Anomalies in Lagrangian field theory

Michael Danos*
Center for Radiation Research, National Bureau of Standards, Gaithersburg, Maryland 20899

Lawrence C. Biedenharn
Department of Physics, Duke University, Durham, North Carolina 27706
(Received 23 October 1986; revised manuscript received 14 July 1987)

We investigate the manner in which anomalies enter the description of experimental data by means of effective Lagrangians, with particular attention to the known result that δ-function-type contact singularities of Feynman integrals do not contribute to the S matrix. The development uses the two-photon decay of singlet positronium as an example. We show that the anomaly becomes essential on replacing positronium effectively by a point particle; the anomaly arises automatically in the resulting effective Lagrangian as a remnant of the underlying QED structure of positronium.

I. INTRODUCTION

It is well known, at the present time, that anomalies play an important, even decisive, role in many physical contexts.[1] In this paper we will demonstrate that anomalies properly belong to effective Lagrangian field theories. By contrast, in renormalizable Lagrangian field theory the existence of any anomalies is considered as a sign of inconsistency; 't Hooft,[2] for example, has given conditions for anomaly cancellation. The main point of the present paper is that the correct treatment of contact singularities in a renormalizable Lagrangian field theory ensures that no anomalies, in fact, exist. The central point for this demonstration is the observation that the essence of the equations of motion is to allow the calculation of the future development of a system from the knowledge of the present state.[3] This is the meaning of the Schrödinger equation (written in the Schrödinger picture)

$$\psi(t+\Delta t)=\psi(t)-i\Delta t H(t)\psi(t) \; ; \tag{1.1}$$

given the wave function at time t, $\psi(t)$, we can compute it at time $t+\Delta t$. Thus, Eq. (1.1) can be integrated forward in time.

On the other hand, if one wants to integrate backward in time one must use instead the equation

$$\psi(t-\Delta t)=\psi(t)+i\Delta t H(t)\psi(t) \; . \tag{1.2}$$

There is no difference between Eqs. (1.1) and (1.2) if one deals with simple functions, e.g., with fixed-particle-number wave functions, as in nonrelativistic quantum mechanics. In that case both Eqs. (1.1) and (1.2) can be written as

$$\frac{\partial}{\partial t}\psi(t)=-iH(t)\psi(t) \; . \tag{1.3}$$

However, one has to return to Eq. (1.1) or (1.2) when dealing with singular quantities. Specifically, the Schwinger-Tomonaga evolution operator $U(t,t_0)$ is highly singular. Therefore one must write

$$U(t+\Delta t,t_0)=U(t,t_0)-i\Delta t H_I(t)U(t,t_0) \tag{1.4}$$

for integration forward in time. [In Eq. (1.4) $H_I(t)$ is the interaction operator in the interaction picture.] We now rewrite Eq. (1.4) by the replacement $t\to t-\Delta t$, and $t+\Delta t\to t$:

$$U(t,t_0)=-i\Delta t H_I(t-\Delta t)U(t-\Delta t,t_0)+U(t-\Delta t,t_0) \tag{1.4a}$$

which means that to compute U at time t we need to know the system at time $t-\Delta t$. To continue, we express $U(t-\Delta t,t_0)$ as

$$U(t-\Delta t,t_0)=-i\Delta t H_I(t-2\Delta t)U(t-2\Delta t,t_0)+U(t-2\Delta t,t_0) \tag{1.4b}$$

and so on, to obtain

$$U(t,t_0)=1-i\sum_{n=1}^{N-1}\Delta t H_I(t-n\Delta t)U(t-n\Delta t,t_0) \; , \tag{1.5}$$

where

$$N\Delta t=t-t_0 \tag{1.6}$$

and where we have used the boundary condition

$$\lim_{t\to t_0} U(t,t_0)=1 \; , \tag{1.7}$$

which yields the 1 in (1.5), and which requires the absence of the term $n=N$ in the sum. The term $n=0$ is absent in view of (1.4a).

We now can safely perform the limit $N\to\infty$ to obtain the Schwinger-Tomonaga integral equation

$$U(t,t_0)=1-i\int_{(t_0)}^{(t)}dt_1 H_I(t_1)U(t_1,t_0) \; , \tag{1.8}$$

where (t) indicates that the integral is to be taken *over the interval open at the limit of integration* $t_1 \to t$, that is, one must use $t_1 \in (t_0,t)$ rather than $T_1 \in [t_0,t]$. In this way, and only in this way, we retain the essential aspect of Eq. (1.5), which is required in order for (1.5) to be integrable, and represent a deterministic equation.

When iterating (1.8) to generate the Neumann series, one finds

$$U(t,t_0)=1-i\int_{(t_0)}^{(t)}dt_1 H_I(t_1)+(-i)^2\int_{(t_0)}^{(t)}dt_1\int_{(t_0)}^{(t_1)}dt_2 H_I(t_1)H_I(t_2) +(-i)^3\int^{(t)}dt_1\int_{(t_0)}^{(t_1)}dt_2\int_{(t_0)}^{(t_2)}dt_3 H_I(t_1)H_I(t_2)H_I(t_3)+\cdots \tag{1.9}$$

which also can be written as[4]

$$t>t_1>t_2>t_3>\cdots \; . \tag{1.10}$$

The crucial point now is that the open interval of integration must be preserved when supplementing the "triangular" region of integration. Thus, for example, for the second-order term,

$$\begin{aligned}\int_{(t_0)}^{(t)}dt_1\int_{(t_0)}^{(t_1)}dt_2 H_I(t_1)H_I(t_2) &= \int_{(t_0)}^{(t)}dt_2\int_{(t_0)}^{(t_2)}dt_1 H_I(t_2)H_I(t_1)\\ &=\tfrac{1}{2}\left[\int_{(t_0)}^{(t)}dt_1\int_{(t_0)}^{(t_1)}dt_2 H_I(t_1)H_I(t_2)+\int_{(t_0)}^{(t)}dt_2\int_{(t_0)}^{(t_2)}dt_1 H_I(t_2)H_I(t_1)\right]\\ &\equiv\tfrac{1}{2}\emptyset\int_{(t_0)}^{(t)}dt_1\int_{(t_0)}^{(t)}dt_2 T(H_I(t_1)H_I(t_2)) \; .\end{aligned} \tag{1.11}$$

The symbol $\emptyset$ means that in performing the integration over the time-ordered product (T product) the points $t_1=t_2$ are to be *outside* of the region of integration. We denoted this by the notation $\emptyset T\int$ in (1.11), and call it an "open integral of a time-ordered product." Accordingly, *contact δ-function-type singularities of the T product do not* contribute to the S matrix (see Appendix A for some further remarks). The consequences of using open integration for the Tomonaga-Schwinger equation for QED were investigated in Ref. 3. Here we would like to analyze the relation of the perturbation theory anomalies to the contact singularities. Before embarking on that task we should, however, make a few remarks.

The term "anomaly" is at the present used in two quite distinct contexts. The one context is in the context of perturbation theory, where the prototype is the Adler-Bell-Jackiw triangle anomaly.[5,6] The discussion of this anomaly will be the subject of the present paper. The other case in which anomalies are employed is in the context of effective Lagrangians.

There exist cases where the effective Lagrangian forbids certain reactions, which are permitted by the underlying more fundamental theory. (Also, the opposite situation is possible.) The appearance of terms in an effective Lagrangian or Hamiltonian which violate some basic symmetries of the complete theory is very well known.

For example, elimination of a part of Hilbert space renders complex the effective Hamiltonian which is valid in the retained part of Hilbert space, resulting in a nonunitary evolution operator—violating both energy and probability conservation. This is well understood and is a useful technique for treating complicated physical systems. Another example, of current interest, is the effective Lagrangian of the [SU(3)] σ model ("Skyrmions"). Witten[7] has shown that this Lagrangian admits a larger symmetry than QCD, and he adds to the Skyrme Lagrangian an anomaly term, removing the unwanted extra symmetry. This type of anomaly, a term added to remedy deficiencies in an effective Lagrangian, will not be elaborated on here.

In Sec. II we reanalyze the structure of the axial-vector-current triangle graph—both in position and momentum space—and show that the associated anomaly is due to a contact singularity. Thus, as discussed above, the singularity lies outside the integration region and does not contribute to the S matrix.

In Sec. III we discuss the example of positronium decay. Section IV contains a summary and our conclusions. Appendix A gives some illustrations of the contradictions one encounters when retaining the contact singularity, while in Appendix B we sketch the character of the contact singularities.

II. THE AXIAL-VECTOR-CURRENT TRIANGLE GRAPH

The characteristics of the axial vertex were carefully analyzed by Adler.[5] We follow closely his presentation.

This triangle graph arises in the context of the standard spinor quantum electrodynamics, when considering the time-ordered products

$$j_\mu^5=\langle T(\psi(x)j_\mu^5(0)\bar\psi(y))\rangle_0 \; , \tag{2.1}$$

$$j^5=\langle T(\psi(x)j^5(0)\bar\psi(y))\rangle_0 \; , \tag{2.2}$$

where

$$j_\mu^5(x)=:\bar\psi(x)\gamma_\mu\gamma_5\psi(x): \; , \tag{2.3}$$

$$j^5(x)=:\bar\psi(x)\gamma_5\psi(x): \; . \tag{2.4}$$

The anomaly arises when evaluating the triangle graph, Fig. 1, where Γ is either γ_5 or $\gamma_\mu\gamma_5$. We assume that

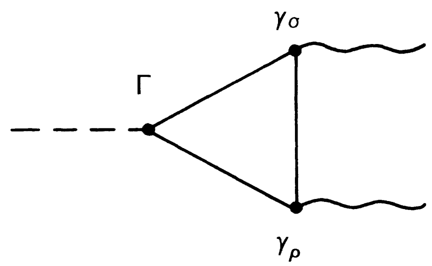

FIG. 1. Axial-vector triangle graph.

the calculation is done in the context of the S matrix, that is, in the description of some physical process. Then (2.1) and (2.2) must be evaluated by open integrals. In position space we have (for $\Gamma=\gamma_\mu\gamma_5$ replace φ by $\partial^\mu\varphi$)

$$M=\emptyset T\int d^4x\, d^4y\, d^4z\, \mathrm{Tr}[\varphi(x)S_F(x-y)\gamma_\sigma A^\sigma(y) \times S_F(y-z)\gamma_\rho A^\rho(z) \times S_F(z-x)\Gamma] \quad (2.5)$$

while in momentum space (for the meaning of open integration in momentum space see Ref. 3)

$$R=\emptyset\int d^4r\, \mathrm{Tr}\left[\frac{1}{\gamma(r+k_1)-m}\gamma_0\frac{1}{\gamma r-m} \times\gamma_\rho\frac{1}{\gamma(r-k_2)-m}\Gamma\right], \quad (2.6)$$

which is, except for a numerical constant, Eq. (16) of Ref. 5. From (2.6) one sees immediately that this *graph* is "superficially" linearly divergent. (It is well known that more than logarithmically divergent Feynman integrals are inherently ambiguous.[8])

In Ref. 5 Adler has carefully analyzed this integral in momentum space. The essential discussion centers around Eq. (13) of Ref. 5. Using the momentum-space-defining equation of the Feynman propagator, Eq. (9) of Ref. 5, to rewrite (2.6), he shows that the anomaly arises from the term

$$A=\int d^4r\, \mathrm{Tr}\left[\gamma_5\gamma^{(1)}\frac{1}{\gamma(r+p_1)-m_0}\gamma^{(2)}\frac{1}{\gamma(r+p_2)-m}-\gamma^5\gamma^{(1)}\frac{1}{\gamma(r+p_1+q)-m_0}\gamma^{(2)}\frac{1}{\gamma(r+p_2+q)-m_0}\right]. \quad (2.7)$$

The other terms satisfy the Ward-Takahashi identity. Upon the shift of the integration variable $r\to r+q$ in the first part, one sees that the two parts of (2.7) cancel. Being associated with a more than logarithmically divergent graph, such a shift "is not necessarily a valid operation." Of course, it is a valid procedure for convergent and only logarithmically divergent integrals.

We now demonstrate that the linearly divergent part arises from a contact singularity and hence does not contribute to the S matrix. That is, if one employs open integrals the singularity is absent in both (2.5) and (2.6).

We now re-do Alder's analysis taking care to incorporate the open integration. The logic of our argument will be as follows. (1) Logarithmically divergent and convergent Feynman integrals are invariant under shifts of the integration variables. (2) The anomaly contained in Eq. (2.6) disappears; i.e., one recovers the nonanomalous Ward-Takahashi identity when shifting the integration variable in (2.7). (3) Hence, the anomaly is associated with the linearly divergent part of (2.6). (4) By considering the position space expression of this Feynman integral, we show that the most singular part of (2.5) indeed is a metadistribution and according to the Jauch-Rohrlich theorem[8] is ambiguous. It is linearly divergent and has the form of a "softened" vacuum-polarization term of QED (multiplied by x and hence linearly and not quadratically divergent), and it is of the contact type, in the same way as is the QED vacuum-polarization term. (5) Since the contact singularity does not contribute to the open integral, only the logarithmically divergent and convergent terms contribute. They can be evaluated by shifting the integration variables in the momentum-space expressions. (6) It has been shown by Adler that the ambiguity associated with the linearly divergent part can be employed to manipulate the magnitude of the anomaly. As this part does not contribute in view of the open integration it remains to be shown that the other parts are anomaly-free; i.e., that the anomaly originates in the contact singularity.

We perform the analysis of the triangle graph in position space. Writing

$$\Gamma^5=\gamma^5+\Lambda^5(x\mid y,z), \quad (2.7a)$$

$$\Gamma^5_\mu=\gamma_\mu\Gamma^5, \quad (2.7b)$$

we have

$$\Lambda^5_{\nu\lambda}(x\mid y,z)=\gamma_5 S_F(x-y)\gamma_\nu S_F(y-z)\gamma_\lambda S_F(z-x). \quad (2.8)$$

Here the Ward-Takahashi identity demands

$$(\gamma_\mu\partial/\partial x_\mu)\Lambda^5_{\nu\lambda}(x\mid y,z)=2im\,\Lambda^5_{\nu\lambda}(x\mid y,z). \quad (2.9)$$

We now check whether Eq. (2.9) holds for the triangle graph. To that end we recall the defining equation of the Feynman propagator:

$$\left[i\gamma_\mu\frac{\partial}{\partial x_\mu}-m\right]S_F(x-y)=S_F(x-y)\left[i\gamma_\mu\frac{\overleftarrow{\partial}}{\partial x_\mu}-m\right]=\delta^4(x-y). \quad (2.10)$$

By means of this equation we can evaluate the four-divergence of $\gamma_\mu \Lambda^5_{\nu\lambda}$:

$$\gamma_\mu \frac{\partial}{\partial x_\mu}\Lambda^5_{\nu\lambda}(x\mid y,z) = -\gamma_5\left[\gamma_\mu \frac{\partial}{\partial x_\mu} S_F(x-y)\right]\gamma_\nu S_F(y-z)\gamma_\lambda S_F(z-x) + \gamma_5 S_F(x-y)\gamma_\nu S_F(y-z)\gamma_\lambda \times\left[-S_F(z-x)\gamma_\mu \frac{\overleftarrow{\partial}}{\partial(-x_\mu)}\right]$$
$$= 2im\,\Lambda^5_{\nu\lambda}(x\mid y,z) + i\gamma_5\delta^4(x-y)\gamma_\nu S_F(y-z)\gamma_\lambda S_F(z-x) + i\gamma_5 S_F(x-y)\gamma_\nu S_F(y-z)\gamma_\lambda\delta^4(z-x)\ . \tag{2.11}$$

The last two terms in (2.11) violate the Ward-Takahashi identity. They seem not to contribute since

$$\mathrm{Tr}\gamma_5\gamma\cdot x\gamma_\nu\gamma\cdot x\gamma_\lambda = 0\ , \tag{2.12}$$

and formally the Ward-Takahashi identity seems to be satisfied. However, this result is not necessarily true owing to the linear divergence of the graph.

We now show that the linearly divergent part of the graph arises from a contact-type singularity. To that end we define $B^\nu_k(y-z)$ by

$$A^\nu(y) = A^\nu(z) + (y-z)^k B^\nu_k(y-z)\ . \tag{2.13}$$

Herewith we obtain (for the pseudoscalar current omit γ_μ and the derivative on ϕ)

$$\int d^4x\,d^4y\,d^4z[\partial_\mu\phi(x)]\gamma_\mu\Lambda^5_{\nu\lambda}(x\mid y,z)A^\nu(y)A^\lambda(z) \to \int d^4x\,d^4y\,d^4z[\partial_\mu\phi(x)]\gamma_\mu\Lambda^5_{\nu\lambda}(x\mid y,z)A^\nu(z)A^\lambda(z)$$
$$+\int d^4x\,d^4y\,d^4z[\partial_\mu\phi(x)]\gamma_\mu\Lambda^5_{\nu\lambda}(x\mid y,z)(y-z)^k B^\nu_k(y-z)A^\lambda(z)$$
$$\equiv \mathrm{I} + \mathrm{II}\ . \tag{2.14}$$

This way we have isolated the linearly divergent part of the graph in the first term; the second term is only logarithmically divergent. We now continue with the first term. There we can perform the folding over d^4y using the identity

$$\int d^4y\, S_F(x-y)\gamma_\sigma S_F(y-z) = i(x_\sigma - z_\sigma)S_F(x-z)\ , \tag{2.15}$$

which is the position space form of the well-known identity [see, e.g., Eq. (8.52), p. 167 of Bjorken and Drell[9]] arising in the context of the Ward-Takahashi identities, to obtain

$$\mathrm{Tr}\int d^4y\,\gamma_\mu\Lambda^5_{\nu\lambda}(x\mid y,z) \to \mathrm{Tr}\gamma_5\gamma_\mu S_F(x-z)(x-z)_\nu\gamma_\lambda S_F(z-x)\ . \tag{2.16}$$

Ignoring for the moment the γ matrices, both for the pseudoscalar and the pseudovector case (2.16) has the analytical form of $(x-z)\Pi(x-z)$, where $\Pi(x-z)$ is the vacuum-polarization tensor of QED. It has been shown in Ref. 3 to have a quadratically divergent contact singularity at $x-z=0$ which is of the form $\delta(x_\mu - z_\mu)\delta(x^\mu - z^\mu)\delta^4(x-z)$. (Since it contains the square of a δ function it is not a distribution; it was called a "metadistribution" in Ref. 3, see Appendix B.) The factor $(x-z)$ changes the quadratic to a linear divergence, but it does not change the character of being a contact singularity. Hence this term does not contribute to the open integral, for both the pseudovector and the pseudoscalar interaction. [A change from the contact to a noncontact type would require the appearance of a derivative operator; e.g., $\gamma\cdot\partial\Pi(x-z)$ would not be of the contact type and would contribute to the open integral.]

We now show that the logarithmically divergent term II satisfies the Ward-Takahashi identity. We have, using (2.10),

$$\gamma_\mu\frac{\partial}{\partial x_\mu}\Lambda^5_{\nu\lambda}(x\mid y,z)_{\mathrm{II}} = \gamma_\mu\frac{\partial}{\partial x_\mu}[\gamma_5 S_F(x-y)S_F(y-z)(y-z)^k\gamma_\lambda S_F(z-x)]$$
$$= -2im\gamma_5 S_F(x-y)\gamma_\nu S_F(y-z)(y-z)^k\gamma_\lambda S_F(z-x) - i\gamma_5\delta^4(x-y)\gamma_\nu S_F(y-z)(y-z)^k\gamma_\lambda S_F(z-x)$$
$$-i\gamma_5 S_F(x-y)\gamma_\nu S_F(y-z)(y-z)^k\gamma_\lambda\delta^4(z-x)\ . \tag{2.17}$$

The last two terms do not contribute owing to Eq. (2.12). Here, however, in contrast with Eq. (2.11), this result is believable since the expressions are only maximally logarithmically divergent and hence give unambiguous results.

The above analysis is based on the validity of Eq. (2.15). This equation is an identity. However, the limits $y \to x$ and $y \to z$ seem to pose problems. Folding is a "benign" operation in that it "softens" the singularity character. Recall that in position space we have, for small $|x|$,

$$S_F(x) = \frac{i}{4\pi^2}\left[\frac{2\gamma x}{(x^2 - i\epsilon)^2} + \frac{m}{2}\frac{1-2m\gamma x}{x^2 - i\epsilon} + \cdots\right]. \tag{2.18}$$

Thus, the Feynman propagator has the singularity structure of a distribution. Hence the folding (2.15) with measure d^4y is perfectly well defined and leads to an unambiguous result, as long as $x - z \neq 0$. Indeed, the highest singularity of (2.18) is $\sim 1/|x|^3$ vs $1/|x|^2$ of (2.15). It is exactly the point $x = z$ (with measure d^4x) which yields the linear divergence. Omitting this point in the open integral validates the use of the result (2.15), and in consequence (2.16).

This way we have seen that the breaking of the Ward-Takahashi identity is a very subtle effect indeed. Considering (2.11) we see that the Ward-Takahashi-violating terms arrive as the indefinite product $0 \times \infty$, where the zero is the result of the trace operation, and the ∞ the linear divergence. A similar remark could be made with respect to the momentum-space expression (2.7). However, we must reemphasize that this term with all its inherent ambiguity is present in the expression of the triangle graph. It is only in the context of the S matrix that it does not contribute in view of its character of being a contact by singularity.

We would like to conclude this section by a caveat: the validity of the Fourier transform has been well established for functions and for distributions. It is an open question what are its characteristics when used for metadistributions.

III. POSITRONIUM DECAY

A transparent example of the triangle anomaly arises in the decay of the singlet positronium state into two photons. This is the QED analogue of the two-photon pion decay; the prototype anomaly. An effective Lagrangian in this case arises by replacing the (interaction picture) "positronium field,"[10,11] written in terms of the (interaction picture) electron-positron fields

$$\Phi(X,\xi) = \int d^3p\, d^3q\, e^{i(px+qy)}\tilde{f}(p,q) \times \bar{v}(p)u(q)d(p)b(q) + \text{H.c.} \tag{3.1}$$

by a local pseudoscalar field $\varphi(x)$, that is, by ignoring the presence of the internal (relative) coordinate ξ. (Note that actually no particle is located at X, which is the center-of-mass coordinate.)

Formally, the difference between the "fundamental" [in terms of $\Phi(X,\xi)$] and the effective Lagrangian [in terms of $\varphi(X)$] description will show up in the form of the replacement of the "true" commutation relations of the "positronium field" by local commutation relations, in our example by the assumption (replacing X by x)

$$[\varphi(x), \psi(y)]_- = [\varphi(x), \bar{\psi}(y)]_- = 0 . \tag{3.2}$$

This assumption of course contradicts the true commutation relation which can be computed from (3.1). We find, for example,

$$[\Phi(x,y,t), \psi_\alpha(z,t)]_- = f_{\alpha\beta}(x-y)\bar{\psi}_\beta(x,t)\delta^3(z-y) , \tag{3.3}$$

which is "canonical" only for large separations $|x-z|$ where $f_{\alpha\beta}$ vanishes. The replacement of (3.3) by (3.2) can lead to important differences in the evaluation of the S matrix.

Owing to the smallness of the binding energy we can safely use free Dirac spinors. Thus we can take, for a positronium at rest, $p_+ = -p_-$,

$$\psi = \int \tilde{f}(p)\bar{v}(p_+)u(p_-)d^3p \tag{3.4}$$

and, since $k_1 = -k_2 \equiv k$, we have, for the two-photon decay (Fig. 1),

$$M = \int d^4p\, \tilde{f}(p)\bar{v}(p_+)\frac{1}{\gamma p - m}\gamma_\mu A_1^\mu \frac{1}{\gamma(p-k) - m} \times \gamma_\nu A_2^\nu \frac{1}{\gamma(p+k) - m}u(p_-) . \tag{3.5}$$

We rewrite (3.5) as

$$M = \text{Tr}\rho(p)\frac{1}{\gamma p - m}\gamma_\mu A_1^\mu \frac{1}{\gamma(p-k) - m} \times \gamma_\nu A_2^\nu \frac{1}{(p+k) - m} , \tag{3.6}$$

while the "effective positronium" decay would be given by (2.6). In (3.6) we have introduced the direct product (N=normalization)

$$\rho(p) = N\tilde{f}(p)u(p_-)\bar{v}(p_+) = N\tilde{f}(p)\begin{bmatrix} \dfrac{\sigma\cdot p_+}{2m} & 1 \\ \dfrac{\sigma\cdot p_+}{2m}\dfrac{\sigma\cdot p_-}{2m} & \dfrac{\sigma\cdot p_-}{2m} \end{bmatrix} . \tag{3.7}$$

When writing ρ in terms of the γ matrices and their products, we recognize that it contains more than γ_5, or $\gamma_\mu\gamma_5$, and in addition, in contrast with (2.6), Eq. (3.7) is burdened by a form factor $\tilde{f}(p)$. At any rate, Eq. (3.7) can be used to arrive at the familiar weak-binding (i.e., nonrelativistic) result of the positronium decay.[12]

A more complete description would be that of Fig. 2, which yields the lowest-order form of the positronium form factor, $\tilde{f}(p)$, as

$$\tilde{f}(p) = \frac{g_{\mu\nu}}{p^2 - i\epsilon} . \tag{3.8}$$

With this form factor the "triangle" graph of the positronium decay is actually "superficially convergent,"

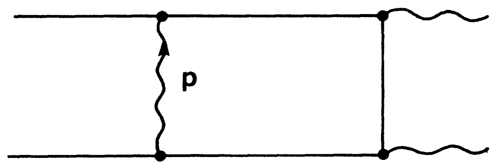

FIG. 2. Lowest-order graph for positronium wave function.

since power counting yields $\int d^4p/p^5$. The replacement of this lowest-order $\tilde{f}(p)$ by the "exact" $\tilde{f}(p)$, i.e., by the "positronium wave function," will not introduce a qualitative change, since QED is a renormalizable theory. Positronium decay is therefore anomaly-free.

IV. DISCUSSION AND CONCLUSIONS

In Sec. III we have demonstrated the manner in which the structure of a pseudoscalar particle, here singlet positronium, enters in the expression for the two-photon decay process. The central expression is (3.7) which gives the structure of the "effective positronium-lepton vertex." If one were to describe positronium by a local field $\varphi(x)$ with the commutation relations (3.2) one would have the choice of (2.1) or (2.2) as "effective positronium-lepton vertex." Then according to the Veltman-Sutherland theorem[13] the photon decay of positronium would not take place. One sees immediately that the reason for this is that the effective positronium field $\varphi(x)$ lacks the internal coordinates, the relative coordinate ξ of (3.1), which is needed to describe the internal structure of positronium. Hence, when working with the effective Lagrangian, which lacks this coordinate, one must add "by hand" an interaction term, in essence the difference $\Delta=(3.7)-(2.1)$, which would then allow the decay of positronium. This difference Δ cannot be derived from the effective Lagrangian; it is the "anomaly" needed to achieve agreement between theory and experiment, if one insists on using the effective Lagrangian. From arguments based on Lorentz and gauge invariance one can write down immediately a list of possible forms of Δ in terms of the fields $\varphi, \mathbf{E}, \mathbf{B}$, the simplest of which is

$$\Delta \simeq C\varphi \mathbf{E}\cdot\mathbf{B} , \tag{4.1}$$

where the structural information (contained in the wave function) appears in the guise of the constant C.

In formal terms, the effective Lagrangian would arise by the replacement (3.1), i.e., by $\Phi(x,\xi)\rightarrow\varphi(x)$. To achieve that, one would have to rewrite the Lagrangian as

$$\begin{aligned} L = {} & L(\psi)+L(A)+L_I(\psi,A)+L(\varphi) \\ & -L(\varphi)+L_I(\varphi,\psi)-L_I(\varphi,\psi) . \end{aligned} \tag{4.2}$$

One would then have

$$\begin{aligned} H = {} & H(\psi)+H(A)+H_I(\psi,A)+H'(\varphi) \\ & +H_I'(\varphi,\psi)-[H'(\varphi)+H_I'(\varphi,\psi)] \\ = {} & H_{\text{eff}}-H' \end{aligned} \tag{4.3}$$

which is still correct. The separation H_{eff} vs H' is, however, only symbolic—it is defined in terms of the matrix elements. (For example, for the positronium decay it would be given by $\langle \Phi | T | \gamma\gamma \rangle - \langle \varphi | T' | \gamma\gamma \rangle$, i.e., by the difference between the "exact" result and the result obtained using the effective Hamiltonian.)

For the effective Lagrangian approach to be a useful procedure one would require that for certain processes, say, for low-energy processes, the matrix elements of the corresponding Hamiltonians cancel, i.e., that for these processes the matrix elements of H' are "small." The danger associated with the introduction of an effective Lagrangian is visible in our example by the fact that this graph which is finite in the full theory turns into a superficially linearly divergent graph in the effective Lagrangian theory.

It is quite possible that the original Hamiltonian contains terms which allow processes which are forbidden in the effective Hamiltonian. In such cases one must retain explicitly the appropriate terms from H' of (4.3), i.e., one must "add them by hand" to H_{eff}.

A case of this kind is the two-photon decay of the pion. In the σ model it is forbidden,[12] in essence by the Noether theorem. To achieve the two-photon decay one thus must augment the σ model by the addition of a term of the form $E\cdot B\varphi$, i.e., of the "anomalous term," to the effective Hamiltonian. And, once this term has been included, the theory is consistent in that it can describe all processes which involve this anomalous interaction.[14]

To summarize, we have seen explicitly how the triangle-anomaly term in an effective theory for two-photon fermion decay arises from eliminating internal degrees of freedom. Similarly, the effective theory for two-photon pion decay arises from eliminating QCD degrees of freedom. Generally speaking, such elimination of degrees of freedom results in terms which cannot be written as local Lagrangian interactions, and can best be expressed topologically. An example is the Wess-Zumino-Witten anomaly[7,15] for SU(3) Skyrmions which is expressed as an integral over a five-dimensional manifold whose boundary is Minkowski space.

A general way of introducing anomalies has been given—albeit from a different point of view—by Wess and Zumino.[15] On the other hand, there may exist theories which at the present time are rejected because of the putative presence of perturbation theory anomalies. This circumstance itself is, however, no reason to reject the theory since such anomalies in fact are not present in the S matrix. It is reasonable to conjecture that all anomalies arise as terms in effective theories resulting from simplifying a more general theory.

ACKNOWLEDGMENTS

We would like to thank S. L. Adler for a very productive discussion. We thank M. Oberguggenberger for very interesting correspondence concerning products of distributions. The work of L.C.B. was supported in part by the National Science Foundation.

APPENDIX A

We give here some illustrations of the difficulties which arise if one includes the contributions from the contact singularity. Recall that the Tomonaga-Schwinger equation arises from the Schrödinger equation for the interaction picture state vector $|S(t)\rangle$,

$$\frac{\partial}{\partial t}|S(t)\rangle=-iH_I(t)|S(t)\rangle \tag{A1}$$

by introducing the evolution operator $U(t,t_0)$ through the defining equation

$$|S(t)\rangle=U(t,t_0)|S(t_0)\rangle \tag{A2}$$

and considering the equation obeyed by U. However, one could instead solve Eq. (A1) directly. Specifically, consider a stationary state, having the state vector $|S_\alpha(t)\rangle$. Being a stationary state it has a time dependence

$$\frac{\partial}{\partial t}|S_\alpha(t)\rangle=-iE_\alpha|S_\alpha(t)\rangle\ , \tag{A3}$$

where E_α is the energy including the shift arising from the radiative corrections. This energy is well defined for a suitably regularized theory. Using (A3), (A1) becomes an eigenvalue equation which, in principle, can be solved. The point of interest in the present context is that the resulting eigenvalue equation contains the interaction only as a single vertex—the solution is determined directly by H_I and the commonly employed interactions involve only well-defined operators, e.g., $\bar{\psi}(x)\psi(x)A(x)$. This is in contrast with the operators arising in the higher-order terms of the time-symmetrized Neumann series, e.g., operators of the form

$$T\int d^4x\, d^4y\ \bar{\psi}(x)\psi(x)A(x)\bar{\psi}(y)\psi(y)A(y)\ .$$

If one would not employ open integrals the contact term $\bar{\psi}(x)\psi(x)A(x)\bar{\psi}(x)\psi(x)A(x)$ would survive; this term is highly singular and in general violates the symmetries of the Lagrangian. Terms of this kind do not arise in (A3).

Another illustration is provided by the Tomonaga-Schwinger differential equation

$$\frac{\partial}{\partial t}U(t,t_0)=-iH(t)U(t,t_0) \tag{A4}$$

which is supposed to be obeyed by the Neumann series (1.9). The difficulty one would generate by ignoring the open integration, here, for example, expressed by replacing the conditions (1.10) by the wrong conditions $t\le t_1\le t_2\le\cdots$, that is by inclusion of the contact points, can be demonstrated as follows. Writing out in detail, for example, the second-order term of (1.9) we have

$$\int dt_1\int dt_2 H_I(t_1)H_I(t_2) =\int dt_1\int dt_2[H_I(t_1)H_I(t_2)|_{t_2\neq t_1} +W(t_1)\delta(t_2-t_1)]\ . \tag{A5}$$

Here the last term is the contact singularity. Performing the time differentiation of (A4) for this term we find

$$\frac{\partial}{\partial t}(\ \ )=H_I(t)\int dt_2 H_I(t_2)+W(t)\ . \tag{A6}$$

The presence of the contact term $W(t)$ indeed would render the expansion (1.9) not to be a solution.

APPENDIX B

Consider the equation

$$x^2f(x)=g(x),\quad g(0)\neq 0\ . \tag{B1}$$

The solution of (B1) is (measure dx implied)

$$f(x)=P\frac{g(x)}{x^2}+\Delta^{(2)}(x)\ . \tag{B2}$$

Here P is "principal value," defined by the average value of the integration along the path above and below $x=0$, while $\Delta^{(2)}(x)$ is defined by

$$x^2\Delta^{(2)}(x)=0\ . \tag{B3}$$

It contains, besides $\delta(x)$ and $\delta'(x)$, also the generalized distribution (called "metadistribution" in Ref. 3)

$$D^{(2)}(x)=\lim_{n_1\to\infty}\lim_{n_2\to\infty} d_{n_1}(x)d_{n_2}(x)\ , \tag{B4}$$

where $d_n(x)$ is a set of "good" functions[16] which has the limit

$$\lim_{n\to\infty} d_n(x)=\delta(x)\ . \tag{B5}$$

The general kth-order metadistributions arise as solutions of an equation analogous to (B1), and obey a condition analogous to (B3), when x^2 is replaced by x^k. The prescription analogous to (B4) then contains k limiting processes.

Owing to the several limiting operations in (B4) the metadistributions are inherently ambiguous, in agreement with the Jauch-Rohrlich theorem[8] concerning the ambiguity of momentum-space Feynman integrals which are more than logarithmically divergent. In fact, the metadistributions arise as the position space form of these integrals. Further details are given in Ref. 3. However, the mathematical properties of the metadistributions, in particular with respect to the Fourier transform, are still largely unknown.

---

*Present address: Duke University, Durham, NC 27706.

[1] *Anomalies, Geometry and Topology,* proceedings of the Symposium, Argonne, Illinois, 1985, edited by W. A. Bardeen and A. R. White (World Scientific, Singapore, 1985).
[2] G. 't Hooft, in *Recent Developments in Gauge Theories,* edited by G. 't Hooft *et al.* (Plenum, New York, 1980).
[3] M. Danos and J. Rafelski, Nuovo Cimento **49A**, 326 (1979).
[4] P. Roman, *Introduction to Quantum Field Theory* (Wiley, New York, 1969).
[5] S. L. Adler, Phys. Rev. **177**, 2426 (1969).
[6] J. S. Bell and R. Jackiw, Nuovo Cimento **60A**, 47 (1969).
[7] E. Witten, Nucl. Phys. **B223**, 422 (1982).

[8]J. M. Jauch and F. Rohrlich, *The Theory of Photons and Electrons* (Addison-Wesley, Reading, MA, 1955).
[9]J. D. Bjorken and S. Drell, *Relativistic Quantum Fields* (McGraw-Hill, New York, 1965).
[10]M. Danos (unpublished).
[11]M. Danos, in *Interaction Studies in Nuclei,* edited by H. Jochim and B. Ziegler (North-Holland, Amsterdam, 1975).
[12]A. Ore and J. Powell, Phys. Rev. **75**, 1696 (1953); **75**, 1963 (1953).
[13]D. G. Sutherland, Nucl. Phys. **B2**, 433 (1967).
[14]S. L. Adler, B. W. Lee, and S. B. Treiman, Phys. Rev. D **4**, 3497 (1971).
[15]J. Wess and B. Zumino, Phys. Lett. **37B**, 95 (1971).
[16]M. J. Lighthill, *An Introduction to Fourier Analysis and Generalized Functions* (Cambridge University Press, Cambridge, England, 1958).