SLAC-PUB-515
October 1968
(TH)

Composite Particles in S-Matrix Theory and in Field Theory *

John M. Cornwall+ and Donald J. Levy‡
Department of Physics
University of California at Los Angeles,
Los Angeles, California 90024

Abstract

We give a model-independent exact solution in S-matrix terms to the problem of relating proper vertex functions and physical scattering amplitudes, a problem studied extensively by Ida. The solution is of Omnès type, supplemented by the solution of some algebraic equations. We use conventional definitions of the renormalization constants Z_1, Z_2 to study the composite limit ($Z_1 = Z_2 = 0$) in S-matrix theory. A model-independent discussion of the same problem is given in terms of the Dyson equations, where essentially the same results are recovered, in a different language. Possible applications to the electromagnetic mass shifts of composite particles are briefly discussed, including problems of gauge invariance.

* Work supported in part by the National Science Foundation, by the Niels Bohr Institute, and by the U.S. Atomic Energy Commission.

+ Alfred P. Sloan Foundation Fellow; now at the Niels Bohr Institute.
++ Now at the Stanford Linear Accelerator Center, Stanford, Calif.

(Submitted to Phys. Rev.)

I. Introduction

Over the last few years, a great many articles have been written concerning the connection between compositeness and vanishing renormalization constants, used in large part upon ideas put forth by Salam[1]. The fundamental ideas in this field are well-known, and we have nothing new to add to them; our concern in the present paper is to state these ideas in a model-independent way, both in S-matrix theory and in field theory, with a view toward possible future applications to such things as the electromagnetic mass shifts of composite particles. We present some new results which connect the S-matrix approach with field theory, and clarify some known results connected with the passage to the composite limit.

There are two overlapping fields in which the compositeness problem can be studied: S-matrix theory, and field theory (as exemplified by the Dyson equations). In S-matrix theory, we have solved a problem first phrased completely by Ida[2,3], which we state as: given the physical S-matrix, say for πN scattering, what is the proper πNN vertex function as a function of one nucleonic mass variable? This is to be contrasted to the essentially trivial problem of constructing the form factor from the physical phase shifts. With this solution in hand, we are in a position to study the composite limit, in much the same way as Kaus and Zachariasen (among others) have done[4,5]. The proper vertex function appears in the decomposition of the partial-wave S-matrix into two parts one of which contains the nucleon pole, and the other which is one-nucleon irreducible but unitary[2,3]. This same decomposition was used in refs. 4 and 5, and our Section II can be considered as an extension of the relevant part of ref.4. We need not repeat in detail the arguments in ref.4 concerning the bootstrap philosophy, or the way in which the vanishing of renormalization constants insures that the "elementary" nucleon drops out of the scattering amplitude, to be replaced by a composite nucleon. What happens, in accordance with ref. 4, is that the unrenormalized proper vertex function and propagator are finite and well-defined in the composite limit, while their renormalized counter parts are not.

This last circumstance is an interesting one to study with the Dyson equations. We have carried out such a study, valid for any finite field theory, in section III and find that the results (so far as field theory and S-matrix theory are comparable) are in agreement with those of section II. We can, in addition, construct formulas for such things as electromagnetic mass shifts of composite particles. The mass-shift formula comes in several guises: one, related to S-matrix theory, in the Dashen-Frautschi[6] formula; two, a Bethe-Salpeter type of formula[7], three, a dispersion formula based on the Källén-Lehmann representation[8]. They are all the same when evaluated exactly; it is when approximations are made that trouble comes in. The Dashen-Frautschi formula is plagued by infra-red problems[9], which are related to gauge invariance problems. Any dispersion integral which saves only certain intermediate states is not gauge-invariant. In an appendix, we show how to select intermediate states in the dispersion integral so that gauge invariance is automatic. This is the analog of Feynman's old proof that we must add photons to the charged legs of a diagram in all possible ways, in order to save gauge invariance.

We have been influenced by a number of authors other than those explicitly cited here; a full list of references would be inordinately lengthy. Hayashi et al[10] have recently published a well-referenced review paper which should be consulted for other publications.

II. S-Matrix Calculation of the Propagator and Vertex Function.

In this section, we restudy the problem of Ida[2,3], which is to calculate the proper vertex function and propagator with S-matrix techniques. Of course, the composite limit is of particular interest; we show how to recover the results of Kaus and Zachariasen[4,5] and others.

1. Kinematics and Definitions

The renormalized propagator we write as $S(\not{p})$, where $\not{p} = p_\mu \gamma^\mu$. Take $W = (p^2)^{\frac{1}{2}}$; then it is convenient to define a function $Z(W)$ by

$$S^{-1}(W) = (W-M)Z(W) \tag{1}$$

By definition $Z(M) = 1$; according to the usual field-theoretic arguments, the nucleon wave-function renormalization constant Z (conventionally written Z_2) is recovered from the asymptotic behaviour of the propagator:

$$\lim_{W\to\infty} Z(W) = Z \tag{2}$$

A dispersion relation for $Z(W)$ can be obtained from the Källen-Lehman representation:

$$Z(W) = 1 + \frac{W-M}{\pi}\int_{-\infty}^{\infty} \frac{\tau(W')dW'}{W'-W} \quad + \text{pole terms, if any} \tag{3}$$

It turns out that $Z(W)$ has at least one pole, for sufficiently small Z. The spectral function τ vanishes in the interval $-(M+\mu) \leq W \leq M+\mu$ (where M is the nucleon mass and μ pion mass), otherwise, $\tau > 0$. If we save only πN intermediate states

$$\tau(W) = 3G^2 \rho(W) \frac{|\Gamma(W)|^2}{(W-M)^2} \qquad \begin{matrix} W > M+\mu, \\ W < -(M+\mu) \end{matrix} \tag{4}$$

where

$$\rho(W) = \left| \frac{k(W)(E-M)}{8\pi W} \right| \tag{5}$$

and k, E are the center of-mass momentum and nucleon energy, respectively, in πN scattering:

$$k(W) = (2W)^{-1}\left[W^2-(M+\mu)^2\right]^{\frac{1}{2}}\left[W^2-(M-\mu)^2\right]^{\frac{1}{2}} \tag{6}$$

$$E-M = (2W)^{-1}\left[(W-M)^2-\mu^2\right] \tag{7}$$

The renormalized proper vertex function $\Gamma(W)$ is normalized so that $\Gamma(M) = 1$ (hence $G^2/4\pi = 14.5$, the πN coupling constant). $\Gamma(W)$ has cuts for $W > M+\mu$, $W < -(M+\mu)$; the phase of Γ on these cuts is related to the $\ell = 1$, $J = 1/2$ πN scattering amplitude (for $W > M + \mu$), or $\ell = 0, J = 1/2$ amplitude (for $W < -(M + \mu)$). To simplify the notations and calculations, we set the 0^+ scattering amplitude to zero (experimentally it is small in the low-energy region), and concentrate on the 1^- amplitude; further, we save only the elastic channel, although the generalization to many channels is straight forward.

2. S-Matrix Approach

Let us define the invariant amplitude for πN scattering with $\ell = 1$, $I = J = 1/2$ by

$$T(W) = \rho(W)^{-1} a_{1-}(W) \tag{8}$$

such that in the elastic region

$$a_{1-}(W) = e^{i\delta} \sin\delta \tag{9}$$

This amplitude T is free of kinematic singularities. If we take the 0^+ amplitude to be zero, then $T(W)$ has a unitary cut only on the right-hand W axis; what would correspond to the normal left-hand cuts in the W^2 plane lie along the imaginary axis.

First, assume that the nucleon is elementary, so that Z_1 and Z are finite. Then $\Gamma(W)$, $Z(W)$ are also finite.

We may write $T(W)$ in terms of its one-particle reducible parts and a remainder $\tilde{T}(W)$:

$$T(W) = \frac{-3G^2\,\Gamma(W)^2}{Z(W)(W-M)} + \tilde{T}(W) \tag{10}$$

The first term on the right comes from the elementary nucleon pole. As Ida [2,3] discusses, the unitarity relation for $\Gamma(W)$ is:

$$\mathrm{Im}\,\Gamma(W) = \Gamma(W)\,\tilde{T}(W)^{*} \qquad W > M+\mu \tag{11}$$

Ida then proves that $\tilde{T}(W)$ is a unitary amplitude:

$$\mathrm{Im}\,\tilde{T}(W) = \rho(W)\,|\tilde{T}(W)|^2 \qquad W > M+\mu \tag{12}$$

It follows that we can write

$$T(W) = N(W)\,D(W)^{-1} \quad , \quad \tilde{T}(W) = \tilde{N}(W)\,\tilde{D}(W)^{-1} \tag{13}$$

where $N, \tilde{N}$ have only "left-hand" cuts, and D, $\tilde{D}$ only right hand cuts. By virtue of the unitarity relation (II) for Γ, we can write

$$\Gamma(W) = \tilde{D}(M)\,\tilde{D}(W)^{-1} \tag{14}$$

If the scattering amplitude decrease sufficiently fast at infinity, we can set $D(\infty) = \tilde{D}(\infty) = 1$. (Strictly speaking, this is in conflict with the LSZ theorem[11], which requires $\Gamma(W)$ to vanish at $W = \infty$, so that Z as calculated from (3) is finite. The required rate of decrease of Γ need only be logarithmic. We cannot treat the large-W region accurately in any event, so we shall ignore the LSZ theorem, and calculate Z from different considerations. These problems do not arise for scalar nucleons). With this normalization, and with some old field theoretic arguments[12], we find:

$$\lim_{W\to\infty} \Gamma(W) = Z_1 = \tilde{D}(M) \tag{15}$$

where Z_1 is the usual vertex function renormalization constant.

The form factor $F(W)$ is defined by $F(W) = \Gamma(W)Z^{-1}(W)$. Since we can write $F(W) = D(M)\ D^{-1}(W)$, it follows that

$$Z(W) = \frac{\tilde{D}(M)\,D(W)}{\tilde{D}(W)\,D(M)} \tag{16}$$

and

$$Z = \frac{\tilde{D}(M)}{D(M)} \quad ; \quad D(M) = Z_1 / Z \tag{17}$$

Thus we find (18)

$$Z(W) = Z \frac{D(W)}{\tilde{D}(W)}$$

Equation (10) becomes

$$T(W) = \frac{N(W)}{D(W)} = -3G^2 \frac{D(M)\tilde{D}(M)}{(W-M)D(W)\tilde{D}(W)} + \frac{\tilde{N}(W)}{\tilde{D}(W)} \tag{19}$$

Observe what happens in the limit $Z_1 = 0$, $Z = 0$: The renormalized quantities $\Gamma(W)$, $Z(W)$ vanish, but the unrenormalized vertex function $\Gamma_u(W) = \Gamma(W)Z_1^{-1}$ and unrenormalized propagator $S_u(W) = ZS(W)$ exist. Further the first term on the right of (19) vanishes (see Kaus and Zachariasen[4] for details). None the less, there is still a particle pole at $W = M$, coming from the vanishing of $\tilde{D}(M)$ in the second term on the right in (19). That $\tilde{D}(M)$ vanish is assured by taking $Z_1 = 0$; if also we take $Z = 0$, we shall later prove that the residue of the pole in T is just $-3G^2$. The elementary particle pole is completely replaced by a composite particle with the same mass and coupling constant.

For sufficiently small Z, Z_1, observe that $\tilde{D}(W)$ has a zero at some position $W = W_R$. This pole in $\Gamma(W)$ and $Z(W)$ was first discussed by Jin and Mac Dowell[13], who proved that the pole at $W = W_R$ does not occur in $T(W)$, since the two terms on the right of (19) cancel each other. $T(W)$, of course, has a pole, whose residue we define as

$$\frac{\tilde{N}(W_R)}{\tilde{D}'(W_R)} = -3g^2 \tag{20}$$

(the prime indicates a differentiation). Obviously, when $W_R \simeq M$, $Z_1 = \tilde{D}(M) \simeq \tilde{D}'(W_R)(M-W_R)$, so Z_1 vanishes as $M \to W_R$. Later, we shall see that for sufficiently small Z_1, $Z \to 1-G^2/g^2$, so as $G \to g$, $Z \to 0$. Conversely, as Z, $Z_1 \to 0$ the elementary nucleon pole, along with one of the Jin-Mac Dowell poles, disappears; the remaining Jin-MacDowell pole in $T(W)$ represents a composite nucleon. A full discussion is given in ref. 4.

3. Solution of Ida's Problem

Our object is to try to solve the equation (19) for specified input forces in either N, or in $\tilde{N}$. One question that might be asked is: given $T(W)$, or alternatively, specified force terms in $N(W)$, calculate $\tilde{T}(W)$(hence $\Gamma(W)$), and the renormalization constants Z and Z_1. (We, of course, believe that the nucleon is composite, and hence $Z = Z_1 = 0$, but it is interesting to study hypothetical worlds where Z and Z_1 are finite, as Ida calculated, especially for the purpose of discussing the passage to the composite limit).

It is heuristically more convenient to solve the converse problem: given $\tilde{T}(W)$, calculate the physical amplitude $T(W)$. It will become clear from the ensuing arguments that the problem of the paragraph above is readily soluble by similar techniques.

We begin with the converse problem, that is, $\tilde{N}(W)$ is given. Please observe that, throughout this paper, all integrals are cut off at a large but finite value of W to avoid convergence problems. Therefore we can write

$$\tilde{D}(W) = 1 - \frac{1}{\pi} \int \frac{\rho(W') \tilde{N}(W') dW'}{W' - W} \tag{21}$$

Let us suppose that the forces are sufficiently strong that there is a Jin-Mac Dowell pole (zero of $\tilde{D}(W)$ at $W = W_R \neq M$. This pole does not appear in $T(W)$, which gives us a single condition. It is convenient to incorporate this condition by writing subtracted dispersion relation for $D(W)$ (the subtraction is actually unnecessary):

$$D(W) = D(W_R) - \frac{1}{\pi} (W - W_R) \int \frac{\rho(W') N(W') \, dW'}{(W' - W)(W' - W_R)} \tag{22}$$

$$D(W_R) = \frac{3G^2 D(M) \tilde{D}(M)}{\tilde{N}(W_R)(W_R - M)} \tag{23}$$

In the limit $Z_1 = 0 (M = W_R)$, (23) gives using (20) and (15):

$$D(M) = \frac{G^2}{g^2} \times D(M)$$

Now for $Z_1 = 0$, $Z = 1 - G^2/g^2$. If $Z \neq 0$, $D(M) = 0$, consistent with (17); if $Z = 0$ and $G^2 = g^2$, $D(M)$ is as yet undetermined. We solve (19) for N:

$$N(W) = \tilde{T}(W) \left[\frac{3G^2 D(M) \tilde{D}(M)}{\tilde{N}(W_R)(W_R - M)} - \frac{3G^2 D(M) \tilde{D}(M)}{\tilde{N}(W)(W-M)} - \frac{W - W_R}{\pi} \int \frac{\rho(W') N(W') dW'}{(W'-W)(W'-W_R)} \right] \tag{24}$$

Let us try the ansatz

$$N(W) = -\frac{3G^2 \tilde{N}(W) D(M)}{\tilde{N}(M)(W-M)} + (W - W_R) H(W) \tilde{T}(W) \tag{25}$$

We find

$$H(W) = J(W) - \frac{1}{\pi} \int \frac{\rho(W') H(W') \tilde{T}(W') dW'}{W'-W} \tag{26}$$

where the integral is over the unitary cut, and

$$J(W) = \frac{-3G^2 D(M)\tilde{D}(M)}{W-W_R} \left\{ \frac{1}{W-M}\left[\tilde{N}(W)^{-1} - \tilde{N}(M)^{-1} \right] - \right.$$

$$\left. - \frac{1}{W_R-M}\left[\tilde{N}(W_R)^{-1} - \tilde{N}(M)^{-1} \right] \right\} \tag{27}$$

It is easy to see that J(W) has no poles; if N(W) goes like W^{-1} at infinity, so does J(W).
The solution to (26) is:

$$H(W) = J(W) - \frac{\tilde{D}(W)}{\pi} \int \frac{\rho(W')\, J(W')\, \tilde{T}(W')\, dW'}{\tilde{D}(W')^{*}(W'-W)} + \lambda \frac{\tilde{D}(W)}{W-W_R} \tag{28}$$

where λ is a number as yet undertermined. Note that H(W)T(W) has no right-hand cut, and that as Z_1(= D(M)) approaches zero, both J(W) and the λ-independent part of H(W) vanish. Later on, we give an exactly soluble (but non-trivial) model in which J(W) vanishes identically; clearly, in the composite limit, certain features of this model must be generally true, since J(W) vanishes in this limit for any S-matrix.

Our solution will be complete, once we have exhibited λ and D(M) in terms of known quantities. This can be done, in general, by writing an unsubtracted dispersion relation for D(W):

$$D(W) = 1 - \frac{1}{\pi} \int \frac{\rho(W') N(W') dW'}{W'-W} \qquad (29)$$

and evaluating (29) at $W = M$. This condition, coupled with the condition (23) that the Jin-Mac Dowell poles cancel in T, furnishes two equations for λ and $D(M)$.

So far, we have only used one-half of the N/D formalism, which expresses unitarity on the right-hand cut via dispersion relation of the type (29). There is, of course, a condition on the"left-hand" cut; from (19), we find

$$\tilde{D}(W) \operatorname{Im} N(W) = D(W) \operatorname{Im} \tilde{N}(W) \qquad (30)$$

We leave it to the reader to show, with the aid of (19),(25), (27) and (28), that (30) is identically satisfied. Finally, the renormalization constants Z_1 and Z can be calculated from (15) and (17).

A number of generally true observations can be made from a remarkably simple model, which posessesses an algebraic solution. Suppose that $N(W)$ has the form:

$$\tilde{N}(W) = \frac{B^2}{W + W_L} \tag{31}$$

It is immediately apparent from (27) that $J(W) \equiv 0$, and we know $H(W)$ directly from (28). The whole model is solved algebraically, and we find:

$$N(W) = \tilde{N}(W) \left[\lambda - \frac{3G^2 D(M)}{\tilde{N}(M)(W-M)} \right] \tag{32}$$

$$D(W) = 1 - \lambda + \lambda \tilde{D}(W) - \frac{3G^2 D(M)}{\tilde{N}(M)(W-M)} \left[\tilde{D}(W) - \tilde{D}(M) \right] \tag{33}$$

(A special case of this solution in potential theory has been given by Kaus and Zachariasen[4].)

From (33), we find:

$$D(M) = \left[1 - \lambda + \lambda \tilde{D}(M) \right] \left[1 + \frac{3G^2 \tilde{D}'(M)}{\tilde{N}(M)} \right]^{-1} \tag{34}$$

Use (33) and (23) to come to

$$D(W_R) = 1 - \lambda + \frac{3G^2 D(M)\tilde{D}(M)}{\tilde{N}(M)(W_R - M)} = \frac{3G^2 D(M)\tilde{D}(M)}{\tilde{N}(W_R)(W_R - M)} \tag{35}$$

It is clear that as $W_R \to M$, $\lambda \to 1$; indeed, it is easy to show that $\lambda = 1 + 0 \quad (W_R - M)^2$ by solving (34) and (35) together. Thus for sufficiently small Z_1 and $W_R - M$, (34) becomes

$$\frac{\tilde{D}(M)}{D(M)} = Z = 1 + 3G^2 \frac{\tilde{D}'(M)}{\tilde{N}(M)} + O\left[(W_R - M)^2\right] \tag{36}$$

(The first equality follows from (17). In the limit $Z_1 = 0$, $W_R = M$, we conclude, with the aid of (20), that

$$Z = 1 - G^2/g^2 \tag{37}$$

In the composite limit $Z_1 = Z = 0$, $\lambda = 1$, we get:

$$N(W) = \left(\frac{W - W_0}{W - M}\right)\tilde{N}(W) \quad , \quad D(W) = \left(\frac{W - W_0}{W - M}\right)\tilde{D}(W) \tag{38}$$

where

$$W_0 = M + 3G^2 \frac{D(M)}{\tilde{N}(M)} \tag{39}$$

Equations (37),(38) and (39) are independent of our special model; see the remarks below (28).

Observe the following: the condition that the composite particle and the eldmentary particle have the same mass is that $Z_1 = 0$; that they have the same couplings requires $Z = 0$. This is perhaps the opposite of one's naive expectations. Furthermore, it is clear that, for sufficiently small Z_1, $N(W)$ always has a zero which approaches W_0 in the composite limit. We shall argue in the next section that W_0 can be identified with the bare mass of the nucleon, M_0. Since $D(M) = Z_1/Z \geq 0$, $N(M) > 0$(for attractive forces, in general), we see from (39) that $M_0 \geq M$, as would be expected in conventional πN field theory. As far as pure S-matrix theory is concerned, Z_1 and Z are logically independent, and $D(M) = Z_1/Z$ can have a wide range of values; therefore, so can W_0. But we always have a finite number (zero in the composite limit) for $Z\delta M$, defined as

$$Z\,\delta M \equiv Z(M_0 - M) = \frac{3G^2 Z_1}{\tilde{N}(M)} \tag{40}$$

The reader must not suppose that Z_1/Z can be chosen completely arbitrarily, at least without the expense of unphysical complications. If $Z_1/Z = D(M)$ is sufficiently small, then it follows by continuity arguments that $D(W)$ has a zero at some point $W = W_1 \simeq M$. By hypothesis, the physical amplitude $T(W)$ has no other poles except the nucleon pole, hence $N(W_1)$ must also be zero, and W_1 is the same zero of $N(W)$ as discussed in the preceding paragraph. A glance at (19) shows that the right-hand side of this equation will have a pole, unless $D(W)$ has a CDD pole at $W = W_1$. Now $D(W)$ vanishes at $W_R \simeq M$ (for small enogh Z_1), and by drawing graphs it will be easy to see that the addition of the CDD pole at $W_1 \simeq M$ produces another zero in $D(W)$, at $W_2 \simeq M$. This can produce a pole in the right-hand side of (19) at $W = W_2$, unless the residue of the pole vanishes (which it must, since $T(W)$ has no such pole). The condition that there be no such pole in (19), along with other previously mentioned conditions, allows one to determine the residue of the CDD pole as well as W_1 and W_2. Everything is well-behaved at the composite limit, when $W_o = W_1 = W_2$ and the CDD pole and extra zero in $D(W)$ disappear(because the residue of the CDD pole vanishes). It is difficult to make any physical sense out of all these zeroes and poles before the composite limit is reached. In the next section, we indicate that field theory probably avoids these complications, by not letting Z_1/Z become too small in the composite limit.

III The Dyson Equations

We now turn to the problem of compositeness as expressed in the Dyson equations. All of the features of the S-matrix theory of Sec.II emerge, as well as some new ones, which involve the nucleonic bare mass, as briefly mentioned in Section II.

To simplify the presentation of this section, we suppose that the only composite particle in the world is the nucleon, all others being elementary; furthermore, we treat all particles as isotopic scalars. . The practical distinction we make between elementary and composite particles is that the <u>renormalized</u> propagators and vector functions of elementary particles are supposed to be well-defined, and to have no extra poles or zeroes. Unrenormalized quantities are distinguished by a subscript u; renormalized quantities have no subscript of this sort.

Treatments similar to ours have appeared in the literature before, but generally based on specific models (e.g., Pradhann and Passi[14] have studied the Lee model with recoil). The only assumptions we make are that the field theory is renormalizable, and that the field-theoretic expressions for Z_1 , Z, etc., are finite in principle, and can be varied by varying renormalized coupling constants and masses. For simplicity, we consider only Yukawa vertices coupling two baryons and a mason.

1. Compositeness via the Dyson Equations

Consider a world in which there are a certain number of pseudoscalar mesons (π, K, ...) and baryons (N, Σ, ...); the nucleon can appear as a bound state in a number of two-body channels (πN, KΣ, ...). There is quite a difference in spirit between a nucleon which is a bound state of itself (and a meson), as in the πN channel, and a nucleon composed of two elementary particles (e.g., K Σ). The latter case is more straightforward but there are no insuperable difficulties for the former case.

Let us begin by defining a renormalized off-shell scattering amplitude for baryons and mesons by

$$\delta(p-p'-q)\, T(p,k,q;\, i\, f) = \frac{-i}{(2\pi)^4} \int \cdots \int e^{-i\left[(p-k)\cdot x + k\cdot y - p'\cdot x' - q\cdot y'\right]} \times$$

$$\times\; K_x \overline{K}_{x'} K_y K_{y'} \langle 0 | \left(\psi(x) \overline{\psi}(x') \phi(y) \phi(y') \right)_+ | 0 \rangle \qquad (41)$$

which describes the scattering of a baryon (of momentum p-k) and meson (k) in the initial channel i, into a baryon (p'=p-q) and meson (q) in the first channel f. The renormalized meson field is ϕ, the renormalized baryon field is ψ, and the K's are free Dirac or Klein-Gordon operators. As before, we set $p^2 = W^2$; when all external momenta are on the mass shell, and we put $\not{p} = W$, the appropriate partial-wave projection of (41) is just the amplitude T(W) introduced in (8). We define a one-nucleon irreducible amplitude $T^1(p,k,q;\, i\ f)$ by subtracting from T(p,k,q; i f) all Feynman graphs which can be separated into two disjoint pieces by cutting a single nucleon line of momentum p. This amplitude yields T(W) on the mass shell. We define the unrenormalized version of T^1 by:

$$T^1_u(p,k,q;\, i\, f) = \prod (Z_j)^{1/2}\, T^1(p,k,q;\, i\, f) \qquad (42)$$

where Z_j are wave-function renormalization constants for the external legs. The fundamental hypothesis we make is that $T^1(p,k,q;i\ f)$ is finite and well-behaved in the limit when the nucleon becomes composite. This is easy to swallow if the channels i,f do not contain the nucleon itself. If they do, a non-perturbative point of view is required, since as a result of our hypothesis we shall prove that all renormalized Yukawa vertices containing a nucleon vanish. Nevertheless, we proceed on this assumption.

We use the notation $\tilde{\Gamma}_u(p,q;f)$ to denote a certain pseudo-proper unrenormalized vertex function, describing a nucleon (of momentum p) going to a channel f consisting of a baryon (p- q) and a meson (q). The vertex function is proper with respect to the nucleon of momentum p, but contains full propagator corrections to the legs of channel f. The Dyson equation for this object is

$$\tilde{\Gamma}_u(p,q;if) = i\gamma_5 \Big\{ H(p,q;f) - \frac{i}{(2\pi)^4}\int d^4k \sum_i \frac{G_i Z_{1i}}{G_f Z_{1f}} \times \qquad (43)$$

$$\times\ Z_{2f} Z_{3f}\, S_0(p-k)\Delta_0(k) T^1(p,k,q;if)\Big\}$$

$$H(p,q;f) = S_u(p-q)S_0^{-1}(p-q)\,\Delta_u(q)\Delta_0^{-1}(q)$$

$$S_0(p-q) = (\not{p}-\not{q}-M)^{-1}\ ;\ \Delta_0(q) = (q^2-\mu^2)^{-1}.$$

In (43) the G_i are renormalized coupling constants coupling a nucleon to channel i, Z_{1i} are vertex renormalization constants, and $Z_{2f}\, Z_{3f}$ is the product of wave-function renormalization constants for the baryon and meson in channel f.

Let us put the particles in channel f on the mass shell, and set $\not{p} = W$ in $\tilde{\Gamma}_u$. Under these circumstances, we have:

$$\tilde{\Gamma}_u(p,q;f) = i\gamma_5 Z_{2f} Z_{3f}\, \Gamma_u(W;f) \tag{44}$$

where $\Gamma_u(W;f)$ is the unrenormalized form of the proper vertex described in section II. From (43), we find (with channel f on shell)

$$\Gamma_u(W;f) = 1 - \frac{i}{(2\pi)^4}\int d^4k\ \sum_i \frac{G_i Z_{1i}}{G_f Z_{1f}} S_0(p-k)\Delta_0(k)\,T'(p,k,q;if) \tag{45}$$

To simplify the notation in what follows, we write explicit formulas as if the nucleon were only coupled to one channel, so that channel labels may be dropped. The reader may convince himself, using (43) and (45), that no real generality is lost in the ensuing discussion.

The idea behind compositeness, as expressed in Section II, is that the amplitude T^1 picks up a pole in the momentum p, as the various forces in the problem are adjusted. Therefore, in the neighborhood of this pole, T^1 has the form (irrelevant Dirac matrices omitted):

$$T^1(p,k,q)\Big|_{\not p=W} = -\frac{R(p,k)\,\bar{R}(p,q)}{W-W_R} + \cdots \tag{46}$$

The residue functions in (46) make sense only at $p^2 = W_R^{\ 2}$. We shall be interested in evaluating functions at W=M; as the composite pole moves close to M, we need only save the pole terms. Insert (46) into the one-channel version of (45) to find:

$$\Gamma_u(W) = 1 + \frac{\hat{R}\,g}{W-W_R} + \cdots \tag{47}$$

$$\hat{R} = \frac{i}{(2\pi)^4}\int d^4k\; i\gamma_5\; S_0(p-k)\,\Delta_0(k)\,R(p,k)\Big|_{\not p=W} \tag{48}$$

g = R(p,q) at this point $p^2=W_R^{\ 2}$, q and p-q on the mass shell. (49)

Of course, we assume that R exists and is not zero; by our previous assumptions about T^1, both R(p,q) and R are well-behaved in the composite limit.

By definition, $Z_1^{-1} = \Gamma_u(M)$. For sufficiently small values of $M-W_R$, we find

$$Z_1 = \frac{M-W_R}{\hat{R}g} + O[(M-W_R)^2] \tag{50}$$

This is in agreement with the general results of section II, that $Z_1 \sim M-W_R$. Observe also that $\Gamma_u(W)$ is well-defined for small Z_1, hence $\Gamma(W) = Z_1\Gamma_u(W)$ vanishes in the composite limit.

We may also calculate the unrenormalized propagator, as usual setting $\not{p} = W$. In the usual language, we have:

$$S_u(W)^{-1} = W - M_0 - \Sigma_u(W) \tag{51}$$

$$\Sigma_u(W) = \frac{i}{(2\pi)^4}\int d^4q \; i\gamma_5 S_0(p-q)\Delta_0(q)\frac{G^2 Z_1^2}{Z_2 Z_3 Z}\,\tilde{\Gamma}_u(p,q)\Big|_{\not{p}=W} \tag{52}$$

In (52), G is a renormalized coupling constant, Z is the nucleon wave-function renormalization constant, and Z_1, Z_2, Z_3 refer to the intermediate state to which the nucleon is coupled. Use (43), (46), and (48) to find the pole contribution to $\Sigma_u(W)$:

$$\Sigma_u(W) = G^2 \frac{Z_1^2}{Z} \frac{\hat{R}^2}{W-W_R} + \cdots \tag{53}$$

By definition,

$$Z^{-1} = 1 - \left.\frac{\partial \Sigma_u(W)}{\partial W}\right|_{W=M} \tag{54}$$

This yields

$$Z = 1 - G^2 \left(\frac{Z_1 \hat{R}}{M-W_R}\right)^2 \tag{55}$$

We note in passing that the mass-shell version of T^1 has a pole at $W = W_R$ of residue $-g^2$ (see (46) and (49)). It is easy to check that the pole at $W = W_R$ of the one-nucleon reducible terms has precisely the opposite residue, so that pole does not appear in T; this is the Jin-Mac Dowell result[13]. Of course, in the composite limit, the one nucleon reducible terms vanish so that $T = T^1$.

For sufficiently small Z_1, we can use (50) to get

$$Z_1 = 1 - G^2/g^2 + \cdots \qquad (56)$$

This is the same result as previously derived in S-matrix theory (see (37)).

It is easy to see that Z_u vanishes like Z_1^2/Z in the composite limit, so that the unrenormalized propagator becomes the bare propagator: $S_u(W) \quad (W-M_o)^{-1}$. We have $S_u(W) = ZS(W)$; using (1), the function $Z(W)/Z$ exists in the limit $Z \to 0$ and is given by

$$\lim_{Z \to 0} \frac{Z(W)}{Z} = \frac{W-M_o}{W-M} \qquad (57)$$

This agrees with the S-matrix theory of section II, provided we identify W_o in (38) with M_o, since $Z(W)/Z = D(W)/\tilde{D}(W)$ from(18). Observe also that, since the propagator is a simple rational function, the phase of the unrenormalized proper vertex function and of the form factor are the same, both being equal to the physical phase shift. In connection with the comparison with S-matrix theory, we remind the reader that equations (37),(38), and (39) are model-independent, although most easily derived from a special model.

We can draw some conclusions about the rate at which Z approaches zero, when Z_1 is already small. The unrenormalized inverse propagator must have a zero at $W = M$; saving only the pole terms gives, with the aid of (51) and (53):

$$O = M - M_0 - G^2 \frac{Z_1^2}{Z} \frac{\hat{R}^2}{M - W_R} + \cdots \tag{58}$$

With the aid of (50) and (51), we find

$$Z = \frac{M - W_R}{M - M_0} \; . \tag{59}$$

By comparing the expressions derived in section II with those of the present section, we derive the following relation which holds in the composite limit:

$$\tilde{D}'(M) = (\hat{R} g)^{-1} \tag{60}$$

Use (17) and (38) to find

$$Z_1 / Z = D(M) = (M - W_0) \tilde{D}'(M) \tag{61}$$

which, with the aid of (50),(60), and the previously mentioned identification $W_o = M_o$, yields (59).

Just as in the S-matrix case, all renormalized vertex functions which involve one (or more) nucleons vanish in the limit. The amplitude T then becomes equal to the amplitude T_1 and it is perfectly straight forward to believe that T_1 is well-behaved when it describes something like $K\Sigma$ elastic scattering, since there is some part of this amplitude which has no nucleon vertices in it at all. But when T_1 describes, e.g., πN scattering, the situation is a little different. Any finite number of Feynman graphs contributing to T_1 must vanish if all vertices involving nucleons vanish. Our hypothesis has been that T_1 does not vanish in the composite limit; this can only be true when an infinite number of graphs contribute. We cannot prove that T_1 does or does not vanish from simple graphical considerations; one must study a set of non-linear integral equations for all T_1's which involve external nucleons to see if they have non-trivial solutions.

2. Electromagnetic Mass Shifts of Composite Particles

We shall be brief in this subsection, because it would take another long article to describe (much less avoid!) all the pitfalls involved in doing a reliable calculation of the neutron-proton mass difference. There are many techniques for evaluating this mass difference: 1) Feynman diagrams, with form factors (the original work is by Feynman and Speisman[15]); 2) the Dashen-Frautschi method[6]; 3) dispersion methods applied to the propagator[6,16]; 4) Bethe-Salpeter method, as well as several other methods. All of these would give the same answer if evaluated exactly, but certain methods give trouble especially when approximations are made. A particularly interesting trouble, from the point of view of compositeness, arises in the work of Fried and Truong[16]: if one sets Z = 0 in their formula, the mass difference vanishes identically. This happens independent of any approximation. The reason for it is that Fried-Truong formula does not take into account the Jin-Mac Dowell pole in the inverse propagator, which must occur for sufficiently small Z.

In section III.1, the nucleon was coupled to a strong interaction channel which made it composite. Let us now add in the photon-nucleon channel, and discuss the situation for small (but finite) Z. Among the T_1's there is a set of photo production amplitude, which for sufficiently small Z, have composite particle poles, as in (46):

$$T'_{\mu}(p,k,q) = -R(p,k)\frac{1}{W-W_R}R_{\mu}(p,q) \qquad (62)$$

Here μ is a 4-vertex index which couples to the photon field. R^{μ} (p,q) is a residue function which gives the total change of the composite nucleon; thus for the proton, ignoring magnetic moments terms:

$$R_{\mu}(p,q) \rightarrow \gamma_{\mu} e \quad (p^2 = W_R^2 , \ p-q \text{ and } q \text{ on shell}) \qquad (63)$$

Just as for the strong interactions, we introduce a quantity

$$\hat{R}_{\gamma p} = \frac{i}{(2\pi)^4} \int d^4k \ \gamma_{\mu} S_o(p-k) \Delta_o^{\mu\nu}(k) R_{\nu}(p,k) \qquad (64)$$

The corresponding quantity for the neutron vanishes, because it has zero charge. In (64) $\Delta_o^{\mu\nu}$ is the free photon propagator. Electromagnetic vertices and corrections to the proton propagator, in the neighborhood of the pole at $W = W_R$, can be expressed with the aid of (64).

To make the point we have in mind, it is simplest to forget the neutron completely, and discuss the electromagnetic mass shift of the proton, as a sort of analogue of the Lamb shift in hydrogen. Adding the neutron merely complicates the writing of formulas. The unrenormalized proton propagator, with electromagnetic corrections included, takes the form

$$S_u^{-1}(W) = W - M_0 - \frac{G^2 \Xi_1^2}{Z} \frac{\hat{R}^2}{W - W_{Rp}} - \frac{eG\Xi_1 \hat{R}\hat{R}_{\gamma p}}{W - W_{Rp}} + \ldots \quad (65)$$

where W_{Rp} is the position of the Jin-MacDowell pole including electromagnetic corrections. In writing (65), we have used the fact that $Z_{1\gamma pp} = Z$.

It is well-known that unrenormalized propagators are not gauge-invariant, because Z is not gauge-invariant, in general. However, an interesting thing happens when Z becomes very small and nucleon approaches compositeness. Gauge-dependent terms can only come from that part of the proton propagator $\Delta_{\mu\nu}(k)$ which go like $k_\mu k_\nu$. It is a consequence of the Ward identity that these gauge-dependent terms must vanish (at least) linearly in W-M, as W approaches M; this must be so, or the electromagnetic mass shift (which comes from evaluating (65) at W = M) would be gauge-dependent. In the composite limit, this factor of W-M will cancel the pole at $W=W_R$, with the result that the gauge-dependent corrections to Z vanish with a higher power of M - W_R than the Feynman gauge electromagnetic corrections. Intuitively, this accords with our belief that the notion of compositeness, as expressed via Z = 0, is gauge-invariant.

We distinguish between M and M_p, the proton mass before and after (respectively) electromagnetism is turned on. There is a similar distinction between W_R and W_{Rp}. Since (by our notion of compositeness) the electromagnetic corrections to W_R are the same as they are to M, we have

$$M_p - W_{Rp} = M - W_R \tag{66}$$

As the composite limit is approached, the wave-function renormalization constant behaves like (see (59)):

$$Z \to \frac{M_p - W_{Rp}}{M_p - M_0} \tag{67}$$

where the higher-order terms contain gauge-dependent terms. If we forget about electromagnetic corrections to the strong vertices and propagators, so that G, Z_1, and R are unchanged, we can use (50), (66) and (67) to evaluate (65), which must vanish at $W = M_p$; the result is the electromagnetic mass shift of the proton without feedback:

$$M_p - M = e\,\hat{R}_{\gamma p} \tag{68}$$

It is easy to check that including feedback would still lead to a finite result at $Z = 0$, in contrast to the Fried-Truong[16] formulas. Clearly, formula (68) is a Bethe-Salpeter type of mass-shift equation; such formulations have already been discussed in the literature[7]. It has an advantage over the Dashen-Frautschi formula that infra-red problems are relatively easily disposed of.

Dispersion relations for $S_u^{-1}(W)$[8,16] are yet another technique for finding mass shifts. The necessary ingredients are things like $\Gamma_\mu(W)$, the NNγ vector with the nucleon off-shell. These are constructed in terms of D(W) (see Section II), which reveals the composite particle pole. This pole is exhibited in $S_u^{-1}(W)$, as discussed by several authors[13,18], by the deformation of an integration contour as the pole moves from the second sheet (for weak forces) to the first sheet (as the strength of the forces is increased). Thereby all the results in Section II and III can be exhibited in an approximation in which only a small number of intermediate states is saved. Unfortunately, not all such approximations are gauge-invariant. In the appendix, we show how to maintain gauge invariance in approximate calculations; for example, if one begins by saving only the πN channel as an intermediate state in the propagator without electromagnetions included, then one must add both the γN and the γπN channels to compute the gauge-invariant propagator. The electromagnetic amplitude must satisfy the relevant Ward identities. That these two channels should be included is intuitively obvious, if one thinks of computing a gauge-invariant set of Feynman graphs by the Cutkosky rules, but the authors are unaware of a detailed discussion in the literature.

References

1. A.Salam, Nuovo Cimento 25,224(1962).
2. Masakuni Ida, Phys. Rev. 135,B499(1964).
3. Masakuni Ida, Phys. Rev. 136,B1767 (1964).
4. P. Kaus and F. Zachariasen, Phys. Rev. 171,1597(1968).
5. P. Kaus and F. Zachariasen, Phys. Rev. 138,B1304(1965)
6. R. Dashen and S.F. Frautschi, Phys. Rev. 135,B1190(1965)
7. John Harte, Phys. Rev. 158,1521(1967).
8. John M. Cornwall and S.H. Patil, Phys.Rev.Letters 18,757(1967)
9. See, e.g., G. Barton, Phys.Rev. 146,1149(1966).
10. K. Hayashi, M. Hirayama, T. Muta, N. Saito, and T. Shirafuji, Fortschr. d.Physik 15, 625 (1967).
11. H. Lehman, K. Symanzik, and W. Zimmermann, Nuovo Cimento 2,425 (1955).
12. G. Källén, Helv. Phys. Acta 25,417 (1952); M.Gell-Mann & F. Zachariasen, Phys.Rev. 123,1065 (1961).
13. Y.S. Jin and S.W. Mac Dowell, Phys.Rev.137,B688(1965).
14. T. Pradhann & J.N. Passi, Phys.Rev. 160,1336(1967).
15. R.F.Feynman and G.S. Speisman, Phys. Rev. 94,[illegible](1954)
16. H. Fried and T. Truong, Phys.Rev. Letters 16,559(1966).
17. H. Pagels, Phys. Rev.144,1261 (1966).
18. T. Saito, Phys. Rev. 145,1302 and 152,1339(1966)

Appendix: On Gauge Invariance

We want to show what combination of intermediate states must be saved in the Lehmann representation of the propagator, in order than the propagator be gauge-invariant. For simplicity, suppose that the charged particles are scalars, called ϕ particles, and that there is a ϕ^3 vertex. We define the renormalized propagator and its Fourier transform by

$$i\Delta_F(x) = \langle 0 | \left(\phi(x)\, \phi^\dagger(0) \right)_+ | 0 \rangle$$

$$\Delta_F(p) = \int d^4x\, e^{ipx} \Delta_F(x) \qquad \text{(A1)}$$

In what follows, we set $p^2 = s$.

Any change in $\Delta_F(p)$ coming from a gauge transformation has two parts : one part from changing the photon propagator, and one part from the phase transformation of the fields ϕ, ϕ^+:

$$\phi(x) \to e^{-iq\Lambda(x)} \phi(x)$$

$$\phi^\dagger(x) \to \phi^\dagger(x)\, e^{iq\Lambda(x)} \qquad \text{(A2)}$$

where q is the charge of the scalar field.

Let us consider an operator gauge function $\Lambda(x)$ which commutes with ϕ, ϕ^+ and is causal (fields commute at space-like separations). To $O(q^2)$, the change is Δ_F coming from (A2) can be written

$$\delta^{(1)}\Delta_F(p) = -\frac{ie^2}{(2\pi)^4}\int d^4k\, \Delta_F(p-k)\, D(k) \quad , \tag{A3}$$

where D(k) is the Fourier transform of

$$D(x) = -i\langle 0|(\Lambda(x)\Lambda(0))_+|0\rangle \tag{A4}$$

To the change (A3) must be added the change coming from the modified photon propagator ($\Delta_{\mu\nu}(k)$), which occurs in the internal lines of $\Delta_F(p)$:

$$\Delta_{\mu\nu}(k) \to \Delta_{\mu\nu}(k) + k_\mu k_\nu D(k) \tag{A5}$$

This induces a change $\delta^{(2)}\Delta_F(p)$. For gauge invariance, the sum of these two changes must add to zero. Although D(k) is rather general, we study the kinematically simple case

$$D(k) = \frac{C}{k^2+i\epsilon} \qquad (C = \text{constant}) \tag{A6}$$

corresponding to the addition of massless scalar ghosts to the electromagnetic field. This D(k) generates divergent expressions in the propagator, but the change in the propagator discontinuity function is finite; studying this discontinuity function is sufficient for our purposes.

Consider the following approximation to the Lehmann representation: only the $\phi\phi$ and $\gamma\phi$ channels are saved. We write:

$$\Delta_F(s) = -\frac{1}{\pi}\int_{\mu^2}^{\infty} \frac{\mathrm{Im}\,\Delta_F(s')\,ds'}{s'-s}$$

$$\mathrm{Im}\,\Delta_F(s) = \pi\,\delta(s-M^2) + G^2\sigma(s) + g^2\tilde{\sigma}(s) \tag{A7}$$

Here

$$\sigma(s) = \frac{1}{16\pi}\left(\frac{s-4M^2}{s}\right)^{1/2} \frac{|F(s)|^2}{(s-M^2)^2} \tag{A8}$$

and (in the Feynman gauge)

$$\tilde{\sigma}(s) = -\frac{1}{16\pi}\left(\frac{s-M^2}{s}\right) g_{\mu\nu}\Gamma^\mu\Gamma^{\nu *}|\Delta_F(s)|^2 \tag{A9}$$

$F(s)$ is the strong ϕ^3 form factor, while Γ^μ is the electromagnetic proper vertex function.

Under a gauge transformation of the photon propagator, $-g_{\mu\nu}$ in (A9) is changed to $-g_{\mu\nu} + C k_\mu k_\nu$, where k_μ is the photon momentum. The Ward identify for Γ_μ reads $k_\mu \Gamma^\mu(s) = \Delta_F(s)^{-1}$. The change (A3) can be analyzed with the aid of (A7) and the usual Cutkosky rules; we have

$$\delta^{(1)} \mathrm{Im}\, \Delta_F(s) = \frac{g^2 C}{16\pi^2} \int ds' \left(\frac{s'-s}{s}\right) \left[\pi\delta(s'-M^2) + \sigma(s')\right] \quad \text{(A10)}$$

The change $\delta^{(2)}$ in $\Delta_F(s)$ is computed from (A9) and the Ward identity

$$\delta^{(2)} \mathrm{Im}\, \Delta_F(s) = \frac{g^2 C}{16\pi} \left(\frac{s-M^2}{s}\right) \quad \text{(A11)}$$

which just cancels off the first term of (A10). It remains to cancel off the second term. This can be done by adding the $\gamma\phi\phi$ intermediate state. We do not require full knowledge of the amplitude V_μ for $\phi \to \gamma\phi\phi$, but only the Ward identity:

$$k_\mu V^\mu(p,k,q) = \frac{F[(p-k)^2]}{(p-k)^2 - M^2} \quad \text{(A12)}$$

where F is the same form factor as in (A8).

We compute the change in Δ_F coming from the photo-production graphs:

$$\delta^{(2)} \mathrm{Im}\, \Delta_F(s) = \frac{g^2}{2(2\pi)^5} \int d^4k\, d^4q\; C\, k_\mu k_\nu V^\mu V^{\nu *} \times$$
$$\times \quad \delta_+(k^2)\, \delta_+(q^2-M^2)\, \delta_+[(p-k-q)^2-M^2] \quad \text{(A13)}$$

With the aid of the Ward identity (A12) we can integrate (A13) into the form:

$$\delta^{(2)} \mathrm{Im}\, \Delta_F(s) = \frac{g^2 e}{256\pi^2} \int ds' \left(\frac{s-s'}{s}\right) \left(\frac{s'-4M^2}{s'}\right)^{1/2} \frac{|F(s')|^2}{(s'-M^2)^2} \qquad (.$$

where we have set $(p-k)^2=s'$. We recognice the appearance of the spectral function $F(s)$ (see (A8)) in (A14). The total change in $\Delta_F(s)$ is given by adding (A10), (A11), and (A14), which gives zero, no matter what $F(s)$ is chosen to be. $F(s)$ itself is gauge-invariant even when electromagnetic corrections are included, because it is (in principle) a measureable quantity.

The general principle is clear: if a particular channel $|n\rangle$ is included in the propagator without electromagnetism, then the channel $|n\gamma\rangle$ must be included in the Lehmann representation with electromagnetic effects included. The Ward identity then insures gauge invariance.

We conclude by observing that it is quite simple to make up approximations to the amplitude $\phi \to \gamma\phi\phi$ which satisfy the Ward identity(A12); for example, the Born graphs with form factors at the vertices, and fully corrected propagators. A somehwat more elaborate analysis is needed when composite particle poles are present, but it is not difficult in principle.