify calculations.

Using the Stückelberg–Bogoliubov–Shirkov–Epstein–Glaser method [5, 17] we make the following perturbative ansatz for $S(g)$:

$$S(g) = \mathbf{1} + \sum_{n=1}^{\infty} \frac{i^n}{n!} \int d^4x_1 \dots d^4x_n T_{j_1 \dots j_n}(x_1, \dots, x_n) g_{j_1}(x_1) \dots g_{j_n}(x_n) \quad \mathbf{g} = (g_j)_{j=0}^G \quad (1)$$

$g_j \in \mathcal{S}(\mathbb{R}^4)$, which is a formal power series in \mathbf{g} . The unknown $T_{j_1 \dots j_n}$ are operator-valued

[†] One introduces them in order to lower the operator dimension of the interaction density W to the renormalizable value of four. This is achieved by decreasing the dimension of the vector potential from two to one as a result of an unphysical cohomological extension of the Wigner one-particle representation theory. After the calculations have been done in the extended setting, one eliminates the ghosts by cohomological descend (BRS invariance).

distributions[†]. They are constructed inductively by means of the following requirements (strongly influenced by the general Wightman framework [42]):

(A) *Specification of the model in first order.* The first-order expressions are the main input of the construction. They specify the model and must be local: $[T_j(x), T_k(y)] = 0$ for $(x - y)^2 < 0$. We assume that $T_0(x) \equiv W_0(x) \equiv W(x) = W(x)^*$ is the physically relevant interaction density in Fock space (i.e. a Poincaré covariant scalar composite described by a Wick polynomial). $T_j \equiv W_j, j = 1, \dots, G$ are auxiliary interactions. The interaction W_j is switched by the spacetime-dependent coupling ‘constant’ $g_j \in \mathcal{S}(\mathbb{R}^4)$. The physically relevant S -matrix is obtained in the adiabatic limit: $g_0 \rightarrow \text{constant}, g_j \rightarrow 0, j = 1, \dots, G$.

(B) *Permutation symmetry.* Due to the ansatz (1) we may require permutation symmetry

$$T_{j_1 \dots j_n}(x_{\pi_1}, \dots, x_{\pi_n}) = T_{j_1 \dots j_n}(x_1, \dots, x_n) \quad \forall \pi \in \mathcal{S}_n. \quad (2)$$

(C) *Causality.*

$$S(\mathbf{g}^{(1)} + \mathbf{g}^{(2)}) = S(\mathbf{g}^{(1)})S(\mathbf{g}^{(2)}) \quad \text{if} \quad \cup_j \text{supp } g_j^{(1)} \cap (\cup_j \text{supp } g_j^{(2)} + \bar{V}_-) = \emptyset \quad (3)$$

where \bar{V}_\mp denotes the closed backward/forward lightcone. This requirement is equivalent to (see the appendix of [16])[‡]

$$T_{j_1 \dots j_n}(x_1, \dots, x_n) = T_{j_1 \dots j_l}(x_1, \dots, x_l) T_{j_{l+1} \dots j_n}(x_{l+1}, \dots, x_n) \quad (4)$$

if

$$\{x_1, \dots, x_l\} \cap (\{x_{l+1}, \dots, x_n\} + \bar{V}_-) = \emptyset.$$

This means that $T_{j_1 \dots j_n}(x_1, \dots, x_n)$ is a (well defined) time-ordered product of $W_{j_1}(x_1), \dots, W_{j_n}(x_n)$. Hence we use the notation

$$T_{j_1 \dots j_n}(x_1, \dots, x_n) = T(W_{j_1}(x_1) \dots W_{j_n}(x_n)). \quad (5)$$

Due to the induction with respect to the order n , the $T_{j_1 \dots j_n}$ are uniquely fixed by causality up to the total diagonal $\Delta_n = \{(x_1, \dots, x_n) | x_1 = x_2 = \dots = x_n\}$. The extension of the $T_{j_1 \dots j_n}$ to the total diagonal is non-unique. It is restricted by the following normalization conditions:

(D) *Poincaré covariance.*

(E) *Unitarity.* $S(\mathbf{g})^{-1} = S(\mathbf{g})^*$ for $\mathbf{g} = (g_0, 0, \dots, 0)$, g_0 real valued.

(F) *Scaling degree.* The degree of the singularity at the diagonal, measured in terms of Steinmann’ scaling degree [6, 40]§, may not be increased by the extension. This ensures *renormalizability by power counting* if the scaling degree (or ‘mass dimension’) of all W_j is ≤ 4 . This degree is a tool which is related to Weinberg’s power counting.

Additional normalization conditions must be imposed, if one wants to maintain further symmetries or relations||, e.g. discrete symmetries (P, C, T), ‘operator gauge invariance’ (29) and (30) or the field equations of the interacting fields ((N4) in [10]), which can be obtained from the functional $S(\mathbf{g})$ (1) by Bogoliubov’s formula (see section 4).

[†] For questions concerning domains we refer to [17]. For all operators which appear in sections 2 and 3 there exists a common dense invariant domain \mathcal{D} and we restrict all operators to this subspace.

[‡] The non-trivial part of this equivalence is that in the n th-order expression of (3) only special test functions appear, whereas (4) holds on $\mathcal{S}(\mathbb{R}^{4n})$.

§ We adopt here the notion of a ‘scaling degree’ to operator-valued distributions by using the strong operator topology. Note that the scaling degree of a Wick monomial agrees with its mass dimension.

|| We consider symmetries and relations which are satisfied away from the total diagonal, due to the causal factorization (5) and the inductive assumption. Poincaré covariance (D) and unitarity (E) are of this type.

The existence of the adiabatic limit restricts the extension additionally: for *pure massive* theories Epstein and Glaser [18] proved that, with correct mass and wavefunction (re)normalization the adiabatic limit of the functional $S(\mathbf{g})$ exists in the strong operator sense and it is this limit which we call ‘ S -matrix’. More precisely setting $\mathbf{g}_\epsilon(x) := (g_0(\epsilon x), 0, \dots, 0)$ the limit

$$S_n \psi \equiv \lim_{\epsilon \rightarrow 0} S_n(\mathbf{g}_\epsilon) \psi \tag{6}$$

exists $\forall \psi \in \mathcal{D}$, where $g_0 \in \mathcal{S}(\mathbb{R}^4)$, $g := g_0(0) > 0$ is the coupling constant and $S_n(\mathbf{g})$ (S_n , respectively) denotes the n th order of the functional $S(\mathbf{g})$ (S -matrix, respectively). It follows that the S -matrix is unitary as an operator-valued formal power series in Fock space [18]: $S = \sum_n S_n$, $S_n \sim g^n$, $S^* S = \mathbf{1} = S S^*$ on \mathcal{D} .

Due to this fact we solely consider models in which all fields are massive. In order that it makes physically sense to consider the S -matrix, we assume that there are no unstable physical particles such as, for example, the W - and Z -bosons in the electroweak theory.

In *vector meson theories* (which are ‘gauge theories’) the crucial problem is the elimination of the unphysical degrees of freedom. In the S -matrix framework this problem turns into the requirement that the S -matrix induces a well defined unitary operator on the space of physical states (this is discussed in detail below). We will see that this condition is very restrictive: it determines the possible interactions to a large extent.

Let us first consider the free incoming fields. We quantize the free vector fields $(A_a^\mu)_{a=1, \dots, M}$ by

$$(\square + m_a^2) A_a^\mu = 0 \quad [A_a^\mu(x), A_b^\nu(y)] = i g^{\mu\nu} \delta_{ab} \Delta_{m_a}(x - y) \quad A_a^{\mu*} = A_a^\mu \tag{7}$$

(where Δ_m is the Pauli–Jordan distribution to the mass m), which corresponds to the ‘Feynman gauge’. The representation of this $*$ -algebra requires an indefinite inner product space. We, therefore, work in a Krein Fock space \mathcal{F} . We denote the scalar product by (\cdot, \cdot) and A^+ is the adjoint of A with respect to (\cdot, \cdot) . Let J be the Krein operator: $J^2 = 1$, $J^+ = J$. Then the indefinite inner product $\langle \cdot, \cdot \rangle$ is defined by

$$\langle a, b \rangle \equiv (a, Jb) \quad a, b \in \mathcal{F} \tag{8}$$

and $*$ denotes the adjoint with respect to $\langle \cdot, \cdot \rangle$:

$$O^* \equiv J O^+ J \quad \langle Oa, b \rangle = \langle a, O^* b \rangle. \tag{9}$$

Let Q be an (unbounded) $*$ -symmetrical nilpotent operator in \mathcal{F}

$$Q = Q^* \quad (\text{on the dense invariant domain } \mathcal{D}) \quad Q^2 = 0. \tag{10}$$

By means of $Q^2 = 0$ one easily finds that \mathcal{D} is the direct sum of three, pairwise orthogonal (with respect to (\cdot, \cdot)) subspaces [22, 23]

$$\mathcal{D} = \text{ran } Q \oplus (\text{ker } Q \cap \text{ker } Q^+) \oplus \text{ran } Q^+ \tag{11}$$

$$\text{ker } Q = \text{ran } Q \oplus (\text{ker } Q \cap \text{ker } Q^+) \quad \text{ker } Q^+ = \text{ran } Q^+ \oplus (\text{ker } Q \cap \text{ker } Q^+). \tag{12}$$

In addition, we assume

$$J|_{\text{ker } Q \cap \text{ker } Q^+} = \mathbf{1} \quad (\text{positivity assumption}). \tag{13}$$

Then the $\langle \cdot, \cdot \rangle$ -product is positive definite on

$$\mathcal{H}_{\text{phys}} \equiv \text{ker } Q \cap \text{ker } Q^+ \tag{14}$$

and $\mathcal{H}_{\text{phys}}$ is interpreted as the *physical subspace* of \mathcal{F} . We denote the projectors on $\text{ran } Q$ ($\mathcal{H}_{\text{phys}}$, $\text{ran } Q^+$, respectively) by P_- (P_0, P_+ , respectively)

$$\mathbf{1} = P_- + P_0 + P_+ \quad (\text{on } \mathcal{D}). \tag{15}$$

Note $Q = P_- Q P_+$. The positivity (13) and $Q = Q^*$ imply [23]

$$P_0 = P_0 J P_0 \quad J = P_0 J P_0 + P_- J P_+ + P_+ J P_- \tag{16}$$

and hence

$$P_0^* = P_0 \quad P_-^* = P_+. \tag{17}$$

Let $S : \mathcal{D} \rightarrow \mathcal{D}$ be the (strong) adiabatic limit (6) of $S(g)$. We define

$$S_{ab} \equiv P_a S P_b \quad a, b \in \{-, 0, +\} \tag{18}$$

and obtain

$$P_a S^* P_b = (S_{(-b)(-a)})^*. \tag{19}$$

For pedagogical reasons we introduce the matrix notation according to the decomposition (11)

$$J = \begin{pmatrix} 0 & 0 & P_- J P_+ \\ 0 & P_0 & 0 \\ P_+ J P_- & 0 & 0 \end{pmatrix} \tag{20}$$

$$S = \begin{pmatrix} S_{--} & S_{-0} & S_{-+} \\ S_{0-} & S_{00} & S_{0+} \\ S_{+-} & S_{+0} & S_{++} \end{pmatrix} \tag{21}$$

and

$$S^* = \begin{pmatrix} (S_{++})^* & (S_{0+})^* & (S_{-+})^* \\ (S_{+0})^* & (S_{00})^* & (S_{-0})^* \\ (S_{+-})^* & (S_{0-})^* & (S_{--})^* \end{pmatrix}$$

by means of (19).

An alternative definition of the physical states is

$$\mathcal{H}'_{\text{phys}} \equiv \frac{\ker Q}{\text{ran } Q} \tag{22}$$

where the scalar product in $\mathcal{H}'_{\text{phys}}$ is defined such that the map

$$\mathcal{H}_{\text{phys}} \rightarrow \mathcal{H}'_{\text{phys}} : \phi \rightarrow [\phi] \tag{23}$$

is a pre-Hilbert space isomorphism. ($[\phi]$ denotes the equivalence class of ϕ .) The first definition (i.e. $\mathcal{H}_{\text{phys}}$) has the advantage that the set of physical states is a *subspace* of the Krein Fock space \mathcal{F} , which has a clear particle interpretation. However, $\mathcal{H}_{\text{phys}}$ is not Lorentz invariant (in contrast to $\mathcal{H}'_{\text{phys}}$). The change of its position inside the total space under Lorentz transformations is of course a result of the lack of Lorentz invariance of J .

To describe *free* spin-1 fields the introduction of a BRST formalism (as we just have done) is not necessary. However, our main topic is to describe the interacting theory in the adiabatic limit. So the incoming and outgoing fields are asymptotically free. Hence, being equipped with a characterization of the asymptotical physical states in terms of a BRST formalism for free fields, our scattering point of view is much simplified.

Let $S^* S = \mathbf{1} = S S^*$ (on \mathcal{D}). We now discuss two different formulations of the physical consistency of the S -matrix.

(i) A consistent S -matrix theory requires

$$P_0 S^* P_0 S P_0 = P_0 = P_0 S P_0 S^* P_0 \iff (S_{00})^* = (S_{00})^{-1} \quad \text{on } \mathcal{H}_{\text{phys}}.$$

(ii) In the framework of the definition (22) of the physical states consistency means that S and $S^{-1} = S^*$ induce well defined operators on the factor space $\mathcal{H}'_{\text{phys}}$ by the definition

$$[\mathcal{O}][\phi] \equiv [\mathcal{O}\phi] \quad \mathcal{O} = S, S^*. \quad (24)$$

This holds true iff

$$\mathcal{O} \ker Q \subset \ker Q \quad \wedge \quad \mathcal{O} \text{ran } Q \subset \text{ran } Q \quad \mathcal{O} = S, S^*. \quad (25)$$

Due to $[\mathcal{O}]^* = [\mathcal{O}^*]$ the physical S -matrix $[S]$ is then unitary.

The following lemma states that (i) is a truly weaker condition than (ii) and gives equivalent formulations of (ii).

Lemma 1. *The following statements (a)–(g) are equivalent and they imply (h). However, (h) does not imply the other statements if $Q \neq 0$.*

- (a) $SS^* = \mathbf{1} = S^*S$ and $S \ker Q \subset \ker Q$ (i.e. $S_{+-} = 0 = S_{+0}$).
- (b) $SS^* = \mathbf{1} = S^*S$ and $[Q, S]|_{\ker Q} = 0$.
- (c) $SS^* = \mathbf{1} = S^*S$ and $S \text{ran } Q \subset \text{ran } Q$ (i.e. $S_{+-} = 0 = S_{0-}$).
- (d) The matrix S (21) has the form

$$S = \begin{pmatrix} S_{--} & S_{-0} & S_{-+} \\ 0 & S_{00} & -S_{00}(S_{-0})^*(S_{--})^{*-1} \\ 0 & 0 & (S_{--})^{*-1} \end{pmatrix} \quad (26)$$

where S_{--} and S_{00} are invertible (on $\text{ran } Q$, $\mathcal{H}_{\text{phys}}$, respectively) and S_{00} , S_{--} , S_{-0} , S_{-+} satisfy

$$(S_{00})^* = (S_{00})^{-1} \quad S_{-+}(S_{--})^* + S_{-0}(S_{-0})^* + S_{--}(S_{-+})^* = 0.$$

- (e) $SS^* = \mathbf{1} = S^*S$ and $S^* \ker Q \subset \ker Q$ (i.e. $(S_{+-})^* = 0 = (S_{0-})^*$).
- (f) $SS^* = \mathbf{1} = S^*S$ and $[Q, S^*]|_{\ker Q} = 0$.
- (g) $SS^* = \mathbf{1} = S^*S$ and $S^* \text{ran } Q \subset \text{ran } Q$ (i.e. $(S_{+-})^* = 0 = (S_{+0})^*$).
- (h) $SS^* = \mathbf{1} = S^*S$ and $(S_{00})^* S_{00} = P_0 = S_{00}(S_{00})^*$.

Proof. (b) \Leftrightarrow (a) \Leftrightarrow (g) and (f) \Leftrightarrow (e) \Leftrightarrow (c) hold true trivially.

(a) \Leftrightarrow (e): (a) implies $S_k \ker Q \subset \ker Q$ for each order S_k of S . Therefore,

$$(S^*)_n = (S^{-1})_n = \sum_{r=1}^n (-1)^r \sum_{\substack{n_1, \dots, n_r \geq 1 \\ n_1 + \dots + n_r = n}} S_{n_1} \dots S_{n_r}$$

maps $\ker Q$ in $\ker Q$ and, hence, this also holds true for $S^* = \sum_n (S^*)_n$. (e) \Rightarrow (a) follows analogously.

(a), (c) \Leftrightarrow (d): by a straightforward calculation one verifies that the equations

$$S_{+-} = 0, S_{+0} = 0, S_{0-} = 0 \quad \sum_c (S_{(-c)(-a)})^* S_{cb} = \delta_{ab} P_a = \sum_c S_{ac} (S_{(-b)(-c)})^* \quad (27)$$

are equivalent to (d).

(a), (c) \Rightarrow (h): choosing $a = 0 = b$ in (27) we obtain $(S_{00})^* S_{00} = P_0 = S_{00}(S_{00})^*$.

To show that (h) does not imply the other statements for $Q \neq 0$ ($\Leftrightarrow \text{ran } Q \neq 0 \Leftrightarrow \text{ran } Q^+ \neq 0$), we give two examples for S which satisfy (h) but not (d).

- The condition (h) is invariant under an exchange of Q and Q^+ . Therefore, there exists a solution S of (h) which maps $\ker Q^+$ in $\ker Q^+$ (and/or $\text{ran } Q^+$ in $\text{ran } Q^+$), i.e. S is a lower triangular matrix.
- The following S -matrix fulfils (h) and $S_{+-} \neq 0 \neq S_{-+}$ (if $cd \neq 0$):

$$S = \begin{pmatrix} ae^{i\alpha} P_- & 0 & ice^{i\alpha} P_- J P_+ \\ 0 & e^{i\phi} P_0 & 0 \\ ide^{i\alpha} P_+ J P_- & 0 & be^{i\alpha} P_+ \end{pmatrix} \quad (28)$$

with $a, b, c, d \in \mathbb{R}$, $\alpha, \phi \in \mathbb{R}$ and $ab + cd = 1$. □

In [2, 30] physical consistency of the S -matrix is satisfied by requiring a perturbative condition which implies $[Q, S] = 0$, namely the ‘free perturbative operator gauge invariance’[†] (or ‘ Q -divergence condition’): let $W \equiv T_0$ be the interaction Lagrangian. Then there exists a Wick polynomial W_1^ν with

$$[Q, W(x)] = i\partial_\nu W_1^\nu(x) \quad (29)$$

and the time-ordered products of W and W_1^ν fulfil

$$[Q, T(W(x_1) \dots W(x_n))] = i \sum_{l=1}^n \partial_\mu^{x_l} T(W(x_1) \dots W_1^\mu(x_l) \dots W(x_n)). \quad (30)$$

Let us assume that (30) holds true to all orders $\leq (n - 1)$. Due to the causal factorization (4) the requirement (30) is then satisfied away from the total diagonal, i.e. on $S(\mathbb{R}^{4n} \setminus \Delta_n)$. Hence, (30) is an additional normalization condition for $T(W \dots W)$ and $T(W \dots W_1^\mu \dots W)$. However, it is a highly non-trivial task to prove that there exists an extension to the diagonal which satisfies (30) and the other normalization conditions (D)–(F).

In contrast to $[Q, S] = 0$ or $[Q, S]|_{\ker Q} = 0$, the Q -divergence condition (29) and (30) also makes sense in models in which the adiabatic limit does not exist, e.g. for massless self-interacting vector fields (‘non-Abelian gauge theories’). For massless $SU(N)$ -Yang–Mills theories it has been proved that the Q -divergence condition (29) and (30) (more precisely the corresponding C -number identities which imply (29) and (30)) can be satisfied to all orders [12, 13], and that these C -number identities imply the usual Slavnov–Taylor identities [14]. In addition, equations (29) and (30) determine to a large extent the possible structure of the model (see below). We emphasize that this is a *pure quantum formulation of gauge invariance*, without reference to classical physics.

In the present case of pure massive vector mesons (or more generally in ‘gauge’ theories in which the strong adiabatic limit (6) of the transition functional $S(g)$ exists) we proceed in an alternative way. *We do not require the Q -divergence condition (29) and (30) as a new physical principle, instead we simply require physical consistency of the S -matrix*, which means

$$[Q, S]|_{\ker Q} = 0 \quad (31)$$

[†] The reason for this name stems from the fact that Q is the generator of the BRST transformation of the free incoming fields. For the present purpose we will simply refer to it as the ‘ Q -divergence condition’ for time-ordered products, since it has nothing to do with the classical notion of gauge invariance in the differential geometric setting of fibre bundles, but secures that the true S -matrix is physical. Roughly speaking it is the off-shell version of physicality of the on-shell S -matrix.

by lemma 1. This is a weaker condition than (29) and (30). However, we will see that it determines the theory to the same extent. Our procedure is similar to Grigore [20]†.

First we construct the operator Q (10) which defines the physical states. For massless vector meson theories the procedure is well known [11, 25]. The nilpotency of Q gives a reason to introduce an anticommuting pair of ghost fields u_a, \tilde{u}_a for each vector field A_a (fermionic ghosts). Then we define‡

$$Q \equiv \int d^3x \sum_a \partial_\nu A_a^\nu(x) \vec{\partial}_0 \tilde{u}_a(x). \quad (32)$$

Turning to massive vector fields A_a (7) we give the fermionic ghost fields the same masses m_a (otherwise the current $\sum_a \partial_\nu A_a^\nu \vec{\partial}_\mu u_a$ or $\sum_a (\partial_\nu A_a^\nu + m_a \phi_a) \vec{\partial}_\mu u_a$ (see below) would not be conserved)

$$\begin{aligned} (\square + m_a^2)u_a &= 0 & (\square + m_a^2)\tilde{u}_a &= 0 \\ \{u_a(x), u_b(y)\} &= 0 & \{\tilde{u}_a(x), \tilde{u}_b(y)\} &= 0 \\ \{u_a(x), \tilde{u}_b(y)\} &= -i\delta_{ab}\Delta_{m_a}(x-y) & u_a^* &= u_a & \tilde{u}_a^* &= -\tilde{u}_a. \end{aligned} \quad (33)$$

If we insert the massive A_a and u_a fields into the formula (32) for Q the nilpotency is lost

$$2Q_{\text{naive}}^2 = \{Q_{\text{naive}}, Q_{\text{naive}}\} = \int d^3x \int d^3y [\partial_\mu A_a^\mu(x), \partial_\nu A_b^\nu(y)] \vec{\partial}_{x_0} \vec{\partial}_{y_0} u_a(x) u_b(y) \neq 0. \quad (34)$$

To restore the nilpotency we proceed as follows [23]: to each vector field A_a we consider a scalar field ϕ_a with the same mass m_a

$$(\square + m_a^2)\phi_a = 0 \quad [\phi_a(x), \phi_b(y)] = -i\delta_{ab}\Delta_{m_a}(x-y) \quad \phi_a^* = \phi_a. \quad (35)$$

We also call the fields ϕ_a ‘ghost fields’ (bosonic ghosts) because they are unphysical (see below). Then, due to $[\partial_\mu A_a^\mu + m_a \phi_a, \partial_\nu A_b^\nu + m_b \phi_b] = 0$, the charge

$$Q \equiv \int d^3x \sum_a (\partial_\nu A_a^\nu(x) + m_a \phi_a(x)) \vec{\partial}_0 u_a(x) \quad (36)$$

is nilpotent and symmetrical. Later we shall see that an additional scalar field H with arbitrary mass $m_H \geq 0$ is needed§

$$(\square + m_H^2)H = 0 \quad [H(x), H(y)] = -i\Delta_{m_H}(x-y) \quad H^* = H. \quad (37)$$

The notation H is reminiscent of the *Higgs* field, but one with no ‘vacuum condensate’ (leaving aside the academic point of whether Higgs idea would allow for our terminology): $\langle 0|H(x)|0\rangle = 0$, where $|0\rangle \in \mathcal{F}$ is the vacuum of the free fields. For the representation of A_a (7), u_a, \tilde{u}_a (33), ϕ_a (35) and H (37) in a Krein Fock space \mathcal{F} and especially the definition of J

† However, there are two conceptual shortcomings in these papers. It is overlooked that most of the trilinear terms in W vanish in the adiabatic limit due to energy–momentum conservation and, hence, to first order the condition (31) yields no information about the trilinear terms in W (see appendix A).

By using the terminology introduced below the second shortcoming can be described as follows: the Higgs field(s) is/are treated as scalar partner(s) (with arbitrary mass $m_H \geq 0$) of the massless vector field(s), which does/do not appear in Q and, hence, is/are physical. By chance this works for the electroweak theory (there is one massless vector field and one Higgs field is needed). However, for example, in the present case of pure massive vector mesons, there would be no Higgs field and such a model is physically inconsistent (31) to second order. This inefficiency does not appear in [2, 30].

‡ The convergence of this integral (and also of the corresponding expression (36) in the massive theory) can be shown by using a method of Requardt [10, 28].

§ For simplicity we only consider models in which *one* additional scalar field suffices for a consistent construction of the S -matrix.

we refer to [23]†. Note that the ghost fields u_a, \tilde{u}_a (33) and ϕ_a (35) are also asymptotic fields, but that asymptotic states containing ghost excitations do not belong to the physical Hilbert space (14). From the commutation relations

$$\begin{aligned} [Q, A_a^\mu] &= i\partial^\mu u_a & [Q, \phi_a] &= im_a u_a & \{Q, u_a\} &= 0 \\ [Q, \tilde{u}_a] &= -i(\partial_\mu A_a^\mu + m_a \phi_a) & [Q, H] &= 0 \end{aligned} \tag{38}$$

we conclude that $\mathcal{H}_{\text{phys}}$ (14) is the linear span of the set of states $B_1 \dots B_l |0\rangle, l \in \mathbb{N}_0$, where B_1, \dots, B_l are transversal vector meson fields (three polarizations) or H -fields. Using this explicit result the positivity assumption (13) can be verified [23]. We emphasize that H is physical, in contrast to the (scalar) bosonic ghost fields ϕ_a .

We are now looking for the possible interactions $W \equiv W_0$ which satisfy the following requirements:

- (a) W is a Wick polynomial in the free incoming fields. Each monomial in W has at least three factors;
- (b) W is invariant with respect to Poincaré transformations;
- (c) the number of u -fields agrees with the number of \tilde{u} -fields in each monomial of W (i.e. the ‘ghost number’ is zero);
- (d) the scaling degree (or mass dimension) of W is ≤ 4 (this is necessary for renormalizability by power counting);
- (e) $W = W^*$ (which yields $S^* = S^{-1}$ if the time-ordered products are suitably normalized);
- (f) physical consistency (31).

First we point out that the requirements (a)–(f) do not fix the interaction uniquely. Even if we replace in (f) the physical consistency (31) by the Q -divergence condition (29) and (30), which is a stronger requirement, the following non-uniqueness is known: considering solely the Q -divergence condition to first order (29), there obviously remains the freedom to add divergence and coboundary couplings to W ,

$$W^{(\beta, \gamma)} = W + \sum_l \beta_l \partial_\nu D_l^\nu + \sum_j \gamma_j \{Q, K_j\} \quad \beta_l, \gamma_j \in \mathbb{R} \tag{39}$$

where D_l^ν, K_j are restricted by (a)–(e). Taking additionally the Q -divergence condition to orders $n \geq 2$ (30) into account, it seems that this freedom can be maintained, due to the following result. It is shown in [15] that the Q -divergence condition (30) can be satisfied to all orders (by choosing suitable normalizations) for any (β, γ) , if the ‘generalized (free perturbative operator) gauge invariance’‡ holds true for $(\beta, \gamma) = (\mathbf{0}, \mathbf{0})$. Moreover, under this

† In [23] it was not realized that without the Higgs field H (or an ingenious substitute) the Q -divergence condition to second order is violated. We represent the H -field similarly to the other scalar fields ϕ_a with $J = \mathbf{1}$ in the H -Fock space.

‡ The ‘generalized (free perturbative operator) gauge invariance’ is the following statement. To the interaction density $W \equiv W_0$ there exist Wick polynomials W_1^ν and $W_2^{\mu\nu}$ with

$$[Q, W_0] = i\partial_\nu W_1^\nu \quad \{Q, W_1^\nu\} = i\partial_\mu W_2^{\mu\nu} \quad [Q, W_2^{\mu\nu}] = 0$$

and the time-ordered products of W_0, W_1^ν and $W_2^{\mu\nu}$ fulfil

$$[Q, T(W_{j_1}(x_1) \dots W_{j_n}(x_n))]_{\mp} = i \sum_{l=1}^n \partial^{x_l} T(W_{j_1}(x_1) \dots W_{j_{l+1}}(x_l) \dots W_{j_n}(x_n)) \quad j_1, \dots, j_n \in \{0, 1, 2\}$$

where we have the anticommutator on the left-hand side iff $(j_1 + \dots + j_n)$ is odd, and where $T(\dots W_3(x_l) \dots) \equiv 0$ by definition. Similarly to (30), the second equation is a normalization condition on the time-ordered products. A proof that it can be satisfied to all orders is still missing in any non-Abelian model. Nevertheless, we strongly presume that this statement holds true for all models which are BRST invariant, especially for the model studied below of three massive, self-interacting vector fields.

assumption, one can prove

$$P_0 T(W^{(\beta,\gamma)}(x_1) \dots W^{(\beta,\gamma)}(x_n)) P_0 = P_0 T(W^{(0,0)}(x_1) \dots W^{(0,0)}(x_n)) P_0 + \text{divergences} \quad (40)$$

(see [15], ‘divergences’ means ‘divergences of local operators’) and hence the physical S -matrix S_{00} is independent from (β, γ) (because the divergences vanish in the adiabatic limit in pure massive theories).

Now we make the most general ansatz for W (up to divergence and coboundary terms) which satisfies (a)–(e)

$$\begin{aligned} W = & f_{abc} : A_{a\mu} A_{b\nu} \partial^\nu A_c^\mu : + f_{abc}^1 : u_a \partial^\mu \tilde{u}_b A_{c\mu} : \\ & + d_{abc} (: A_a^\mu \phi_b \partial_\mu \phi_c : - : A_a^\mu \partial_\mu \phi_b \phi_c :) + e_{abc} : A_a^\mu A_{b\mu} \phi_c : \\ & + h_{abc} : \tilde{u}_a u_b \phi_c : + j_{abc} : \phi_a \phi_b \phi_c : + k_{ab} (: H A_a^\mu \partial_\mu \phi_b : - : \partial_\mu H A_a^\mu \phi_b :) \\ & + l_{ab} : A_a^\mu A_{b\mu} H : + p_{ab} : H \tilde{u}_a u_b : + q_{ab} : H \phi_a \phi_b : + r_a : H^2 \phi_a : + s : H^3 : \\ & + (\text{quadrilinear terms, e.g. } \sim : AAAA : , : AAu\tilde{u} : , : AA\phi\phi : , : H^4 :) \quad (41) \end{aligned}$$

where $f_{abc}, f_{abc}^1, d_{abc}, e_{abc}, h_{abc}, j_{abc}, k_{ab}, l_{ab}, p_{ab}, q_{ab}, r_a, s \in \mathbb{R}$ are arbitrary constants (which are introduced without any knowledge about a gauge group). One can show that physical consistency (31) requires that f_{abc} be totally antisymmetric (i.e. this part of the result (46) below is independent of the restriction to *three* vector fields which is made in the following). Hence the simplest non-trivial model of self-interacting massive ($m_a > 0 \forall a$) vector fields is that of *three* fields: $a = 1, 2, 3^\dagger$. From now on we specialize to this case. In particular, a single massive spin-1 field (corresponding to the $U(1)$ -case) cannot be self-interacting.

Our aim is to determine the parameters in W (41) and the tree-normalization terms (see (49) below) by the consistency condition (31). The hope is that one can conclude from

$$0 = [Q, S_n]_{\ker Q} = \frac{i^n}{n!} \int dx_1 \dots dx_n [Q, T(W(x_1) \dots W(x_n))]_{\ker Q} \quad (42)$$

that $[Q, T(W(x_1) \dots W(x_n))]$ must be a sum of divergences of local operators and that then one can follow the calculations in [2].

However, there is a difficulty at first order, which is explained in detail in appendix A. The adiabatic limit of the trilinear terms in W vanishes (except possibly the H -couplings) due to energy–momentum conservation. This is just the fact that for stable particles there are no S -matrix elements with three particle legs; the lowest tree contributions involve \geq four legs. Hence, we get no information about these terms from $[Q, S_1]_{\ker Q} = 0$. For the quadrilinear terms $W^{(4)}$ in W the procedure works (see also appendix A): we obtain that $[Q, W^{(4)}]$ must be the divergence of a Wick polynomial. In [20] it is shown that this implies $W^{(4)} = 0$ except for a term $\sim : H^4 :$. However, the latter term will be excluded by consistency (31) to second order. (The well known $: H^4 :$ -coupling is of second order in g , it appears in the framework of causal perturbation theory as a tree-normalization term of $T(W(x_1)W(x_2))$, see below.)

To determine the parameters in the trilinear terms $W^{(3)}$ of W (41) we compute the tree diagrams in $T(W^{(3)}(x_1)W^{(3)}(x_2))$ (which depend on these parameters) and require

$$\int d^4x_1 d^4x_2 [Q, T(W^{(3)}(x_1)W^{(3)}(x_2))|_{\text{tree}}]_{\ker Q} = 0. \quad (43)$$

Instead of the long explicit calculation we give a *heuristic* description.

\dagger The resulting model is usually called ‘ $SU(2)$ Higgs–Kibble model’. It is obtained from the electroweak theory, which is studied in detail in [2], by setting the Weinberg angle $\Theta_W = 0$ and omitting the photon field, which decouples in this case.

- For the terms with $x_1 \neq x_2$ the T -product factorizes, e.g. for $x_1 \notin x_2 + \bar{V}^-$ we have

$$[Q, T(W^{(3)}(x_1)W^{(3)}(x_2))|_{\text{tree}}] = [Q, W^{(3)}(x_1)]W^{(3)}(x_2)|_{\text{tree}} + W^{(3)}(x_1)[Q, W^{(3)}(x_2)]|_{\text{tree}}. \quad (44)$$

This makes it plausible that the cancellation of these terms[†] yields the same restrictions as the Q -divergence condition to first order (29): $[Q, W^{(3)}] = i\partial_\nu W_1^{(3)\nu}$ for some Wick monomial $W_1^{(3)\nu}$ [‡]. In this way we obtain

$$\begin{aligned} f_{abc} &= -f_{abc}^1 & \text{and} & & f_{abc} & \text{is totally antisymmetric} \\ d_{abc} &= f_{abc} \frac{m_b^2 + m_c^2 - m_a^2}{4m_b m_c} & e_{abc} &= f_{abc} \frac{m_b^2 - m_a^2}{2m_c} \\ h_{abc} &= f_{abc} \frac{m_a^2 + m_c^2 - m_b^2}{2m_c} & j_{abc} &= 0 & r_a &= 0 \end{aligned} \quad (45)$$

and some relations for k_{ab}, l_{ab}, p_{ab} and q_{ab} [2, 20, 30]. These results in no restriction on s . One easily verifies that these values of the parameters are not only necessary for the cancellation of the terms $x_1 \neq x_2$ in (43) and for (29), they are also sufficient. By absorbing a constant factor in g we obtain

$$f_{abc} = \epsilon_{abc}. \quad (46)$$

The latter are the structure constants of $su(2)$. So the gauge group structure is not put in, it arises as a consequence of physical consistency and the ansatz (41) for W . For more complicated models (i.e. more than three vector fields) this conclusion is impossible at this stage, because one does not know that the f_{abc} 's satisfy the Jacobi identity. However, then the latter is obtained in the next step (see below). So far it is possible that all couplings involving the H -field vanish, i.e. the Higgs field is not yet needed.

- The remaining terms in (43) come from the diagonal $x_1 = x_2$. Their cancellation cannot be achieved without H -couplings. This condition yields important results[§].

* The f_{abc} 's must fulfil the Jacobi identity. (In our simple model this is already known (46).)^{||}

* The masses must agree

$$m \equiv m_1 = m_2 = m_3. \quad (47)$$

* The H -coupling parameters take the values

$$k_{ab} = \frac{1}{2}\kappa\delta_{ab} \quad l_{ab} = -\frac{1}{2}\kappa m\delta_{ab} \quad p_{ab} = \frac{1}{2}\kappa m\delta_{ab} \quad q_{ab} = \frac{\kappa m_H^2}{4m}\delta_{ab} \quad (48)$$

where $\kappa \in \{-1, 1\}$. The parameter s is still free.

* There is no term $\sim : H^4 :$ in W (i.e. in first order in g).

[†] The fact that the terms $x_1 \neq x_2$ cannot be cancelled by diagonal terms $x_1 = x_2$ becomes clear from the explicit expressions.

[‡] It is obvious that the Q -divergence condition to first order implies the cancellation of the terms $x_1 \neq x_2$ in (43), but here we proceed in the opposite direction.

[§] These results agree precisely with those derived from the Q -divergence condition (for second-order tree diagrams) in [2, 30] and with [20].

^{||} Stora [41] found (for an arbitrary number of massless self-interacting vector fields) that the Q -divergence condition to first order implies that the coupling parameters are totally antisymmetric and that the Q -divergence condition for second-order tree diagrams yields the Jacobi identity.

* In $T(W(x_1)W(x_2))|_{\text{tree}}$ the C -number distributions are Feynman propagators with derivatives: $\Delta^F(x_1 - x_2)$, $\partial_\mu \Delta^F(x_1 - x_2)$ and $\partial_\nu \partial_\mu \Delta^F(x_1 - x_2)$. The first two extend uniquely to the diagonal $x_1 = x_2$ and the last one has a distinguished extension, namely $\partial_\nu \partial_\mu \Delta^F(x_1 - x_2)^\dagger$. We denote this extension by $T(W(x_1)W(x_2))|_{\text{tree}}^0$. So-called ‘tree-normalization terms’

$$N(x_1, x_2) = C\delta(x_1 - x_2) : B_1(x_1)B_2(x_2)B_3(x_3)B_4(x_4) : \quad C \in \mathbb{R} \text{ or } i\mathbb{R} \quad (49)$$

($B_1, \dots, B_4 \in \{A^\mu, u, \tilde{u}, \phi, H\}$) can be added to $T(W(x_1)W(x_2))|_{\text{tree}}^0$, if they satisfy the properties (b) (Poincaré covariance), (c) (ghost number), (d) (scaling degree) and (e) (unitarity) which are required above for W (here they restrict $N(x_1, x_2)$). These tree-normalization terms (49) correspond to the quadrilinear terms of order g^2 (g denotes the coupling constant) in the interaction Lagrangian of the conventional theory, they have the same influence on the perturbation series of the S -matrix. *The cancellation of the terms $x_1 = x_2$ in (43) fixes the possible tree-normalization terms (i.e. the constants C in (49)) uniquely (in terms of s), except for \ddagger*

$$N_{H^4}(x_1, x_2) = \lambda\delta(x_1 - x_2) : H^4(x_1) : \quad \lambda \in \mathbb{R}. \quad (50)$$

- The parameters s and λ , which are still free, are determined by physical consistency (42) for the tree diagrams to *third* order (analogously to [2]):

$$s = \frac{m_H^2}{4m} \quad \lambda = -\frac{m_H^2}{16m^2}. \quad (51)$$

So W and the tree diagram normalizations to second order are completely determined (up to the sign κ , which is conventional) and these terms agree precisely with the interaction Lagrangian obtained by the Higgs mechanism. We have made the assumption that there is at most one physical scalar field H . This assumption is most probably not necessary. By specializing the results of [30] to our model (of three self-interacting massive vector fields) one finds that there is no solution of the Q -divergence condition with more than one H -field, provided the H -fields couple to the vector or ghost fields.

The Higgs potential is not put in here, it is derived from physical consistency. Spontaneous symmetry breaking plays no role in this approach, because we *start with the massive free incoming fields*. A proof that physical consistency can be satisfied to all orders (by choosing suitable normalizations) is missing up to now, but we are convinced that this holds true.

3. Renormalizability and ghosts

The formulation of massive self-interacting vector mesons of the previous section will now serve as a point of departure for a more fundamental conceptual discussion. Similar to Weinberg (Weinberg’s old work on Feynman rules for higher spin) we start with Wigner’s theory of particle representations; in our case because we want to avoid any parallelism to (quasi)classical systems (quantization) and (for reasons which will gradually become clear to the reader) develop local quantum physics from an intrinsic point of view as much as possible. In particular, we would like to understand the curious phenomenon that, contrary to the classical situation, the possibilities of perturbative renormalizable QFT are more restrictive, the higher the spin. Whereas the classical theory is in need of an additional selection principle (the gauge

\dagger The general extension which is Poincaré covariant (D) and does not increase the scaling degree (F) reads $\partial_\nu \partial_\mu \Delta^F(x_1 - x_2) + Cg_{\nu\mu}\delta(x_1 - x_2)$ $C \in \mathbb{C}$ arbitrary.

\ddagger For the expert we mention that tree-normalization terms with B_1, \dots, B_4 exclusively scalar fields (e.g. N_{H^4}) are required. In contrast to the other tree-normalization terms (i.e. the tree-normalization terms with vector field factors) they violate the normalization condition (N3) in [10] (or (43) in [17]). However, this causes no harm.

principle in the case of zero mass), local quantum physics for spin ≥ 1 is more restrictive: *the particle content and renormalizability fix the vector meson theory* (where the possible triviality for spin > 1 is a special case[†]). It is this result read backward into (quasi)classical field theory, which, in the spirit of Bohr's correspondence principle gives a fundamental physical support for the *classical gauge selection principle* (which then gives strong link with the mathematical–aesthetical appeal to fibre bundles). As a side result we will learn that the massive theory fulfils the Schwinger–Swieca [38, 43] screening mechanism, and that (similar to the gauge interpretation in terms of a Higgs field) the theory has more physical degrees of freedom than the massive vector mesons from which we started in zeroth order, namely consistent perturbation theory requires the introduction of a scalar H -field (the Higgs field without Higgs condensate[‡]) of the previous section.

It is well known that the step from the Wigner representation theory of particles (positive-energy irreducible representations of the Poincaré group with finite spin/helicity) to local free fields is described in terms of intertwiners u, v [§]

$$\psi^{[n_+, n_-]}(x) = \int \sum_{s_3=-s}^s \{e^{-ipx} u(\vec{p}, s_3) a(\vec{p}, s_3) + e^{ipx} v(\vec{p}, s_3) b^*(\vec{p}, s_3)\} \frac{d^3 p}{2\omega} \quad (52)$$

which intertwine the Wigner representation matrices $D^{(s)}(R(\Lambda, p))$ with matrices of the covariant Lorentz group representation $D^{[n_+, n_-]}(\Lambda)$

$$D^{[n_+, n_-]}(\Lambda)u(p) = u(\Lambda p)D^{(s)}(R(\Lambda, \Lambda p)) \quad (53)$$

where for convenience we have collected the $(2s + 1) u^{(n_+, n_-)}(p)$ mixed (un)dotted $(2n_+ + 1)(2n_- + 1)$ component u -spinors into a rectangular $(2n_+ + 1)(2n_- + 1) \times (2s + 1)$ matrix $u(p)$ and similar for the v 's.

Let us first note that the covariantized inner product in the *one-particle Wigner space* for $s \geq 1$ contains necessarily first or higher powers of momenta. Associated with this is the fact that the intertwiners u, v have a dimension ≥ 1 which immediately translates into an operator dimension of the field $\dim \psi \geq 2$. Since interaction densities $W_0 \equiv W$ are at least trilinear in free fields and since the smallest possible operator dimension is one (for scalar fields), it is impossible to satisfy the renormalizability condition $\dim W \leq 4$ within the Stückelberg–Bogoliubov–Epstein–Glaser operator approach. Inspired by the idea of a cohomological representation of the physical space and the physical observables in the previous section one looks for a cohomological extension of the Wigner space for massive vector mesons in order that the associated two-point function (or propagation kernel) has a milder (renormalizable) high-momentum behaviour. The results of the previous section also suggest the simplest possibility to achieve that, namely to use three additional indefinite metric scalar wavefunctions (where on the Wigner level the ‘statistics’ is yet undetermined). More precisely we form an extended Hilbert space H_{ext} by the multicomponent wavefunctions $(A_\mu, \omega, \bar{\omega}, \varphi)$ defined on the mass shell. H_{ext} has in addition to a positive-definite inner product another one which does not have this property and corresponds to the Krein structure of the previous section. The latter is relevant in order to have at least some pseudo-unitary Lorentz covariance and a definition of

[†] In a recent paper Scharf and Wellmann [31] have shown that there exists no renormalizable theory for $s = 2$ which satisfies the free perturbative operator gauge invariance.

[‡] We use the notation H -field only in order to avoid any association with Higgs condensates.

[§] The letter u , which was used in the previous section for a ghost field, means here an intertwiner. The ghost fields in the Wigner one-particle space will be denoted by $(\omega, \bar{\omega}, \varphi)$. The corresponding Fock space fields are the fields (u, \tilde{u}, ϕ) of the previous section.

modular localized subspaces. On H_{ext} we define a BRS-like operator by

$$s_W \begin{pmatrix} A_\mu^a(p) \\ \omega_a(p) \\ \bar{\omega}_a(p) \\ \varphi_a(p) \end{pmatrix} = \begin{pmatrix} p_\mu \omega_a(p) \\ 0 \\ -p^\mu A_\mu^a(p) - im_a \varphi_a(p) \\ im_a \omega_a(p) \end{pmatrix}. \tag{54}$$

The so defined s_W -operation defines a differential space since the definition easily leads to $s_W^2 = 0$. One then uses this s_W in order to write the following cohomological representation for the physical (Wigner) Hilbert space H_W in terms of the above extended space H_{ext} :

$$H_W = \frac{\ker s_W}{\text{ran } s_W} \tag{55}$$

$$= cl \left\{ A_\mu(p) \mid p^\mu A_\mu(p) = 0, - \int A_\mu(p) A^\mu(p) \frac{d^3 p}{2\omega} < \infty \right\}$$

$$= cl \left\{ A_\mu(p) \mid - \int A_\mu(p) \left(g^{\mu\nu} - \frac{p^\mu p^\nu}{m^2} \right) A_\nu(p) \frac{d^3 p}{2\omega} < \infty \right\}$$

i.e. we obtain the L^2 -closure of the space of transversal vector wavefunctions which in terms of the associated fields (52) had the high dimension $\dim A = 2$ which was responsible for the lack of renormalizable interactions within the original (non-extended) formulation. (The transversality condition does not lower the dimension of the vector wavefunctions or the corresponding fields.) On the other hand, the extended Hilbert space has no transversality condition and obeys the classical assignment of dimensions (i.e. $\dim A = 1$ in H_{ext}). This is due to the fact that ghost contributions damp the high-momentum behaviour of the associated two-point function[†].

The extension has, however, an influence on the modular localization theory in Wigner space where the latter is (via the CCR/CAR functor) the pre-empted locality in Fock space. For example the wedge (Rindler or Bisognano–Wichmann) localized subspace involves instead of the simple complex conjugation in addition the Krein operator η , i.e. one has for the action of the pre-Tomita operator[‡]:

$$(S\psi)(p) = (\eta\bar{\psi})(p) \tag{56}$$

where ψ is the multi-component wavefunction involving the ghosts in addition to the vector potential. The standard modular localization theory can be found in [33] and the adaptation to the present extended ‘pseudo-modular’ case will be treated in a separate paper. The physical Wigner subspace (more precisely it is a cohomologically defined factor space) is precisely characterized by the validity of the ‘correct’ modular localization associated with the Tomita theory. Although our ghost extension of the $(m, s = 1)$ Wigner representation is not uniquely

[†] In the Lagrangian framework there is an alternative method to lower the dimension of the massive vector field A from two to one. In contrast to massless vector fields, the Proca field ($\mathcal{L} = -\frac{1}{4}F^2 + \frac{1}{2}m^2 A^2$, $F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu$) has a well defined propagator without introducing a ‘gauge-fixing term’, but it has dimension $\dim A = 2$. Stückelberg’s trick is to add a ‘gauge-fixing term’ $\mathcal{L}_1 = -\frac{1}{2}\lambda(\partial \cdot A)^2$, $\lambda > 0$ (this misleading terminology is chosen because \mathcal{L}_1 has the same form as the gauge-fixing Lagrangian in the massless case), which damps the high-momentum behaviour of the propagator such that $\dim A = 1$. In this Stückelberg formalism, the necessity to introduce ghosts shows up in the fact that without ghosts it is impossible to define a *stable* physical subspace or factor space. For $\lambda = 1$ (Feynman gauge) (7) one obtains the same free vector fields as in the quantization of the cohomological extension of the Wigner one-particle space. We prefer the latter method, because we do not want to rely on a Lagrangian framework, instead we want to be close to the particle picture.

[‡] The pre-Tomita operator S has nothing to do with the above nilpotent operators s_W (54) (in the one-particle Wigner space) or s (57) in Fock space.

fixed (we chose a ‘minimal’ extension) we believe that any other cohomological extension which also lowers the $\dim A = 2$ down to its classical value $\dim A = 1$ will contain the minimal and possible additional pieces which do not change the physical content[†].

The next step from particles to fields is the answer to the question of what is the action of s_W on the *multi-particle tensor space* [9]. From the usual Fock space formalism we are used to the following action of derivations δ on tensor products:

$$\delta(\psi_1 \otimes \psi_2) = \delta\psi_1 \otimes \psi_2 + \psi_1 \otimes \delta\psi_2.$$

It is easy to see that the tensor product action of s_W must include a grading in order to maintain the nilpotency,

$$s(\psi_1 \otimes \psi_2) = s\psi_1 \otimes \psi_2 + \psi_1(-)^{\text{degree } \psi_1} \otimes s\psi_2 \tag{57}$$

where s denotes the Fock space version of s_W [‡]. So, two of the three scalar ghost fields (u, \tilde{u}, ϕ) [§], which are companions of each massive vector meson field, are required to be graded fermionic fields and the third one must be bosonic.

In this way we obtain the Fock space formalism of the previous section with $[Q, \cdot]$ being the implementation of s in Fock space. Whereas the ghost formalism can be pursued back into the Wigner one-particle theory, the necessity to choose trilinear couplings at first order in g with coefficients fulfilling group-theoretical symmetry, as well as the necessity of enlargement of the vector meson set-up by additional physical degrees of freedom (whose simplest and perhaps only realization are the scalar H -fields) only shows up as a consistency requirement above the zeroth order.

4. Determination by field content

In the second section physical consistency was formulated for the S -matrix (31). We are now looking for a corresponding condition in terms of *interacting fields*. Such fields (including composites) are defined by means of the Bogoliubov transition functional $S(g)$ (1) in Fock space as formal power series in g_0 . The interacting field $W_{j\text{int}}(x; g_0 W)$ due to the interaction $W \equiv W_0$ and corresponding to the Wick polynomial $W_j, j = 1, \dots, G$ of free fields, is defined by

$$W_{j\text{int}}(x; g_0 W) \equiv \frac{\delta}{i\delta g_j(x)} S(g_0, 0, \dots, 0)^{-1} S(g_0, 0, \dots, g_j, 0, \dots)|_{g_j=0}. \tag{58}$$

By inserting (1) one obtains the perturbative expansion of the interacting fields

$$W_{j\text{int}}(x; g_0 W) = W_j(x) + \sum_{n=1}^{\infty} \frac{i^n}{n!} \int d^4x_1 \dots d^4x_n g_0(x_1) \dots g_0(x_n) R(W(x_1) \dots W(x_n); W_j(x)) \tag{59}$$

with the ‘totally retarded products’

$$R(A_1(x_1) \dots A_n(x_n); A(x)) \equiv \sum_{I \subset \{1, \dots, n\}} (-1)^{|I|} \bar{T}(A_i(x_i), i \in I) T(A_k(x_k), k \in I^c, A(x)) \tag{60}$$

[†] The geometrical Faddeev–Popov method can also be considered as a minimal extension of the functional measure. It owes its unique appearance more to geometric than quantum physical reasoning.

[‡] From the context it should be clear whether we mean by the letter s the present nilpotent Fock space operator s or the spin.

[§] We recall that (u, \tilde{u}, ϕ) are the Fock space fields of the previous section and correspond to $(\omega, \bar{\omega}, \varphi)$ in the Wigner one-particle space.

where A_1, \dots, A_n, A are Wick polynomials, $I^c \equiv \{1, \dots, n\} \setminus I$ and \bar{T} denotes the ‘antichronological product’. The corresponding generating functional is $S(g)^{-1}$. The antichronological products can be obtained uniquely from the time-ordered products by the usual inversion of a formal power series

$$\bar{T}(A_1(x_1) \dots A_n(x_n)) = \sum_{P \in \text{Part} \{1, \dots, n\}} (-1)^{|P|+n} \prod_{p \in P} T(A_i(x_i), i \in p). \quad (61)$$

By means of causality (4) one easily finds that the R -products (60) have totally retarded support with respect to the distinguished coordinate x ,

$$\text{supp } R(A_1(x_1) \dots A_n(x_n); A(x)) \subset \{(x_1, \dots, x_n; x) \mid x_i \in x + \bar{V}_-, \forall i = 1, \dots, n\}. \quad (62)$$

In a pure massive theory the strong adiabatic limit of the interacting fields exists as a formal power series in $g := g_0(0)$:

$$\begin{aligned} W_{j \text{ int}}(x)\psi &\equiv W_{j \text{ int}}(x; W)\psi \equiv \lim_{\epsilon \rightarrow 0} W_{j \text{ int}}(x; g_0\epsilon W)\psi \\ \text{where } g_0\epsilon(x) &\equiv g_0(\epsilon x) \quad \psi \in \mathcal{D} \end{aligned} \quad (63)$$

because this holds true for $S(g)$ (6) [18].

A $*$ -symmetrical interacting field in the adiabatic limit (i.e. $\phi_{\text{int}}(x) \equiv \phi_{\text{int}}(x; W) = \lim_{\epsilon \rightarrow 0} \phi_{\text{int}}(x; g_0\epsilon W)$, $\phi_{\text{int}}^* = \phi_{\text{int}}$) is called *physical* (or an *observable*) if it induces a well defined operator on the factor space $\mathcal{H}'_{\text{phys}}$ (22). This holds true iff

$$\phi_{\text{int}}(f; W) \ker Q \subset \ker Q \quad \wedge \quad \phi_{\text{int}}(f; W) \text{ ran } Q \subset \text{ran } Q \quad \forall f \in \mathcal{S}(\mathbb{R}^4) \quad (64)$$

(cf (25)). This is equivalent to

$$[Q, \phi_{\text{int}}(f; W)]|_{\ker Q} = 0 \quad \forall f \in \mathcal{S}(\mathbb{R}^4). \quad (65)$$

(The non-trivial part of this statement is that (65) implies $\phi_{\text{int}}(f) \text{ ran } Q \subset \text{ran } Q$. However, using the notation of section 2, the condition (65) is equivalent to $\phi_{\text{int}}(f)_{+-} = 0 = \phi_{\text{int}}(f)_{+0}$. Hence $\phi_{\text{int}}(\bar{f})_{0-} = (\phi_{\text{int}}(f)_{+0})^* = 0$. Together we obtain $\phi_{\text{int}}(f) \text{ ran } Q \subset \text{ran } Q$.)

Let $F_a^{\mu\nu} \equiv \partial^\mu A_a^\nu - \partial^\nu A_a^\mu$. We now require that (for each $a = 1, \dots, M$) there exists a physical field $\mathcal{F}_{a \text{ int}}^{\mu\nu}$, i.e.

$$[Q, \mathcal{F}_{a \text{ int}}^{\mu\nu}(x; W)]|_{\ker Q} = 0 \quad (66)$$

with the following additional properties.

- (i) $\mathcal{F}_{a \text{ int}}^{\mu\nu} = F_a^{\mu\nu} + \sum_k c_k \psi_{ak \text{ int}}^{\mu\nu}$, where the c_k are formal power series of (constant) complex numbers and $\mathcal{F}_{a \text{ int}}^{\mu\nu}$ agrees in zeroth order with the free $F_a^{\mu\nu}$, i.e. the c_k vanish to zeroth order: $c_k^{(0)} = 0$;
- (ii) the zeroth order $\psi_{ak \text{ int}}^{\mu\nu}$ of $\psi_{ak \text{ int}}^{\mu\nu}$ is a Wick monomial which has precisely one factor F or A , the other factors are ghost or scalar fields, and $\psi_{ak}^{\mu\nu} \neq F_b^{\mu\nu}, \forall a, b \in \{1, \dots, M\}$;
- (iii) the scaling degree (or mass dimension) of ψ_{ak} is ≤ 4 ;
- (iv) $\mathcal{F}_{a \text{ int}}^{\mu\nu}$ is a Lorentz tensor of second rank and is antisymmetrical in (μ, ν) ;
- (v) the ghost number of $\mathcal{F}_{a \text{ int}}^{\mu\nu}$ is zero;
- (vi) $\mathcal{F}_{a \text{ int}}^{\mu\nu}(x; W)^* = \mathcal{F}_{a \text{ int}}^{\mu\nu}(x; W)$.

The requirement (ii) is badly motivated (except for the demand that ψ_{ak} must be a Wick monomial), its main purpose is to shorten the calculations. However, the condition $\psi_{ak}^{\mu\nu} \neq F_b^{\mu\nu}$ in (ii) is necessary for the uniqueness of \mathcal{F}_{int} : the fields

$$\left(1 + \sum_{k=1}^{\infty} b_k g^k\right) \mathcal{F}_{a \text{ int}}^{\mu\nu} \quad b_k = \text{constant} \in \mathbb{R}$$

satisfy all other requirements if \mathcal{F}_{int} does so. It is an interesting question as to how far (ii) can be weakened such that the uniqueness of \mathcal{F}_{int} does not get lost.

Due to the normalization condition (F) (scaling degree) for the time-ordered products, property (iii) implies $\dim \mathcal{F}_{a \text{int}}^{\mu\nu} \leq 4$. To specify (iv)–(vi) note that they must be fulfilled in particular by the Wick monomials ψ_{ak} (in the case of (vi) the coefficients c_k are also involved: $c_k^* \psi_{ak}^* = c_k \psi_{ak}$). However, these three requirements also restrict the normalization of the higher orders of $\psi_{ak \text{int}}^{\mu\nu}$ and $F_{a \text{int}}^{\mu\nu}$. It is easy to see that these additional normalization conditions can be fulfilled (e.g. by antisymmetrization in (μ, ν) of an arbitrary Poincaré covariant extension) and we assume that the normalizations are always done in such a way.

The most general ansatz which is compatible with (i)–(vi) reads

$$\begin{aligned}
\mathcal{F}_{d \text{int}}^{\mu\nu} = & F_{d \text{int}}^{\mu\nu} + t_{dab}(F_a^{\mu\nu} \phi_b)_{\text{int}} + \tilde{t}_{dab}(A_a^\mu \partial^\nu \phi_b - A_a^\nu \partial^\mu \phi_b)_{\text{int}} \\
& + v_{dabc}(: F_a^{\mu\nu} \phi_b \phi_c :)_{\text{int}} + \tilde{v}_{dabc}(: A_a^\mu \partial^\nu \phi_b \phi_c : - : A_a^\nu \partial^\mu \phi_b \phi_c :)_{\text{int}} \\
& + w_{dabc}(: F_a^{\mu\nu} u_b \tilde{u}_c :)_{\text{int}} + \tilde{w}_{dabc}(: A_a^\mu u_b \partial^\nu \tilde{u}_c : - : A_a^\nu u_b \partial^\mu \tilde{u}_c :)_{\text{int}} \\
& + w'_{dabc}(: A_a^\mu \partial^\nu u_b \tilde{u}_c : - : A_a^\nu \partial^\mu u_b \tilde{u}_c :)_{\text{int}} \\
& + x_{da}(F_a^{\mu\nu} H)_{\text{int}} + \tilde{x}_{da}(A_a^\mu \partial^\nu H - A_a^\nu \partial^\mu H)_{\text{int}} \\
& + y_{dab}(F_a^{\mu\nu} \phi_b H)_{\text{int}} + \tilde{y}_{dab}(A_a^\mu \partial^\nu \phi_b H - A_a^\nu \partial^\mu \phi_b H)_{\text{int}} \\
& + y'_{dab}(A_a^\mu \phi_b \partial^\nu H - A_a^\nu \phi_b \partial^\mu H)_{\text{int}} \\
& + z_{da}(: F_a^{\mu\nu} H^2 :)_{\text{int}} + \tilde{z}_{da}(: A_a^\mu \partial^\nu H H : - : A_a^\nu \partial^\mu H H :)_{\text{int}} \tag{67}
\end{aligned}$$

where $t_{dab}, \dots, w'_{dabc}, x_{da}, \dots, \tilde{z}_{da}$ are (arbitrary) constants which are formal power series in \mathbb{R}

$$t_{dab} = \sum_{k=1}^{\infty} t_{dab}^{(k)} g^k \quad x_{da} = \sum_{k=1}^{\infty} x_{da}^{(k)} g^k \quad \text{etc.}$$

We define

$$\mathcal{F}_d^{\mu\nu(k)} \equiv t_{dab}^{(k)} F_a^{\mu\nu} \phi_b + \dots + x_{da}^{(k)} F_a^{\mu\nu} H + \dots \quad \forall k \geq 1$$

hence[†]

$$\mathcal{F}_{d \text{int}}^{\mu\nu} = F_{d \text{int}}^{\mu\nu} + \sum_{k=1}^{\infty} \mathcal{F}_{d \text{int}}^{\mu\nu(k)} g^k. \tag{68}$$

By definition we may assume $v_{dabc} = v_{dacb}$.

We require that the interaction density W satisfies the properties (a)–(e) listed in section 2. The physical consistency of the S -matrix $[Q, S]|_{\ker Q} = 0$ is omitted here, we replace it by $[Q, \mathcal{F}_{\text{int}}(x; W)] = 0$ (see (69) below). In particular, we require $\dim W \leq 4$ which ensures renormalizability. Additionally, we assume that W contains no quadrilinear terms. Most probably this assumption is not necessary, i.e. the other requirements (including the physicality (66)) exclude the quadrilinear terms. However, this still needs to be checked and will be left open here.

We now replace the requirement (66) by

$$[Q, \mathcal{F}_{a \text{int}}^{\mu\nu}(x; W)] = 0. \tag{69}$$

[†] By definition the map $\phi \rightarrow \phi_{\text{int}}$ (ϕ is a Wick monomial) is linear with C^∞ -functions as coefficients (here the coefficients are constants).

This may be justified in the following way[†]. Since we are working in the adiabatic limit our Q (36) (constructed in terms of incoming free fields) agrees with the Kugo–Ojima operator Q_{int} [10, 25] which implements the BRST transformation of the interacting fields. Hence $[Q, \mathcal{F}_{a\text{int}}^{\mu\nu}(x; W)]$ is again a local operator. However, by the Reeh–Schlieder theorem [27] such an operator is zero if it vanishes on the vacuum, which is an element of the kernel of Q . Thus the requirements (66) and (69) are equivalent[‡]. The main (completely new) result of this section is that the *physicality* (69) fixes the parameters $t_{dab}, \dots, \tilde{z}_{da}$ in the ansatz (67) and the interaction W (including the tree-normalization terms to second order (49)) up to the same non-uniqueness as in section 2, namely the addition of divergence and coboundary couplings to W (39). In particular, we will see that *an additional physical degree of freedom is required: the Higgs field H is needed in the interaction W as well as in the field $\mathcal{F}_{d\text{int}}^{\mu\nu}$.*

We verify this statement in appendix B by explicit calculation of the tree diagrams to lowest orders. We do this only for the simplest non-trivial model, namely three self-interacting massive vector fields ($m_a > 0, a = 1, 2, 3$) as in section 2. We make the same ansatz for W as in section 2, but without the quadrilinear terms. Up to divergence and coboundary couplings (39) this is the most general trilinear ansatz which fulfils the requirements (a)–(e) of section 2.

The parameters in the ansatz for W and $\mathcal{F}_{d\text{int}}^{\mu\nu}$ (67) are determined by inserting these expressions into the physicality requirement (69). Here we only state the results:

- for W we obtain precisely the same expression as in section 2,
- we have computed the parameters in \mathcal{F}_{int} (which are formal power series) up to second order in g ,

$$\begin{aligned} \mathcal{F}_{d\text{int}}^{\mu\nu} = & F_{d\text{int}}^{\mu\nu} - \frac{g}{m} \epsilon_{abc} (F_b^{\mu\nu} \phi_c)_{\text{int}} + \frac{g}{m} (F_d^{\mu\nu} H)_{\text{int}} \\ & - \frac{g^2}{4m^2} [\delta_{da} \delta_{bc} - (\delta_{dc} \delta_{ba} + \delta_{db} \delta_{ca})] (: F_a^{\mu\nu} \phi_b \phi_c :)_{\text{int}} \\ & - \frac{g^2}{2m^2} \epsilon_{abc} (F_b^{\mu\nu} \phi_c H)_{\text{int}} + \frac{g^2}{4m^2} (: F_d^{\mu\nu} H^2 :)_{\text{int}} + \mathcal{O}(g^3). \end{aligned} \quad (70)$$

To gain a better understanding of the latter result we identify these physical fields $\mathcal{F}_{d\text{int}}^{\mu\nu}$ (70) as gauge-invariant fields within the framework of spontaneous symmetry breaking of the $SU(2)$ gauge symmetry. In this semiclassical picture the scalar fields ϕ_a (35) and H (37) form two $SU(2)$ doublets

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_2 + i\phi_1 \\ v + H - i\phi_3 \end{pmatrix} \quad (71)$$

and

$$\tilde{\Phi} = \frac{1}{\sqrt{2}} \begin{pmatrix} v + H + i\phi_3 \\ -\phi_2 + i\phi_1 \end{pmatrix} \quad (72)$$

where v is the vacuum expectation value of the original (i.e. non-shifted) field $\tilde{H} = v + H$. v is proportional to the gauge boson mass

$$m = \frac{1}{2} g v. \quad (73)$$

The composite fields

$$\mathcal{G}_3^{\mu\nu} := \tilde{\Phi}^* F^{\mu\nu} \tilde{\Phi} = -\Phi^* F^{\mu\nu} \Phi = - \sum_{a=1,2,3} F_a^{\mu\nu} \Phi^* \sigma_a \Phi \quad (74)$$

[†] We thank Klaus Fredenhagen for bringing this argument to our attention.

[‡] This argument cannot be applied to $[Q, S]|_{\ker Q} = 0$ (31), because the S -matrix is non-local.

(σ_a are the Pauli matrices) and

$$\mathcal{G}_1^{\mu\nu} := \frac{1}{2}(\tilde{\Phi}^* F^{\mu\nu} \Phi + \Phi^* F^{\mu\nu} \tilde{\Phi}) \quad \mathcal{G}_2^{\mu\nu} := \frac{1}{2}i(\tilde{\Phi}^* F^{\mu\nu} \Phi - \Phi^* F^{\mu\nu} \tilde{\Phi}) \quad (75)$$

are $SU(2)$ gauge invariant, or equivalently they are invariant with respect to the classical† BRST transformation

$$s(F_a^{\mu\nu}) = ig\epsilon_{abc}F_b^{\mu\nu}u_c \quad s(\phi_a) = imu_a + \frac{1}{2}ig(u_a H + \epsilon_{abc}\phi_b u_c) \quad s(H) = -\frac{1}{2}ig u_a \phi_a. \quad (76)$$

By multiplying out the matrices in $\mathcal{G}_d^{\mu\nu}$ (74) and (75) we find that (with a suitable normalization) the corresponding quantum fields are proportional to $\mathcal{F}_{d\text{int}}^{\mu\nu}$, so far as we have determined the constant power series $t_{dab}, \dots, \tilde{z}_{da}$ (67), i.e.

$$\mathcal{G}_{d\text{int}}^{\mu\nu}(x; W) = \frac{1}{2}v^2 \mathcal{F}_{d\text{int}}^{\mu\nu}(x; W) + \mathcal{O}(g^3). \quad (77)$$

Finally, some more remarks on the validity of the Schwinger–Swieca screening property are in order. Since the physical field strength is an operator in Hilbert space (a positive-definite metric) which fulfils a Maxwell-type equation and since the particle spectrum has an isolated mass hyperboloid for the vector meson, the theorem of Swieca [43] is applicable and hence the charge defined in terms of the large surface integral over the physical field strength vanishes. This charge screening is more interesting if additional spinor matter is present.

5. Renormalizability for $s \geq 1$ without ghosts?

We have seen that the cohomological representation of the Wigner theory for vector mesons achieves the magic trick of rescuing renormalizability, whereas the naive application of the standard causal perturbation theory based on interactions W_{phys} in terms of local Wick polynomials in physical (Wigner) field coordinates, which inevitably leads to $\dim W_{\text{phys}} \geq 5$, fails on the count of renormalizability‡. The trick employs ghosts in intermediate steps in a way which is reminiscent of a catalyser in chemistry, that is to say a physical problem of perturbatively coupled massive spin-1 interactions, which *a priori* has nothing to do with ghosts, had to be cohomologically extended, because that was apparently the only way to reconcile the standard perturbative machinery (of deformation of free theories by W 's) with the short-distance renormalizability requirement. However, at the end, after the cohomological descend, one obtains local physical vector meson fields and a physical S -matrix (both renormalizable) within a physical Fock space of massive spin-1 particles, together with new physical degrees of freedom. This is a physical result which totally conceals the intermediate presence of ghosts. Although the final result agrees formally with (the gauge-invariant part of) renormalized gauge theory, the underlying physical idea and the words used to describe it are quite different. Instead of the Higgs mechanism, which generates the $s = 1$ mass through an additionally introduced (by hand) scalar field with non-vanishing vacuum expectation value, and which has no visible intrinsic (gauge-invariant) meaning, the vector meson mass in the present approach is directly linked with the aforementioned Schwinger–Swieca screening mechanism. The latter, together with the renormalizability requirement, demands the presence of additional degrees of freedom which we realized as scalar particles and identified with the Higgs field without its vacuum condensate.

† To avoid problems of defining products of interacting fields (this can be done by means of (58)) and the BRST transformation thereof, we only consider *classical* fields here.

‡ This is a well known limitation in Weinberg's programme of using Wigner's particle theory as a starting point for perturbative renormalization for massive vector mesons and higher-spin particles.

Although we have neither demonstrated the uniqueness of the ghosts nor of the new physical degree of freedom, we believe, that as already mentioned in section 3, our minimal solution is unique in the sense that any other solution involving higher-spin ghost and induced physical objects always contains our minimal solution (plus possible additional couplings with coupling parameters which may be set to zero).

Between the two formulations for interacting vector mesons, once as the quantized version of classical gauge theories with the Higgs mechanism and, on the other hand, the ‘ghost catalyser’ to implement perturbative renormalization we prefer the latter because it underlines the preliminary aspect of our understanding of higher-spin interactions more clearly. In this sense it upholds the Bohr–Heisenberg maxim that even in cases of discoveries via non-observable and formal constructs, one should always aim for a reduction to observable concepts in order to obtain a profound conceptual understanding.

The present formulation also gives a more detailed picture of physical fields (cf (70)). Whereas their locality and spectral properties are as in a renormalizable lower-spin $s < 1$ coupling, their operator dimension generally does not have the form of logarithmic corrections on canonical values of Lagrangian fields. Rather they involve higher Wick polynomials up to operator dimension four i.e. are composites in the unphysical fields which interpolate ‘elementary’ (in the sense of a perturbative description only) physical particles. This is the consequence of the fact that the cohomological descent renders the standard formulae in terms of retarded R -products invalid. Instead of Lagrangian interpolating fields with their canonical dimensions and logarithmic radiative corrections we have to work with the physical composites for the interpolation of the physical particles. The Q -invariant linear combinations which represent the physical fields within the extended Wick formalism have no intrinsic meaning or in other words, the present formalism has not supplied us with a formalism which leads to the physical vacuum expectation values in a more direct and less *ad hoc* manner. Whereas symmetries of a free-field theory can be used for the selection of a natural invariant subalgebra, the selection of a Q -invariant subalgebra of Wick polynomials remains a rather unnatural procedure. Whereas an observable subalgebra resulting as the fixed-point algebra of an internal symmetry contains all the structure which via the superselection theory allows an intrinsic reconstruction of the algebra of fields, there is no intrinsic natural way to reconstruct the ghost extended algebra from the physical algebra. It also indicates that even though the final physical degrees of freedom are local and can be described in terms of covariant point-like fields, the formalism deviates in ‘some way’ from the standard local behaviour in that the perturbation cannot be interpreted as a deformation on the original local class (Borchers class) of local physical fields but rather appears as a subclass of a deformed unphysical extended theory with artificial looking rules concerning its position within the extended local class. In this way the violation of the Bohr–Heisenberg maxim about observables in the present description of higher-spin interaction becomes most evident. Before we present some ideas about an alternative approach avoiding non-observable aspects, it is helpful to recall how the Wigner theory deals with the vector potential in the case of zero mass, i.e. for photons.

On a very formal level (ignoring regularizations of line integrals) we could, in fact, also have obtained the lowering of the operator dimension $\dim A$ from two to one without any cohomological extension by allowing non-point-like and non-covariant vector potentials which are interpolating the same particles as the covariant ones. Although on-shell objects such as the S -matrix of the final QFT are unaffected, as long as those fields remain almost local in the sense of [21], it is not known how to deal with such objects in a causal approach based on interaction polynomials W and transition operators $S(g)$.

It is interesting to note that the Wigner theory of free photons does not allow the introduction of any local covariant vector potential, thus raising the suspicion that massless

spin-1 interactions formulated in a physical Hilbert space lead to some sort of clash between renormalizability (which requires vector potentials of $\dim A_\mu = 1$) and the locality and covariance of interaction densities W . This problem with vector potentials comes from the non-compact structure of the little group (stability group) of a light-like vector say $p_R = (1, 0, 0, 1)$, which is the twofold covering of the Euclidean group $E(2)$ (and is denoted by $\tilde{E}(2)$). Whereas the rotation has the interpretation as a helicity rotation (rotation around the third axis), the ‘translations’ are Lorentz transformations which tilt the t - z wedge, leaving its upper light-like vector unchanged. As far as the two transversal coordinates are concerned they behave like 2-parametric Galilei velocity transformations (i.e. ‘lightcone translations’ without the energy positivity property) with the two longitudinal lightcone translations playing the role of the Hamiltonian, respectively, the central mass in the quantum mechanical representation theory of the Galilei group. The embedding of $\tilde{E}(2)$ into $SL(2, \mathbb{C})$ for the above choice of reference vector is

$$\alpha(\rho, \theta) = \begin{pmatrix} e^{i\frac{1}{2}\theta} & \rho \\ 0 & e^{-i\frac{1}{2}\theta} \end{pmatrix} \quad p_R \sim \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \quad (78)$$

where θ is the angle for the rotations around the 3-axis and $\rho = \rho_1 + i\rho_2$ parametrizes the Euclidean translations by the vector (ρ_1, ρ_2) . The unitary representation theory of this non-compact group is somewhat more complicated than that of $SU(2)$. However, it is obvious that the representations fall into two classes; the ‘neutrino–photon’ class with $U(\alpha(\rho, 0)) = 1$, i.e. a trivial representation of the Euclidean translations, and the remaining faithful ‘continuous spin’ (infinite-dimensional) representations with $U(\alpha(\rho, 0)) \neq 1$. A more detailed analysis shows that the latter lacks the strong localization requirements which one must impose on those positive-energy representations which are used for the description of particles[†]. Hence the Wigner theory does not allow one to describe photon operators in terms of covariant vector fields. On the other hand, a covariant field strength $F_{\mu\nu}$ has the following intertwiner representation in terms of the Wigner annihilation/creation operators for circular polarized photons $a_\pm^\#(k)$:

$$F_{\mu\nu}(x) = \frac{1}{(2\pi)^{3/2}} \int \left\{ e^{-ikx} \sum_{\pm} u_{\mu\nu}^{(\pm)}(k) a_\pm(k) + \text{h.c.} \right\} \frac{d^3k}{2|k|} \quad (79)$$

$$u_{\mu\nu}^{(\pm)}(k) \simeq k_\mu e_\nu^{(\pm)}(k) - k_\nu e_\mu^{(\pm)}(k). \quad (80)$$

Here $e_\mu^{(\pm)}(k)$ are the polarization vectors which are obtained by application of the following Lorentz transformation to the standard reference vectors $\frac{1}{\sqrt{2}}(0, \pm 1, i, 0)$ (which are $\perp (1, 0, 0, 1)$): a rotation of the z -axis into the momentum direction $\vec{n} \equiv \frac{\vec{k}}{\omega}$ (fixed uniquely by the standard prescription in terms of two Euler angles) and a subsequent Lorentz boost along this direction which transforms $(1, \vec{n})$ into $k = \omega(1, \vec{n})$. It is these vectors that do not behave as covariant under those Lorentz transformations which involve the above ‘little group translations’ but rather produce an affine transformation law

$$G(\rho) e^{(\lambda)}(p_R) = e^{(\lambda)}(p_R) + \begin{cases} -\frac{1}{2}(\bar{\rho}, 0, 0, \bar{\rho}) & \lambda = + \\ +\frac{1}{2}(\rho, 0, 0, \rho) & \lambda = - \end{cases} \quad \rho = \rho_1 + i\rho_2 \quad (81)$$

($G(\rho)$ is the Minkowski space representation of the Euclidean translations by (ρ_1, ρ_2)), whereas under x - y rotations the $e^{(\lambda)}$ picks up the standard Wigner phase factor. The polarization vectors

[†] For the non-faithful representation ($m = 0, h = \text{semi-integer}$) ($U(\alpha(\rho, 0)) \neq 1$) one finds that the infinite set of intertwiners in (52) is restricted to $u^{(n_+, n_-)}$ with $n_- = n_+ \pm h$ [46].

do not behave as 4-vectors since they are not invariant under the Euclidean translations in $\tilde{E}(2)$, as one would have expected for a (non-existent!) *bona fide* intertwiner from the $(0, h = 1)$ Wigner representation to the $D^{[\frac{1}{2}, \frac{1}{2}]}$ covariant representation. Rather the intertwiner only has Lorentz covariance up to additive gauge transformations, i.e. up to affine longitudinal terms. For general Lorentz transformations this affine law reads

$$(U(\Lambda)e)_\mu(k) = \Lambda_\mu^\nu e_\nu(\Lambda^{-1}k) + k_\mu H(\Lambda, k). \quad (82)$$

This peculiar manifestation of the $(0, h = 1)$ little group $\tilde{E}(2)$ is the cause for the appearance of the local gauge issue in local quantum physics. Unfortunately, this quantum origin is somewhat hidden in the quantization approach, where it remains invisible behind the geometrical interpretation in terms of fibre bundles. In terms of the potential in the physical Fock space we have

$$U(\Lambda)A_\mu(x)U^*(\Lambda) = \Lambda_\mu^{-1\nu}A_\nu(\Lambda x) + \partial_\mu H \quad (83)$$

where H is a concrete operator involving the $a_\pm^\#$ and e_\pm from (80). Formally, the string-like localized vector potential (with localization chosen in the space-like n -direction) may be written as a line integral over the field strength

$$\begin{aligned} \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x) &= F_{\mu\nu}(x) \\ A_\nu &= \frac{1}{n \cdot \partial} n^\mu F_{\mu\nu} \curvearrowright n \cdot A = 0. \end{aligned} \quad (84)$$

The regularization needed in order to make mathematical sense out of these string-like objects and related technical problems will not be discussed in this qualitative presentation. The physical manifestation of the existence of the somewhat non-local vector potential is the breakdown of the additivity of the algebras generated by the field strength for non-simply connected regions [21]. The natural localization regions for vector potentials are space-like cones or wedges which contain such cones. On the other hand, the causal perturbation theory (and any other formulation using Lagrangian quantization) requires point-like fields.

Comparing this with the massive case, one notes that the problematic aspects of the use of vector potentials in the local description of $s = 1$ show up in the quantum properties of free Wigner photons before the renormalization procedure, whereas for massive vector mesons it appears only in facing the issue of renormalizability (dimensional counting) of interactions. What is common to both cases is the fact that no existing formulation of renormalized perturbation theory is capable of dealing with $\dim A_\mu = 1$ non-local vector potentials within the framework of causal Wick polynomials W . Both cases can be dealt with in terms of the same remedy namely cohomological extension by ghosts within a BRS-like formalism; however, as a result of the availability of arguments based on scattering theory, the BRS formalism is conceptually much simpler in the massive case, whereas in the massless theory one must use the full BRS formalism which suffers a contribution from interactions and requires additional concepts for the separation of observable algebras from physical states [10].

It should be clear from these remarks that in order to follow the Bohr–Heisenberg maxim and remove the ‘ghost catalysers’ in favour of a ghost-free formulation, one has to go significantly outside the present perturbative framework. A hint where to look comes from the observation that the ghosts have been introduced to lower the operator dimension of A_μ from two to one, while still maintaining the (formal) locality and the covariance of the operators. This is clearly an off-shell short-distance argument. So this observation suggests

to look for an on-shell formulation. On-shell quantities are the true S -matrix and formfactors of physical operators A between multi-particle scattering states:

$$S |p_1, \dots, p_n\rangle^{\text{out}} = |p_1, \dots, p_n\rangle^{\text{in}} \quad (85)$$

$$\text{out} \langle p'_1 \dots p'_m | A |p_1, \dots, p_n\rangle^{\text{in}} .$$

Note that we defined the formfactors in such a way that the matrix elements of S themselves correspond to the formfactor of the identity operator. So the first question is: can one compute the perturbative expansion of S while staying on-shell all the time. The causal approach used before does not fulfil this requirement, since the Bogoliubov transition operator $S(g)$ is off-shell (in the sense of this paper) and only approaches the mass shell in the adiabatic limit[†]. In fact it is not possible to formulate causality which involves products of fields directly on-shell, rather the on-shell substitute is the notoriously elusive crossing symmetry together with some inexorably linked subtle analytic continuation properties. So the question is: does there exist a formulation of quantum field theory which, different from quantization, uses such on-shell concepts in its construction.

Fortunately, the beginning of such a new approach to QFT already exists [32, 36]. This approach can be viewed as a generalization of the Wigner theory in the presence of interactions. It bypasses completely the use of point-like fields and aims directly at the wedge algebras and the associated net of smaller localized algebras obtained by intersecting wedge algebras.

It is deeply satisfying that the modular theory used in this construction directly converts the Wigner one-particle theory into the interaction-free nets of algebras without using fields and their equivalence classes in intermediate steps. The generalization to interacting models is given by the modular interpretation of the S -matrix as a relative modular invariant which ‘measures’ the interaction for wedge-localized algebras. The naturalness of the wedge region appeared for the first time long ago in Unruh’s Minkowski space analogue of the Hawking thermal aspects of black holes. Indeed, its use for a generalization of the Wigner approach in the presence of interactions leads rather swiftly to a profound understanding of two aspects: the relation of the thermal aspect of modular (wedge) localization with the aforementioned elusive crossing symmetry of particle physics, and the existence of generators of the wedge algebra which, applied to the vacuum, create one-particle states free of particle–antiparticle polarization clouds, despite the presence of interactions. The crossing is known to be closely related to the TCP theorem; in fact, it is a kind of TCP property for individual particles in formfactors whereby the particle in an incoming configuration is flipped into an (analytically continued) outgoing antiparticle. The existence of (vacuum) polarization-free generators (abbreviated as ‘PFG’) for wedge localized algebras, on the other hand, is a new discovery. Their correlation functions can be expressed in terms of ‘nested’ products of S -matrices [36]. A closer look has revealed that in the special case of $d = 1 + 1$ purely elastic S -matrices (‘factorizing models’ [24, 39]) the PFGs, in fact, obey the Zamolodchikov–Faddeev algebra, and that the latter receives in turn for the first time a physical spacetime interpretation which transcends its formal use [33–36].

The potential relevance of these remarks for a ghost-free formulation of higher-spin interactions becomes clearer if one notes that the reasons for their introduction is that the point-like interaction densities W_{phys} in terms of physical higher-spin fields have a short-distance behaviour beyond that allowed by the formal renormalizability criterion. If we could avoid

[†] The difference between on- and off-shell looks quite innocuous in momentum space, however, the spacetime aspects could be different. In order to illustrate this point, just look at the massive Thirring model. The $S(g)$ (respectively, the field correlation functions) show the full field-theoretic vacuum and one-particle polarization structure (virtual particle structure), whereas the S -matrix and the closely related PFG wedge generators (which are introduced below) are ‘quantum mechanical’, i.e. obey particle number conservation.

such off-shell objects as point-like fields and their correlation functions in favour of formfactors of localized operators (85) then the existence of the theory would not be threatened by the short-distance behaviour in the construction of the theory. Rather it would come out from the definition of formfactors of products of such operators by summing over the complete set of intermediate particle states. The first step in such a new construction programme namely the construction of the PFGs belonging to the model is purely on-shell and only involves physical particles. The correlation functions of PFGs are on-shell objects, in the old fashioned language they contain a ‘natural cut-off’ as a result of their somewhat delocalized nature, but unlike the brutal force cut-offs in the Lagrangian approach, these objects coexist in the same Hilbert space together with sharper localized operators which belong to intersections of wedge algebras. The non-triviality check in this approach, which is expected to replace the good short-distance behaviour in the standard renormalization theory, is the non-triviality of algebras localized in double cones obtained by intersecting wedges

$$\mathcal{A}(\mathcal{O}) = \bigcap_{\mathcal{O} \subset W} \mathcal{A}(W). \quad (86)$$

It may not be widely known, but it is nevertheless true that all physical information is contained in local algebras[†] and that there is no necessity from a conceptual point of view to use point-like field coordinates apart from certain distinguished point-like currents related to symmetries and possibly associated order/disorder fields.

The first step, namely the construction of the on-shell PFGs is reminiscent of attempts in the 1960s to construct unitary crossing symmetric S -matrices. This attempt, even in its limited perturbation version, was ill-fated and did not lead to tangible results. It remained as a very fundamental problem in particle physics closely related to the inverse problem, i.e. the question to what extent a physically admissible (on-shell) S -matrix determines uniquely an (off-shell) algebraic net. The modular approach to wedge algebras and the concept of PFGs incorporates and enriches this old S -matrix programme, so that it becomes the first step in a construction of local nets. The hope that in this new version it becomes at least susceptible to a kind of on-shell perturbation theory rests on the new field-theoretic setting together with rich new concepts which were absent in the days of the old bootstrap programme. In fact, the success of this programme for $d = 1 + 1$ factorizing models [24, 39] (in which case such an S -matrix which avoids the standard construction exists even in a summed up non-perturbative form) may be taken as an encouragement for a future exploration of this new approach to QFT.

One expects that this new approach reproduces the perturbative results of the standard renormalization theory wherever the latter is applicable. The innovative power of the new approach is expected to show up in situations where one has reason to believe that Lagrangian quantization does not provide physical field coordinates of sufficient low operator dimension related to good short-distance properties needed for the W 's in the causal approach. There is no *a priori* reason why the Lagrangian fields should be those with the lowest short-distance dimensions in the local equivalence class of all local fields of one model. In the case of spin-1, one can still overcome such obstacles against the standard formulations by the *ad hoc* ghost trick which is suggested by the quantization of gauge theories. However, could there also be (e.g. higher-spin) cases with a finite number of coupling parameters which are declared ‘non-renormalizable’ in the standard approach, even though they have a well defined perturbation theory in the new sense?

[†] This independence of the physical content from field coordinatizations is the main reason for the interest in the algebraic approach (i.e. local quantum physics).

6. Future perspectives

The history of gauge theory and its connection with higher-spin renormalizable QFT is one of the theoretically most fascinating and experimentally most significant developments in the 20th century. Contrary to a widely held opinion this is still an unfinished story with possible future surprises.

The usefulness of the role of the ‘gauge principle’ as a selection principle in the classical setting and its attractive geometric appeal in the semiclassical realm of quantum matter coupled to external fields (minimal electromagnetic coupling) is universally recognized. Acknowledging its important role in the discovery of the renormalization of interactions between massive vector mesons, we nevertheless tried to argue in this paper that further progress in this area requires new ideas beyond those of standard gauge theory, which are of a more algebraic (or local quantum physical) than differential geometric kind. As a compromise between the standard approach and a future ghost-free framework we presented a formulation of interactions of massive vector mesons using ghosts which extend the Wigner one-particle states of the vector meson. In this formulation the necessity of the presence of further Higgs-like physical degrees of freedom is most easily recognizable, even though they have been stripped of their role of ‘fattening’ the photons/gluons into massive vector mesons via Higgs condensates since their mass was assumed to be non-vanishing from the outset. It is interesting to note that even without this role their presence is nevertheless necessary for perturbative consistency and the Schwinger–Swieca charge-screening mechanism. Although we hope that this formulation will only be intermediary and replaced eventually by a ghost-free modular approach (the remarks in the previous section) in our view it is already, with all its shortcomings, more suitable than the standard gauge approach to highlight those points which suggest the necessity of a radically different formulation.

The viewpoint in this paper is in some ways opposite to the standard one. Whereas generally the massless case is used as the basic reference and the mass of vector mesons is attributed to the Higgs mechanism in gauge theories which follows more the geometric logic of classical field theory, we have favoured a particle viewpoint which emphasizes the physical particle interpretation and leaves the more difficult off-shell infrared rearrangements in the massless limit to future more detailed investigations. Whether a completely ghost-free formulation (which fully explains the gauge-restricted structure of interacting higher-spin objects as consequences of the same basic local quantum physical properties which hold for low spins) is feasible, only the future can tell.

We also mentioned various historical precedents to our point of view. In this context it is interesting to add a quotation from a paper of Anderson [1]. Concerning the possible non-intrinsic aspect of the Higgs mechanism for massive photons in comparison to plasmons he writes: ‘How, then, if we were confined to a plasma as we are to the vacuum and could only measure renormalized quantities, might we try to determine whether, before turning on the effects of electromagnetic interaction, A (the vector potential) had been a massless gauge field and ...? As far as we can see this is not possible.’ The lack of intrinsic physical meaning of Higgs condensates was also the main issue of the well known Elitzur theorem.

If by our observations on massive vector mesons we succeed in redirecting some of the attention which has been given to geometric/mathematical aspects towards the many open and interesting conceptual local quantum problems, the time it took for writing this paper was well spent.

Acknowledgments

The authors thank Klaus Fredenhagen for helpful suggestions and one of us (MD) is indebted to Günter Scharf for several fruitful discussions. MD was supported by the Alexander von Humboldt Foundation and the Deutsche Forschungsgemeinschaft.

Appendix A. The S -matrix to first order

First we give an incomplete proof of the following statement (which is of general interest).

Conjecture. *Let $V(x)$ be an arbitrary Wick polynomial with derivatives (i.e. an arbitrary element of the Borchers class of the free fields) and let all free fields be massive. Then the relation*

$$\lim_{\epsilon \rightarrow 0} \int d^4x V(x)g(\epsilon x) = 0 \quad (g \in \mathcal{D}(\mathbb{R}^4), g(0) = 1) \quad (\text{A1})$$

implies that there exists another Wick polynomial (with derivatives) W_1^μ such that

$$V = \partial_\mu W_1^\mu + V_0. \quad (\text{A2})$$

Thereby V_0 is the sum of all Wick monomials $C : \partial^{a_1} A_{j_1}(x) \dots \partial^{a_n} A_{j_n}(x) : (C = \text{constant}, A_j$ is a free field with mass m_j) in V for which the masses m_{j_1}, \dots, m_{j_n} are such that the solution for $(k_1, \dots, k_n) \in \mathbb{R}^{4n}$ of

$$k_1 + \dots + k_n = 0 \quad \bigwedge \quad k_l^2 = m_{j_l}^2 \quad \forall l = 1, \dots, n \quad (\text{A3})$$

contains no non-empty open subset of the manifold $k_l^2 = m_{j_l}^2, \forall l = 1, \dots, n$.

A simple example for a term in V_0 is : $\varphi(x)^3$:, where $(\square + m^2)\varphi = 0, m > 0$.

Incomplete proof†. To simplify the notation we assume that the free fields are bosonic scalars, the generalization to other types of fields is obvious. Global factors 2π are omitted in the whole proof. We write V in the form

$$V(x) = \sum_n \sum_{j_1, \dots, j_n} P_{j_1 \dots j_n}(\partial^{x_1}, \dots, \partial^{x_n}) : A_{j_1}(x_1) \dots A_{j_n}(x_n) : |_{x_1 = \dots = x_n = x} \quad (\text{A4})$$

where the sum over n is finite, the A_j 's are free fields and $P_{j_1 \dots j_n}(\partial^{x_1}, \dots, \partial^{x_n})$ is a polynomial in the partial derivatives $\partial_\mu^{x_l}, 1 \leq l \leq n$. Obviously, we may replace $P_{j_1 \dots j_n}(\partial^{x_1}, \dots, \partial^{x_n})$ by

$$\bar{P}_{j_1 \dots j_n}(\partial^{x_1}, \dots, \partial^{x_n}) \equiv \frac{1}{n!} \sum_{\pi \in \mathcal{S}_n} P_{j_{\pi(1)} \dots j_{\pi(n)}}(\partial^{x_{\pi(1)}}, \dots, \partial^{x_{\pi(n)}}) \quad (\text{A5})$$

in (A4). In terms of annihilation and creation operators the free fields read

$$\begin{aligned} A_l(y) &= \int d^4p \delta(p^2 - m_l^2) \bar{A}_l(p) e^{ipy} \\ \bar{A}_l(p) &:= [\Theta(-p^0) a_l(-\vec{p}) + \Theta(p^0) b_l^+(\vec{p})] \end{aligned} \quad (\text{A6})$$

† Klaus Fredenhagen gave us the main idea of the proof. The method is strongly influenced by the proof of lemma 2 (including appendix B) in [7].

($a_l = b_l$ is possible), where $[b_l(\vec{p}), b_j^+(\vec{k})] = \delta_{lj} 2\sqrt{\vec{p}^2 + m_l^2} \delta^3(\vec{p} - \vec{k})$ and similar for $[a_l, a_j^+]$. Now we insert (A4)–(A6) into (A1)

$$\begin{aligned} 0 &= \lim_{\epsilon \rightarrow 0} \sum_n \int dx g(\epsilon x) \int dx_1 \dots dx_n \delta(x_1 - x, \dots, x_n - x) \\ &\quad \times \sum_{j_1, \dots, j_n} \bar{P}_{j_1 \dots j_n}(\partial^{x_1}, \dots, \partial^{x_n}) : A_{j_1}(x_1) \dots A_{j_n}(x_n) : \\ &= \sum_n \int d^4 p_1 \dots d^4 p_n \sum_{j_1, \dots, j_n} \bar{P}_{j_1 \dots j_n}(i p_1, \dots, i p_n) \delta^4(p_1 + \dots + p_n) \\ &\quad \times : \prod_{l=1}^n \delta(p_l^2 - m_{j_l}^2) [\Theta(-p_l^0) a_{j_l}(-\vec{p}_l) + \Theta(p_l^0) b_{j_l}^+(\vec{p}_l)] :. \end{aligned} \tag{A7}$$

(Here we have exchanged the order of the integrations and the limit $\epsilon \rightarrow 0$. This can be justified by considering matrix elements between wavepackets.) Taking improper matrix elements $\langle b_{r_1}^+(\vec{k}_1) \dots b_{r_s}^+(\vec{k}_s) \Omega | \dots a_{r_{s+1}}^+(-\vec{k}_{s+1}) \dots a_{r_t}^+(-\vec{k}_t) \Omega \rangle$ of (A7), where Ω is the vacuum of the free fields, we obtain†

$$0 = \bar{P}_{r_1 \dots r_t}(i k_1, \dots, i k_t) \delta^4(k_1 + \dots + k_t) \quad \text{where} \quad k_l = (\pm \sqrt{\vec{k}_l^2 + m_{r_l}^2}, \vec{k}_l) \quad \forall l \tag{A8}$$

more precisely $k_l^0 > 0$ for $l = 1, \dots, s$ and $k_l^0 < 0$ for $l = s+1, \dots, t$. If r_1, \dots, r_t are such that the corresponding term belongs to $(V - V_0)$, then the condition (A8) restricts the polynomial $\bar{P}_{r_1 \dots r_t}$ in the following way. We make a Taylor expansion with respect to $k = (k_1 + \dots + k_t)$ at $k = 0$:

$$\bar{P}_{r_1 \dots r_t}(i k_1, \dots, i k_t) = \bar{P}_{r_1 \dots r_t}(i k_1, \dots, i k_{t-1}, -i(k_1 + \dots + k_{t-1})) + i k_\mu \tilde{Q}_{r_1 \dots r_t}^\mu(k_1, \dots, k_t) \tag{A9}$$

where $\tilde{Q}_{r_1 \dots r_t}^\mu$ is also a polynomial. From (A8) we conclude (for all (r_1, \dots, r_t) which appear in $(V - V_0)$) that $\bar{P}_{r_1 \dots r_t}(i k_1, \dots, i k_{t-1}, -i(k_1 + \dots + k_{t-1}))$ vanishes on a non-empty open subset of the manifold $k_i^2 = m_{r_i}^2, \forall i = 1, \dots, t - 1$ and $(k_1 + \dots + k_{t-1})^2 = m_{r_t}^2$. We strongly presume that this implies that there exist other polynomials $P^{(l)}(k_1, \dots, k_{t-1})$ ($l = 1, \dots, t$) (for simplicity we omit the indices r_1, \dots, r_t) such that

$$\begin{aligned} \bar{P}(i k_1, \dots, i k_{t-1}, -i(k_1 + \dots + k_{t-1})) &= \sum_{l=1}^{t-1} (k_l^2 - m_{r_l}^2) P^{(l)}(k_1, \dots, k_{t-1}) \\ &\quad + ((k_1 + \dots + k_{t-1})^2 - m_{r_t}^2) P^{(t)}(k_1, \dots, k_{t-1}). \end{aligned} \tag{A10}$$

For vanishing masses this conjecture has been proved by Buchholz and Fredenhagen (in appendix (b) of [7])‡. By means of (A10) we conclude

$$\begin{aligned} \bar{P}(i k_1, \dots, i k_t) \prod_{l=1}^t \delta(k_l^2 - m_{r_l}^2) &= [((k_1 + \dots + k_{t-1})^2 - k_t^2) P^{(t)}(k_1, \dots, k_{t-1}) \\ &\quad + i k_\mu \tilde{Q}^\mu(k_1, \dots, k_t)] \prod_{l=1}^t \delta(k_l^2 - m_{r_l}^2) \\ &= i(k_1 + \dots + k_t)_\mu Q^\mu(i k_1, \dots, i k_t) \prod_{l=1}^t \delta(k_l^2 - m_{r_l}^2) \end{aligned} \tag{A11}$$

† Usually there are several possibilities to contract the creation and annihilation operators. All give the same contribution due to the permutation symmetry of $\bar{P}_{r_1 \dots r_t}$ (A5).

‡ Unfortunately, a straightforward generalization of their proof does not work. However, for our purpose we do not need the validity of (A10).

where $Q^\mu(i k_1, \dots, i k_t) := -i(k_1 + \dots + k_{t-1} - k_t)^\mu P^{(t)}(k_1, \dots, k_{t-1}) + \tilde{Q}^\mu(k_1, \dots, k_t)$. Next we symmetrize Q^μ according to (A5) and denote the result by \bar{Q}^μ . Equation (A11) still holds true if we replace Q^μ by \bar{Q}^μ . Summing up we obtain

$$\begin{aligned} (V - V_0)(x) &= \sum_n \sum_{j_1, \dots, j_n} \int d^4 p_1 \dots d^4 p_n i(p_1 + \dots + p_n)_\mu \bar{Q}_{j_1 \dots j_n}^\mu(ip_1, \dots, ip_n) \\ &\quad \times e^{i(p_1 + \dots + p_n)x} : \prod_{l=1}^n \delta(p_l^2 - m_{j_l}^2) [\Theta(-p_l^0) a_{j_l}(-\vec{p}_l) + \Theta(p_l^0) b_{j_l}^+(\vec{p}_l)] : \\ &= \partial_\mu^x \sum_n \sum_{j_1, \dots, j_n} \bar{Q}_{j_1 \dots j_n}^\mu(\partial^{x_1}, \dots, \partial^{x_n}) : A_{j_1}(x_1) \dots A_{j_n}(x_n) : |_{x_1 = \dots = x_n = x} \end{aligned}$$

(we set $\bar{Q}_{r_1 \dots r_t}^\mu := 0$ if (r_1, \dots, r_t) belongs to V_0) which is the assertion (A2). \square

We do not need the general result (A1)–(A3), we only deal with very special situations. Physical consistency (31) to first order means

$$\lim_{\epsilon \rightarrow 0} \int d^4 x g(\epsilon x) [Q, W(x)]|_{\ker Q} = 0. \quad (\text{A12})$$

Let us first consider the *quadrilinear* terms in $V(x) := [Q, W(x)]$, i.e. $[Q, W(x)^{(4)}]$, where $W^{(4)} \sim : AAAA : , \sim : AAu\bar{u} : , \sim : u\bar{u}u\bar{u} : , \sim : \phi\phi\phi\phi : , \sim : \phi\phi AA : , \sim : \phi\phi u\bar{u} : , \sim : \phi\phi H^2 : , \sim : AAH^2 : , \sim : u\bar{u}H^2 : , \sim : H^4 :$ (the upper index (4) refers to the quadrilinear terms). They satisfy the kinematical equations (A3) for a sufficiently large set of momenta, i.e. they belong to $(V - V_0)^\dagger$. The polynomials $P_{j_1 \dots j_n}$ (A4) (and hence also \bar{P} (A5)) are of the form

$$\bar{P}(ik_1, \dots, ik_4) = c_0 \quad \text{or} \quad \bar{P}^\mu(ik_1, \dots, ik_4) = \sum_{i=1}^4 c_i k_i^\mu$$

where c_0 and c_i are constants. Taking matrix elements of (A12) with improper two-particle states, where the incoming state is an element of $\ker Q$, we obtain (A8), which implies $\bar{P} = 0$ or $\bar{P}^\mu = c(k_1^\mu + \dots + k_4^\mu)$ ($c = \text{constant}$). Hence (A12) implies that there exists a Wick polynomial $W_1^{(4)v}$ such that

$$[Q, W(x)^{(4)}] = i\partial_\nu W_1^{(4)v}(x).$$

This is the Q -divergence condition to first order (29). In addition $W^{(4)}$ must satisfy the requirements (a)–(e) listed in section 2. The only solution is $\sim : H^4 : [20]$, but such a term will be excluded by physical consistency (31) to second order.

A *trilinear* term in $W(x)$ (with masses m_1, m_2 and m_3) fulfils the kinematical equations (A3) for a sufficient large set iff there exists a permutation $\pi \in \mathcal{S}_3$ such that

$$m_{\pi(3)} > m_{\pi(1)} + m_{\pi(2)}. \quad (\text{A13})$$

However, physical consistency to higher orders will require $m_1 = m_2 = m_3 =: m$ apart from the H -couplings ($m_H > 0$ will not be restricted; see section 2). Neglecting this exception we conclude from (A7) that the trilinear part $S_1^{(3)}$ of the S -matrix to first order vanishes in the adiabatic limit, i.e.

$$S_1^{(3)} = \lim_{\epsilon \rightarrow 0} \int d^4 x g(\epsilon x) W^{(3)}(x) = 0 \quad (\text{A14})$$

and $[Q, S_1^{(3)}] = 0$ (31) is trivially satisfied. To get restrictions on $W^{(3)}$ we must consider physical consistency (31) to orders $n \geq 2$, especially for second-order tree diagrams.

\dagger For a free field φ note that $[Q, \varphi(x)]$ satisfies the same field equation as φ and hence has the same mass.

Appendix B. Determination of the parameters in W and \mathcal{F}_{int} by requiring $[Q, \mathcal{F}_{\text{int}}(x; W)] = 0$

The calculation is lengthy, hence we concentrate on the essential points. We determined the parameters in the ansatz for W (41) and \mathcal{F}_{int} (67) by inserting these expressions into the physicality requirement (69) to lowest perturbative orders. Once a parameter is fixed, its value will be used in the following calculations without mentioning it.

- To zeroth order the condition is trivially fulfilled due to[†]

$$[Q, F_d^{\mu\nu}] = 0. \quad (\text{B1})$$

- To first order the equation

$$[Q, \mathcal{F}_d^{(1)}(x)] + i \int dx_1 [Q, R(W(x_1); F_d(x))] = 0 \quad (\text{B2})$$

is required. The second term contains tree diagrams only. By means of (60) they can be written in the form

$$R(W(x_1); F_d^{\mu\nu}(x)) = \Theta(x^0 - x_1^0) [F_d^{\mu\nu}(x), W(x_1)] + \delta(x - x_1) i C \frac{1}{2} (f_{bcd} - f_{cbd}) : A_b^\mu(x_1) A_c^\nu(x_1) : \quad (\text{B3})$$

where C is an arbitrary constant. The terms $\sim \Theta(x^0 - x_1^0)$ have propagators $\Delta^{\text{ret}}(x - x_1)$, $\partial\Delta^{\text{ret}}(x - x_1)$, $\partial\partial\Delta^{\text{ret}}(x - x_1)$, where Δ^{ret} is the retarded fundamental solution of the Klein–Gordon equation, $\text{supp } \Delta^{\text{ret}} \subset \bar{V}^+$. The term $\sim C\delta(x - x_1)$ is due to the non-uniqueness of the extension of $T(W(x_1)F_d^{\mu\nu}(x))$ to the diagonal $x_1 = x_2$. This is similar to the tree-normalization terms (49). The tensor f_{bcd} which is the same as in the ansatz for W^\ddagger is antisymmetrized according to the normalization condition $\mathcal{F}_{d\text{int}}^{\mu\nu} = -\mathcal{F}_{d\text{int}}^{\nu\mu}$ (iv). Now let ψ_1, ψ_2 be arbitrary free fields and $y_1, y_2 \in \mathbb{R}^4$, $(y_j - x)^2 < 0$ ($j = 1, 2$). By commuting (B2) with $\psi_1(y_1)$ and $\psi_2(y_2)$ we obtain

$$0 = \int dx_1 [[[Q, [F_d(x), W(x_1)]], \psi_1(y_1), \psi_2(y_2)] \Theta(x^0 - x_1^0). \quad (\text{B4})$$

We conclude that for $x_1 \in (x + \bar{V}^-) \setminus \{x\}$ the expression $[Q, [F_d(x), W(x_1)]]$ must be a divergence of a local operator with respect to x_1 , i.e. there must exist a two-point distribution $R_{1d}^\tau(x_1, x)$ with

$$\Theta(x^0 - x_1^0) [Q, [F_d(x), W(x_1)]] = \partial_{x_1}^{x_1} R_{1d}^\tau(x_1, x) + \delta(x - x_1) P_d(x) \quad (\text{B5})$$

where $P_d(x)$ is a Wick polynomial (with constant coefficients) in free fields (possibly with a derivative). In contrast to the adiabatic limit of W (A1) there is no kinematical restriction (in the sense of (A3)) here, because one of the free-field operators in $W(x_1)$ is contracted with $F(x)$ and, hence, the corresponding momentum must not be on-shell. We have not worked out a general proof (in the style of appendix A) of the step from (B4) to

[†] If we had not have made the assumption $c_k^{(0)} = 0$ (i), we would find here that most of the zeroth-order coefficients must vanish. It would remain $\mathcal{F}_d^{\mu\nu(0)} = F_d^{\mu\nu} + x_{da}^{(0)} F_a^{\mu\nu} H + z_{da}^{(0)} : F_a^{\mu\nu} H^2 :$ only. $x^{(0)}$ and $z^{(0)}$ are then forced to vanish by higher orders of (69). However, this complicates the calculation a lot and the assumption $c_k^{(0)} = 0$ is reasonable. So we do not make this effort here.

[‡] Here we use an additional normalization condition, namely that the term $\sim C\delta(x - x_1) : AA :$ has the same colour tensor as the corresponding non-local term (i.e. the term in $\Theta(x^0 - x_1^0) [F^{\mu\nu}(x), W(x_1)]$ with the same external legs).

[§] Note that every operator-valued distribution $G(x_1, \dots, x_n)$, $(x_j \in \mathbb{R}^4)$ can be written as divergence of a *non-local* operator, e.g. $G(x_1, \dots, x_n) = \partial_{x_1}^{x_1} \int dy \partial^{\nu} D^{\dots}(x_1 - y) G(y, x_2, \dots, x_n)$, where D^{\dots} is a fundamental solution of the wave equation.

(B5). However, the explicit calculations show that this conclusion is correct[†]. Inserting (B5) and (B3) into (B2) it results

$$[Q, \mathcal{F}_d^{\mu\nu(1)}(x)] + iP_d^{\mu\nu}(x) - C\frac{1}{2}(f_{bcd} - f_{cbd})[Q, :A_b^\mu(x)A_c^\nu(x):] = 0 \tag{B6}$$

which means that the ‘local terms’[‡] must satisfy the condition (B2) separately. Later when we shall know more about W (which determines P_d by (B5)), we will compute most of the coefficients in $\mathcal{F}_d^{(1)}$ (67) and the normalization constant C from (B6). Inserting (B1) into (B5) we find that W must fulfil

$$[F_d(x), [Q, W(x_1)]] = \partial_\tau^{x_1} R_{1d}^\tau(x_1, x) \quad \text{for } x_1 \in (x + \bar{V}^-) \setminus \{x\}. \tag{B7}$$

Obviously, this condition is truly weaker than $[Q, W] =$ (divergence of a Wick polynomial) (29). However, the higher orders of the physicality requirement (69) yield more information about W .

- To second order (69) reads

$$[Q, \mathcal{F}_d^{(2)}(x)] + i \int dx_1 [Q, R(W(x_1); \mathcal{F}_d^{(1)}(x))] + \frac{1}{2}i^2 \int dx_1 dx_2 [Q, R(W(x_1)W(x_2); F_d(x))] = 0. \tag{B8}$$

The tree diagrams in the third term have the structure[§]

$$R(W(x_1)W(x_2); F_d(x))|_{\text{tree}} = \sum \mathcal{D}_1^{\text{ret}}(x - x_1)\mathcal{D}_2^{\text{ret}}(x_1 - x_2) : B_1(x_1)B_2(x_2)B_3(x_2) : + (x_1 \leftrightarrow x_2) \tag{B9}$$

where $\mathcal{D}_j^{\text{ret}} = \Delta^{\text{ret}}, \partial\Delta^{\text{ret}}, (\partial\partial\Delta^{\text{ret}} + C_j\delta)$ (C_j are arbitrary constants) and with free fields B_k (possibly with derivative), i.e. $B_k \in \{A^\mu, \partial A, \partial\partial A, (\partial)u, (\partial)\tilde{u}, (\partial)\phi, (\partial)H\}$. All other diagrams in (B8) have less than three legs at the vertex/vertices $\neq x$. Now let $y_l \in \mathbb{R}^4, l = 1, 2, 3$, with $(y_l - x)^2 < 0$ and let $\psi_l, l = 1, 2, 3$, be arbitrary free fields. In the triple commutator of (B8) with $\psi_1(y_1), \psi_2(y_2)$ and $\psi_3(y_3)$ only the terms (B9) survive (the terms $\sim C_1\delta(x - x_1)$ in (B9) do not contribute either)

$$0 = \int dx_1 dx_2 [[[[Q, R(W(x_1)W(x_2); F_d(x))|_{\text{tree}}, \psi_1(y_1)], \psi_2(y_2)], \psi_3(y_3)] - W(x_2)[F_d(x), W(x_1)]|_{\text{tree}}, \psi_1(y_1)], \psi_2(y_2)], \psi_3(y_3)] \tag{B10}$$

where we have inserted (60), (61) and the causal factorization (4) of the time-ordered products due to $x \notin \{x_1, x_2\} + \bar{V}^-$. Similarly to the step from (B4) to (B5) we conclude that $[Q, R(W(x_1)W(x_2); F_d(x))|_{\text{tree}}]$ must be a sum of divergences (with respect to x_1 or x_2) of local operators for $x_1, x_2 \in (x + \bar{V}^-) \setminus \{x\}$. Now we transform (B10) by means of identities of the kind

$$[Q, (M(x)N(y))|_{\text{tree}}] = [Q, M(x)N(y)]|_{\text{tree}} \tag{B11}$$

$$= ([Q, M(x)]N(y))|_{\text{tree}} + (M(x)[Q, N(y)])|_{\text{tree}} \tag{B12}$$

[†] This remark also concerns the analogous (more complicated) steps from (B10) to (B13) and from (B29) to (B30) in second and third order.

[‡] We set quotation marks because the splitting of a distribution in a local and a non-local part is non-unique.

[§] Note that totally retarded products (60) contain connected diagrams only.

where M and N are arbitrary Wick polynomials. Then using (B1) and (B7) we find the condition

$$[F_d(x), [Q, T(W(x_1)W(x_2))]]|_{\text{tree}} - [Q, W(x_1)][F_d(x), W(x_2)]|_{\text{tree}} - [Q, W(x_2)][F_d(x), W(x_1)]|_{\text{tree}} = \text{div}_{x_1} + (x_1 \leftrightarrow x_2) \quad (\text{B13})$$

for $x_1, x_2 \in (x + \bar{V}^-) \setminus \{x\}$, where div_y means some divergence of a local operator with respect to y . Next we specialize to the subregion $x_1^0 > x_2^0$, where $T(W(x_1)W(x_2))$ factorizes. By means of (B7) again we see that W must satisfy

$$[[F_d(x), W(x_1)], [Q, W(x_2)]]|_{\text{tree}} = \text{div}_{x_1} + \text{div}_{x_2} \quad (\text{B14})$$

for $x_1, x_2 \in (x + \bar{V}^-) \setminus \{x\}, x_1^0 > x_2^0$.

Inserting the ansatz for W into this condition one finds that (B14) yields the same restrictions on the parameters in W as the Q -divergence condition to first order (29):

$$[Q, W] = \text{divergence of a Wick polynomial.} \quad (\text{B15})$$

In other words, we obtain the same results for the parameters in W as in (46), especially $f_{abc} = \epsilon_{abc}$. We recall that these values of the parameters are not only necessary for (B15), they are also sufficient. Hence (B7) is also fulfilled. By inserting (B15) into (B13) we find the condition

$$[F_d(x), [Q, T(W(x_1)W(x_2))]]|_{\text{tree}} = \text{div}_{x_1} + (x_1 \leftrightarrow x_2) \quad \text{for } x_1, x_2 \in (x + \bar{V}^-) \setminus \{x\}. \quad (\text{B16})$$

This is a necessary but not sufficient condition for $[Q, T(W(x_1)W(x_2))]|_{\text{tree}} = \text{div}_{x_1} + (x_1 \leftrightarrow x_2)$. To gain full information on this latter condition we need to go to third order.

- However, first we return to the ‘local terms’ to first order (B6), because we shall need the validity of this equation. We explicitly calculate the terms on the left-hand side of (B5), transform them into divergence form (as far as possible) and obtain[†]

$$P_d^{\mu\nu} = \epsilon_{bcd}[F_b^{\mu\nu} u_c - A_b^\mu \partial^\nu u_c + A_b^\nu \partial^\mu u_c]. \quad (\text{B17})$$

Inserting this into (B6) we find that (B6) is fulfilled iff the normalization

$$C = -1 \quad (\text{B18})$$

is chosen and the first-order coefficients in $\mathcal{F}_{d\text{int}}^{\mu\nu}$ take the values

$$\begin{aligned} t_{dbc}^{(1)} &= -\frac{1}{m_c} \epsilon_{bcd} & \tilde{t}^{(1)} &= 0 & v^{(1)} &= 0 & \tilde{v}^{(1)} &= 0 \\ w^{(1)} &= 0 & \tilde{w}^{(1)} &= 0 & w'^{(1)} &= 0 & & \\ \tilde{x}^{(1)} &= 0 & y^{(1)} &= 0 & \tilde{y}^{(1)} &= 0 & y'^{(1)} &= 0 & \tilde{z}^{(1)} &= 0. \end{aligned} \quad (\text{B19})$$

$x^{(1)}$ and $z^{(1)}$ are still arbitrary. Note that (B7), (B18) and (B19) are not only necessary for (B2), together they are also sufficient.

[†] By the experience of [2] we know that all contributions to P_d come from the contraction of $F_d(x)$ with $\partial A(x_1)$ in the first term of $W(x_1)$ (41).

- We now consider the condition (69) to third order, i.e. the analogous equation to (B2) and (B8). By commuting this equation with arbitrary free fields $\psi_1(y_1), \psi_2(y_2), \psi_3(y_3), \psi_4(y_4), (y_l - x)^2 < 0 \forall l$, we obtain

$$\begin{aligned}
 0 = & \frac{i^2}{2!} \int dx_2 dx_3 [[[[[[Q, R(W(x_2)W(x_3); \mathcal{F}^{(1)}(x))], \psi_1(y_1)], \psi_2(y_2)], \psi_3(y_3)], \psi_4(y_4)] \\
 & + \frac{i^3}{3!} \int dx_1 dx_2 dx_3 [[[[[[Q, R(W(x_1)W(x_2)W(x_3); F(x))], \psi_1(y_1)], \\
 & \dots \psi_4(y_4)]. \tag{B20}
 \end{aligned}$$

Only tree diagrams of the following types contribute (we use the same notation as in (B9)): in the first term

$$\sum \mathcal{D}_1^{\text{ret}}(x - x_2) \mathcal{D}_2^{\text{ret}}(x - x_3) : B_0(x) B_1(x_2) B_2(x_2) B_3(x_3) B_4(x_3) : + (x_2 \leftrightarrow x_3) \tag{B21}$$

(where $B_0(x)$ is a free field or $\equiv 1$), in the second term

$$\sum \mathcal{D}_3^{\text{ret}}(x - x_1) \mathcal{D}_4^{\text{ret}}(x_1 - x_2) \mathcal{D}_5^{\text{ret}}(x_1 - x_3) : B_5(x_2) B_6(x_2) B_7(x_3) B_8(x_3) : + \dots \tag{B22}$$

and

$$\sum \mathcal{D}_6^{\text{ret}}(x - x_1) \mathcal{D}_7^{\text{ret}}(x_1 - x_2) \mathcal{D}_8^{\text{ret}}(x_2 - x_3) : B_9(x_1) B_{10}(x_2) B_{11}(x_3) B_{12}(x_3) : + \dots \tag{B23}$$

where the dots denote terms obtained by cyclic permutations. In the first term in (B20) we take (60) and the causal factorization based on $x \notin (\{x_2, x_3\} + \bar{V}^-)$ into account. By means of (B15) it results in

$$\begin{aligned}
 \frac{i^2}{2!} \int dx_2 dx_3 [[[[[[[[Q, \mathcal{F}^{(1)}(x)], T(W(x_2)W(x_3))] - W(x_2) [[Q, \mathcal{F}^{(1)}(x)], W(x_3)] \\
 - W(x_3) [[Q, \mathcal{F}^{(1)}(x)], W(x_2)]]]]]]]_{\text{tree}, \psi_1(y_1)} \dots, \psi_4(y_4)]. \tag{B24}
 \end{aligned}$$

Additionally, we have used the fact that due to (B21) only the disconnected diagram of $T(W(x_2)W(x_3))$ contributes, hence $[\dots [\mathcal{F}^{(1)}(x), [Q, T(W(x_2)W(x_3))]]] \dots, \psi_4(y_4)] = \text{div}_{x_2} + \text{div}_{x_3}$ by (B15).

In the second term in (B20) the situation is more complicated. The diagrams of the type (B23) obey the causal factorization based on $x \notin (\{x_1, x_2, x_3\} + \bar{V}^-)$. However, for the diagrams of the type (B22) we only know $x \notin (\{x_2, x_3\} + \bar{V}^-)$ (or promulgated configurations). To fix the position of the third vertex we consider two smooth functions h_1, h_2 with

$$\begin{aligned}
 1 = h_1(y) + h_2(y) \quad \forall y \in x + \bar{V}^- \quad h_1 \in \mathcal{D}(\mathbb{R}^4) \quad x \in \text{supp } h_1 \\
 \text{supp } h_2 \cap x + \bar{V}^+ = \emptyset \quad (y_l - z)^2 < 0 \quad \forall z \in \text{supp } h_1 \quad l = 1, 2, 3, 4. \tag{B25}
 \end{aligned}$$

With that the fact that the second term in (B20) can be written in the form

$$\begin{aligned}
 \frac{i^3}{3!} \int dx_1 dx_2 dx_3 (h_1(x_1)h_2(x_2)h_2(x_3) + h_2(x_1)h_1(x_2)h_2(x_3) + h_2(x_1)h_2(x_2)h_1(x_3) \\
 + h_2(x_1)h_2(x_2)h_2(x_3)) [\dots [Q, R(W(x_1) \dots; F(x))]_{\text{tree}}] \dots, \psi_4(y_4)]. \tag{B26}
 \end{aligned}$$

In the terms with a factor $h_1(x_l)$ only the diagrams of the type (B22) contribute, and we may insert the causal factorization due to $\{x, x_l\} \cap (\{x_j, x_k\} + \bar{V}^-) = \emptyset$ (where

$\{l, j, k\} = \{1, 2, 3\}$). So the contribution of $(h_1(x_1)h_2(x_2)h_2(x_3) + h_2(x_1)h_1(x_2)h_2(x_3) + h_2(x_1)h_2(x_2)h_1(x_3))$ in (B26) reads

$$\begin{aligned} & \frac{i^3}{2!} \int dx_1 dx_2 dx_3 h_1(x_1)h_2(x_2)h_2(x_3) [\dots \{[Q, R(W(x_1); F(x))], T(W(x_2)W(x_3))\} \\ & \quad - W(x_2)[[Q, R(W(x_1); F(x))], W(x_3)] \\ & \quad - W(x_3)[[Q, R(W(x_1); F(x))], W(x_2)]\}_{|_{\text{tree}} \dots, \psi_4(y_4)} \end{aligned} \tag{B27}$$

where we have used $[Q, T(W(x_2)W(x_3))] = \text{div}_{x_2} + \text{div}_{x_3}$, which holds true here for the same reason as in (B24). From (B2) and (B3) we know

$$\begin{aligned} & [Q, \mathcal{F}^{(1)}(x)] + i \int dx_1 h_1(x_1)[Q, R(W(x_1); F(x))] \\ & = -i \int dx_1 h_2(x_1)[Q, R(W(x_1); F(x))] \\ & = -i \int dx_1 h_2(x_1) \partial_\tau^{x_1} [F(x), W_1^\tau(x_1)] \end{aligned} \tag{B28}$$

where we have used $[Q, W] = \partial_\tau W_1^\tau$ (B15). We now insert (B27), (B26) and (B24) into (B20). By means of (B28) we obtain

$$\begin{aligned} 0 = & \frac{i^3}{3!} \int dx_1 dx_2 dx_3 h_2(x_1)h_2(x_2)h_2(x_3) [\dots \{[Q, R(W(x_1)W(x_2)W(x_3); F(x))]\}_{|_{\text{tree}}} \\ & \quad - \partial_\tau^{x_1} ([F(x), W_1^\tau(x_1)], T(W(x_2)W(x_3))) \\ & \quad - W(x_2)[[F(x), W_1^\tau(x_1)], W(x_3)] - W(x_3)[[F(x), W_1^\tau(x_1)], W(x_2)]\}_{|_{\text{tree}}} \\ & \quad - \partial^{x_2}(\dots) - \partial^{x_3}(\dots)\}_{\dots, \psi_4(y_4)} \end{aligned} \tag{B29}$$

where $\partial^{x_2}(\dots)$ and $\partial^{x_3}(\dots)$ are obtained from $\partial^{x_1}(\dots)$ by cyclic permutation. The next step is analogous to the step from (B4) to (B5), but more complicated because one has to care about the cancellation of the boundary terms. Taking the freedom in the choice of h_2 into account, equation (B29) is equivalent to the existence of a local operator $R_3^\tau(x_1; x_2, x_3; x)$ with

$$\begin{aligned} & [Q, R(W(x_1)W(x_2)W(x_3); F(x))]\}_{|_{\text{tree}}} = \partial_\tau^{x_1} R_3^\tau(x_1; x_2, x_3; x) + \text{cyclic permutations} \\ & \quad \forall x_1, x_2, x_3 \in (x + \bar{V}^-) \setminus \{x\} \end{aligned} \tag{B30}$$

because the contribution from, for example, $\partial^{x_1}(\dots)$ in (B29) is equal to

$$-\frac{i^3}{3!} \int dx_1 dx_2 dx_3 (\partial_\tau h_2)(x_1)h_2(x_2)h_2(x_3) [\dots R_3^\tau(x_1; x_2, x_3; x) \dots, \psi_4(y_4)]. \tag{B31}$$

Let us explain this latter statement. First note that due to (62) and (B25) there is only a contribution in (B31) for $x_1, x_2, x_3 \in (x + \bar{V}^-) \setminus \{x\} \wedge x_1 \notin (\{x_2, x_3\} + \bar{V}^-)$. In this region we have

$$\begin{aligned} & R(W(x_1)W(x_2)W(x_3); F(x)) = [[F(x), W(x_1)], T(W(x_2)W(x_3))] \\ & \quad - W(x_2)[[F(x), W(x_1)], W(x_3)] - W(x_3)[[F(x), W(x_1)], W(x_2)] \end{aligned} \tag{B32}$$

due to causal factorization. Now we take into account that (in this region) $\partial_\tau^{x_1} R_3^\tau(x_1; x_2, x_3; x)$ comes from the $[Q, W(x_1)]$ terms in $[Q, R(WWW; F)]$ (where

$R(WWW; F)$ is given by (B32)). However, by replacing $W(x_1)$ by $\partial_\tau W_1^\tau(x_1)$ in (B32) we obtain exactly the $\partial^{x_1}(\dots)$ term in (B29).

By means of causal factorization and (B1), (B15) and (B16), the condition (B30) can be written in the form

$$\begin{aligned} & \{[F(x), [Q, T(W(x_1)W(x_2)W(x_3))]] + [Q, T(W(x_2)W(x_3))][W(x_1), F(x)] \\ & \quad + [Q, T(W(x_1)W(x_2))][W(x_3), F(x)] \\ & \quad + [Q, T(W(x_1)W(x_3))][W(x_2), F(x)]\}_{\text{tree}} \\ & = \text{div}_{x_1} + \text{cyclic permutations} \end{aligned} \tag{B33}$$

for $x_1, x_2, x_3 \in (x + \bar{V}^-) \setminus \{x\}$. Next we specialize to the subregion $x_1 \notin (\{x_2, x_3\} + \bar{V}^-)$. Again by causal factorization and (B16) we find that $T(W(x_2)W(x_3))$ must fulfil

$$[[F(x), W(x_1)], [Q, T(W(x_2)W(x_3))]]_{\text{tree}} = \text{div}_{x_1} + \text{div}_{x_2} + \text{div}_{x_3} \tag{B34}$$

for $x_1, x_2, x_3 \in (x + \bar{V}^-) \setminus \{x\} \wedge x_1 \notin (\{x_2, x_3\} + \bar{V}^-)$. Inserting the explicit expression for W (with the values of the parameters obtained so far) one sees that this condition is equivalent to

$$[Q, T(W(x_2)W(x_3))|_{\text{tree}}] = \text{div}_{x_2} + (x_2 \leftrightarrow x_3). \tag{B35}$$

As we mentioned in section 2 the latter requirement is satisfied iff the masses agree $m \equiv m_1 = m_2 = m_3$ (47), the parameters of the H -coupling take the values (48)†, there is no term $\sim : H^4 :$ in W and if the normalization constants in the tree-normalization terms (49) are suitably chosen (up to λ (50) they are uniquely fixed by (B35)). So far there remain two free parameters in W and $T(WW)|_{\text{tree}}$: s (41) (cf (48)) and λ (50). In section 2 they have been determined by

$$[Q, T(W(x_1)W(x_2)W(x_3))|_{\text{tree}}] = \text{div}_{x_1} + \text{cyclic permutations}. \tag{B36}$$

By inserting (B35) into (B33) we find the weaker requirement

$$[F(x), [Q, T(W(x_1)W(x_2)W(x_3))|_{\text{tree}}]] = \text{div}_{x_1} + \text{cyclic permutations} \tag{B37}$$

for $x_1, x_2, x_3 \in (x + \bar{V}^-) \setminus \{x\}$. However, the latter condition does not determine s and λ . One needs to consider the physicality condition (69) to fourth order. There the values (51) for s and λ are obtained, as can be seen by an analogous procedure. So we obtain exactly the same results for the parameters in W and the tree normalizations in $T(WW)$ as in section 2. We recall that they agree completely with the interaction Lagrangian obtained by the Higgs mechanism.

- Now we have explained how to handle the essential difficulties which appear in the determination of the parameters in W (41) and in \mathcal{F}_{int} (67) by the physicality (69). We do not give further details. Instead we give a heuristic summary and the remaining results. To n th order the requirement (69) reads

$$0 = [Q, \mathcal{F}_d^{(n)}(x)] + \sum_{l=1}^n \frac{i^l}{l!} \int dx_1 \dots dx_l [Q, R(W(x_1) \dots W(x_l); \mathcal{F}_d^{(n-l)}(x))] \tag{B38}$$

where $\mathcal{F}_d^{(0)} \equiv F_d$.

† In the following calculations the undetermined sign κ in (48) is chosen to be $\kappa = 1$.

- (I) In the ‘non-local terms’ ($x_j \neq x$ for at least one j) we can apply the causal factorization (4) of the time-ordered products. Then the non-local terms in (B38) cancel due to

$$[Q, T(W(x_1) \dots W(x_k))] = \text{sum of divergences of local operators} \quad 1 \leq k \leq n \quad (\text{B39})$$

and the physicality (B38) to lower orders $< n$ (see, e.g., (B28) and (B29)). We have seen that for the tree diagrams to lowest orders the condition (B39) is not only sufficient for the cancellation, it is also necessary. However, the lowest-order tree diagrams of (B39) fix the parameters in W and $T(WW)|_{\text{tree}}$, as was shown in [2].

For the cancellation of the ‘non-local terms’ in (B38) it is not important which kind of physical fields we require to exist, the properties (i)–(vi) of \mathcal{F}_{int} (67) are essentially not needed. Hence we conjecture that already the existence of *any* non-trivial observable fixes the interaction.

- (II) The cancellation of the remaining terms in (B38), which are ‘local’ (i.e. their support is the total diagonal $x_j = x, \forall j$), requires a suitable normalization of the time-ordered products $T(W(x_1) \dots W(x_l); \mathcal{F}_d^{(n-l)}(x))$ (see, e.g., (B18)) and a suitable choice of the parameters in $\mathcal{F}_d^{(n-l)}, l = 0, 1, \dots, n-1$. In this way the latter parameters are *uniquely* determined (at least to lowest orders).

So the cancellation of the ‘local terms’ to second order requires

$$\begin{aligned} t^{(2)} = 0 \quad \tilde{t}^{(2)} = 0 \quad v_{dabc}^{(2)} &= -\frac{1}{2m^2} [(\delta_{da} - \frac{1}{2}mx_{da}^{(1)})\delta_{bc} - \frac{1}{2}(\delta_{dc}\delta_{ba} + \delta_{db}\delta_{ca})] \\ \tilde{v}^{(2)} = 0 \quad w^{(2)} = 0 \quad \tilde{w}^{(2)} = 0 \quad w'^{(2)} = 0 \quad \tilde{x}^{(2)} = 0 & \quad (\text{B40}) \\ y_{dbc}^{(2)} = \frac{1}{2m^2}\epsilon_{dbc} - \frac{1}{m}x_{da}^{(1)}\epsilon_{abc} \quad \tilde{y}^{(2)} = 0 \quad y'^{(2)} = 0 \quad \tilde{z}^{(2)} = 0 \quad z^{(1)} = 0. \end{aligned}$$

The very last equation results from the terms $\sim g^2 : F(x)u(x)H^2(x) :$. In this case the cancellation is of a special kind as explained below. The other equations are obtained analogously to (B6) and (B19). They come from bilinear and trilinear terms. $x^{(1)}, x^{(2)}$ and $z^{(2)}$ are still arbitrary.

To third order the terms $\sim g^3 : F(x)u(x)\phi(x)\phi(x) :$, ($\sim g^3 : F(x)u(x)H^2(x) :$, respectively) cancel iff

$$x_{da}^{(1)} = \frac{1}{m}\delta_{da} \quad \left(z_{da}^{(2)} = \frac{1}{4m^2}\delta_{da} \text{ respectively} \right). \quad (\text{B41})$$

Similarly one finds to fourth order that $x^{(2)}$ must vanish.

Summing up (B19), (B40) and (B41) we obtain the result (70). However, we have not checked that the cancellation of the ‘local terms’ in (B38) can be achieved for loop diagrams and for tree diagrams to higher orders. For the tree diagrams with more than three external legs (e.g. the very last equation in (B40) and the terms considered in (B41)) there is no contribution from $\mathcal{F}_d^{(n)}$ in (B38) due to our ansatz (67). Hence only the parameters in $\mathcal{F}_d^{(n-l)}, l \geq 1$ are available (besides normalization constants). However, most of the latter parameters have already been fixed (or restricted) by (B38) to lower orders. (For example, to achieve the cancellation of the terms considered in (B41), it is important that $x^{(1)}$ ($z^{(2)}$, respectively) has not been fixed previously.) So we have shown the existence of observables $\mathcal{F}_{d\text{int}}^{\mu\nu}, d = 1, 2, 3,$

† Mainly parameters in $\mathcal{F}_d^{(n)}$ are determined; most of the parameters in $\mathcal{F}_d^{(n-l)}, 1 \leq l \leq n-1$, have already been fixed in earlier steps of the induction.

only partially. However, there is a strong hint that this holds true for all terms to all orders from the identification (given in the main text) of our physical fields $\mathcal{F}_{d\text{int}}^{\mu\nu}$ (70) as gauge-invariant fields in the framework of spontaneous symmetry breaking of the $SU(2)$ gauge symmetry.

References

- [1] Anderson P W 1963 Plasmons, gauge invariance and mass *Phys. Rev.* **130** 439
- [2] Aste A, Dütsch M and Scharf G 1997 On gauge invariance and spontaneous symmetry breaking *J. Phys. A: Math. Gen.* **30** 5785
Dütsch M and Scharf G 1999 Perturbative gauge invariance: the electroweak theory *Ann. Phys., Lpz.* **8** 359
Aste A, Dütsch M and Scharf G 1999 Perturbative gauge invariance: the electroweak theory II *Ann. Phys., Lpz.* **8** 389
- [3] Becchi C, Rouet A and Stora R 1975 Renormalization of the Abelian Higgs–Kibble model *Commun. Math. Phys.* **42** 127
Becchi C, Rouet A and Stora R 1976 Renormalization of gauge theories *Ann. Phys., NY* **98** 287
- [4] Bell J S 1973 High energy behavior of tree diagrams in gauge theories *Nucl. Phys. B* **60** 427–36
- [5] Bogoliubov N N and Shirkov D V 1959 *Introduction to the Theory of Quantized Fields* (New York: Interscience)
- [6] Brunetti R and Fredenhagen K 1997 Interacting quantum fields in curved space: renormalization of ϕ^4 *Proc. Conf. Operator Algebras and Quantum Field Theory (Accademia Nazionale dei Lincei, Rome, July 1996)* Preprint gr-qc/9701048
Brunetti R and Fredenhagen K 2000 *Commun. Math. Phys.* **208** 623–61
(Brunetti R and Fredenhagen K 1999 Preprint math-ph/9903028)
- [7] Buchholz D and Fredenhagen K 1977 A note on the inverse scattering problem in quantum field theory *Commun. Math. Phys.* **56** 91
- [8] Cornwall J M, Levin D N and Tiktopoulos G 1974 Derivation of gauge invariance from high-energy unitarity bounds on the S matrix *Phys. Rev. D* **10** 1145
- [9] Dubois-Violette M Private communication
- [10] Dütsch M and Fredenhagen K 1999 *Commun. Math. Phys.* **203** 71
Dütsch M and Fredenhagen K 1999 *Commun. Math. Phys.* **203** 71–105
(Dütsch M and Fredenhagen K 1998 Preprint hep-th/9807078)
- [11] Dütsch M, Hurth T, Krahe F and Scharf G 1993 Causal construction of Yang–Mills theories. I. *Nuovo Cimento A* **106** 1029
- [12] Dütsch M, Hurth T and Scharf G 1995 Causal construction of Yang–Mills theories. Unitarity IV *Nuovo Cimento A* **108** 737
- [13] Dütsch M 1996 On gauge invariance of Yang–Mills theories with matter fields *Nuovo Cimento A* **109** 1145
- [14] Dütsch M 1997 Slavnov–Taylor identities from the causal point of view *Int. J. Mod. Phys. A* **12** 3205
- [15] Dütsch M 1996 Non-uniqueness of quantized Yang–Mills theories *J. Phys. A: Math. Gen.* **29** 7597
- [16] Dütsch M 1989 Causal functional integral for the S -matrix of QED *Nuovo Cimento A* **102** 1469
- [17] Epstein H and Glaser V 1973 The role of locality in perturbation theory *Ann. Inst. H Poincaré A* **19** 211
- [18] Epstein H and Glaser V 1976 Adiabatic limit in perturbation theory *Renormalization Theory* ed G Velo and A S Wightman pp 193–254
- [19] Faddeev L D and Popov V N 1967 Feynman diagrams for the Yang–Mills field *Phys. Lett. B* **25** 29
- [20] Grigore D R 1998 On the uniqueness of the non-Abelian gauge theories in Epstein–Glaser approach to renormalization theory Preprint hep-th/9806244
Grigore D R 1998 The standard model and its generalizations in Epstein–Glaser approach to renormalization theory Preprint hep-th/9810078
Grigore D R 1999 The standard model and its generalizations in Epstein–Glaser approach to renormalization theory II: the fermion sector and the axial anomaly Preprint hep-th/9903206
- [21] Haag R 1992 *Local Quantum Physics* (Berlin: Springer)
- [22] Horuzhy S S and Voronin A V 1989 Remarks on mathematical structure of BRST theories *Commun. Math. Phys.* **123** 677
- [23] Krahe F 1996 A causal approach to massive Yang–Mills theories *Acta Phys. Polon. B* **27** 2453
- [24] Karowski M and Weisz P 1978 Exact form factors in $(1+1)$ -dimensional field theoretic models with soliton behavior *Nucl. Phys. B* **139** 455
- [25] Kugo T and Ojima I 1979 Local covariant operator formalism of non-Abelian gauge theories and quark confinement problem *Prog. Theor. Phys. Suppl.* **66** 1

- [26] Lewellyn Smith C H 1973 *Proc. 5th Hawaii Topical Conf. Partical Physics* ed P N Dobson *et al* (Honolulu: University of Hawaii Press)
- [27] Reeh H and Schlieder S 1961 Bemerkungen zur Unitäräquivalenz von Lorentzinvarianten Feldern *Nuovo Cimento* **22** 1051
- [28] Requardt M 1976 Symmetry conservation and integrals over local charge densities in quantum field theory *Commun. Math. Phys.* **50** 259
- [29] Sakurai J J 1960 Theory of strong interactions *Ann. Phys., NY* **11** 1
- [30] Scharf G 1999 General massive gauge theory *Nuovo Cimento A* **112** 619
(Scharf G 1999 *Preprint* hep-th/9901140)
- [31] Scharf G and Wellmann M 1999 Spin-2 gauge theories and perturbative gauge invariance *Preprint* hep-th/9903055
- [32] Schroer B and Wiesbrock H-W 2000 Modular theory and geometry *Rev. Math. Phys.* **12** 139–581
Schroer B and Wiesbrock H-W 2000 Modular constructions of quantum field theories with interactions *Rev. Math. Phys.* **12** 301
(Schroer B and Wiesbrock H-W 1998 *Preprint* hep-th/9812251)
- [33] Schroer B 1997 Motivations and physical aims of algebraic QFT *Ann. Phys.* **255** 270
- [34] Schroer B 1998 QFT at the turn of the century: old principles with new concepts (an essay on local quantum physics) *J. Math. Phys* at press (*Preprint* FU-SFB288)
- [35] Schroer B 1999 Modular wedge localization and the $d = 1 + 1$ formfactor program *Ann. Phys., NY* **275** 190
- [36] Schroer B 1999 New concepts in particle physics from solution of an old problem *Preprint* hep-th/9908021
- [37] Schweber 1961 *An Introduction to Relativistic Quantum Field Theory* part 4 (New York: Harper and Row)
- [38] Schwinger J 1962 Gauge invariance and mass *Phys. Rev.* **125** 397
- [39] Smirnov F A 1992 *Formfactors in Completely Integrable Models of Quantum Field Theory (Advanced Series Mathematical Physics vol 14)* (Singapore: World Scientific)
- [40] Steinmann O 1971 *Perturbation Expansions in Axiomatic Field Theory (Lecture Notes in Physics vol 11)* (Berlin: Springer)
- [41] Stora R 1997 Local gauge groups in quantum field theory: perturbative gauge theories, talk given at the workshop *Local Quantum Physics at the Erwin-Schrödinger-Institute (Vienna)*
- [42] Streater R F and Wightman A S 1964 *PCT, Spin and Statistics and All That* (New York: Benjamin)
- [43] Swieca J A 1976 Charge screening and mass spectrum *Phys. Rev. D* **13** 312
- [44] 't Hooft G 1998 The glorious days of physics, renormalization of gauge theories *Preprint* hep-th/9812203
- [45] 't Hooft G and Veltman M T 1972 Combinatorics of gauge fields *Nucl. Phys. B* **50** 318
- [46] Weinberg S 1996 *The Theory of Quantized Fields* vol I (Cambridge: Cambridge University Press)