[1]For the companion article, see L. C. McDermott and P. S. Shaffer, "Research as a guide for curriculum development: An example from introductory electricity, part I: Investigation of student understanding," Am. J. Phys. **60**, 994–1002 (1992).
[2]L. C. McDermott *et al.*, *Physics by Inquiry* (Physics Education Group, University of Washington, Seattle, WA, 1982–1992), soon to be available through a commercial publisher.
[3]L. C. McDermott *et al.*, *Tutorials in Physics* (Physics Education Group, University of Washington, Seattle, WA, 1991–1992). This is the tentative title of a set of instructional materials currently under development.
[4]A. B. Arons, *The Various Language* (Oxford, New York, 1977).
[5]See, for example, *Elementary Science Study* (Education Development Center, Newton, MA, 1968); *Science Curriculum Improvement Study* (University of California, Berkeley, CA, 1968).
[6]L. C. McDermott, "What we teach and what is learned—Closing the gap," Am. J. Phys. **59**, 301–315 (1991).
[7]For an alternative method of introducing the concept of potential difference without the use of meters, see J. Evans, "Teaching electricity with batteries and bulbs," The Phys. Teach. **15** (1), 15–22 (1978).
[8]See, for example, D. M. Shipstone and R. F. Gunstone, "Teaching children to discriminate between current and energy," in *Aspects of Understanding Electricity*, Proceedings of an International Workshop, Ludwigsburg, Germany, edited by R. Duit, W. Jung, and C. von Rhöneck (Verlag Schmidt & Klaunig, Kiel, Germany, 1984), pp. 287–297; M. S. Steinberg, "Reinventing electricity," in *Misconceptions in Science and Mathematics*, Proceedings of the first international seminar, edited by H. Helm and J. D. Novak (Cornell U. P., Ithaca, NY, 1983) pp. 388–401; P. Licht, "Teaching electrical energy, voltage and current: an alternative approach," Phys. Educ. **26**, 272–277 (1991).
[9]The use of operational definitions is discussed in A. B. Arons, *A Guide to Introductory Physics Teaching* (Wiley, New York, 1990).
[10]The use of interactive lectures in mechanics is discussed in A. Van Heuvelen, "Learning to think like a physicist: A review of research-based instructional strategies," Am. J. Phys. **59**, 891–897 (1991).
[11]Private communication, P. Laws, Dickinson College (1992).
[12]Private communication, D. Sokoloff, University of Oregon (1991).
[13]Private communication, S. Sharpe, University of Washington (1991).
[14]E. Mazur, Department of Physics, Harvard University, "Improving student understanding in introductory science courses: The peer instruction method," 1992 (draft article).
[15]For other examples of how research has been used by the Physics Education Group to guide curriculum development, see M. L. Rosenquist and L. C. McDermott, "A conceptual approach to teaching kinematics," Am. J. Phys. **55**, 407–415 (1987), L. C. McDermott and M. D. Somers, "Building a research base for curriculum development: an example from mechanics," in *Research in Physics Learning: Theoretical Issues and Empirical Studies*, Proceedings of an International Workshop, Bremen, Germany, edited by R. Duit, F. Goldberg, and H. Niedderer (IPN/Institute for Science Education, Kiel, Germany, 1992), pp. 330–355.
[16]P. S. Shaffer, Ph. D. dissertation, Department of Physics, University of Washington, 1992 (unpublished).

Renormalization of a model quantum field theory

Per Kraus[a)] and David J. Griffiths
Physics Department, Reed College, Portland, Oregon 97202

(Received 9 December 1991; accepted 30 April 1992)

Renormalization is the technique used to eliminate infinities that arise in quantum field theory. This paper shows how to renormalize a particularly simple model, in which a single mass counterterm of second order in the coupling constant suffices to cancel all divergences. The model serves as an accessible introduction to Feynman diagrams, covariant perturbation theory, and dimensional regularization, as well as the renormalization procedure itself.

I. INTRODUCTION

Quantum field theory—in particular, quantum electrodynamics—has produced by far the most exacting predictions in all of physics. The magnetic dipole moment of the electron, for example, has been calculated to 14 significant digits, and the result confirmed in the laboratory with exquisite precision.[1] And yet, a straightforward application of the basic rules leads to nonsensical infinities which must be circumvented before intelligible results can be obtained. This process, which is known as "renormalization," stands as one of the greatest triumphs of theoretical physics.[2] Unfortunately, renormalization is generally considered too difficult and sophisticated for most graduate students, let alone undergraduates. Part of the problem is that "realistic" field theories are burdened by distracting features such as spin and gauge invariance, and even artificial "textbook" examples (the so-called ϕ^3 and ϕ^4 theories)[3] involve diverting technical complications.

Our aim in this paper is to present a reasonably complete and self-contained treatment of renormalization[4] for a very simple model: the "ABC theory."[5] We hope that this study will be accessible to advanced undergraduates and to nonspecialists who would like to know (in something more than a merely qualitative sense) what renormalization is all

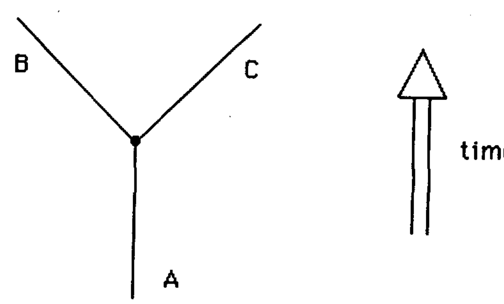

Fig. 1. The basic vertex in ABC theory.

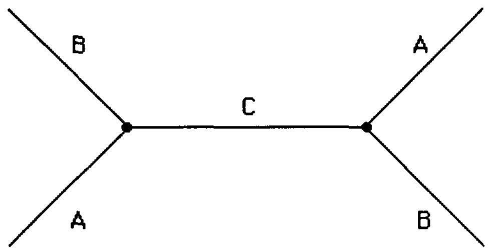

Fig. 2. AB scattering: $A+B \rightarrow A+B$.

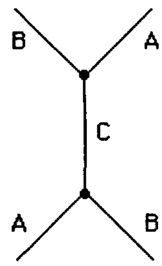

Fig. 4. Another diagram representing AB scattering.

about, and at the same time will serve as a useful introduction to the subject for graduate students embarking on a first course in quantum field theory.

In Sec. II we describe the model and develop the "Feynman Rules" for calculating lifetimes and scattering cross sections (the quantities of preeminent interest to particle physicists).[6] In Sec. III we discover that straightforward application of the Feynman Rules leads to a divergent integral, and we explore the physical significance of this result. In Sec. IV we exploit the method of "dimensional regularization" to isolate and "tame" the infinity, and in Sec. V we show how an astute modification of the Feynman Rules eliminates it altogether. Finally, in Sec. VI we demonstrate that repeated application of the same technique removes *all* the infinities that arise in the ABC model.

II. THE FEYNMAN RULES FOR ABC THEORY

Imagine a world populated by just three kinds of particles: the A (mass m_A), the B (mass m_B), and the C (mass m_C). They carry spin zero, and each is its own antiparticle (like the π^0 meson).[7] Their only interaction is represented diagrammatically in Fig. 1; by convention, time runs upward on the page,[8] so this diagram reads "A converts into B plus C:"

$$A \rightarrow B+C.$$

If the A is heavy enough ($m_A > m_B + m_C$), this describes the decay of the A. More complicated processes are obtained by making replicas of the fundamental vertex, twisting them around, and linking them together—always joining two lines of the same type. Thus, Fig. 2 describes the collision (scattering) of an A and a B:

$$A+B \rightarrow A+B.$$

(An A and a B went *in*, an A and a B came *out*; the C in this case is an unobserved "virtual" particle, which serves only to "mediate" the interaction between A and B.) Similarly, Fig. 3 represents the scattering of two A's:

$$A+A \rightarrow A+A$$

(this time there are two virtual B's and two virtual C's).

But Fig. 2 is not the only diagram corresponding to the process $A+B \rightarrow A+B$; Figs. 4 and 5 also do the job. Now, every "Feynman diagram" (as these figures are called) stands for a particular (complex) *number*: the "amplitude," $\mathscr{M}$, for the process in question. We will see in a moment how this amplitude is actually *calculated*; once this has been done we *add* the amplitudes for all the distinct diagrams with the appropriate set of external legs. For example, if we wish to determine the cross section for AB scattering, we add the amplitudes for Figs. 2, 4, 5,...everything we can make by gluing together replicas of the fundamental vertex (Fig. 1) with the generic structure of Fig. 6. Of course, there are infinitely many such diagrams. However, as we shall see, each vertex introduces a factor of the "coupling constant," g (the parameter which fixes the overall strength of the interaction), and if g is small, the more complicated diagrams, incorporating more and more vertices, contribute less and less to the total. In this case it is reasonable to hope that a good approximation to the AB cross section would be obtained by including only diagrams 2 and 4, since these are the only two-vertex diagrams for the process. If a more precise answer is required, we would include the four-vertex diagrams (such as Fig. 5). Thus the Feynman diagrams with N vertices generate the N^{th}-order term of a perturbation expansion in the coupling constant g.

The protocol for determining the amplitude $\mathscr{M}$ associated with a given Feynman diagram is as follows:[6]

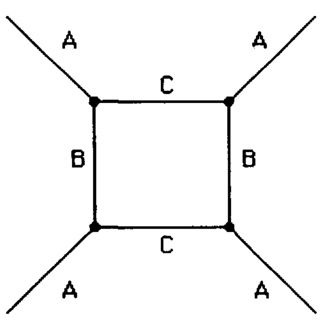

Fig. 3. AA scattering: $A+A \rightarrow A+A$.

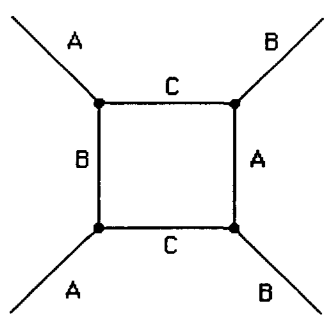

Fig. 5. A fourth-order diagram for AB scattering.

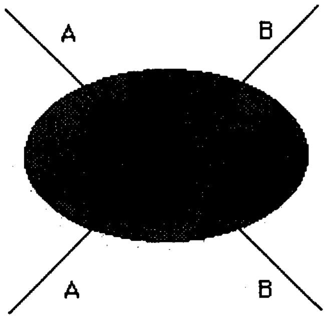

Fig. 6. Generic diagram for AB scattering.

(1) *Notation.* Label the incoming and outgoing 4-momenta[9] $p_1, p_2, \cdots, p_n$ (Fig. 7). Label the internal momenta $q_1, q_2, \ldots$. Put an arrow on each line, to indicate the "positive" direction (arbitrarily assigned, in the case of internal lines).

(2) *Vertex Factor.* For each vertex, write down a factor of $-ig$, where g is the coupling constant.[10]

(3) *Propagator.* For each internal line, write a factor

$$\frac{1}{q_j^2 - m_j^2 c^2},$$

where q_j is the 4-momentum of the (virtual) particle, and m_j is its mass.[11]

(4) *Conservation of Energy and Momentum.* For each vertex, include a factor

$$(2\pi)^4 \delta^4(k_1 + k_2 + k_3),$$

where the k's are the three 4-momenta coming *into* the vertex (if the arrow points *outward*, then k is *minus* the 4-momentum). The delta-function[12] imposes conservation of energy and momentum at the vertex—it vanishes unless the sum of the incoming momenta equals the sum of the outgoing momenta.[13]

(5) *Integration over Internal Momenta.* For each internal line, write down a factor

$$\frac{1}{(2\pi)^4} d^4 q_j$$

and integrate over all internal momenta.[14]

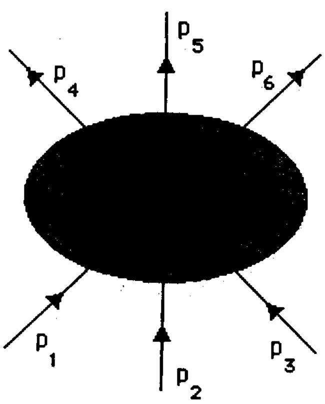

Fig. 7. Labeling a diagram.

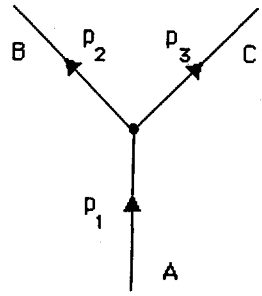

Fig. 8. The simplest diagram, with momenta labeled.

(6) *Cancel the Overall Delta-Function.* The result will contain a factor

$$(2\pi)^4 \delta^4(p_1 + p_2 + \cdots - p_n)$$

enforcing overall conservation of energy and momentum. Erase this factor, and what remains is $-i\mathscr{M}$.

Consider, for example, the simplest diagram of all, Fig. 1. Labeling the lines, in accordance with Rule 1, yields Fig. 8. There are no internal lines; there is one vertex, from which we pick up a factor $-ig$ (Rule 2), and a delta-function

$$(2\pi)^4 \delta^4(p_1 - p_2 - p_3)$$

(Rule 4), which we promptly discard (Rule 6), obtaining $-i\mathscr{M} = -ig$, or

$$\mathscr{M} = g. \tag{1}$$

A more substantial example is AB scattering (Fig. 9). In this case there are two vertices (hence two factors of $-ig$), one internal line, with the propagator

$$\frac{i}{q^2 - m_C^2 c^2},$$

two delta-functions

$$[(2\pi)^4 \delta^4(p_1 - p_3 - q)][(2\pi)^4 \delta^4(p_2 + q - p_4)],$$

and one integration

$$\frac{1}{(2\pi)^4} d^4 q.$$

Rules (1) through (5), then, yield

$$-i(2\pi)^4 g^2 \int \frac{1}{q^2 - m_C^2 c^2} \delta^4(p_1 - p_3 - q) \times \delta^4(p_2 + q - p_4) d^4 q.$$

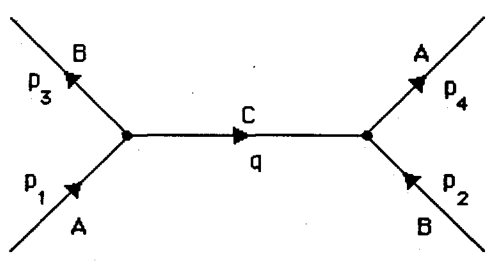

Fig. 9. Momentum-labeled diagram for AB scattering.

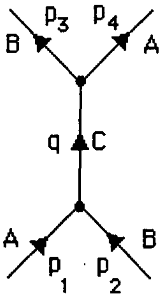

Fig. 10. The other second-order diagram for AB scattering.

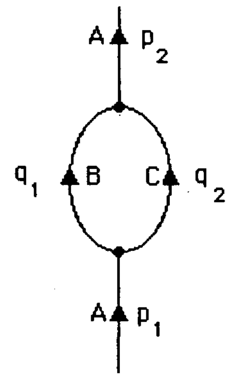

Fig. 11. A simple loop diagram.

The first delta-function serves to pick out the value of everything else at the point $q=p_1-p_3$, leaving

$$-ig^2\frac{1}{(p_1-p_3)^2-m_C^2c^2}(2\pi)^4\delta^4(p_1+p_2-p_3-p_4).$$

As promised, there is one remaining delta-function, which we erase (Rule 6); the result is

$$\mathscr{M}_1=\frac{g^2}{(p_1-p_3)^2-m_C^2c^2}.$$

But that's not the whole story. As we mentioned earlier, there is another second-order diagram (Fig. 10), for which we obtain

$$\mathscr{M}_2=\frac{g^2}{(p_1+p_2)^2-m_C^2c^2}.$$

Evidently the *total* amplitude for AB scattering, to order g^2, is

$$\mathscr{M}=\mathscr{M}_1+\mathscr{M}_2$$
$$=g^2\left(\frac{1}{(p_1-p_3)^2-m_C^2c^2}+\frac{1}{(p_1+p_2)^2-m_C^2c^2}\right). \quad (2)$$

For the sake of completeness we indicate briefly how one gets from the amplitude ($\mathscr{M}$) to the quantities of physical interest: the lifetime (τ) or the scattering cross section (σ). The connection is provided by Fermi's "Golden Rule;" schematically,

$$\text{transition rate}=\frac{2\pi}{\hbar}|\mathscr{M}|^2\rho,$$

where ρ is the density of final states, or "phase space" available.[15] Specifically, if particle 1 (at rest) disintegrates into particles 2, 3,...,n:

$$1\to2+3+\cdots+n$$

the decay rate is given by[16]

$$\Gamma=\frac{S}{2\hbar m_1}\int|\mathscr{M}|^2(2\pi)^4\delta^4(p_1-p_2-p_3-\cdots-p_n)$$
$$\times\prod_{j=2}^{n}\left(2\pi\delta(p_j^2-m_j^2c^2)\theta(p_j^0)\frac{d^4p_j}{(2\pi)^4}\right), \quad (3)$$

and the lifetime is $\tau=1/\Gamma$. If there are just *two* particles in the final state, the integrals can be carried out explicitly:[17]

$$\Gamma=\frac{S|\mathbf{p}|}{8\pi\hbar m_1^2c}|\mathscr{M}|^2, \quad (4)$$

where $\mathbf{p}$ is the 3-momentum of either outgoing particle. For example, if $m_B=m_C=0$ in the ABC model, the lifetime of the A, to lowest order [Eq. (1)], is

$$\tau=\frac{16\pi\hbar m_A}{g^2}. \quad (5)$$

Similarly, if two particles collide

$$1+2\to3+4+\cdots+n$$

the cross section is given by[16]

$$\sigma=\frac{S\hbar^2}{4\sqrt{(p_1\cdot p_2)^2-(m_1m_2c^2)^2}}$$
$$\times\int|\mathscr{M}|^2(2\pi)^4\delta^4(p_1+p_2-p_3-p_4-\cdots-p_n)$$
$$\times\prod_{j=3}^{n}\left(2\pi\delta(p_j^2-m_j^2c^2)\theta(p_j^0)\frac{d^4p_j}{(2\pi)^4}\right). \quad (6)$$

Again, the general formula simplifies considerably if there are only two particles in the final state; in this case the differential scattering cross section, in the center-of-momentum frame is[17]

$$\frac{d\sigma}{d\Omega}=\left(\frac{\hbar c}{8\pi}\right)^2\frac{S|\mathscr{M}|^2}{(E_1+E_2)^2}\frac{|\mathbf{p}_f|}{|\mathbf{p}_i|}, \quad (7)$$

where $\mathbf{p}_f$ is either outgoing momentum and $\mathbf{p}_i$ is either incoming momentum. For example, if $m_B=m_C=0$ and the incident momentum is small compared to m_Ac, the cross section for AB scattering, to lowest order [Eq. (2)], is

$$\sigma=\frac{1}{\pi}\left(\frac{\hbar g^2}{2m_A^3c^3}\right)^2. \quad (8)$$

III. LOOP DIAGRAMS

So far, so good. The difficulty arises when we attempt to evaluate more complicated diagrams, containing loops. For example, consider the diagram in Fig. 11.[18] Rules (1)–(5) yield

$$(-ig)^2\int\frac{i}{q_1^2-m_B^2c^2}\frac{i}{q_2^2-m_C^2c^2}(2\pi)^4\delta^4(p_1-q_1-q_2)$$

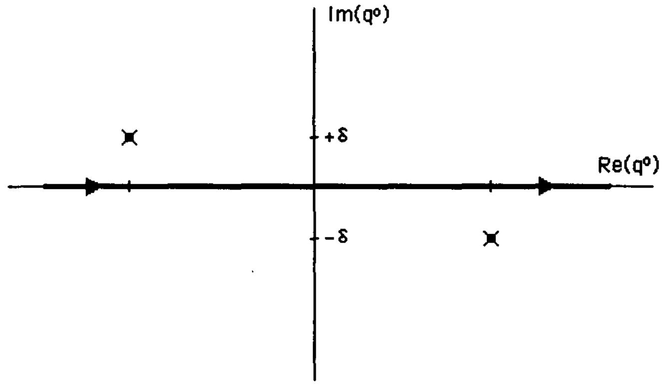

Fig. 12. The complex q^0 plane.

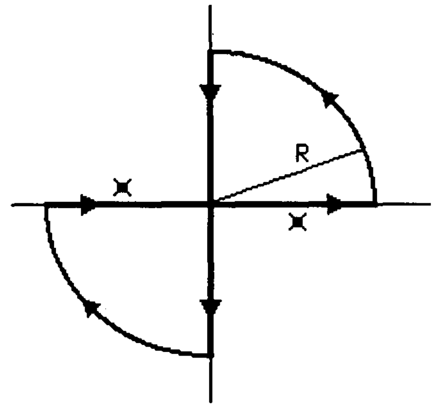

Fig. 13. "Figure-8" contour.

$$\times(2\pi)^4\delta^4(q_1+q_2-p_2)\frac{d^4q_1}{(2\pi)^4}\frac{d^4q_2}{(2\pi)^4}$$
$$=g^2\delta^4(p_1-p_2)\int\frac{1}{q_1^2-m_B^2c^2}\frac{1}{(p_1-q_1)^2-m_C^2c^2}d^4q_1, \tag{9}$$

and hence

$$\mathscr{M}(p)=\frac{ig^2}{(2\pi)^4}\int\frac{1}{(q^2-m_B^2c^2)[(p-q)^2-m_C^2c^2]}d^4q. \tag{10}$$

[The delta-function in Eq. (9) fixes $p_2=p_1$, and there is only one q left, so we can drop the subscripts with impunity.]

Integrals of this kind are easiest to compute in four-dimensional polar coordinates (one radius and three angles). Unfortunately, $q^2=(q^0)^2-\mathbf{q}^2$ is not the square of the radius—in fact, it need not even be positive. There is a trick (called "Wick rotation") for handling this problem; we illustrate the method first on a somewhat simpler integral:

$$I\equiv\int\frac{1}{q^2-m^2c^2+i\varepsilon}d^4q \tag{11}$$

(at this point it is important to restore the small imaginary term in the propagator).[11] The q^0 integral is

$$\int_{-\infty}^{\infty}\frac{1}{(q^0)^2-\mathbf{q}^2-m^2c^2+i\varepsilon}dq^0. \tag{12}$$

Its integrand has poles, in the complex q^0-plane, at

$$q^0=\pm\sqrt{\mathbf{q}^2+m^2c^2-i\varepsilon}\cong\pm[\sqrt{\mathbf{q}^2+m^2c^2}-i\delta] \tag{13}$$

where

$$\delta=\frac{\varepsilon}{2\sqrt{\mathbf{q}^2+m^2c^2}}$$

is a small positive quantity. These singularities are shown in Fig. 12, together with the path of integration (the real axis). Now consider for a moment the "figure-8" contour in Fig. 13. Cauchy's theorem tells us that since this path encloses no singularities the integral must be zero. If we allow the radius R of the two quarter-circles to increase without limit, their contribution to the integral must go to zero (the arc length increases like R, but the integrand goes like $1/R^2$). It follows that

$$\int_{-\infty}^{\infty}\frac{dq^0}{q^2-m^2c^2+i\varepsilon}=\int_{-\infty i}^{\infty i}\frac{dq^0}{q^2-m^2c^2+i\varepsilon}. \tag{14}$$

Now let $q^0=iq^4$, so that

$$\int_{-\infty i}^{\infty i}\frac{dq^0}{(q^0)^2-\mathbf{q}^2-m^2c^2+i\varepsilon}$$
$$=\int_{-\infty}^{\infty}\frac{idq^4}{-(q^4)^2-\mathbf{q}^2-m^2c^2+i\varepsilon}$$
$$=-i\int_{-\infty}^{\infty}\frac{dq^4}{\bar{q}^2+m^2c^2},$$

where $\bar{q}^2\equiv(q^1)^2+(q^2)^2+(q^3)^2+(q^4)^2$ and we drop the ε, because we are no longer integrating in the vicinity of a pole. Thus the integral [Eq. (11)] is equal to

$$I=-i\int\frac{d^4\bar{q}}{\bar{q}^2+m^2c^2} \tag{15}$$

(we use the overbar to indicate that the coordinates are labeled 1, 2, 3, 4, and the metric is Euclidean; $\bar{q}^2$ really *is* the square of the radial coordinate).

The same Wick rotation technique, applied to the integral in Eq. (10), yields[19]

$$\mathscr{M}=-\frac{g^2}{(2\pi)^4}\int\frac{1}{(\bar{q}^2+m_B^2c^2)[(\overline{p-q})^2+m_C^2c^2]}d^4\bar{q}, \tag{16}$$

where $p^4=-ip^0$. Unfortunately, it is now clear that this integral blows up.[20] For in polar coordinates $d^4\bar{q}=r^3drd^3\Omega$ (where $d^3\Omega$ stands for the four-dimensional solid angle), and the denominator goes like r^4 at large r. Thus the upper end of the radial integral has the form

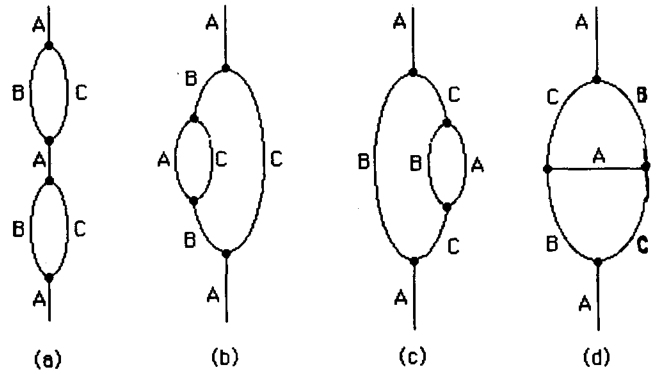

Fig. 14. Fourth-order diagrams for A→A.

$$\int^{\infty} \frac{1}{r^4} r^3 dr = \ln r|^{\infty} \to \infty. \qquad (17)$$

In the following sections we shall see how to handle the divergence in Eq. (16). But before we come to that we need to explore the physical significance of diagrams like Fig. 11, which have only two external lines. An A comes in, and an A goes out:

$A \to A.$

Nothing much of interest here, it would appear. However, this diagram does suggest that the A particle—even if it is not actually colliding with anything—is not quite the same as it would have been absent any coupling to the B and C. In colorful language, it has the possibility of splitting into a B and a C, which then recombine in flight. This process (which occurs even if $m_A < m_B + m_C$, since the B and C are virtual particles) *changes the effective mass* of the A, as we will see in a moment.

Of course, Fig. 11 is not the *only* diagram that contributes to A→A. There are, for example, four 4th-order diagrams (Fig. 14), and an infinite hierarchy of even more complicated higher-order ones. In this collection, some diagrams, like the first one in Fig. 14, have the property that snipping a single internal line breaks them into two separate pieces; we call these "one-particle reducible" diagrams. Suppose we let a shaded bubble (Fig. 15) represent the set of all one-particle *ir*-reducible diagrams (not counting the unadorned line itself), and let $\Sigma(p)$ stand for the sum of all their amplitudes (putting aside for a moment the inconvenient fact that most of these are infinite!). Then the collection of *all* A→A diagrams can be represented schematically by Fig. 16. In particular, if the A line is internal to some larger diagram, then the simple propagator $i/(p^2 - m_A^2 c^2)$ is embellished by all these higher-order processes to become

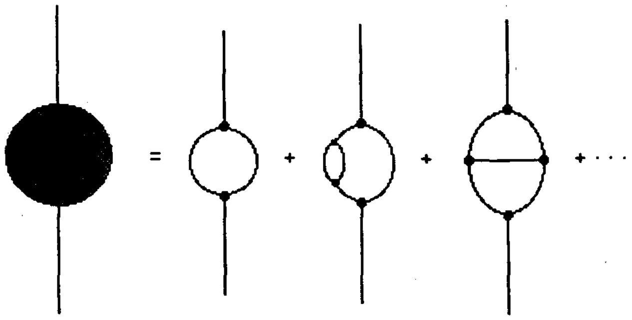

Fig. 15. Bubble representing all one-particle-irreducible diagrams.

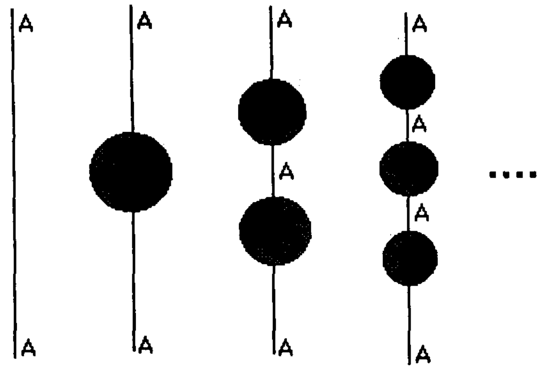

Fig. 16. Set of all diagrams representing A→A.

$$\frac{i}{p^2 - m_A^2 c^2} + \frac{i}{p^2 - m_A^2 c^2}(-i\Sigma)\frac{i}{p^2 - m_A^2 c^2} + \frac{i}{p^2 - m_A^2 c^2}$$
$$\times (-i\Sigma)\frac{i}{p^2 - m_A^2 c^2}(-i\Sigma)\frac{i}{p^2 - m_A^2 c^2} + \cdots$$
$$= \frac{i}{p^2 - m_A^2 c^2} \sum_{n=0}^{\infty} \left(\frac{\Sigma}{p^2 - m_A^2 c^2}\right)^n$$
$$= \frac{i}{p^2 - m_A^2 c^2} \frac{1}{1 - \Sigma/(p^2 - m_A^2 c^2)} = \frac{i}{p^2 - m_A^2 c^2 - \Sigma}. \qquad (18)$$

The effect of all these diagrams, then, is to replace the "bare" mass (m_A) in the propagator with the "renormalized" mass[18]

$$m_A'(p) = \sqrt{m_A^2 + \Sigma(p)/c^2}. \qquad (19)$$

Here m_A' is the mass one actually measures in the laboratory; notice that it is a function of p(!)—though as it turns out this variation is typically rather slight.[21] At this stage it is also *infinite*, but we are going to take care of that in a moment.[22]

IV. REGULARIZATION

Now that we understand the physical significance of diagram 11 (and its higher-order cousins), let us return to the essential problem: the divergence of the associated amplitude, Eq. (16). Our first task is to "regularize" the integral: rewrite it in such a way as to isolate the infinity. Since the trouble occurs at large r, we might, for example, truncate the radial integral at some "cutoff" value R, separate out the R-dependent part of the answer, and save for later the question of what to do as $R \to \infty$.[23] A more elegant way of accomplishing essentially the same purpose is to multiply the integrand by $R^2/(\vec{q}^2 + R^2)$; this gives us two extra factors of r in the denominator of Eq. (17), and thus renders the integral finite. Again, the "correct" (infinite) value would be the limit $R \to \infty$. But there is a much nicer way to regularize integrals of this kind, which was introduced by 't Hooft and Veltman in 1972[24] and is called "dimensional regularization:" we do the calculation in $4 - \epsilon$ dimensions, and at the very end of the problem, when we have isolated and neutralized the term that blows up, take the limit $\epsilon\to 0$.[25]

To begin with it is useful to get the angle dependence (in the term $\mathbf{p}\cdot\mathbf{q}$) out of the integrand by exploiting a famous trick of Feynman's. Noting that

$$\frac{1}{\mathrm{AB}}=\int_0^1 \frac{dx}{[\mathrm{A}(1-x)+\mathrm{B}x]^2} \tag{20}$$

(which is easily proved by direct integration) we obtain

$$\frac{1}{(\vec{q}^2+m_{\mathrm{B}}^2c^2)[(\overline{p-q})^2+m_{\mathrm{C}}^2c^2]}$$

$$=\int_0^1 \frac{dx}{\{(\vec{q}^2+m_{\mathrm{B}}^2c^2)(1-x)+[(\overline{p-q})^2+m_{\mathrm{C}}^2c^2]x\}^2}$$

$$=\int_0^1 \frac{dx}{(\vec{q}'^2+N^2)^2} \tag{21}$$

where $\vec{q}'\equiv\vec{q}-\vec{p}x$ and

$$[N(x)]^2\equiv\vec{p}^2x(1-x)+m_{\mathrm{B}}^2c^2(1-x)+m_{\mathrm{C}}^2c^2x. \tag{22}$$

Thus the amplitude [Eq. (16)] can be written[26]

$$\mathscr{M}=-\frac{g^2}{(2\pi)^4}\int_0^1 dx\int\frac{d^4\vec{q}'}{(\vec{q}'^2+N^2)^2}. \tag{23}$$

We propose to evaluate the amplitude in $(4-\epsilon)$ dimensions. This means replacing

$$\frac{d^4\vec{q}'}{(2\pi)^4}=\frac{1}{(2\pi)^4}r^3dr\,d^3\Omega$$

by

$$\frac{d^{4-\epsilon}\vec{q}'}{(2\pi)^{4-\epsilon}}=\frac{1}{(2\pi)^{4-\epsilon}}r^{3-\epsilon}dr\,d^{3-\epsilon}\Omega. \tag{24}$$

Meanwhile,[27] g carries the dimensions of momentum to the power $(1+\epsilon/2)$; to keep track of this, let us write

$$g=g_0(Mc)^{1+\epsilon/2}, \tag{25}$$

where M is some convenient reference mass (say, the average of m_{A}, m_{B}, and m_{C}), and g_0 is now a *dimensionless* coupling constant. Then

$$\mathscr{M}_\epsilon=-\frac{g_0^2(Mc)^{2+\epsilon}}{(2\pi)^{4-\epsilon}}\int_0^1 dx\int_0^\infty\frac{r^{3-\epsilon}dr}{(r^2+N^2)^2}\int d^{3-\epsilon}\Omega. \tag{26}$$

Now, the integral over all angles in n dimensions is well known:[28]

$$\int d^{n-1}\Omega=\frac{2\pi^{n/2}}{\Gamma(n/2)}. \tag{27}$$

(For example, in two dimensions we have $\int_0^{2\pi}d\phi=2\pi$, and in three dimensions $\int_0^{2\pi}d\phi\int_0^\pi\sin\theta\,d\theta=4\pi$.) That, of course, is for *integer* values of n, but we shall extend the formula to the case of $(4-\epsilon)$ dimensions [indeed, this—together with Eqs. (24) and (25)—is what we *mean* by evaluating the amplitude in $(4-\epsilon)$ dimensions]:

$$\int d^{3-\epsilon}\Omega=\frac{2\pi^{2-\epsilon/2}}{\Gamma(2-\epsilon/2)}. \tag{28}$$

As for the radial integral, it can be found in any table of definite integrals; assuming $\epsilon>0$,

$$\int_0^\infty\frac{r^{3-\epsilon}}{(r^2+N^2)^2}dr=\frac{1}{2N^\epsilon}\Gamma(\epsilon/2)\Gamma(2-\epsilon/2). \tag{29}$$

Thus we are left with

$$\mathscr{M}_\epsilon=-g_0^2\frac{(Mc)^{2+\epsilon}}{(4\pi)^{2-\epsilon/2}}\Gamma(\epsilon/2)\int_0^1 N^{-\epsilon}dx. \tag{30}$$

We are interested eventually in the limit $\epsilon\to 0$, so let us expand in powers of ϵ. The only singular term is the gamma function:[29]

$$\Gamma(z)\approx\frac{1}{z}-\gamma+(\ )z+(\ )z^2+\cdots \tag{31}$$

(where $\gamma=0.577216$...is the Euler–Mascheroni constant). For the rest we use the expansion

$$z^\epsilon=e^{\epsilon\ln z}=1+\epsilon\ln z+\frac{1}{2}\epsilon^2(\ln z)^2+\cdots. \tag{32}$$

Collecting like powers of ϵ, we find:

$$\mathscr{M}_\epsilon=-\left(\frac{g_0Mc}{4\pi}\right)^2\left[\frac{2}{\epsilon}-\gamma-\int_0^1\ln\left(\frac{N^2}{4\pi(Mc)^2}\right)dx+(\ )\epsilon+(\ )\epsilon^2+\cdots\right]. \tag{33}$$

The x integral can be evaluated in closed form (though the result is not terribly illuminating):

$$X\equiv\int_0^1\ln\left(\frac{\vec{p}^2x(1-x)+m_{\mathrm{B}}^2c^2(1-x)+m_{\mathrm{C}}^2c^2x}{4\pi(Mc)^2}\right)dx$$

$$=-2+\ln\left(\frac{m_{\mathrm{B}}m_{\mathrm{C}}}{4\pi M^2}\right)+\frac{(m_{\mathrm{B}}^2-m_{\mathrm{C}}^2)c^2}{\vec{p}^2}\ln(m_{\mathrm{C}}/m_{\mathrm{B}})$$

$$-\frac{a}{\vec{p}^2}\ln\left(\frac{\vec{p}^2+(m_{\mathrm{B}}^2+m_{\mathrm{C}}^2)c^2-a}{2m_{\mathrm{B}}m_{\mathrm{C}}c^2}\right), \tag{34}$$

where

$$a\equiv\sqrt{\vec{p}^4+2\vec{p}^2(m_{\mathrm{B}}^2+m_{\mathrm{C}}^2)c^2+(m_{\mathrm{B}}^2-m_{\mathrm{C}}^2)^2c^4}.$$

Regularization does not *change* anything—the amplitude [Eq. (33)] is still infinite in the real world of 4 dimensions $(\epsilon\to 0)$. But it does cleanly *isolate* the singularity: it's the $2/\epsilon$ term in the square brackets.[30] Now our job is to eliminate this singularity altogether.

V. RENORMALIZATION

Suppose we introduce a new "self-interaction" of the form $\mathrm{A}\to\mathrm{A}$ (Fig. 17),[31] with the associated "vertex factor"

$$-i\left(\frac{g_0Mc}{4\pi}\right)^2\frac{2}{\epsilon}, \tag{35}$$

and agree to do all calculations in $(4-\epsilon)$ dimensions, taking the limit $\epsilon\to 0$ only at the very *end* of the process, *after* adding all amplitudes of a given order. Notice that this "mass insertion" (or "mass counterterm") is *second* order in the coupling constant—the same as $\mathscr{M}_\epsilon$ [Eq. (33)]; whenever an A line contains a loop of the form Fig. 11, there will be another diagram of the same order, with a "×" in place of the loop. The combined amplitude is the *sum* of the two, and the coefficient in Eq. (35) has (obviously) been carefully chosen so as to cancel the singular

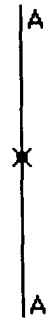

Fig. 17. Mass insertion.

term coming from the loop [Eq. (33)]. We can now take the limit $\epsilon \to 0$, and the result is perfectly finite. In particular, for A→A, we have (to second order),

$$\mathscr{M}=\left(\frac{g_0 Mc}{4\pi}\right)^2\left[\gamma+\int_0^1 \ln\left(\frac{N^2}{4\pi(Mc)^2}\right)dx\right]. \quad (36)$$

In Fig. 18 we have plotted the resulting renormalized mass [Eq. (19)] as a function of momentum.[32]

From a technical point of view removing the infinity in Eq. (33) has been absurdly easy—we simply introduced a new coupling tailor-made to eliminate the offending term. But how are we to *justify* this seemingly ad hoc procedure? The essential idea is as follows: The parameter m_A enters the theory via the propagator (Rule 3), but although we have *called* it "the mass of the A particle," there is no *a priori* guarantee that it represents the *physical* mass of the A, as we would measure it in the laboratory—on the contrary, we have already seen [Eq. (19)] that it is *not.* By introducing the mass counterterm (Fig. 17) we have in effect added a constant (dependent on ϵ, to be sure, but not on the momentum)[33] to the mass of the A. Since the "bare" mass (m_A) is not observable anyway, it can hardly matter how we apportion it between the propagator and the counterterm—all that matters is the final *physical* mass, and that includes both of them. Of course, it is awkward that the counterterm blows up as $\epsilon \to 0$, but then, exactly the same thing happens in classical electrodynamics, where the infinite potential energy of a point charge obliges us to introduce a (negative!) infinite counterterm to make the physical mass come out finite.[34]

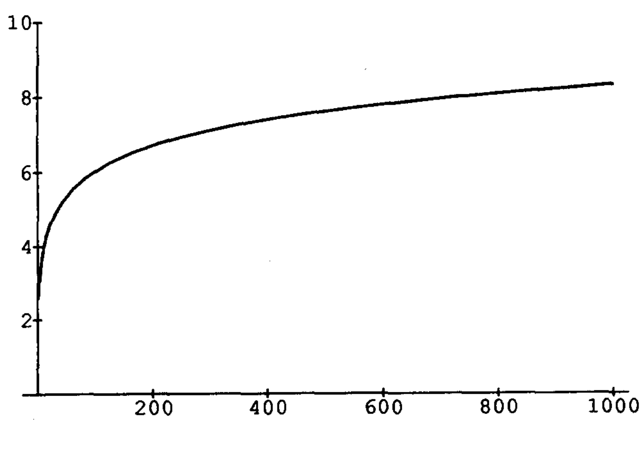

Fig. 18. Renormalized mass as a function of momentum, for the case $m_B = m_C$. Specifically, what is plotted is $y(x)$, where $(m_A')^2 = (\text{constant}) - (Mg_0/4\pi)^2 y(x)$ and $x \equiv \vec{p}^2/(2m_B c)^2$.

VI. OTHER DIVERGENT DIAGRAMS

But we are not quite done yet, for the simple loop (Fig. 11) is not the only divergent diagram in ABC theory. The integrals we encounter [such as Eqs. (10) and (11)] are of the form

$$\int F(q^2)d^4q \to \int F(-r^2)r^3 dr d\Omega \to \int^{\infty} r^D dr. \quad (37)$$

The integrand goes like some power (D) of r, at large r, and the integral is convergent only if D is less than -1. [In Eq. (11), $D=+1$, and the integral is quadratically divergent; in Eq. (10), $D=-1$, and the integral is logarithmically divergent.) Now, the generic Feynman diagram has E external lines, I internal lines, and V vertices. Every external line attaches to one vertex, every internal line connects two vertices, and every vertex joins three lines, so

$$E+2I=3V. \quad (38)$$

There is a propagator ($\propto q^{-2}$) associated with each internal line, so these contribute $-2I$ to D. Each internal momentum introduces a factor of d^4q, but each vertex brings in a delta-function, and this enables us to perform $V-1$ of the integrals (the last delta involves external momenta only, so it does not count). Thus integration over the internal momenta contributes $3[I-(V-1)]$ to D. All told, then, $D=-2I+3(I-V+1)$, or, using Eq. (38) to eliminate I,

$$D=3-\frac{3V}{2}-\frac{E}{2}. \quad (39)$$

If $D \geqslant -1$, the diagram is certainly divergent; on the other hand, if $D<-1$ the diagram is not necessarily convergent, for we have not distinguished the different integration variables q_i that go into calculating the amplitude. It may be that *one* of the integrals diverges ($D_i \geqslant -1$) even though the *total* D is in the convergent range. In general, a Feynman diagram is convergent if *it and all its subdiagrams* have $D<-1$; otherwise it is divergent.[35] (A subdiagram is any piece of the original that can be separated out by cutting lines.) But according to Eq. (39), $D \geqslant -1$ implies

$$3V+E \leqslant 8, \quad (40)$$

and since E is at least 2, this means $V \leqslant 2$. Apart from the trivial case $V=1$, $E=3$ (which is the primitive vertex itself, and is certainly not divergent),[36] the only solution is $V=E=2$, which is precisely the one-loop diagram we studied earlier. So as it turns out the *only* divergent diagrams in ABC theory are those that contain simple loops (Fig. 11) as subdiagrams.[37] As indicated earlier, all such diagrams are rendered finite by the associated mass counterterms.

It is of interest to see explicitly how this mechanism works in higher orders. Consider, for example, the fourth-order contributions to A→A (Fig. 14)—to which we must now add five mass-counterterm diagrams (Fig. 19). Diagram (d) is finite (it contains no simple loops or mass insertions). The infinities in (a) are exactly canceled by those in (e), (f), and (i) as we now demonstrate. The amplitude for (a) breaks into a product,

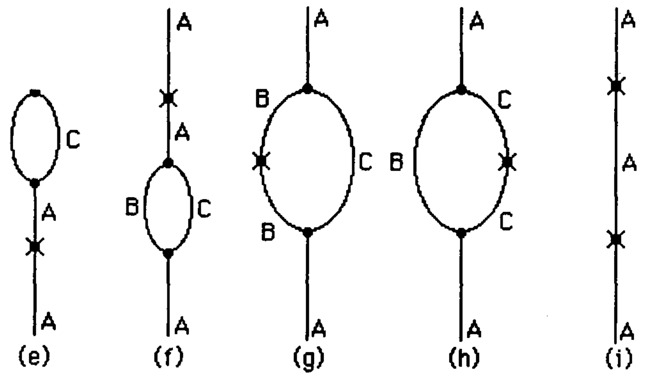

Fig. 19. Fourth-order mass counterterm diagrams for A→A.

$$\mathcal{M}_a=\frac{\mathcal{M}^2}{p^2-m_A^2c^2}, \tag{41}$$

where $\mathcal{M}$ is the single-loop amplitude Eq. (10), which we evaluated in Eq. (33):

$$\mathcal{M}=-\left(\frac{g_0Mc}{4\pi}\right)^2\left[\frac{2}{\epsilon}-J\right], \tag{42}$$

and $J\equiv\gamma+X+(\,)\epsilon+\cdots$ is finite (as $\epsilon\to 0$). The amplitudes for (e) and (f) are equal:

$$\mathcal{M}_e=\mathcal{M}_f=\left(\frac{g_0Mc}{4\pi}\right)^2\frac{2}{\epsilon}\,\frac{\mathcal{M}}{p^2-m_A^2c^2}, \tag{43}$$

and the amplitude for (i) is

$$\mathcal{M}_i=\left(\frac{g_0Mc}{4\pi}\right)^4\frac{4}{\epsilon^2}\,\frac{1}{p^2-m_A^2c^2}. \tag{44}$$

In the sum,

$$\mathcal{M}_a+\mathcal{M}_e+\mathcal{M}_f+\mathcal{M}_i=\left(\frac{g_0Mc}{4\pi}\right)^4\left[\left(\frac{2}{\epsilon}-J\right)^2-\frac{4}{\epsilon}\left(\frac{2}{\epsilon}-J\right)+\frac{4}{\epsilon^2}\right]\frac{1}{p^2-m_A^2c^2} \tag{45}$$

the terms in $1/\epsilon$ and $1/\epsilon^2$ cancel out, as promised.

This leaves only the pairs [(b), (g)] and [(c), (h)] to worry about. The amplitude for (g) is easily calculated:

$$\mathcal{M}_g=i\left(\frac{2}{\pi^2\epsilon}\right)\left(\frac{g_0Mc}{4\pi}\right)^4Q(p), \tag{46}$$

where

$$Q(p)\equiv\int\frac{d^4q}{(q^2-m_B^2c^2)^2[(p-q)^2-m_C^2c^2]}. \tag{47}$$

Notice that this integral (with $D=-3$) is finite.[38] Meanwhile, diagram (b) yields

$$\mathcal{M}_b=-\frac{1}{\pi^4}\left(\frac{g_0Mc}{4\pi}\right)^4\int\frac{d^4q_1}{(q_1^2-m_B^2c^2)^2[(p-q_1)^2-m_C^2c^2}\times\left(\int\frac{d^4q_2}{(q_2^2-m_A^2c^2)[(q_1-q_2)^2-m_C^2c^2]}\right). \tag{48}$$

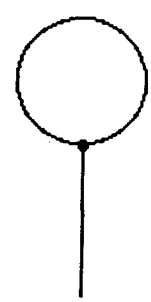

Fig. 20. Tadpole diagram in ϕ^3 theory.

The q_2 integral is, of course, the one we encountered earlier—it comes from the simple loop subdiagram. According to Eq. (10) this integral is equal to $[(2\pi)^4/ig^2]\mathcal{M}(q_1)$, where $\mathcal{M}$ is given by Eq. (33) with m_A in place of m_B; thus

$$\mathcal{M}_b=\frac{i}{\pi^2}\left(\frac{g_0Mc}{4\pi}\right)^2\int\frac{\mathcal{M}(q)d^4q}{(q^2-m_B^2c^2)^2[(p-q)^2-m_C^2c^2]}. \tag{49}$$

Isolating the divergent term in $\mathcal{M}$ [as in Eq. (42)] we obtain

$$\mathcal{M}_b=-i\left(\frac{2}{\pi^2\epsilon}\right)\left(\frac{g_0Mc}{4\pi}\right)^4Q(p)+\frac{i}{\pi^2}\left(\frac{g_0Mc}{4\pi}\right)^4\int\frac{J(q)d^4q}{(q^2-m_B^2c^2)^2[(p-q)^2-m_C^2c^2]}. \tag{50}$$

The first term cancels $\mathcal{M}_g$ [Eq. (46)], and the second is finite.[39] Thus the mass counterterm, which was designed to remove this infinity in second order, automatically removes all infinities in fourth order as well.

In other theories the situation if not so simple, because there exist two or more species of divergent (sub)diagrams. For example, in six dimensions diagram (d) in Fig. 14 is divergent, and one must introduce another counterterm (of order g^4) to kill it. In ϕ^3 theory (which amounts to making the A, B, and C particles identical) there is a divergent "tadpole" diagram (Fig. 20). In quantum electrodynamics the triangle diagram (Fig. 21) is divergent (electron propagators go like $1/q$, not $1/q^2$); elimination of this infinity requires renormalization of the coupling con-

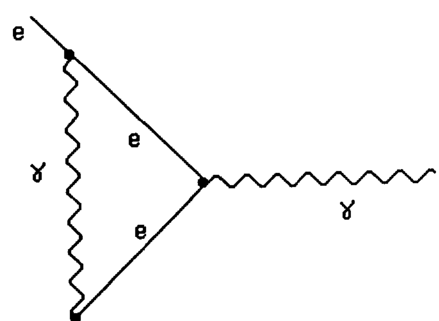

Fig. 21. Triangle diagram in QED.

stant. More subtle and insidious things can happen, too. In Eq. (50) the second term is finite, even though $J(q)$ itself blows up at large q [it goes like $\ln(q)$, but there are enough powers of q in the denominator to overcome the logarithm].[39] The reason is that the infinity in (b) comes from the simple loop, and that is covered by the first term in Eq. (50). But if we had a diagram with $D > -1$, which also contained a divergent subdiagram (something that cannot occur in ABC theory, but does in other models), then the analog to the second term in Eq. (50) would itself blow up, and we would need a further counterterm (of order g^4) to fix it.

Typically, then, the mass counterterm becomes a complicated infinite series in powers of g, and the coupling constant itself must be renormalized in a similar manner. It may even happen that *no* set of counterterms renders the theory finite (this is the case, for example, for ABC theory in more than six dimensions). In that event the theory is said to be "nonrenormalizable;" such theories are presumably unacceptable as models of the real world, since they do not yield intelligible predictions. The model we have explored here (ABC theory in four dimensions) is the *simplest possible* example of renormalization, in the sense that a single mass counterterm, of order g^2, suffices to render the theory finite.[40]

ACKNOWLEDGMENTS

This work was completed while D.J.G. was on sabbatical at Lawrence Berkeley Laboratory, Berkeley, CA 94720, and was supported in part by the Director, Office of Energy Research, Office of High Energy and Nuclear Physics, Division of High Energy Physics of the U. S. Department of Energy under Contract No. DE-AC03-76SF00098, and in part by a Faculty Sabbatical Fellowship from Associated Western Universities.

[a]Present address: Physics Department, Princeton University, Princeton, NJ 08544.

[1]T. Kinoshita and W. B. Lindquist, "Eighth-order magnetic moment of the electron," Phys. Rev. D **42**, 636–655 (1990); R. S. Van Dyck, Jr., P. B. Schwinberg, and H. G. Dehmelt, "New high-precision comparison of electron and positron g factors," Phys. Rev. Lett. **59**, 26–29 (1987).

[2]In addition to the texts cited below (Refs. 3 and 6), there are extensive bibliographies (together with very useful commentary) in George Leibbrandt, "Introduction to the technique of dimensional regularization," Revs. Mod. Phys. **47**, 849–876 (1975); T. P. Cheng and L. F. Li, "Resource Letter GI-1: Gauge Invariance," Am. J. Phys. **56**, 586–600 (1988). For a fascinating personal account of the early history see F. J. Dyson, *Disturbing the Universe* (Harper and Row, New York, 1979), Chaps. 5 and 6.

[3]The ϕ^3 theory is discussed (for example) in J. Collins, *Renormalization* (Cambridge University Press, Cambridge, 1984); ϕ^4 is treated in P. Ramond, *Field Theory: A Modern Primer* (Benjamin/Cummings, Reading, 1981); and L. H. Ryder, *Quantum Field Theory* (Cambridge University Press, Cambridge, 1985).

[4]A similar motivation underlies M. Hans, "An electrostatic example to illustrate dimensional regularization and renormalization group technique," Am. J. Phys. **51**, 694–698 (1983). In this charming paper the same basic inspiration is applied to the problem of calculating the potential of a uniformly charged infinite straight wire, where a somewhat analogous divergence arises in a much more transparent context. Another illuminating model is provided by L. R. Mead and J. Godines, "An analytical example of renormalization in two-dimensional quantum mechanics," Am. J. Phys. **59**, 935–937 (1991).

[5]This model was introduced to us by Max Dresden. It is discussed in D. Griffiths, *Introduction to Elementary Particles* (Wiley, New York, 1987), Chap. 6.

[6] For the purposes of this paper the Feynman rules will be taken as axiomatic. They can be derived from the underlying quantum field theory, once the Lagrangian for the system has been specified. See C. Itzykson and J. B. Zuber, *Quantum Field Theory* (McGraw-Hill, New York, 1980), Chap. 6, or Ref. 5, Chap. 11. The Feynman rules are somewhat different, of course, for other theories (quantum electrodynamics, weak interactions, chromodynamics, etc.), but their basic structure is similar.

[7]It goes without saying that the ABC particles have no electric charge; their *only* interaction is the one represented by Figure 1.

[8]Many authors let time run to the right; it is a matter of taste. By the way, the horizontal dimension (for us) has no spatial connotation; for example, Figure 2 does *not* mean to suggest that the force is repulsive—maybe it is, and maybe it isn't.

[9]The four-momentum of a particle of (rest) mass m and velocity $\mathbf{v}$ is $p^\mu = (p^0,p^1,p^2,p^3) = (E/c,p_x,p_y,p_z)$, where $E = \gamma mc^2$ is the relativistic energy of the particle and $\mathbf{p} = (p_x,p_y,p_z)$ is its relativistic momentum: $\mathbf{p} = \gamma m\mathbf{v}$. Here $\gamma = 1/\sqrt{1-(v/c)^2}$ and c is the speed of light.

[10]In ABC theory the coupling constant has the dimensions of momentum; a *dimensionless* coupling parameter can be defined by dividing through by Mc, where M is some convenient reference mass (say, the average of m_A, m_B, and m_C).

[11]The "square" of a four-vector is defined as follows: $q^2 \equiv (q^0)^2 - (q^1)^2 - (q^2)^2 - (q^3)^2$. It is positive for a "timelike" four-vector, and negative for a "spacelike" four-vector. For a *real* particle, the square of the four-momentum must obey the equation $p^2 = (E/c)^2 - \mathbf{p}\cdot\mathbf{p} = (mc)^2$. However, "virtual" particles (which begin and end within a Feynman diagram) do not satisfy this constraint, and hence the propagator—in spite of appearances—does not imply division by zero. When necessary to resolve ambiguities of integration a term $+i\varepsilon$ is added to the denominator, where ε is a small positive number that is taken to zero at the end of the calculation.

[12]The Dirac delta function is discussed in Appendix A of Ref. 5. For our present purposes it is enough to know that $\int f(x)\delta(x-a)dx = f(a)$, for any function $f(x)$, provided the domain of integration includes the point a. The four-dimensional delta-function is simply the product of four one-dimensional deltas: $\delta^4(k) = \delta(k^0)\delta(k^1)\delta(k^2)\delta(k^3)$.

[13]Notice that the Feynman rules rigorously enforce conservation of energy and momentum at every step. The only sense in which virtual particles are kinematically anomalous is that they do not "lie on their mass shell"—see Ref. 11.

[14]This is a fourfold integral over each internal momentum: $d^4q \equiv dq^0dq^1dq^2dq^3$, and all integrals run from $-\infty$ to $+\infty$. Notice that every δ is accompanied by a factor of (2π), and every d carries a factor of $(2\pi)^{-1}$.

[15]J. J. Sakurai, *Advanced Quantum Mechanics* (Addison-Wesley, Reading, 1967), pp. 41 and 312; Ref. 5, Sec. 6.2.

[16]S is a product of statistical factors: $(1/j!)$ for each group of j identical particles in the final state. The step function, $\theta(x)$, is 0 (when $x<0$) and 1 (when $x>0$). The "dot product" of two 4-vectors is $p\cdot q = p^0q^0 - \mathbf{p}\cdot\mathbf{q}$.

[17]The details are given in Ref. 5, examples 6.6 and 6.7. If there are three or more particles in the final state, the outgoing momenta are not kinematically determined, and we need to know the specific functional form of $\mathscr{M}$ before the integration can be performed.

[18]We shall for the moment think of Fig. 11 as a free-standing Feynman diagram, with p_1 and p_2 as external momenta, constrained to lie on the mass shell of particle A [i.e., $p_1^2 = p_2^2 = (m_Ac)^2$]. Later on, however, the same figure will be used to represent a piece of a larger diagram, in which case p_1 and p_2 become unconstrained internal momenta, perhaps even spacelike (if the figure appears in the horizontal orientation).

[19]The integrand in Eq. (10) has poles at $q^0 \cong \pm[\sqrt{\mathbf{q}\cdot\mathbf{q}+(m_Bc)^2} - i\delta_1]$ and at $q^0 \cong p^0 \pm [\sqrt{(\mathbf{p}-\mathbf{q})\cdot(\mathbf{p}-\mathbf{q})+(m_Cc)^2} - i\delta_2]$, where $\delta_1 \equiv \epsilon/[2\sqrt{\mathbf{q}\cdot\mathbf{q}+(m_Bc)^2}]$ and $\delta_2 \equiv \epsilon/[2\sqrt{(\mathbf{p}-\mathbf{q})\cdot(\mathbf{p}-\mathbf{q})+(m_Cc)^2}]$ are small positive numbers. To ensure that all four poles lie in the second and fourth quadrants (as required for the Wick rotation argument) we shall assume that $p^0 = 0$. More precisely, we assume that p^μ is spacelike, and adopt the Lorentz frame in which $p^0 = 0$ (see Ref. 18). The case of timelike p^μ is a good deal more complicated, and we shall not pursue it here. (Incidentally, there is a sound physical reason why the timelike case is more difficult: at sufficiently high energy it is possible to create

real B and C particles, not merely virtual ones, and this gives the amplitude a much richer analytic structure.)

[20]Of course, the integral I [in Eq. (11)] also diverges, and hence the Wick rotation argument [leading to Eq. (15)] is suspect. If this disturbs you, cube the integrand in Eq. (11) and run the same argument with the resulting finite integral. The point is that if the original (Minkowski metric) integral is finite, then it is equal to the Wick-rotated (Euclidean metric) version, and conversely, if the latter diverges, so too must the former. But it is easier to test for convergence in the Euclidean form—indeed, it can usually be done by inspection, simply counting the powers of q in the numerator and the denominator.

[21]The renormalized mass is not to be confused with the so-called "relativistic mass" ($m_{\rm rel}\equiv\gamma m$), which involves a purely kinematic dependence of inertia on velocity. It is somewhat analogous to the "effective mass" of electrons in a solid; see D. Park, *Introduction to the Quantum Theory* (McGraw–Hill, New York, 1974), 2nd ed., Sec. 13.2. A similar dynamical mass correction occurs in classical electrodynamics, where the energy associated with the electromagnetic fields of a charged particle contributes (via Einstein's formula $E=mc^2$) to its inertia; see D. J. Griffiths and R. Owen, "Mass renormalization in classical electrodynamics," Am. J. Phys. **51**, 1120–1126 (1983).

[22]Incidentally, the mass correction for a point charge in classical electrodynamics (Ref. 21) is also infinite, because the electrostatic potential energy diverges. Perhaps what this means is that there is no such thing as a point charge in classical electrodynamics.

[23]There is an "uncertainty principle" sense in which large momenta correspond to short distances. One might argue that the infinity in Eq. (10) is an artifact of our ignorance about the small-scale structure of the particles, and the cutoff is simply a crude way of acknowledging this ignorance. Such an interpretation is close in spirit to the classical resolution proposed in Ref. 22.

[24]G. 't Hooft and M. Veltman, "Regularization and renormalization of gauge fields," Nucl. Phys. B **44**, 189–213 (1972). See also Refs. 2, 3, and 4.

[25]In terms of Eq. (17), the three regularization procedures suggested amount, respectively, to (i) lowering the upper limit of the integral, (ii) increasing the power of r in the denominator, and (iii) reducing the power of r in the numerator.

[26]In the change of variables ($q\to q'$ we have shifted the q^4 axis up to $(q')^4=q^4+ip^0x$, which is permissible only if the integrand has no singularities in the intervening strip. As before (see Ref. 19) we avoid the problem here by assuming $p^0=0$. This also guarantees that N^2 is positive definite.

[27]A simple way to determine the units of g is to recalculate the lifetime of the A(A→B+C) in $4-\epsilon$ dimensions. The amplitude is still g [Eq. (1)], but in the Golden Rule (3) $\delta^4(p)\to\delta^{4-\epsilon}(p)\sim p^{-(4-\epsilon)}$ and $d^4p\to d^{4-\epsilon}p\sim p^{4-\epsilon}$, whereas $\delta(p^2)\sim p^{-2}$ is unchanged. Thus $1/\tau=\Gamma\sim(g^2/\hbar m)(mc)^{-\epsilon}$, and hence $g^2\sim(\hbar m/\tau)mc^{\epsilon}\sim(mc)^{2+\epsilon}$.

[28]The volume of an n sphere is calculated in many calculus books; for example, T. M. Apostol, *Calculus, Vol. II* (Wiley, New York, 1969), 2nd ed., p. 411. The surface area can be found by differentiating this formula with respect to the radius, and the solid angle is then obtained by setting the radius equal to 1.

[29]For a derivation of this formula see Ryder (Ref. 3), Appendix 9B.

[30]The singularity in $\mathscr{M}_\epsilon$ *looks* like a simple pole at $\epsilon=0$, but this is deceptive. Equation (29) is false for negative ϵ; in fact, the integral diverges for $\epsilon<0$, and the amplitude is infinite—as was clear already from the power-counting argument in Eq. (17).

[31]The Feynman rules for this self-interaction are the same as before, except that the delta-functions enforcing conservation of energy and momentum are of the form $\delta^4(k_1+k_2)$. Incidentally, we must of course do exactly the same thing for B and C as we do here explicitly for A.

[32]$\mathscr{M}$ [in Eq. (36)] is the lowest-order contribution to Σ; our result, therefore, is correct only to second-order in g_0. Note that $p^2=(p^0)^2-\mathbf{p}\cdot\mathbf{p}=-\vec{p}^2$ is a negative number, because we are assuming the momentum is spacelike (see Ref. 19).

[33]We could "soak up" the other constant terms in Eq. (33), while we are at it, by adding the appropriate constants to the $2/\epsilon$ in Eq. (35). But we *cannot* add anything that depends on p, for that would not amount to reapportioning the mass, but would describe an entirely different theory. (Actually, that last statement is a little too strong: we *could* add a term proportional to p^2, for this can be combined with the *first* term in the propagator, and the resulting overall factor absorbed into a renormalized coupling constant. In more complicated theories this is in fact necessary, but we do not need such a term in the ABC model.)

[34]See Ref. 22 and the last citation in Ref. 21. Michael Hans (Ref. 4) likens renormalization to changing the reference point for electrostatic potential, in those situations for which using the "point at infinity" gives rise to a divergent integral, and we are obliged (formally) to subtract an infinite constant in order to compensate for our poor initial choice of reference point.

[35]This is a particular instance of "Weinberg's theorem," which is discussed in Ref. 3.

[36]It is true that $D=0$ for this case, but since there is no integration involved, the amplitude is still finite.

[37]We have not included mass counterterm diagrams in this discussion, for the obvious reason: *any* diagram containing such a term is divergent, because of the factor $1/\epsilon$ in the coupling [Eq. (35)].

[38]We are only concerned with the infinite terms, at this point, so we have used $g=g_0Mc$ and done the integral (47) in four dimensions.

[39]In truth, there is a lacuna in this argument, since Eq. (50) obliges us to integrate over *all* q^μ, and we have only worked out $J(q)$ for *spacelike* q^μ.

[40]A theory in which a finite number of counterterms is needed, each of finite order in the coupling constant, is sometimes called "super-renormalizable"—see Collins, Ref. 3, section 5.7.3. The ABC model (requiring just one) is about as super-renormalizable as a theory can get.

SCIENCE AND LITERATURE

The widely prevalent opinion that almost any literary work, even if it amounts to no more than writing advertising copy or a book review, not to mention that Ph.D. thesis on 'Some little known laundry bills of George Moore', is intrinsically superior to almost any scientific activity is not one with which a scientist can be expected to sympathise.

Peter Medawar, *Pluto's Republic (Oxford, New York, 1984), p. 278.*