SLAC-PUB-733
April 1970
(TH)

A MODEL OF COUPLING CONSTANT RENORMALIZATION*

Kenneth G. Wilson

Laboratory of Nuclear Studies,
Cornell University, Ithaca, New York 14850†

and

Stanford Linear Accelerator Center,
Stanford University, Stanford, California 94305

(Submitted to Phys. Rev.)

* Work supported by the U. S. Atomic Energy Commission.

† Permanent address (after September 1970).

ABSTRACT

The charged scalar theory of pi-mesons interacting with a fixed nucleon source is truncated as follows: π-mesons are permitted to exist only in a set of discrete states $\psi_m(k)$ such that k is of order Λ^m in the state $\psi_m(k)$; Λ is an arbitrary constant above 4×10^6. Also two mesons of the same charge cannot occupy the same state. The resulting Hamiltonian can be solved by a perturbation expansion in Λ^{-1} provided there are only a finite number M of states ψ_m. When $M \to \infty$ the renormalized coupling constant and ground state energy diverge in perturbation theory (in the coupling constant). If the unrenormalized coupling constant is allowed to go to infinity as $M \to \infty$ it is proven that the renormalized theory exists (without ghost states) for any value of the renormalized coupling constant. The proof uses the perturbation analysis in Λ^{-1} carried to all orders. This analysis leads to the definition of a transformation T which eliminates one meson degree of freedom from any given Hamiltonian, replacing it by an effective Hamiltonian with one less degree of freedom. The effective Hamiltonian gives exactly all energy levels of the original Hamiltonian except those with mesons in the removed degree of freedom. The renormalizability of the theory is proven using topological properties of T. In particular there is a sub-transformation T_A with a nontrivial fixed point P_c whose properties determine the principal features of the renormalized theory. The idea of the fixed point is a generalization of the Gell-Mann-Low eigenvalue condition for the bare coupling constant of quantum electrodynamics.

I. INTRODUCTION

The problem of renormalization has been remarkably unimportant in the study of pure strong interactions (i.e., strong interactions without radiative or weak corrections). The ideas developed since 1954 — dispersion relations, Regge poles, current algebra, and pole dominance — all can be formulated and applied without encountering any of the divergences that occur in unrenormalized perturbation theory. As a result one gets the impression that renormalization is no more than a technical modification which one makes on closed loop Feynman graphs when very accurate perturbation formulae are needed, as for the electron magnetic moment. This impression has encouraged the idea that Lagrangian models of current algebra, such as field algebra and the Quark model, can be analyzed for their equal time commutators as if renormalization were unnecessary.[1]

An entirely opposite picture results from exactly soluble models of field theories with interaction. There are two known model theories which require wave function or charge renormalization, namely, the Lee model[2] and the Thirring model.[3] It is well known that the renormalized Lee model has a ghost state. The Thirring model involves the Fermi interaction for a zero mass spinor field in one space and one time dimension. The model has a solution after renormalization, but the solution has radically different behavior at short distances than one would expect from a canonical Lagrangian picture. The renormalized spinor field does not satisfy canonical commutation relations.[3] More generally, the renormalized theory is scale invariant, as one would have predicted from the Lagrangian (there are no dimensional parameters in the Thirring model, the only parameter being a dimensionless coupling constant). However the renormalized fields (but not the conserved currents) have different scaling properties from those one predicts from the canonical commutation rules. The dimension of the spinor field (which determines its scaling properties) depends on the coupling constant and can vary from 1/2 to ∞.[4]

The only known relativistic theories where renormalization does not affect the short distance behavior appreciably are the "superrenormalizable" theories which may require mass renormalization but do not require infinite coupling constant or wave function renormalization in perturbation theory.[5] In these theories the short distance behavior is close to the free field behavior. Unfortunately there are no acceptable four-dimensional superrenormalizable theories.

In a recent paper, it was proposed that there would be nontrivial renormalization effects in strong interactions.[6] It was postulated that these effects would have the same form as in the Thirring model, namely scale invariance would be valid at short distances but the dimensions of local fields would be different from any free field model (except for the currents of current algebra whose dimensions are fixed by the algebra). It was shown that renormalization effects could account for a universal $\Delta I = 1/2$ rule in weak interactions and could determine the convergence or divergence of some of the Weinberg sum rules.

The fact that the $\Delta I = 1/2$ rule might be explained by renormalization effects means that renormalization can be of great practical importance. One would like to understand renormalization better. The Lee model and the Thirring model fall far short of providing the depth of understanding required. The reason is that both models have very special features and the renormalization of these models may simply reflect these special features. The Lee model is special because of the decoupling of the N-θ channel from the many-particle channels. This decoupling is the simplification that makes solution of the Lee model possible. The Thirring model is special for many reasons, but in particular the electromagnetic current of the Thirring model satisfies a free field equation which is the starting point for solving the model. Also there is no coupling constant renormalization in the Thirring model. If there had been coupling constant renormalization in the Thirring model it might have shown the same diseases as the Lee model which does involve coupling constant renormalization.

The purpose of this paper is to define and solve a new model of coupling constant renormalization. The new model is a cousin of the Lee model but its renormalization is very different from that of the Lee model. The new model is a derivative of the charged scalar theory of pions coupled to a fixed nucleon source. The model Hamiltonian is obtained essentially by projecting the Hamiltonian of the charged scalar theory onto a specially constructed subspace of the original Hilbert space. The result of renormalizing the model is that the renormalized theory exists without ghosts, the renormalized coupling constant is arbitrary but the unrenormalized coupling constant is infinite.

The model of this paper cannot be solved in closed form. To make it soluble by series expansions a large parameter Λ is introduced artificially into the model; the model is then solved by an expansion in Λ^{-1}. The way Λ is introduced is by restricting the pi mesons of the model to be in one of a discrete set of wave functions $\psi_m(k)$, where the mean momentum of $\psi_m(k)$ is Λ^m (in units of the pion mass). Thus instead of the pion energy being continuously variable from 1 to ∞, it is restricted to the discrete values 1, Λ, Λ^2, etc. This means the Hamiltonian has some terms of order 1, some terms of order Λ, etc., so one can do perturbation theory when Λ is large. This idea was explained in an earlier paper[7] where a more complicated version of the model was proposed.

Because the model cannot be solved in closed form the renormalization analysis is much more complex than for either the Lee model or the Thirring model. The analysis is further complicated because one cannot simply study the lowest order term in the Λ^{-1} expansion. To prove the renormalizability of the theory one must show that the expansion in Λ^{-1} of the renormalized theory is finite to all orders and that the sum of the series converges. To prove this a rigorous analysis of the model is given using formal techniques of analysis in Hilbert space plus some topological methods. The formal analysis is possible because the model is specially constructed to involve only bounded operators. To ensure that no unbounded operators occur, the number of π mesons per state ψ_m is limited to one of each charge, and the total number of states ψ_m is cut off at m=M. One investigates the limit for $M \rightarrow \infty$, but for any finite M one has bounded operators.

The author recommends that the papers of Lee[2] (on the Lee model) and Johnson[3] (on the Thirring model) be read before attacking the present paper. They provide some background on exact solutions of renormalizable theories and are very much simpler to read.

There are three interesting features in the model of this paper. The first is simply that a finite renormalized theory exists. Actually, all that is proved is that the renormalized energy levels exists. Because there are no continuum (momentum) states open to pions there is no scattering in the model; all energy levels are discrete and hence calculating the energy levels is the most important problem in the model. The theory is found to be free of ghosts. No matrix elements of operators other than the Hamiltonian are discussed. In particular the nucleon isospin operators are not examined, which means we cannot compute the renormalized coupling constant as conventionally defined. The reason these operators are not considered is that the analysis that would be required exceeds the author's patience.

The second feature of the model is that scale invariance is preserved in the renormalized theory for energies large compared to the pion mass. The unrenormalized Hamiltonian of the full charged scalar theory is scale invariant in the limit of zero pion mass. This invariance is preserved in the unrenormalized Hamiltonian at the model except that it is a discrete invariance: only scale transformations which take wave functions $\psi_m(k)$ into wave functions $\psi_{m+\ell}(k)$ occur in the model. The renormalized energy levels exhibit scale invariance when the energies are large, but the scaling law is different from what one predicts from the unrenormalized Hamiltonian. To be precise the unrenormalized Hamiltonian H_0 goes into $\Lambda^{-1}H_0$ when $\psi_m \rightarrow \psi_{m+1}$, apart from terms of order 1, but the renormalized Hamiltonian H_R goes into $\Lambda^{-1}\beta H_R$ where β is a constant (about 1/2). So the model of this paper supports the hypothesis that renormalization can preserve scale invariance at large energies but will change the scaling laws of operators.

The third feature of the model, and probably the most important, is that in order to prove the renormalizability of the model it is necessary to define and study a topological transformation T acting on a space S of cutoff Hamiltonians. The space S contains the unrenormalized cutoff Hamiltonians for any cutoff M. However, it also contains cutoff Hamiltonians involving arbitrarily complicated interactions involving products of arbitrarily many meson creation and destruction operators. In other words the space S includes nonrenormalizable interactions of arbitrarily complicated structure. The transformation T takes a Hamiltonian with cutoff M into a Hamiltonian with cutoff M-1 without changing the physics of these Hamiltonians. To be precise, the original Hamiltonian and the transformed

Hamiltonian have exactly the same energy levels except for those energy levels with mesons explicitly present in the state ψ_M; such levels are not present in the transformed Hamiltonian. The transformation defines how the coupling constants of all possible interactions must change as the cutoff M changes in order to keep the energy levels of the theory fixed. Having very many coupling constants all changing as the cutoff changes is analogous to having an infinite number of counter terms in a renormalization analysis in ordinary perturbation theory. One has an infinite number of counter terms when one tries to renormalize a non-renormalizable theory. This is customarily regarded as a disaster, for one presumes that for every infinite counter term there is an arbitrary finite counter term, leading to an infinite number of parameters. This disaster does not occur in the model. The reason is that strict bounds on the coupling constants will be included in the definition of S, and one cannot introduce extra free parameters without violating these bounds. What actually happens is that the possible renormalizable theories of the model are described by effective cutoff Hamiltonians obtained by applying T an infinite number of times to the original unrenormalized uncutoff Hamiltonian. This means that the renormalized Hamiltonians must lie in a subspace R_S of S, where R_S is the limit of the subspaces $T^m(S)$ for $m \to \infty$. The space R_S is found to be a three-dimensional space for given cutoff M. Hence there are only three adjustable parameters in the renormalized Hamiltonian they are a scale factor, an additive constant, and the renormalized coupling constant (suitably defined).

If one is interested only in the first two features of the model one can probably skim the hard parts (Section V and Appendix B). One would read these sections in detail only to check for mistakes. However to understand the transformation T one must study the whole paper in detail; it is hard to have a clear understanding of the role of the transformation T without studying the spaces $T^m(S)$; one must see how these spaces shrink with m to the limiting space R_S, and one must understand in practice the relevance of these spaces to the renormalization problem. At present the only way to get the necessary practice is to work through the model of this paper.

Gell-Mann and Low have given a general discussion of nonperturbative renormalization theory, using quantum electrodynamics as an example.[8] The relation of their work to the type of model considered here is discussed in Section VII. The idea of a transformation T in which an infinite set of coupling constants are transformed as the cutoff M is reduced is a generalization of Gell-Mann and Low's idea of a cutoff-dependent electromagnetic coupling $e(\Lambda)$.

In the author's previous paper on model Hamiltonians,[7] a more complicated model was discussed, in which π mesons were allowed to have any momentum in the intervals $0 < k < k_0$, $\Lambda/2 < k < \Lambda$, $\Lambda^2/2 < k < \Lambda^2$, etc., where k_0 was a constant. This meant the meson creation and destruction operators were continuum creation and destruction operators, which are hardly suitable for rigorous analyses. The Λ^{-1} expansion was proposed but only carried out in lowest order. Even the lowest order calculation was complicated by the fact that the unperturbed Hamiltonians were themselves insoluble field theoretic Hamiltonians. One had to guess the qualitative structure of their solution. Furthermore as the cutoff M went to infinity the coupling constant in the unperturbed Hamiltonian had to become large resulting in closely spaced isobar states, which interfered with the perturbation calculation in Λ^{-1}. None of these difficulties are present in the model of this paper. The meson creation and destruction operators of this paper are defined to be discrete and bounded. The unperturbed Hamiltonians are finite dimensional and diagonalizable in closed form (cf. Table I). The energy level spacing of the unperturbed Hamiltonian does not become small for large coupling — the isobars in the previous theory involved many mesons in a single quantum state and this is forbidden in the present model. This means the present model lacks much of the physics of the full charged scalar theory; but it still illustrates the renormalization problem, which is its only purpose.

This paper divides into three stages. The first stage consists of Sections II, III, and IV. In Section II, the Hamiltonian of the model is defined. In Section III the perturbation expansion in Λ^{-1} is formulated for the cutoff Hamiltonian and some properties of the expansion are worked out in low orders. In Section IV a perturbation formula is defined which allows the Λ^{-1} expansion to be defined to all orders in a convenient form. The second stage consists of Sections V and VI. In Section V the transformation T is defined. Its principal properties are stated (Theorems 1-4, the proofs of these theorems are in Appendix B). Then the topological analysis required to prove renormalizability is carried through. Finally, the renormalized Hamiltonian is defined for any given renormalized coupling constant. In Section VI scale transformations are defined, and the scaling properties of the renormalized energy levels are computed. The third stage consists of Section VII, where it is shown that the transformation T is more than a technical device to prove the existence of the renormalized theory. Specifically it is shown that the renormalized theories are not the unique solution of any uncutoff Hamiltonian; instead the transformation T is involved in the definition of the renormalized theory, and this definition is most simply stated in terms of one of the fixed points of the transformation. We also relate the renormalization program of this paper to conventional renormalization theory and especially to the Gell-Mann-Low analysis.

II. THE MODEL HAMILTONIAN

The unrenormalized Hamiltonian of the model is as follows:

$$H = \sum_{m=0}^{\infty} \Lambda^m \left\{ \left(a_m^+ a_m + b_m^+ b_m - 1 \right) + g_0 \left(a_m + b_m^+ \right) \tau^+ + g_0 \left(a_m^+ + b_m \right) \tau^- \right\} \tag{II.1}$$

where g_0 and Λ are constants, τ^+ and τ^- are the isospin raising and lowering operators for the nucleon, and the operators a_m^+ and b_m^+ are π^+ and π^- creation operators respectively for the state ψ_m. The subtraction -1 is included for irrelevant reasons. The constant Λ must be large ($> 4\times 10^6$ in the rigorous analysis). To prevent two π^+ or two π^- from occupying the same state, the operators a_m, a_m^+, b_m, and b_m^+ are assigned the commutation relations of a set of Pauli spin operators:

$$\left\{ a_m, a_m^+ \right\} = \left\{ b_m, b_m^+ \right\} = 1 \tag{II.2}$$

$$a_m^2 = \left(a_m^+ \right)^2 = b_m^2 = \left(b_m^+ \right)^2 = 0 \tag{II.3}$$

$$\left[a_m, b_m \right] = \left[a_m, b_m^+ \right] = \left[a_m, a_n \right] = 0 \quad \text{(etc.)} \quad (m \neq n) \tag{II.4}$$

where $[\,]$ is a commutator, $\{\,\}$ is an anticommutator. The Hilbert space on which H acts is a product space. The components of the product are, first, the two-dimensional nucleon space with the bare proton state $|p\rangle$ and bare neutron state $|n\rangle$ as a basis. Secondly, for each wave function ψ_m there is a component space of four dimensions. A basis for each such component consists of a vacuum state, a π^+ state, a π^- state, and a $\pi^+\pi^-$ state, each meson being in the state ψ_m.

The model Hamiltonian can be arrived at starting from the full Hamiltonian of the charged scalar fixed source theory[9] if one replaces the fixed momentum creation operators $a^+_{\underset{\sim}{k}}$ and $b^+_{\underset{\sim}{k}}$ of the mesons by

$$a^+_{\underset{\sim}{k}} \rightarrow \sum_m a^+_m \psi_m(\underset{\sim}{k}) \tag{II.5}$$

$$b^+_{\underset{\sim}{k}} \rightarrow \sum_m b^+_m \psi_m(\underset{\sim}{k}) \quad . \tag{II.6}$$

After these substitutions are inserted in the full Hamiltonian, one must drop any off diagonal products such as $a^+_n a_m$ $(n \neq m)$ and replace integrals such as $\int_{\underset{\sim}{k}} \omega_{\underset{\sim}{k}} |\psi_m(\underset{\sim}{k})|^2$ or $\int_{\underset{\sim}{k}} (2\omega_{\underset{\sim}{k}})^{-1/2} \psi_m(\underset{\sim}{k})$ (where $\omega_{\underset{\sim}{k}}$ is $(1+\underset{\sim}{k}^2)^{1/2}$) by order of magnitude estimates, assuming the functions $\psi_m(\underset{\sim}{k})$ are normalized to one and vanish unless $\underset{\sim}{k} \sim \Lambda^m$. There is no need for the model Hamiltonian to have any connection with the fixed source theory, because the model will be studied on its own merits. The connection with the fixed source theory is used only to provide a language to describe the operators a_m, etc. Likewise, the wave functions $\psi_m(\underset{\sim}{k})$ play no role in the analysis of the model; their only purpose is to give an intuitive meaning to the operators a_m, etc.

One can cut off the Hamiltonian by restricting the sum over m to a finite range, say $0 \leq m \leq M$. Then the Hamiltonian becomes a finite bounded matrix; in this case it is diagonalizable without renormalization. The problem of renormalization arises when one tries to let $M \rightarrow \infty$. Then one has an infinite number of degrees of freedom, which is well known to be a source of difficulties.[10] To compound the situation the scale of energy associated with the m^{th} degree of freedom increases as Λ^m, so that the most important degrees of freedom are those with $m \sim M$ instead of small m. Clearly one has difficulties in the limit $M \rightarrow \infty$ regardless of what happens in perturbation theory; but it is still worth showing that in perturbation theory one has a problem specifically with coupling constant renormalization. Let $|P\rangle$ and $|N\rangle$ be the normalized physical proton and neutron states, i.e., the ground states of H. The renormalized coupling constant is

$$g_R = g_0 \langle P | \tau^+ | N \rangle \tag{II.7}$$

using the definition analogous to that used in the full charged scalar theory. The matrix element $\langle P | \tau^+ | N \rangle$ can be computed to second order in g_0 by straightforward perturbation theory. If the cutoff M is finite then

$$g_R = g_0 - g_0^3 \ (M + 1) \ + \ O\left(g_0^5\right) . \tag{II.8}$$

The cutoff momentum k_M is of order Λ^M so M is proportional to log k_M; hence g_R is logarithmically divergent as in the full charged scalar theory. The divergence for $M \rightarrow \infty$ is directly due to there being an infinite number of degrees of freedom in the no-cutoff limit.

The structure of the energy level spectrum of the cutoff Hamiltonian can be seen by a qualitative analysis. It is convenient to call a meson in a state $\psi_m(k)$ an "m-meson". Let the cutoff Hamiltonian be denoted H_M. It has the structure

$$H_M = \sum_{m=0}^{M} \Lambda^m O_m \tag{II.9}$$

where O_m is independent of Λ and involves only m-meson operators and the nucleon operators τ^+ and τ^-. The smallest part of H_M is O_0. This is the only part of H_M involving 0-mesons, and for Λ large O_0 is a perturbation on the rest of the Hamiltonian. The remainder of the Hamiltonian has energies of order Λ or larger so should have energy level spacings of order Λ; each level is four-fold degenerate (at least) because each level is independent of the presence or absence of 0-mesons. Adding O_0 splits these levels, with the splitting being of order 1. Next one can discuss the effect of the term ΛO_1; clearly this should lead to a gross spacing of order Λ, neglecting fine structure due to O_0. But ΛO_1 can itself be regarded as a perturbation; there exist (neglecting ΛO_1 and O_0) a spacing of order Λ^2, then a spacing of order Λ^3, etc.

The problem of renormalization is the problem of computing the ground state and those excited states which have a finite energy above the ground state in the limit $M \to \infty$. This means calculating states with an energy of order Λ^m above the ground state, for any m, but with m held fixed when $M \to \infty$. In practice one calculates only energy differences between the ground state and various excited states. The ground state energy itself diverges for $M \to \infty$. An energy difference of order Λ^m is much smaller than the basic energy scale Λ^M, when M is large, so a very precise calculation is required to give these energy differences accurately. This fact plus the fact that the model cannot be solved exactly, and must be solved as a perturbation expansion in Λ^{-1}, is the reason this paper is so long.

The model Hamiltonian is invariant to three symmetries: charge conservation, charge conjugation, and time reversal. The charge Q is

$$Q = \sum_m \left(a_m^+ a_m - b_m^+ b_m\right) + 1/2\,(\tau_z + 1) \tag{II.10}$$

where $1/2\,\tau_z$ is the z component of the nucleon isospin; Q commutes with H. The charge conjugation transformation interchanges π^+ with π^-, p with n. Let U_c be the unitary transformation giving these interchanges; then

$$U_c^+ a_m U_c = b_m \tag{II.11}$$

$$U_c^+ b_m U_c = a_m \tag{II.12}$$

$$U_c^+ \tau^+ U_c = \tau^- \tag{II.13}$$

$$U_c^+ H U_c = H \quad . \tag{II.14}$$

The time reversal transformation is an anti-linear unitary transformation U_T with the properties

$$U_T^+ a_m U_T = a_m^* \tag{II.15}$$

$$U_T^+ b_m U_T = b_m^* \tag{II.16}$$

$$U_T^+ \tau^+ U_T = (\tau^+)^* \tag{II.17}$$

$$U_T^+ H U_T = H^* \tag{II.18}$$

III. PRELIMINARY ANALYSIS OF THE MODEL HAMILTONIAN

In order to solve the renormalization problem, one must first be able to solve the cutoff Hamiltonian for arbitrarily large cutoff M. In this part, we give a preliminary discussion of the solution of the cutoff Hamiltonian for large M.[11] The constant Λ is also large, but held fixed and M can be arbitrarily large even compared to Λ. The cutoff Hamiltonian naturally separates into an unperturbed Hamiltonian and a perturbation:

$$H_M = H_{0M} + H_{IM} \tag{III.1}$$

where

$$H_{0M} = \Lambda^M O_M \tag{III.2}$$

$$H_{IM} = \sum_{m=0}^{M-1} \Lambda^m O_m \tag{III.3}$$

and

$$O_m = a_m^+ a_m + b_m^+ b_m - 1 + g_0 \left(a_m + b_m^+\right) \tau^+ + g_0 \left(a_m^+ + b_m\right) \tau^- \tag{III.4}$$

The operator O_M is easily diagonalized. One can ignore the mesons in states other than ψ_M, then O_M acts on the eight-dimensional Hilbert space involving the nucleon and mesons in the state ψ_M. Due to charge conservation the matrix for O_M separates into submatrices of size 3×3 at most. The eigenstates of O_M are given in Table I [the variables (m,g) of Table I must be replaced by $(1, g_0)$]. It has two degenerate ground states: a state $|P\rangle$ of charge 1 and a state $|N\rangle$ of charge 0. The ground state becomes highly degenerate when mesons in other states ψ_m are considered, since one can add such mesons to the states $|P\rangle$ and $|N\rangle$ without changing the eigenvalue of O_M.

The Hamiltonian H_{0M} has an energy level spacing of order Λ^M ($g_0 \Lambda^M$ if g_0 is large) while H_{IM} is at most of order Λ^{M-1} ($g_0 \Lambda^{M-1}$ for g_0 large). Hence one is allowed to treat H_{IM} as a perturbation when Λ is large, for any value of g_0. However one must carry the perturbation expansion out to order M at least because one ultimately is interested in energy level spacings which may be of order 1. In the lowest order of degenerate perturbation theory, the ground state of H and excited states at energy Λ^{M-1} or less above the ground state are given by an effective Hamiltonian

$$H_{eff} = E_{0M} + PH_{IM} P \tag{III.5}$$

where E_{0M} is the ground state energy of H_{0M} and P is a projection operator on the ground states of H_{0M}. H_{eff} acts on a product space whose components are the two-dimensional space with basis $|P\rangle$ and $|N\rangle$ and the meson space for the states ψ_m, $0 \le m \le M-1$. One can introduce isospin raising and lowering operators $\tau_R^{\pm}$ for $|P\rangle$ and $|N\rangle$; then H_{eff} involves a set of operators $\left[\tau_R^{\pm}, a_m \right.$ $\left.(0 \le m \le M-1), \text{etc.}\right]$ which are equivalent to the operators of H_{M-1}. The only way P affects the operator H_{IM} is through the nucleon operators τ^+ and τ^-; the meson operators in H_{IM} are unaffected. To express H_{eff} in terms of $\tau_R^{\pm}$ one must express $P\tau^+P$ and $P\tau^-P$ in terms of $\tau_R^{\pm}$ and the meson operators. The operator $P\tau^+P$ affects only the states $|P\rangle$ and $|N\rangle$ not the meson states, and because it increases the charge by one unit, $P\tau^+P$ must be proportional to τ_R^+. The proportionality constant Z is found from Table I to be (using the constants of Table I)

$$Z = \langle P|\tau^+|N\rangle = (m^2 + g^2)(m^2 + 2g^2)^{-1} \tag{III.6}$$

With m=1 and $g=g_0$ this is

$$Z(g_0) = \left(1 + g_0^2\right) \left(1 + 2g_0^2\right)^{-1} \quad . \tag{III.7}$$

Likewise $P\tau^- P$ is $Z(g_0)\ \tau_R^-$. Hence $PH_{IM}P$ has the same form as H_M itself except that M is replaced by M-1 and $g_0 \tau^{\pm}$ is replaced (in Eq. III.4) by $g_{M-1} \tau_R^{\pm}$, with

$$g_{M-1} = h(g_0) \tag{III.8}$$

$$h(g) = g(1 + g^2)\ (1 + 2g^2)^{-1} \tag{III.9}$$

When degenerate perturbation theory is carried to higher orders, one still computes an effective Hamiltonian H_{eff} which acts on the space of ground states of H_{0M}. The effective Hamiltonian is no longer just $PH_{IM}P$ but contains higher order terms in H_{IM}, for example, the second order term is $PH_{IM}(1-P)(E_{0M} - H_{0M})^{-1}\ H_{IM}P$. The term of n^{th} order involves products of n interaction Hamiltonians and n-1 energy denominators. It is useful to discuss in a schematic way the types of terms generated in the higher order calculation. Let x_m stand for an operator of the form $a_m^+ a_m + b_m^+ b_m - 1$, $a_m + b_m^+$, or $a_m^+ + b_m$. Let τ stand for any nucleon operator, τ_R for any operator acting on $|P\rangle$ and $|N\rangle$. Let x_m^2 stand for operators made of any product of operators of type x_m. One can easily make a table of the type of operators that can occur in H_{eff} for a given order in Λ, remembering that H_{eff} involves H_{IM} times products of $(E_{0M} - H_{0M})^{-1}\ H_{IM}$, the whole product being projected with P. The results are shown in Table II.

The formulae for the higher order terms of the degenerate perturbation calculation are too complicated to quote explicitly. Fortunately they are not needed; it is sufficient to have upper bounds for each type of term and these can be obtained. Table II gives the order in Λ of each term and all that remains is to obtain numerical upper bounds. This will be done in Section V.

The Hamiltonian H_{eff} has a basic energy scale $\sim\Lambda^{M-1}$ which is still much larger (for large M) than the energy scales of interest. H_{eff} can again be analyzed by perturbation theory. One writes

$$H_{eff} = H_{0eff} + H_{Ieff} \tag{III.10}$$

The unperturbed Hamiltonian is

$$H_{0eff} = \Lambda^{M-1} O_{M-1} (g_{M-1}) \tag{III.11}$$

where $O_{M-1}(g_{M-1})$ is the same as O_{M-1} except that g_0 is replaced by g_{M-1} and $\tau^{\pm}$ by $\tau_R^{\pm}$. All other terms in H_{eff} form the perturbation H_{Ieff}, which is at most of order Λ^{M-2}. The eigenstates of $O_{M-1}(g_{M-1})$ can be determined from Table I; like O_M it has two degenerate ground states $|P'\rangle$ and $|N'\rangle$ if mesons in states other than ψ_{M-1} are ignored. One can use degenerate perturbation theory starting from the states $|P'\rangle$ and $|N'\rangle$ to determine the eigenstates of H_{eff} of energy Λ^{M-2} or less above the ground state. Again one must calculate the perturbation analysis to many orders, in order to keep terms with energies of order 1 or larger. The result is a second effective Hamiltonian H'_{eff} involving meson operators a_m etc. for $m \leq M - 2$ and isospin operators $\tau_R^{\pm\prime}$ connecting the states $|P'\rangle$ and $|N'\rangle$.

One can determine the type of operators that occur in H'_{eff} for each order in Λ. The basic operators are operators acting on $|P'\rangle$ and $|N'\rangle$, denoted τ'_R, and meson operators of type x_m for $m \leq M - 2$. The results are shown in Table III. In constructing Table III, one uses the fact that operators of the form $(x_{M-1})\,\tau_R$ and $(x_{M-1})^2\,\tau_R$ in H_{Ieff} are reduced to the form τ'_R in H'_{eff}. Furthermore, the symmetries of the theory ensure that an operator of the form τ'_R in H_{eff} not multiplied by a meson operator can only be a constant. The important result illustrated by Table III is the following. To compute H'_{eff}, one must compute many orders in a perturbation treatment of H_{Ieff}. H_{Ieff} itself divides into two parts. The simple part of H_{Ieff} are the terms coming from $PH_{IM}P$; these terms have the structure $x_{M-2}\,\tau_R$, $x_{M-3}\,\tau_R$, etc., and depend only on the single constant g_{M-1}. The complex part of H_{Ieff} comes from the higher order terms in H_{IM}, and includes all terms of type $(x_{M-2})^2\,\tau_R$, $(x_{M-2})\,(x_{M-3})\,\tau_R$, etc. In computing H'_{eff}, even the simple part of H_{Ieff} generates all types of terms in H'_{eff}, through terms of order $(H_{Ieff})^2$, $(H_{Ieff})^3$, etc. The important fact is for a given term in H'_{eff}, say $(x_{M-2})^2\,\tau'_R$, its coefficient comes predominately from the simple part of H_{Ieff}, and hence the coefficient is primarily determined by the constant g_{M-1}. H_{Ieff} also has an $(x_{M-2})^2\,\tau_R$ term but this affects the coefficient of $(x_{M-2})^2\,\tau'_R$ only in order Λ^{M-4} whereas the dominant part of the coefficient is of order Λ^{M-3}. Because of this result one can give bounds on the complex terms like $(x_{M-2})^2\,\tau'_R$ in H'_{eff} which depend on g_{M-1} only and do not involve the size of the corresponding term in H_{eff}. These bounds are of crucial importance for the rigorous analysis; they ensure that the complex interactions cannot increase without bound as one repeats the perturbation analysis many times. Furthermore, it means that Table III has the same form it would have had if one had started with the cutoff M-1, and obtained H'_{eff} by solving H_{M-1}. The only exception is the constant in Table III of order Λ^M.

One can repeat the perturbation analysis many times generating a sequence of Hamiltonians which will be denoted $H_N(M)$. The Hamiltonian $H_{M-1}(M)$ is H_{eff}, $H_{M-2}(M)$ is H'_{eff}. In general $H_N(M)$ is the effective Hamiltonian after M-N perturbation calculations; $H_N(M)$ involves the meson operators a_m, etc. for $m \leq N$, and isospin operators analogous to $\tau_R^{\pm}$ or $\tau_R^{\pm\prime}$. The operator $H_N(M)$ gives the energy levels of H with energies of order Λ^N or less above the ground state. For each operator $H_N(M)$ one can give a classification table analogous to Tables II and III; the result is Table II with M replaced by N, except for constant terms. The unperturbed part of $H_N(M)$ would appear to be just $\Lambda^N O_N(g_N)$ where

$$g_N = h\,(g_{N+1}) \tag{III.12}$$

This is what one gets if the unperturbed Hamiltonian is defined as the term of order Λ^N in $H_N(M)$. However, to ensure that the perturbation is small even when M-N is much larger than Λ, the unperturbed part of $H_N(M)$ will be defined to include other terms of the form $\left(a_N + b_N^+\right)\tau^+$, or $\left(a_N^+ + b_N\right)\tau^-$, regardless of their order in Λ. The unperturbed Hamiltonian still has the form $\Lambda^N O_N(g_N)$, but g_N differs from $h(g_{N+1})$ in order Λ^{-1}. Since one has to compute a whole sequence of constants g_N ($N = M-1$, $M-2$, etc.) the small differences between g_N and $h(g_{N+1})$ for each N can build up to a macroscopic effect when M-N is large.

To compute an eigenstate of energy Λ^m above the ground state of H_M, one must take the effective Hamiltonian $H_m(M)$ and solve for states corresponding to excited states of the unperturbed part of $H_m(M)$. One could set up a perturbation method for computing these states. It will not be necessary for the purposes of this paper to discuss these states in detail, so the perturbation method will not be developed here.

IV. A PERTURBATION FORMULA

There are various standard formulae for the effective Hamiltonian that results when a perturbation H_I is treated to all orders. They all have drawbacks, so a suitable formula will be derived here. The formula obtained below has two properties: The effective Hamiltonian is hermitian, and involves only unperturbed energies in energy denominators. The second property is useful because the unperturbed energies are known explicitly. The first property is obviously useful, and is not true of many standard formulae.

Let $H = H_0 + H_I$ and let P be the projection operator on the ground states of H_0. Let $|\psi\rangle$ be any eigenstate of H with an energy E close to the ground state energy E_0 of H_0. It is convenient to have an operator R which gives the part of $|\psi\rangle$ outside the space projected by P in terms of the part of $|\psi\rangle$ inside the space. That is

$$(1-P)|\psi\rangle = RP|\psi\rangle \tag{IV.1}$$

Such an operator can be defined as follows. The eigenvalue equation has two parts:

$$E(1-P)|\psi\rangle = (1-P)\,H(1-P)|\psi\rangle + (1-P)\,H_IP|\psi\rangle \tag{IV.2}$$

$$EP|\psi\rangle = PH_I\,(1-P)|\psi\rangle + PHP|\psi\rangle \tag{IV.3}$$

If an operator R exists satisfying Eq. (IV.1) one can multiply the second equation (Eq. (IV.3)) by R and subtract from the first, giving

$$0 = \left\{(1-P)\,HR + (1-P)\,H_I - RPH_IR - RPH\right\} P|\psi\rangle \tag{IV.4}$$

Equation (IV.4) will certainly be satisfied if we demand that

$$(1-P)HR + (1-P)H_IP - RPH_IR - RPHP = 0 \tag{IV.5}$$

This equation can be cast in a form suitable for iteration in H_I. From the original definition of R, it should take states within the subspace projected by P into states orthogonal to this subspace; we can also require that R gives zero acting on states outside the subspace. This means that

$$R = RP \tag{IV.6}$$

$$R = (1-P)\ R \tag{IV.7}$$

Assuming this, and using the fact that $PH_0P = E_0P$, one can rewrite Eq. (IV.5) as

$$(E_0-H_0)R = (1-P)\ H_IP + (1-P)\ H_IR - RH_IP - RH_IR \tag{IV.8}$$

or

$$R = (E_0-H_0)^{-1}\ (1-P-R)\ H_I(P+R) \tag{IV.9}$$

This equation can be solved iteratively to give R as a power series in H_I. It is easily seen that the expansion satisfies the assumptions of Eqs. (IV.6) and (IV.7).

The argument so far does not prove that any operator R satisfying Eq. (IV.9) will also satisfy Eq. (IV.1), but this will be established later if H_I is sufficiently small.

One can now write Eq. (IV.3) as

$$EP|\psi\rangle = \left\{PH_0P + PH_IP + PH_IR\right\} P\,|\psi\rangle \tag{IV.10}$$

One could therefore define H_{eff} to be $H_0 + PH_IP + PH_IR$ except that PH_IR is not hermitian. The reason for this is that although two eigenstates $|\psi_1\rangle$ and $|\psi_2\rangle$ with distinct eigenvalues are orthogonal, the corresponding projected states $P\,|\psi_1\rangle$ and $P|\psi_2\rangle$ will probably not be orthogonal, and therefore cannot be distinct eigenstates of a hermitian operator. To remedy the situation, one notes that

$$\langle\psi_1\,|\,\psi_2\rangle = \langle\psi_1|P\,|\,\psi_2\rangle + \langle\psi_1|R^+R\,|\,\psi_2\rangle \tag{IV.11}$$

This suggests replacing the projected states $P|\psi_1\rangle$ and $P|\psi_2\rangle$ by the states $(1+R^+R)^{1/2}\, P|\psi_1\rangle$ and $(1+R^+R)^{1/2}\, P|\psi_2\rangle$, which are still states in the subspace projected by P but have the same scalar product as $|\psi_1\rangle$ and $|\psi_2\rangle$. The operator $(1+R^+R)^{1/2}$ is well defined as a power series in R^+R when H_I is small. To obtain H_{eff}, write the eigenvalue equation as

$$(E-H)\,(P+R)\,P|\psi\rangle = 0 \tag{IV.12}$$

and multiply by $(P+R^+)$:

$$E(P+R^+)\,(P+R)\,P|\psi\rangle = (P+R^+)\,H(P+R)\,P|\psi\rangle \tag{IV.13}$$

Now

$$(P+R^+)\,(P+R) = P+R^+R = (1+R^+R)\,P \tag{IV.14}$$

using Eqs. (IV.6) and (IV.7). Hence, multiplying Eq. (IV.13) by $(1+R^+R)^{-1/2}$ gives

$$E\,|\phi\rangle = H_{eff}\,|\phi\rangle \tag{IV.15}$$

where

$$|\phi\rangle = (1+R^+R)^{1/2}\,P|\psi\rangle \tag{IV.16}$$

$$H_{eff} = (1+R^+R)^{-1/2}\,(P+R^+)\,H(P+R)\,(1+R^+R)^{-1/2} \tag{IV.17}$$

The formula for H_{eff} is evidently hermitian.

The above argument is not rigorous, so it must now be proven that the eigenvalues of H_{eff} are the eigenvalues of H near E_0, and that eigenstates $|\phi\rangle$ of H_{eff} become eigenstates of $|\psi\rangle$ of H through the formula

$$|\psi\rangle = (P+R)\,(1+R^+R)^{-1/2}\,|\phi\rangle \tag{IV.18}$$

Assume that R is defined by Eq. (IV.9) solved by iteration assuming H_I is small. It is shown in Appendix A that the iteration converges if H_I is sufficiently small. The solution satisfies Eqs. (IV.6) and (IV.7). From these and Eq. (IV.9) one obtains

$$(1-P-R)\ H(P+R) = 0 \qquad \text{(IV.19)}$$

which is essentially Eq. (IV.5). Also

$$(1-P-R)\ (P+R) = 0 \qquad \text{(IV.20)}$$

This is because

$$(1-P-R)\ (P+R) = (1-P)(1-R)(1+R)\ P = (1-P)\ R^2 P \qquad \text{(IV.21)}$$

and

$$R^2 = RP(1-P)\ R = 0 \qquad \text{(IV.22)}$$

Let $|\phi\rangle$ be an eigenstate of H_{eff} in the subspace projected by P, and let E be its eigenvalue. Define $|\psi\rangle$ by Eq. (IV.18).

One can write

$$(1+R^+R)^{1/2}\ (E-H_{eff})\ P|\phi\rangle = 0 \qquad \text{(IV.23)}$$

Using Eqs. (IV.17), (IV.18) and (IV.14), Eq. (IV.23) may be rewritten

$$(P+R^+)\ (E-H)|\psi\rangle = 0 \qquad \text{(IV.24)}$$

This equation cannot be used to infer that $(E-H)|\psi\rangle = 0$ because $P+R^+$ projects onto a subspace and does not have an inverse. However, from Eqs. (IV.19) and (IV.20) one can obtain

$$(1-P-R)\ (E-H)|\psi\rangle = (1-P-R)\ (E-H)\ (P+R)\left\{(1+R^+R)^{-1/2}|\phi\rangle\right\} = 0 \qquad \text{(IV.25)}$$

Adding Eqs. (IV.24) and (IV.25) gives

$$(1+R^+-R)\ (E-H)\ |\psi\rangle = 0 \qquad \text{(IV.26)}$$

It is shown in Appendix A that $(1+R^{+}-R)$ has an inverse (for sufficiently small H_I) so this equation does imply that $|\psi\rangle$ is an eigenstate of H with eigenvalue E. The Hamiltonian H_{eff} has matrix elements equal to zero except within the subspace projected by P. Within this subspace H_{eff} has d orthogonal eigenstates, where d is the dimension of the subspace. These eigenstates define (through Eq. (IV.18)) d orthogonal eigenstates of H (orthogonality is easily verified). The energies of these eigenstates are close to E_0 because H_{eff} is approximately PHP when H_I is small so that R is small.

An alternative form of H_{eff} is obtained as follows. Write

$$H_{eff} = E_0 P + (1+R^{+}R)^{-1/2} (P+R^{+}) (H_I + H_0 - E_0) (P+R) (1+R^{+}R)^{-1/2} \tag{IV.27}$$

Using $(H_0 - E_0) P = 0$ and Eq. (IV.9), one can rewrite this as

$$H_{eff} = E_0 P + (1+R^{+}R)^{-1/2} (P+R^{+}) \left\{ H_I(P+R) - (1-P-R) H_I(P+R) \right\} (1+R^{+}R)^{-1/2} \tag{IV.28}$$

Using Eq. (IV.14), this simplifies to

$$H_{eff} = E_0 P + P (1+R^{+}R)^{1/2} H_I (P+R) (1+R^{+}R)^{-1/2} \tag{IV.29}$$

This formula is not manifestly hermitian, but H_{eff} is still hermitian since it is still defined by Eq. (IV.17).

V. EXACT PERTURBATION ANALYSIS OF THE MODEL

The outline of a method of solving the cutoff model Hamiltonian H_M has been given in Section III. One uses the definition of H_{eff} given in Section IV in each degenerate perturbation calculation. The result is that starting from H_M, for any M, one defines a sequence of effective Hamiltonians denoted $H_N(M)$ involving meson operators a_m, a_m^+, b_m, b_m^+ for $0 \leq m \leq N$ and isospin operators which will be denoted $\tau^{\pm}$ regardless of what states they act on ($|p\rangle$, $|n\rangle$, or $|P\rangle$, $|N\rangle$, or $|P'\rangle$, $|N'\rangle$, etc.). The effective Hamiltonians involve very complicated interactions, not just the O_m terms of the original model. From the analysis of Section III, one can expect to get upper bounds on these terms such that a $\Lambda^N O_N$ term is the dominant term in $H_N(M)$ provided an appropriate coupling constant replaces g_0 in O_N. The Hamiltonians $H_N(M)$ give the energies of the ground state of H_M and the excited states of H_M in which only the first N degrees of freedom are excited. If the energy levels are counted from the lowest level up, the ground state being number one, then $H_N(M)$ describes the first 2^{2N+3} levels of H_M.

The limit of no cutoff, that is the $M \to \infty$ limit, can be studied by studying the limits of $H_N(M)$ for fixed N, as $M \to \infty$. This means one is studying a fixed number of energy levels as M increases. It will be proven in this section that the limit of $H_N(M)$ for $M \to \infty$ exists provided one makes the renormalizations one expects from ordinary perturbation theory. This means that before letting $M \to \infty$ one must first subtract a constant E_M from $H_N(M)$ and allow the bare coupling constant g_{0M} to vary with M. The variation will be such that $g_{0M} \to \infty$ as $M \to \infty$, i.e., the interaction term in H_M swamps the free meson energy in the limit $M \to \infty$. The proof requires that Λ be larger than 4×10^6. The limit may exist for smaller Λ but in this case the upper bounds used in the proof no longer apply.

The Hamiltonians of this paper involve only bounded operators: the operators a_m, a_m^+, τ^+, τ^-, etc. All have operator bounds of order 1. Anyone with experience in rigorous quantum mechanics knows the joys of having only bounded operators. This ensures that terms that look small by a power of Λ will indeed be small if Λ is large enough; for finite M the perturbation expansions in Λ^{-1} will be easily proven to converge and one can concentrate on the problems of the $M \to \infty$ limit.

The analysis of the limit for $M \to \infty$ is still very complex; it will be presented here in a formal and not well-motivated manner. Before presenting the procession of theorems and definitions, the basic problem involved will be sketched briefly. The essential problem is to have a bound on the difference $\| H_N(M, g_{0M}) - E_M - H_N(L, g_{0L}) + E_L \|$ where $\| \; \|$ is the ordinary operator bound, and the dependence of $H_N(M)$ on g_{0M} has been noted explicitly. One must be able to show that this bound goes to zero as M and L go to ∞, provided the sequences $\{g_{0M}\}$ and $\{E_M\}$ have been chosen appropriately. The crucial step in establishing such a bound will be to show that the difference $H_N(M) - H_N(L) - E$ is arbitrarily small when M and L are large provided E is properly chosen and provided the terms of order $\Lambda^N O_N$ which dominate $H_N(M)$ and $H_N(L)$ have identical effective coupling constants (see Theorem 10). This condition will force one to have different bare coupling constants; $g_{0M} \neq g_{0L}$. As a preliminary to proving this theorem it will be proved (Theorem 1) that $H_N(M)$ is dominated by a term of the structure $\Lambda^N O_N$ with an appropriate effective coupling constant in O_N. This proof is necessary because otherwise one might worry that terms nominally of order Λ^{N-1} or less would be multiplied by powers of M, which would dominate the Λ^N term when $M \gg \Lambda$.

In order to clarify the calculation of bounds some topological language will be used. A space S of Hamiltonians will be defined which includes the effective Hamiltonians $H_N(M)$ as special cases. The perturbation analysis which defines $H_{N-1}(M)$ given $H_N(M)$ defines a transformation T on the space S. The space S will be defined so that T(S) is contained in S. A metric will be defined on S, and convergence questions discussed in terms of this metric. The Hamiltonians $H_N(M)$, considered for all possible values of g_0, define "curves" in S.

The exact and rigorous analysis of the renormalization problem begins here. The first step is to define the space S of Hamiltonians. It is convenient to adopt a specific way of representing the Hamiltonians that will be included in S. Let H be any Hamiltonian involving the meson degrees of freedom 0-N plus nucleon operators, for example $H_N(M)$ for some M. It will be convenient to renumber the meson operators, making the switches $a_0, b_0 \longleftrightarrow a_N, b_N, a_1, b_1 \longleftrightarrow a_{N-1}, b_{N-1}$, etc. In the new numbering a_m^+ creates a meson in the state ψ_{N-m}. This is to be true for all N, so the state associated with a_m^+ is different for different N. It is also convenient to separate an additive and a multiplicative factor from H, writing

$$H = J\mathscr{H} + \mathscr{E} \tag{V.1}$$

where J and $\mathscr{E}$ are constants. A normalization condition will be imposed on $\mathscr{H}$, determining J, but the separation of $\mathscr{E}$ from $J\mathscr{H}$ will be left indeterminate. (The transformation T will be defined to determine J, $\mathscr{H}$ and $\mathscr{E}$ separately.) One now lets $\mathscr{H}$ have the following structure:

$$\mathscr{H} = m V_{01} + \sqrt{2}\, g V_{02}\, \tau^+ + \sqrt{2}\, g V_{03}\, \tau^- + \sum_{k=1}^{N} \underset{\sim}{V}_k \cdot \underset{\sim}{A}_{k-1} + \sum_{k=0}^{N} C_k \tag{V.2}$$

where $\underset{\sim}{V}_k$ is a vector (V_{k1}, V_{k2}, V_{k3}) and

$$\underset{\sim}{V}_k = \sum_{m=k}^{N} \underset{\sim}{T}_m \tag{V.3}$$

$$T_{m1} = \Lambda^{-m}\left(a_m^+ a_m + b_m^+ b_m - 1\right) \tag{V.4}$$

$$T_{m2} = (1/\sqrt{2})\ \Lambda^{-m}\left(a_m + b_m^+\right) \tag{V.5}$$

$$T_{m3} = (1/\sqrt{2})\ \Lambda^{-m}\left(a_m^+ + b_m\right) \tag{V.6}$$

and A_{k1}, A_{k2}, A_{k3}, and C_k are operators which depend only on $\tau^{\pm}$ and the meson operators numbered from 0 to k. The vector notation $\underset{\sim}{A}_k$, $\underset{\sim}{S}_k$, etc. is used purely for convenience. The constants m and g will be required to satisfy a normalization condition:

$$m^2 + 2\,g^2 = 1 \tag{V.7}$$

To ensure this normalization condition, m and g will be represented as

$$m = \cos\,\theta \tag{V.8}$$

$$g = (1/\sqrt{2})\,\sin\,\theta \tag{V.9}$$

The set of parameters J, $\mathscr{E}$, N, and θ, and the operators $\underset{\sim}{A}_k$ and C_k will be called the "decomposition" of H. The representation is highly redundant; for example, C_N is by itself totally arbitrary. The reason for using this redundant representation is the following: One can see from Table II that the operators a_k, b_k for large k (new numbering) appear in any effective Hamiltonian $H_N(M)$ predominantly in terms such as $\underset{\sim}{V}_1 \cdot \underset{\sim}{A}_0$ or $\underset{\sim}{V}_2 \cdot \underset{\sim}{A}_1$. Terms which must go into C_k (the x_k^2 terms of Table II) have much smaller coefficients. Hence by making the separation one can put stringent upper bounds on the operators C_k.

The space S will be defined in two steps, the first step being to define a subsidiary space S_A.

<u>Definition.</u> A point $P_A \in S_A$ consists of an angle θ and an infinite set of operators $\underset{\sim}{A}_k$ and C_k $(0 \le k < \infty)$. The angle θ is restricted to the range $0 \le \theta \le \pi/2$. The operators $\underset{\sim}{A}_k$ and C_k can depend only on the nucleon isospin operators $\tau^{\pm}$ and the meson operators a_m, a_m^+, b_m, and b_m^+ for $0 \le m \le k$. The dependence on these operators is arbitrary except as follows. The operators $\underset{\sim}{A}_k$ and C_k must satisfy the following operator bounds:

$$\|A_{k1}\| \le 200\ mg^2\ \Lambda^{-k-1} \tag{V.10}$$

$$\|A_{k2}\| \le 200\ \sqrt{2}\ g^3\ \Lambda^{-k-1} \tag{V.11}$$

$$\|A_{k3}\| \le 200\ \sqrt{2}\ g^3\ \Lambda^{-k-1} \tag{V.12}$$

$$\|C_k\| \le 200\ g^2\ \Lambda^{-2k-1} \tag{V.13}$$

where $m = \cos\theta$, $g = (1/\sqrt{2})\sin\theta$. Secondly, the operators $\underset{\sim}{A}_k$ and C_k must satisfy symmetry requirements: A_{k1} and C_k must carry charge 0 while A_{k2} creates one unit of charge and A_{k3} destroys a unit of charge. Under charge conjugation $A_{k1} \to A_{k1}$, $A_{k2} \leftrightarrow A_{k3}$, and $C_k \to C_k$. Under time reversal $\underset{\sim}{A}_k \to \underset{\sim}{A}_k^*$ and $C_k \to C_k^*$. Also A_{k1} and C_k must be hermitian, while $A_{k3} = A_{k2}^+$. These requirements ensure that $\mathscr{H}$ (defined by Eq. (V.2)) is hermitian and invariant to the symmetries. The parameter θ and the operators $\underset{\sim}{A}_k$ and C_k will be called the decomposition of P_A.

The powers of Λ in these bounds are what one would expect from Table II; the coefficients are hindsight bounds. It is convenient for the following analysis to insist that an infinite set of $\underset{\sim}{A}_k$ and C_k be specified even if a particular Hamiltonian involves only a finite subset of them. The superfluous $\underset{\sim}{A}_k$ and C_k can be chosen arbitrarily subject to the restrictions of the definition of S_A.

The space S is defined as follows:

<u>Definition</u>. A point $P \in S$ consists of three constants J, $\mathscr{E}$, and N, and a point $P_A \in S_A$. The four objects J, $\mathscr{E}$, N, and P_A will be called the decomposition of P. N must be an integer, J must be positive, but $\mathscr{E}$ is arbitrary. There are no upper bounds on J, $|\mathscr{E}|$, or N.

Next the transformation T acting on S will be defined. Many details of the definition are handled in Appendix B, only an outline is given here. Any Hamiltonian H in S has a dominant term of the form

$$H_0 = \mathscr{E} + J\left\{m\left(a_0^+ a_0 + b_0^+ b_0 - 1\right) + g\left(a_0 + b_0^+\right)\tau^+ + g\left(a_0^+ + b_0\right)\tau^-\right\} \tag{V.14}$$

The remaining terms in H form a perturbation H_I:

$$H_I = H - H_0 \tag{V.15}$$

From the definitions (V.1) and (V.2) and the bounds (V.10) - (V.13), H_I is of order $J\Lambda^{-1}$ or less and therefore can be treated as a perturbation relative to H_0. In particular, one can use the formulae of Section IV to define a new Hamiltonian H_{eff} whose eigenvalues are the eigenvalues of H near the ground state energy of H_0.

Suppose H has a decomposition (J, $\mathscr{E}$, N, P_A) (with P_A in S_A). The Hamiltonian H_{eff} can also be decomposed in the form (J', $\mathscr{E}'$, N', P_A') with P_A' in S_A, that is H_{eff} can be written in the form defined in Eqs. (V.1) - (V.9). (The resulting operators A_k', etc. satisfy the bounds of Eqs. (V.10) - (V.13); see Theorem 1.) Specific formulae for J', $\mathscr{E}'$, N', and P_A' (i.e., g', m', $\underset{\sim}{A}_k'$, and C_k') are obtained in Appendix B. (Cf., Eqs. (B.20) - (B.24).) The general form of these formulae is as follows:

$$N' = N-1 \tag{V.16}$$

$$J' = \Lambda^{-1} J\, T_B(P_A) \tag{V.17}$$

and

$$\mathscr{E}' = \mathscr{E} + J\, T_C\,(P_A) \tag{V.18}$$

$$P'_A = T_A\,(P_A) \tag{V.19}$$

where $T_B\,(P_A)$ and $T_C\,(P_A)$ are functions depending on P_A but not N, J, or $\mathscr{E}$, and T_A is a transformation on the space S_A, independent of N, J, or $\mathscr{E}$. It is clear that J and $\mathscr{E}$ will be multiplicative and additive factors in H_{eff} so do not effect T_A, T_B, or T_C. It is less obvious that T_A, T_B, and T_C can be defined to be independent of N; this result is proven in Appendix B. Equations (V.16) - (V.19) define the transformation T.

The reason for defining the subsidiary space S_A is that the transformation T_A acts on this space, and it is convenient to do much of the topological analysis on the transformation T_A rather than on T itself. The space S_A is a continuous closed space; in particular, it does not involve the discrete variable N.

The unrenormalized cutoff Hamiltonians H_M are all in S. The decomposition of H_M can be defined to be

$$J = \Lambda^M \left(1 + 2\, g_0^2\right)^{1/2} \tag{V.20}$$

$$\mathscr{E} = 0 \tag{V.21}$$

$$\theta = \tan^{-1}\,(\sqrt{2}\, g_0) \tag{V.22}$$

$$m = \left(1 + 2\, g_0^2\right)^{-1/2}, \quad g = g_0 \left(1 + 2\, g_0^2\right)^{-1/2} \tag{V.23}$$

$$\underset{\sim}{A}_k = C_k = 0 \qquad \text{(all k)} \tag{V.24}$$

g_0 must be positive so that θ lies between 0 and $\pi/2$. Note that $m \leq 1$ and $g \leq (1/\sqrt{2})$, this is required by the normalization condition (V.7).

In Appendix B several theorems about the transformation T_A are proven. These theorems will be quoted below and are the basis for the analysis in this section.

Theorem 1. If $P_A \in S_A$ then $T_A(P_A)$ is also in S_A, i.e.,

$$T_A(S_A) \subset S_A \tag{V.25}$$

Theorem 2. Let $P_A \in S_A$ have a decomposition $(\theta, \underset{\sim}{A}_k, C_k)$, and let $T_A(P_A)$ have the decomposition θ', $\underset{\sim}{A}'_k$, and C'_k. Let $m = \cos\theta$ and $g = (1/\sqrt{2}) \sin\theta$. Then

$$\tan\theta' = \sqrt{2}\, g''/m'' \tag{V.26}$$

$$T_B(P_A) = (m''^2 + 2\, g''^2)^{1/2} \tag{V.27}$$

where

$$|m'' - m| < .01\, m g^2 \tag{V.28}$$

$$|g'' - g(1 - g^2)| < .01\, g^3 \tag{V.29}$$

Also

$$|T_C(P_A) + 1| < .01 \tag{V.30}$$

$$(1 - .51 \sin^2\theta) \tan\theta \le \tan\theta' \le (1 - .48 \sin^2\theta) \tan\theta \tag{V.31}$$

Theorem 3. Let P_A and $T_A(P_A)$ have the decompositions defined in Theorem 2. Let the component A_{k1} of $\underset{\sim}{A}_k$ vanish for all k. Then

$$A'_{k1} = 0 \qquad \text{(all k)} \tag{V.32}$$

$$m'' = m \tag{V.33}$$

where m" is the constant in Theorem 2.

The significance of these theorems is essentially as follows. Theorem 1 ensures that if the decomposition of P_A satisfies the bounds (V.10) - (V.13), then so does the decomposition of $T_A(P_A)$. A consequence of Theorem 1 is that the effective Hamiltonians $H_N(M)$ are in S for any N, any M, and any value of g_0. Theorem 2 gives limits on the values of $T_B(P_A)$, $T_C(P_A)$, and θ' which depend only on m and g, not on $\underset{\sim}{A}_k$ and C_k. The constants m'' and g'' appear in an intermediate stage in the calculation of H_{eff}. To lowest order in Λ^{-1}, g'' is equal to $g(1-g^2)$; this follows from Eq. (III.6) using Eq. (V.7). The bounds in Eqs. (V.28) - (V.30) were originally of order Λ^{-1}, but were replaced by numerical bounds (valid for $\Lambda > 4 \times 10^6$) for convenience. Theorem 3 shows that A_{k1} will vanish for the effective Hamiltonians $H_N(M)$. It was not obvious (to the author, at least) that this would be so.

Before presenting Theorem 4, a metric must be defined in the space S_A. Let $P_A = (\theta, \underset{\sim}{A}_k, C_k)$ and $P'_A = (\theta', \underset{\sim}{A}'_k, C'_k)$ be two points in S_A. It is convenient to define two distances in S_A, one being a distance between θ and θ', the other a distance between the operators $\left\{\underset{\sim}{A}_k, C_k\right\}$ and the operators $\left\{\underset{\sim}{A}'_k, C'_k\right\}$. It is also convenient to use the notation $|P_A - P'_A|$ for the pair of distances (d_1, d_2).

<u>Definition</u>. Let $P_A = (\theta, \underset{\sim}{A}_k, C_k)$ and $P'_A = (\theta', \underset{\sim}{A}'_k, C'_k)$ be in S_A. Then $|P_A - P'_A| = (d_1, d_2)$ with

$$d_1 = 2\,|\sin 1/2\,(\theta - \theta')| \tag{V.34}$$

$$d_2 = \mathrm{Max}\left\{\sqrt{2}\,\Lambda^{k+1}\,\|A_{ki} - A'_{ki}\|,\ \Lambda^{2k+1}\,\|C_k - C'_k\|\right\} \tag{V.35}$$

where the maximum is over all possible values of k and i.

The distance d_1 is more transparent if written in terms of m, g, m', and g':

$$d_1 = \left\{(m-m')^2 + 2\,(g-g')^2\right\}^{1/2} \tag{V.36}$$

No <u>a priori</u> rationale for these definitions of d_1 and d_2 will be given. A certain amount of experimentation was required to determine how to define these distances; the above formulae turned out to be useful. It is clear from Eqs. (V.35) and (V.36) that the metric satisfies the triangle inequality and that $|P_A - P'_A| = (0, 0)$ only if $P_A = P'_A$.

Theorem 4. Let P_A and P'_A be in S_A. Let $|P_A - P'_A| = (d_1, d_2)$ and

$|T_A(P_A) - T_A(P'_A)| = (d'_1, d'_2)$.

Then

$$.38\, d_1 - 10^{-5}\, d_2 \le d'_1 \le 20\, d_1 + 10^{-5}\, d_2 \tag{V.37}$$

$$d'_2 \le 1100\, d_1 + .06\, d_2 \tag{V.38}$$

The coefficients 10^{-5} and .06 are numerical upper bounds to terms behaving as Λ^{-1}. These bounds are valid for $\Lambda > 4 \times 10^6$. The first set of bounds force d'_1 to be of order d_1 unless $d_2 \gg d_1$; d'_1 cannot be much less or much greater than d_1 unless $d_2 \gg d_1$. The second bound is a straight upper bound on d'_2. In particular if $d_1 = 0$ then d'_2 is smaller than d_2. Hence as long as $\theta = \theta'$, the transformation T_A brings the points P_A and P'_A closer together.

The four theorems stated above are proved in Appendix B. The only assumption is $\Lambda > 4 \times 10^6$. The remainder of the analysis of this section is self-contained. The next stage is a set of topological theorems and definitions. First one defines a set of curves Q_L in the space S_A. They are generated by the effective Hamiltonians $H_N(M)$ as a function of the coupling constant g_0. The curves turn out to depend only on the difference $L = M - N$, not M or N separately. It is convenient to parameterize these curves by their θ coordinate rather than by the unrenormalized coupling constant. The parameter in these curves will be denoted t. Let the decomposition of $Q_L(t)$ be written

$$(\theta_L(t), \underset{\sim}{A}_{Lk}(t), C_{Lk}(t)) .$$

Definition. $Q_0(t)$ is the curve

$$\theta_0(t) = t \tag{V.39}$$

$$\underset{\sim}{A}_{0k}(t) = C_{0k}(t) = 0 \tag{V.40}$$

Definition. The curve Q_L is defined iteratively for $L > 0$ by the relation

$$Q_L = T_A (Q_{L-1}) \tag{V.41}$$

If one were parameterizing using the unrenormalized coupling constant, one would have had $Q_L(t) = T_A (Q_{L-1}(t))$. With the alternative parameterization, $Q_L(t)$ must still be the transform of some point on Q_{L-1}. This point can be denoted $Q_{L-1}(F_L(t))$:

$$Q_L(t) = T_A \left(Q_{L-1} \left[F_L(t) \right] \right) \tag{V.42}$$

Definition. The parameterization $Q_L(t)$ of Q_L is to be chosen so that

$$\theta_L(t) = t \qquad \text{(all L)} \tag{V.43}$$

In practice this definition defines the function $F_L(t)$.

We shall also be interested in the inverse function $f_L(t)$ to $F_L(t)$. This function satisfies

$$Q_L (f_L(t)) = T_A \ (Q_{L-1}(t)) \tag{V.44}$$

Since the θ coordinate of $Q_L (f_L(t))$ is $f_L(t)$, one has

$$f_L(t) = \theta \text{ coordinate of } T_A (Q_{L-1}(t)) \tag{V.45}$$

The next theorem gives several properties of $Q_L(t)$, $f_L(t)$, and $F_L(t)$. These properties will be established simultaneously in a proof by induction.

Theorem 5.

a. $Q_L(t)$ is a single-valued function of t defined for $0 \le t \le \pi/2$.

b. $f_L(t)$ is a continuous single-valued function of t defined for $0 \le t \le \pi/2$ satisfying

$$f_L(0) = 0 \tag{V.46}$$

$$f_L(\pi/2) = \pi/2 \tag{V.47}$$

$$0 < f_L(t) < t \qquad \text{for} \qquad 0 < t < \pi/2 \tag{V.48}$$

c. $F_L(t)$ is a continuous single-valued function of t defined for $0 \le t \le \pi/2$ satisfying

$$F_L(0) = 0 \tag{V.49}$$

$$F_L(\pi/2) = \pi/2 \tag{V.50}$$

$$t < F_L(t) < \pi/2 \qquad \text{for } 0 < t < \pi/2 \tag{V.51}$$

d. Consider any pair of numbers t and t' in the range 0 to $\pi/2$. Let $|Q_L(t) - Q_L(t')| = (d_1, d_2)$. Then

$$d_2 \le 4000\, d_1 \tag{V.52}$$

$$|f_L(t) - f_L(t')| \le 40\, |t - t'| \tag{V.53}$$

$$|F_L(t) - F_L(t')| \le 40\, |t - t'| \tag{V.54}$$

Part a is the crucial part of the theorem. It states that the curve Q_L, projected on the θ axis, covers the full range $0 \le \theta \le \pi/2$ once and only once. If, for example, the curve Q_L covered only part of this range, the theory would not be renormalizable. This point will be discussed later.[12]

Proof of Theorem 5. The property a and Eq. (V.52) hold for L = 0. That is, $Q_0(t)$ satisfies a from its definition, and $|Q_0(t) - Q_0(t')| = (d_1, 0)$ for all t and t' so satisfies (V.52). Suppose property a and Eq. (V.52) are true of $Q_{L-1}(t)$. We prove a - d for Q_L, F_L, and f_L. Equations (V.46) - (V.48) are consequences of the inequalities (V.31) (remember that the θ coordinate of $Q_{L-1}(t)$ is t). Now let t'' and t''' be two parameters in the range 0 to $\pi/2$. Let $|Q_{L-1}(t'') - Q_{L-1}(t''')| = (d_1, d_2)$, and let $|T_A(Q_{L-1}(t'')) - T_A(Q_{L-1}(t'''))| = (d_1', d_2')$. These distances must satisfy the inequalities of Theorem 4, and d_2 satisfies Eq. (V.52) by assumption. These equations can be combined to give inequalities not involving d_2:

$$.34\, d_1 \le d_1' \le 21\, d_1 \tag{V.55}$$

$$d_2' \le 1340\, d_1 \tag{V.56}$$

Note that

$$d_1 = 2\,|\sin 1/2\,(t'' - t''')| \tag{V.57}$$

$$d_1' = 2\left|\sin 1/2\left[f_L(t'') - f_L(t''')\right]\right| \tag{V.58}$$

Because t'', t''', $f_L(t'')$, and $f_L(t''')$ all lie between 0 and $\pi/2$, the arguments of the sines lie between $-\pi/4$ and $\pi/4$. For angles ϕ in this range

$$(2\sqrt{2}/\pi)\,|\phi| \leq |\sin\phi| \leq |\phi| \tag{V.59}$$

One deduces from Eqs. (V.55) and (V.57) - (V.59) that

$$|f_L(t'') - f_L(t''')| \leq (21\,\pi/2\sqrt{2})\,|t'' - t'''| \tag{V.60}$$

This proves that $f_L(t)$ is continuous; it also proves Eq. (V.53). Since $f_L(t)$ is continuous and satisfies Eqs. (V.46) and (V.47), there must be at least one root t' to the equation $t = f_L(t')$ for any t between 0 and $\pi/2$. This equation cannot have two roots t'' and t''' for if $t = f_L(t'') = f_L(t''')$ then $d_1' = 0$; by Eq. (V.55), d_1 must also be zero which means that $t'' = t'''$. Finally if $t = f_L(t'')$ then $t \leq t'' \leq \pi/2$ (using Eq. (V.48)). Hence $F_L(t)$ (the inverse function to $f_L(t)$) satisfies c. Now let t and t' be arbitrary parameters in the range 0 to $\pi/2$. Let $t'' = F_L(t)$ and $t''' = F_L(t')$. Then $t = f_L(t'')$, $t' = f_L(t''')$. Using Eqs. (V.55) and (V.57) - (V.59), one gets

$$|F_L(t) - F_L(t')| < \pi(.68\sqrt{2})^{-1}\,|t - t'| \tag{V.61}$$

which proves Eq. (V.54). Furthermore, the inequalities (V.55) and (V.56) give $d_2' \leq 4000\, d_1'$ which proves Eq. (V.52). Finally, a is a consequence of c, using Eq. (V.42) and the continuity of T_A (Theorem 4).

The next problem is to discuss the limit of the curve Q_L for $L \rightarrow \infty$. Determining the limit of $Q_L(t)$ for $L \rightarrow \infty$ with t held fixed is equivalent to determining the limit of the Hamiltonians $H_N(M)$ for $M \rightarrow \infty$ holding the effective coupling constant in $H_N(M)$ fixed. It is convenient to introduce subsets S_L of S_A which contain Q_L. The set S_0 is the set S_A itself, the definition of S_L is

<u>Definition</u>. S_L for $L > 0$ is the set

$$S_L = T_A\,(S_{L-1}) \tag{V.62}$$

S_L consists of all points in S_A which can be obtained by applying the transformation T_A L times to some point in S_A. Evidently all points in S_L also are in S_{L-1}:

<u>Theorem 6</u>.

$$S_L \subset S_{L-1} \quad \text{for} \quad L \geq 1 \tag{V.63}$$

The following theorem gives an upper bound on the "cross-sectional size" of S_L for given angle θ:

<u>Theorem 7</u>. Let P_A and P'_A be any pair of points in S_L. Let $|P_A - P'_A| = (d_1, d_2)$.

Then

$$d_2 \leq 4000\, d_1 + 300 \times (.2)^L \tag{V.64}$$

The cross section is the maximum value of d_2 for $d_1 = 0$. Theorem 7 states that the cross section goes to zero as $L \to \infty$; the spaces S_L shrink to a single curve as $L \to \infty$ (see below).

<u>Proof of Theorem 7</u>. The proof is by induction. For $L = 0$ the theorem is true simply because the bounds (V.10) - (V.13) force d_2 to be less than 300 for any pair of points in S_A. Suppose the theorem is true for S_{L-1}. Let P_A and P'_A be two points in S_L. Let $|P_A - P'_A|$ be (d'_1, d'_2). There must exist (by definition of S_L) two points P_B and P'_B in S_{L-1} with $P_A = T_A(P_B)$, $P'_A = T_A(P'_B)$. Let $|P_B - P'_B|$ be (d_1, d_2). Then the distances d_1, d_2, d'_1, and d'_2 satisfy Eqs. (V.37) and (V.38). Also d_2 satisfies the inequality (V.64) with L-1 substituted for L.

Combining these inequalities gives

$$.34\, d_1 - .003 \times (.2)^{L-1} \leq d_1' \tag{V.65}$$

$$d_2' \leq 1340\, d_1 + 18 \times (.2)^{L-1} \tag{V.66}$$

These inequalities can be combined to give[13]

$$d_2' \leq 4000\, d_1' + 300 \times (.2)^{L} \tag{V.67}$$

Q.E.D.

The next three theorems will be used to show that the curves $Q_L(t)$ have a limit curve R(t) for $L \rightarrow \infty$. The curve R is the limit of the subsets S_L for $L \rightarrow \infty$. The curve R has the property $T_A(R) = R$: it is an invariant subspace of the transformation T_A.

<u>Theorem 8</u>. Let $\{P_L\}$ be a sequence of points with $P_L \in S_L$. Denote the θ coordinate of P_L by θ_L. Assume that θ_L approaches a limit θ for $L \rightarrow \infty$. Define P_L' to be $P_L' = T_A(P_L)$. Denote the θ coordinate of P_L' by θ_L'. Then

a. $\lim\limits_{L \rightarrow \infty} P_L$ exists: call this limit R

b. $\lim\limits_{L \rightarrow \infty} \theta_L' = \theta'$ exists

c. $\lim\limits_{L \rightarrow \infty} P_L' = T_A(R)$

<u>Proof of Theorem 8</u>. Let L be large and K be even larger. Because $S_K \subset S_L$, both P_L and P_K are in S_L. Let $|P_K - P_L| = (d_1, d_2)$. Then

$$d_1 = 2\,|\sin 1/2\, (\theta_K - \theta_L)| \tag{V.68}$$

and by Theorem 7

$$d_2 \leq 4000\, d_1 + 300 \times (.2)^{L} \tag{V.69}$$

One can make both d_1 and d_2 arbitrarily small by choosing L and K large enough. This is true of d_1 by the assumption that θ_L approaches a limit for $L \to \infty$. It is true of d_2 from Eq. (V.69). Hence by the Cauchy criterion the sequence P_L has a limit R. That is, if P_L has a decomposition θ_L, $\underset{\sim}{A}_{Lk}$, C_{Lk}, then θ_L, $\underset{\sim}{A}_{Lk}$, and C_{Lk} all have limits for $L \to \infty$, and the limits θ, $\underset{\sim}{A}_k$, and C_k define the point R.

To prove b and c consider the distances $|P_L - R| = (d_1', d_2')$ and $|P_L' - T_A(R)| = (d_1'', d_2'')$. Since Eqs. (V.68) and (V.69) hold for any K, they hold for the limit $K \to \infty$, giving

$$d_1' = 2\,|\sin 1/2\,(\theta_L - \theta)| \tag{V.70}$$

$$d_2' \leq 4000\, d_1' + 300 \times (.2)^L \tag{V.71}$$

One can make d_1' and d_2' arbitrarily small by making L large enough. Therefore, due to inequalities of Theorem 4, one can also make d_1'' and d_2'' small enough by making L large enough. Hence c is true, and b is a corollary of c.

<u>Theorem 9</u>. Let $\{P_L\}$ and $\{P_L''\}$ be any two sequences satisfying $P_L \in S_L$ and $P_L'' \in S_L$. Let the θ coordinates of P_L and P_L'' be θ_L and θ_L'' respectively. Assume that the sequences θ_L and θ_L'' approach the same limit θ as $L \to \infty$. Then

$$\lim_{L\to\infty} P_L = \lim_{L\to\infty} P_L'' \tag{V.72}$$

The proof is simple. Let $|P_L - P_L''| = (d_1, d_2)$. Then since P_L and P_L'' are in S_L

$$d_1 = 2\,|\sin 1/2\,(\theta_L - \theta_L'')| \tag{V.73}$$

$$d_2 \leq 4000\, d_1 + 300 \times (.2)^L \tag{V.74}$$

As $L \to \infty$, $d_1 \to 0$ and hence $d_2 \to 0$ also. Q.E.D.

Theorem 10.

a. $\lim_{L \to \infty} Q_L(t) = R(t)$ exists for all t in the range $0 \le t \le \pi/2$.

b. $\lim_{L \to \infty} f_L(t) = f(t)$ exists $(0 \le t \le \pi/2)$.

c. $T_A(R(t)) = R(f(t))$ (V.75)

d. $f(0) = 0$ (V.76)

$f(\pi/2) = \pi/2$ (V.77)

$0 < f(t) < t \qquad (0 < t < \pi/2)$ (V.78)

e. $\lim_{L \to \infty} F_L(t) = F(t)$ where F is the inverse function to f; also both F(t) and f(t) are continuous single-valued functions of t defined for $0 < t < \pi/2$.

f. $F(0) = 0$ (V.79)

$F(\pi/2) = \pi/2$ (V.80)

$t < F(t) < \pi/2 \qquad (0 < t < \pi/2)$ (V.81)

Proof of Theorem 10. Part a is a consequence of Theorem 8a. Now let $P_L = Q_L(t)$ be a sequence as in Theorem 8; define $P'_L = T_A(P_L)$ as in Theorem 8. Then θ'_L is

$$\theta'_L = f_{L+1}(t) \tag{V.82}$$

By Theorem 8b, θ'_L has a limit; this is true for any t so the function $f_L(t)$ has a limit f(t) for $L \to \infty$. This proves b. To prove c, compare the sequence $\{P'_L\}$ with the sequence $P''_L = Q_L(f(t))$. These two sequences satisfy the assumptions of Theorem 9. Hence they have the same limit point. By Theorem 8c, P'_L has the limit $T_A(R(t))$. By Theorem 10a P''_L has the limit $R(f(t))$. This proves c. To prove d one uses c and the inequality (V.31) (note that the θ coordinate of R(t) is t since the θ coordinate of $Q_L(t)$ is t for all L).

To prove e let t be arbitrary in the range $0 \leq t \leq \pi/2$ and define the sequence $t_L = F_L(t)$. Let L and K (K > L) be large. Then $t = f_L(t_L) = f_K(t_K)$. Therefore

$$0 = f_L(t_L) - f_K(t_K) = \left[f_L(t_L) - f_K(t_L)\right] + \left[f_K(t_L) - f_K(t_K)\right] \tag{V.83}$$

Therefore

$$|f_K(t_L) - f_K(t_K)| = |f_L(t_L) - f_K(t_L)| \tag{V.84}$$

Now use Theorem 5d:

$$|t_L - t_K| = |F_K(f_K(t_L)) - F_K(f_K(t_K))| \leq 40\,|f_K(t_L) - f_K(t_K)| =$$

$$= 40\,|f_L(t_L) - f_K(t_L)| \tag{V.85}$$

The function $f_L(t)$ approaches $f(t)$ for $L \to \infty$ on the closed interval $0 \leq t \leq \pi/2$. Hence this limit is uniform in t. Hence, $|f_L(t_L) - f_K(t_L)|$ is arbitrarily small for sufficiently large L and K irregardless of the value of t_L. This means $|t_L - t_K| \to 0$ as L and K approach ∞ : hence the sequence t_L has a limit for $L \to \infty$. This is true for any t so $F_L(t)$ has a limit $F(t)$. Since $F_L(t)$ is the inverse to $f_L(t)$, and since both F_L and f_L are continuous uniformly in L by Theorem 5d, F(t) is the inverse to f(t) and both are continuous. Also since F_L and f_L are single-valued, so are F and f. Finally f is a consequence of d and e.

Armed with Theorems 1 - 10, one can now attack the renormalization problem. One starts with a sequence of unrenormalized cutoff Hamiltonians H_M. The bare coupling constant g_0 is permitted to vary with M and is denoted g_{0M}. In addition H_M is permitted to have an additive constant $\mathscr{E}_{0M}$ also varying with M. The renormalization problem is to choose the sequences g_{0M} and $\mathscr{E}_{0M}$ so that H_M has a finite limit for $M \to \infty$. Since the number of degrees of freedom changes as $M \to \infty$ one has to specify what one means by the limit. To be precise we demand that each energy level, counting in order of increasing energy, has a finite limit.

This is equivalent to demanding that the energy levels of the effective Hamiltonians $H_N(M)$ have limits as $M \to \infty$ keeping N fixed, since the effective Hamiltonians $H_N(M)$ describe the first 2^{2N+3} energy levels of H_M. The limit of $H_N(M)$ for N fixed is a simpler limit since now the number of degrees of freedom is fixed. It will be found that $H_N(M)$ has a limit as an operator for $M \to \infty$ (the limit will be denoted H_{RN}), which ensures that the eigenvalues of $H_N(M)$ have limits. There are other parts to the renormalization problem, namely computing matrix elements of the operators $\tau^+ \tau^-$, a_m, etc. between eigenstates of the renormalized Hamiltonian. These other problems will not be discussed.

The effective Hamiltonians $H_N(M)$ (with $H_M(M)$ defined to be H_M) are all in the space S. Denote the decomposition of $H_N(M)$ by $(J_N(M), \mathscr{E}_N(M), N, P_N(M))$ where, in turn, $P_N(M)$ is a point in S_A with the decomposition $(\theta_N(M), \underset{\sim}{A}_{kN}(M), C_{kN}(M))$. Denote the decomposition of the original cutoff Hamiltonians H_M by $(J_{0M}, \mathscr{E}_{0M}, M, P_{0M})$; the decomposition of P_{0M} is $(\theta_{0M}, 0, 0)$ and J_{0M} and θ_{0M} are

$$J_{0M} = \Lambda^M \left(1 + 2\, g_{0M}^2\right)^{1/2} \tag{V.86}$$

$$\theta_{0M} = \tan^{-1} (\sqrt{2}\, g_{0M}) \tag{V.87}$$

Since $H_N(M)$ is defined as the transform by T of $H_{N+1}(M)$ one has

$$P_N(M) = T_A\, (P_{N+1}(M)) \tag{V.88}$$

Since $P_N(N)$ lies on the curve Q_0, this means $P_N(M)$ is on Q_{M-N}:

$$P_N(M) = Q_{M-N}\, (\theta_N(M)) \tag{V.89}$$

Also one has

$$J_{N-1}(M) = \Lambda^{-1}\, J_N(M)\, T_B \left[P_N(M)\right] \tag{V.90}$$

and

$$\mathscr{E}_{N-1}(M) = \mathscr{E}_N(M) + J_N(M)\, T_C \left[P_N(M)\right] \tag{V.91}$$

from Eqs. (V.17) and (V.18). Finally, one has from Eqs. (V.89), (V.45), and (V.42)

$$\theta_N(M) = f_{M-N}\left[\theta_{N+1}(M)\right] \tag{V.92}$$

$$\theta_{N+1}(M) = F_{M-N}\left[\theta_N(M)\right] \tag{V.93}$$

The condition $H_M(M) = H_M$ means $J_M(M) = J_M$ and $\theta_M(M) = \theta_{0M}$.

One wants to choose the sequences θ_{0M} and $\mathscr{E}_M$ so that the Hamiltonians $H_N(M)$ have a limit for $M \to \infty$. Customarily one would fix θ_{0M} and $\mathscr{E}_M$ by requiring that the renormalized coupling constant and the ground state energy be fixed independent of M. We cannot calculate the renormalized coupling constant since this requires knowing the ground state matrix element of $\tau^{\pm}$, and these matrix elements are not discussed in this paper. So a more ad hoc procedure will be used. Clearly if $H_N(M)$ is to approach a limit for $M \to \infty$, the sequences $\mathscr{E}_N(M)$ and $\theta_N(M)$ must approach limits as $M \to \infty$. The simplest way to ensure this is for $\mathscr{E}_N(M)$ and $\theta_N(M)$ to be independent of M. This cannot be true for all N, but it can be arranged for one value of N, say N = 0. So let $\theta_0(M)$ be a constant θ_R (between 0 and $\pi/2$) and let $\mathscr{E}_0(M)$ be 0.

Given $\theta_0(M) = \theta_R$ and $\mathscr{E}_0(M) = 0$, for all M, one can reconstruct the complete double sequence $H_N(M)$. First one computes $\theta_N(M)$, for all M and $1 \le N \le M$ using Eq. (V.93). Secondly one computes $g_{0M} = (1/\sqrt{2}) \tan \theta_M(M)$ and J_{0M} from Eq. (V.86). Third, one computes all the $J_N(M)$ $(0 \le N < M)$ from $J_M(M) = J_{0M}$ and Eq. (V.90). Finally one computes $\mathscr{E}_N(M)$ $(1 \le N \le M)$ from Eq. (V.91). The points $P_N(M)$ are given by Eq. (V.89).

Now one can consider the limit for $M \to \infty$ of $H_N(M)$. The results are stated in Theorem 11.

Theorem 11. Assume $\theta_R \neq \pi/2$. Then

a. $\lim_{M \to \infty} H_N(M) = H_{RN}$ exists for all N

Let H_{RN} have the decomposition $(J_{RN}, \mathscr{E}_{RN}, N, P_{RN})$ and let the θ coordinate of P_{RN} be θ_{RN}. Then

b. $$P_{RN} = R(\theta_{RN}) \tag{V.94}$$

c. $$H_{RN} = T\left[H_{RN+1}\right], \text{ i.e.} \tag{V.95}$$

$$\theta_{RN} = f(\theta_{RN+1}) \tag{V.96}$$

$$J_{RN} = \Lambda^{-1} J_{RN+1} T_B(P_{RN+1}) \tag{V.97}$$

$$\mathscr{E}_{RN} = \mathscr{E}_{RN+1} + J_{RN+1} T_C \left[P_{RN+1}\right] \tag{V.98}$$

$$P_{RN} = T_A \left[P_{RN+1}\right] \tag{V.99}$$

Proof of Theorem 11. The first step uses induction in N. For $N = 0$, $\theta_0(M)$ has a limit θ_R for $M \to \infty$ by definition. Hence the sequence $P_0(M)$ satisfies the assumptions of Theorem 8. Hence $P_0(M)$ has a limit for $M \to \infty$; from Theorem 10, this limit is $R(\theta_R)$. Now suppose that $\theta_N(M)$ and $P_N(M)$ have limits θ_{RN} and $P_{RN} = R(\theta_{RN})$ respectively. Consider the sequence $\theta_{N+1}(M)$ as a function of M. It is given by Eq. (V.93). Since $\theta_N(M)$ has a limit θ_{RN}, since the function $F_{M-N}(\theta)$ has a limit $F(\theta)$, and since $F_{M-N}(\theta)$ is continuous in θ uniformly in M (see Theorem 5d), the sequence $\theta_{N+1}(M)$ must have a limit θ_{RN+1}. Also $\theta_{RN+1} = F(\theta_{RN})$. Hence θ_{RN} satisfies c. Since $\theta_{N+1}(M)$ has a limit, $P_{N+1}(M)$ has a limit (Theorem 8); the limit is $R(\theta_{RN+1})$ (Theorems 9 and 10). Because θ_{RN} is $f(\theta_{RN+1})$, one has

$$R(\theta_{RN}) = T_A \left[R(\theta_{RN+1})\right] \tag{V.100}$$

(Theorem 10c). By induction one has established limits for $\theta_N(M)$ and $P_N(M)$ for all N, as $M \to \infty$. The limit P_{RN} satisfies b and c and θ_{RN} satisfies c. The next step is to look at the scale factors $J_N(M)$. We use Theorem 3. The points of the curve Q_0 satisfy the prerequisites of Theorem 3; hence all the curves Q_L have the property that A_{k1} vanishes for all k at any point on the curve. In particular $A_{kN1}(M)$ vanishes for all k. Look at $T_B\left[P_N(M)\right]$. Let $P_N(M)$ be the point P_A of Theorem 2. Using the notation of Theorem 2 and the result of Theorem 3,

$$T_B(P_N(M)) = (m^2 + 2\, g''^2)^{1/2} \tag{V.101}$$

$$\tan \theta' = \sqrt{2}\, g''/m \tag{V.102}$$

Note that θ' is $\theta_{N-1}(M)$, θ (notation of Theorem 2) is $\theta_N(M)$, and $m = \cos \theta$. One can eliminate g'' to obtain

$$T_B(P_N(M)) = \cos \theta_N(M)/\cos \theta_{N-1}(M) \tag{V.103}$$

Using Eqs. (V.86), (V.87), and (V.90), one obtains[14]

$$J_N(M) = \Lambda^N \left[\cos \theta_N(M)\right]^{-1} \tag{V.104}$$

Since $\theta_N(M)$ has a limit θ_{NR} for $M \to \infty$, so does $J_N(M)$, provided θ_{NR} is not $\pi/2$. But from c and Theorem 10f, one sees that $\theta_{NR} < \pi/2$ if $\theta_R < \pi/2$. So $J_N(M)$ has a limit J_{RN}.

$$J_{RN} = \Lambda^N \left[\cos \theta_{NR}\right]^{-1} \tag{V.105}$$

Using Eq. (V.91), one can now show that $\mathscr{E}_N(M)$ has a limit $\mathscr{E}_{RN}$ for $M \to \infty$. It is easily seen that J_{RN} and $\mathscr{E}_{RN}$ satisfy c. This completes the proof of Theorem 11.

The existence of the renormalized energies has now been proved. The renormalized theory is defined by the sequence of renormalized cutoff Hamiltonians

H_{RN}. Because of Theorem 11c, this sequence has a common set of eigenvalues: H_{RN} describes the first 2^{2N+3} of these. The complete set of eigenvalues defines the complete renormalized Hamiltonian H_R . Unlike the renormalized Lee model, the present renormalized theory has no ghost states: the bare coupling constants g_{0M} are real for all M and all the Hamiltonians H_M (and $H_N(M)$ and H_{RN}) are hermitian. The limit of g_{0M} for $M \to \infty$ is ∞; this is proven in Section VII.

To conclude this section, it will be shown that the set of renormalized Hamiltonians H_{RN} is independent of the choice of the unrenormalized cutoff Hamiltonians H_M, in the following sense.

Theorem 12. Suppose that the cutoff Hamiltonians H_M have the decomposition $(J_{0M}, \mathscr{E}_{0M}, M, P_{0M})$ where P_{0M} lies on a one parameter curve $Q_0'(t)$:

$$P_{0M} = Q_0' (\theta_{0M}) \tag{V.106}$$

Suppose that the curve $Q_0'(t)$ is any curve in the space S_A defined for $0 \leq t \leq \pi/2$, such that t is the θ coordinate of $Q_0'(t)$ and the bound (V.52) of Theorem 5d is satisfied by Q_0'.

Construct the sequence of effective Hamiltonians $H_N(M)$ starting from H_M, and let $H_N(M)$ have the decomposition $(J_N(M), \mathscr{E}_N(M), N, P_N(M))$. The points $P_N(M)$ lie on curves $Q'_{M-N}(t)$ defined by analogy with $Q_L(t)$. Let $\theta_N(M)$ be the θ coordinate of $P_N(M)$.

Let $J_N(M)$, $\mathscr{E}_N(M)$, and $\theta_N(M)$ be determined by the boundary conditions

$$\theta_0(M) = \theta_R \tag{V.107}$$

$$\mathscr{E}_0(M) = 0 \tag{V.108}$$

$$J_0(M) = (\cos \theta_R)^{-1} \tag{V.109}$$

Then Theorem 11 holds for these $H_N(M)$ and the limiting Hamiltonians H_{RN} are independent of the choice of the curve Q_0'.

To prove Theorem 12 one first rederives Theorems 5 - 11 with Q_L' replacing Q_L; the arguments are unchanged except in Theorem 11 where the scale factors $J_N(M)$ are computed using a different boundary condition. To show that $J_N(M)$ has a limit as $M \to \infty$ one must show that $T_B(P)$ is a continuous function of P. This is true; the proof will be omitted.

To show that the limiting Hamiltonians H_{RN} are independent of the starting curve Q_0', we show that the limiting Hamiltonians H_{RN} are uniquely determined by their properties, as specified in Theorem 11, plus the boundary conditions. Using Theorem 11c, one finds

$$\theta_{RN+1} = F(\theta_{RN}) \tag{V.110}$$

So one can compute θ_{RN} for all N given $\theta_{R0} = \theta_R$. Then by 11b, P_{RN} is determined. Then one can use 11c to determine J_{RN} and $\mathcal{E}_{RN}$ starting from the boundary conditions (V.108) and (V.109).

The scale factors $J_0(M)$ were specified in this discussion instead of $J_M(M)$ simply to ensure that H_{RN} would be independent of the choice of curve Q_0'.

VI. APPROXIMATE SCALE INVARIANCE IN THE RENORMALIZED THEORY

When a quantum theory is invariant to the orientation of the coordinate system, it must be rotationally symmetric — that is, there must exist unitary operators R which generate rotations and which commute with the Hamiltonian. One can then diagonalize the generators of infinitesimal rotations simultaneously with the Hamiltonian; one can classify the resulting eigenstates by angular momentum eigenvalues, etc.

Likewise, when a quantum theory contains no parameters with the dimensions of energy, it must be invariant to a choice of energy scale. This immediately implies that the theory is invariant to a set of unitary operators U(s) which change all energies by a scale factor s. The Hamiltonian H is not invariant to U(s), since H is itself an energy; instead, one has

$$U^{+}(s)\ H\ U(s) = sH \tag{VI.1}$$

There will be an infinitesimal generator D which generates infinitesimal scale transformations (a transformation with $s = 1 + \epsilon$ where ϵ is infinitesimal). However, D does not commute with H and cannot be simultaneously diagonalized with H. Instead, scale invariance is used to generate a set of energy levels with any energy sE given a level with energy E.

In field theoretic problems there are usually mass parameters in the theory, but sometimes these parameters become negligible at high energies or short distances. For example the propagator of a free scalar or spinor field at short distances is independent of the free field mass and is equal to the propagator of the zero mass theory. The free zero mass scalar and spinor field theories are scale invariant.[15] The standard interacting field theories (quantum electrodynamics or pseudoscalar meson theory) have only masses as dimensional parameters, but when solved in perturbation theory they do not become scale invariant at short distances (large momenta). The propagators of the interacting theories involve logarithms of (q^2/m^2) where m is a renormalized mass and q the argument of the propagator. However, if one holds the renormalized coupling constant e fixed then at very large q^2 the logarithmic terms become so large that the terms of order $e^{2k} \log^k (q^2/m^2)$ in the perturbation expansion are much larger than the Born approximation. To determine the propagator for this range of q^2, in particular in the limit $q^2 \to \infty$ one must sum the complete perturbation expansion. There are presently no methods for doing this (see especially the remarks of Bogoliubov and Shirkov[16]). There is then a question of whether the mass dependence will disappear at values of q^2 so large that the complete perturbation expansion has to be summed. The best analysis of this problem in relativistic theory is that of Gell-Mann and Low.[8]

In the model, what happens is this. The energy levels of order Λ^n expanded in powers of the renormalized coupling constant g_R have terms of order $n^k g_R^k$ which prevent any scaling laws from holding. But when n is so large that $ng_R \gg 1$ the complete series in g_R must be summed, and then the theory becomes scale invariant, in a manner to be explained below. If g_R itself is of order 1 rather than small, then scale invariance sets in for much smaller n; the only requirement is $n \gg 1$.

There is a feature of scale transformations which distinguishes them in a very fundamental way from all other symmetries of the theory. The other symmetries (charge symmetries, etc.) are well defined in the presence of the cutoff M of the model. The scale transformations are not. The scale transformations of the model are transformations U_ℓ which take the creation and destruction operators a_m^+, a_m, b_m^+, and b_m, for any m, into the operators $a_{m+\ell}^+$, $a_{m+\ell}$, $b_{m+\ell}^+$, and $b_{m+\ell}$. (Because the momentum continuum has been replaced by a discrete index m the scale transformations are labeled by a discrete variable ℓ instead of a continuous variable s.) Since the creation and destruction operators satisfy the same commutation relations for any m, the transformation should exist and be unitary, except for endpoint effects. Namely in the cutoff theory there are no operators a_m, etc., with $m > M$, or $m < 0$. So the operators a_M, etc. cannot be transformed. To have scale transformations well defined one must have operators a_m, etc. defined for $-\infty < m < \infty$. But this raises a new problem: if there are an infinite number of a_m, then they act in an infinite product Hilbert space, which is inseparable and therefore hard to work with.[10] This problem has not been mentioned up to now since it was evident once the unrenormalized Hamiltonian was defined that one could only solve it by introducing a cutoff M. Furthermore, when the limit $M \to \infty$ was defined in Section V, it was defined only for the effective Hamiltonians $H_N(M)$ for fixed N, which act on Hilbert spaces with a fixed and finite number of meson degrees of freedom.

The natural way to show that a theory has an approximate symmetry is to show that it departs only a small amount from a theory with the exact symmetry. In the present example of scale invariance, this would require constructing a version of the model which is exactly scale invariant. But this is very difficult precisely because of the problems of the infinite number of degrees of freedom. The problem is not the problem of keeping the pions with arbitrary large m. It was shown at the end of Section V that one could define a renormalized Hamiltonian H_R which includes all the renormalized energy levels including those involving m-mesons with arbitrarily large m. The set of such energy levels can be ordered by their energy and therefore form a countable set of states, which one can think of as defining a separable subspace of the original inseparable space. The problem is that the exactly scale invariant theory would have to include degrees of freedom m with $m \to -\infty$. With such terms present there would be on every gross energy level an infinite sequence of fine structure, hyperfine structure, hyper-hyperfine structure, etc., with the net result that in a finite energy interval there would be an uncountable number of distinct energy levels. These would not form a continuum because each energy must be the sum of terms of order Λ^{-1}, Λ^{-2}, Λ^{-3}, etc. with coefficients of order 1.

Rather than try to develop a formalism for handling the difficulties of the inseparable space of energy levels of a scale invariant theory, we will define approximate scale invariance to mean simply that for each energy level of H_R of sufficiently large energy, there is another energy level which is approximately a factor s_0 larger in energy. The factor s_0 will be determined below; it will be of order Λ. The correspondence will not be one to one; for an energy level of energy E, there will be four energy levels of approximately energy s_0E due to the fact that the energy levels of energy s_0E involve one more meson degree of freedom.

One can try to predict the value of s_0 by considering the unrenormalized Hamiltonians H_M. If one applies the scaling operator U_1 to H_M one gets

$$U_1^+ H_M(g_0) U_1 = \Lambda^{-1} H_{M+1}(g_0) - \Lambda^{-1} O_0 \qquad \text{(VI. 2)}$$

where U_1 is the operator that takes a_m into a_{m+1}, and O_0 is the term of order 1 in H_{M+1}. Since the eigenvalues of $(\Lambda^{-1}H_{M+1}(g_0) - \Lambda^{-1}O_0)$ differ in order Λ^{-1} from the eigenvalues of $\Lambda^{-1}H_{M+1}(g_0)$, it follows that $H_M(g_0)$ and $\Lambda^{-1}H_{M+1}(g_0)$ have the same eigenvalues except for fine structure of order Λ^{-1}.

Suppose that $H_M(g_0)$ had a well-defined limit as $M \to \infty$ for fixed g_0. Then in particular the energy levels of $H_M(g_0)$ and $H_{M+1}(g_0)$ would be the same for sufficiently large M (excluding energies of order the cutoff, that is energies of order

Λ^M). But then a given energy level of H_M would be Λ^{-1} times the energy of a level of H_{M+1}, from Eq. (VI.2). For sufficiently large M this level of H_{M+1} is also a level of H_M. Hence for every energy level of H_M of energy E, there would be another level with energy ΛE. So the factor s_0 would be Λ.

The prediction is wrong; s_0 is not Λ. The reason for the failure is that the renormalized energy levels are obtained by solving Hamiltonians $H_M(g_{0M})$ where g_{0M} changes with M. It will be shown later that $g_{0M} \to \infty$ as $M \to \infty$ so even for large M, g_{0M} is not constant.

The idea that operators do not scale as predicted from an unrenormalized theory was used in a recent discussion of approximate scale invariance in strong interactions.[6] However, the analogy to the model of this paper is inexact since in the strong interaction problem, the scaling law for the Hamiltonian is fixed by general arguments; it is the other fields in the theory, such as the pion field, whose scaling laws (dimensions) were permitted to be arbitrary.

The remainder of this section is devoted to the technical problem of computing the nature of the energy levels with energies of order Λ^n with n large, and extracting the scale factor s_0. It will be shown not only that these energies scale by a factor $\Lambda\beta^{-1}$, where β is approximately 1/2; it will also be shown that the error to this scaling law itself scales like Λ, as if the Hamiltonian consisted of two terms, one scaling as $\Lambda\beta^{-1}$, the other as Λ under a scale transformation (cf., Eqs. (VI.24) and (VI.25)).

In the following it is assumed that the function f(t) and the "curve" R(t) defined in Section V are differentiable. I have not proved this.

The renormalized theory is defined by a sequence of Hamiltonians H_{RN}. These Hamiltonians are determined by three parameters J_{RN}, $\mathscr{E}_{RN}$, and θ_{RN}. We must study H_{RN} when N is large. This requires knowledge of J_{RN}, $\mathscr{E}_{RN}$, and θ_{RN} for large N.

First look at the sequence $\left\{\theta_{RN}\right\}$. Since θ_{RN} is the θ coordinate of P_{RN}, and since P_{RN} is the transform T_A of P_{RN+1}, one can apply the inequality (V.31) to obtain

$$\left\{1-.51 \sin^2 \theta_{RN+1}\right\} \tan \theta_{RN+1} \leq \tan \theta_{RN} \leq \left\{1-.48 \sin^2 \theta_{RN+1}\right\} \tan \theta_{RN+1} \tag{VI.3}$$

First of all this implies that

$$\theta_{RN} < \theta_{RN+1} < \pi/2 \tag{VI.4}$$

(we assume $\theta_{R0} < \pi/2$ which then forces θ_{RN} to be less than $\pi/2$: see the proof of Theorem 11). So $\left\{\theta_{RN}\right\}$ is an increasing and bounded sequence. Therefore it has a limit for $N \to \infty$. The limit must be $\pi/2$. The reason is that since $\theta_{RN} = f(\theta_{RN+1})$ the limit θ must satisfy $\theta = f(\theta)$. Also $\theta_{R0} < \theta \leq \pi/2$. But from Theorem 10d, the only such θ is $\theta = \pi/2$. Therefore, when N is sufficiently large, θ_{RN} is approximately $\pi/2$. Write

$$\theta_{RN} = \pi/2 - \phi_N \tag{VI.5}$$

When ϕ_N is small the inequality (VI.3) is approximately

$$.49\,(\phi_{N+1})^{-1} \leq \phi_N^{-1} \leq .52\,(\phi_{N+1})^{-1} \tag{VI.6}$$

e.g.

$$\phi_{N+1} \approx 1/2\ \phi_N \tag{VI.7}$$

To be more precise, consider the formula $\theta_{RN} = f(\theta_{RN+1})$ and expand in powers of ϕ_{N+1}:

$$\pi/2 - \phi_N = f(\pi/2 - \phi_{N+1}) = f(\pi/2) - \phi_{N+1}\, f'(\pi/2) + O(\phi_{N+1}^2) \tag{VI.8}$$

Since $f(\pi/2)$ is $\pi/2$, one gets

$$\phi_N = f'(\pi/2)\ \phi_{N+1} + O(\phi_{N+1}^2) \tag{VI.9}$$

and Eq. (VI.7) shows that $f'(\pi/2)$ is approximately 2.

Let

$$\beta = \left\{f'(\pi/2)\right\}^{-1} \simeq .5 \tag{VI.10}$$

One can rewrite Eq. (VI.9) to read

$$\phi_{N+1} = \beta\phi_N + O(\phi^2_{N+1}) \tag{VI.11}$$

An analysis of this equation shows that

$$\phi_N = a\,\beta^N + O(\phi^2_N) \tag{VI.12}$$

where a is a constant (a will depend on θ_{R0}).

Now look at J_{RN}, $\mathscr{E}_{RN}$, and P_{RN}. From Eq. (V.105), assuming N is large, one has

$$J_{RN} = \Lambda^N (\cos\,\theta_{RN})^{-1} \simeq \Lambda^N\,\phi_N^{-1} \simeq \Lambda^N a^{-1} \beta^{-N} \tag{VI.13}$$

To compute P_{RN} one must study the curve R(t). One has

$$R(\theta_{RN}) \simeq R(\pi/2) - \phi_N\,R'(\pi/2) \tag{VI.14}$$

Let

$$P_c = R(\pi/2) \tag{VI.15}$$

$$P_d = -\,R'(\pi/2) \tag{VI.16}$$

Then

$$P_{RN} \simeq P_c + a\,\beta^N\,P_d \tag{VI.17}$$

Finally, from Theorem 11c one has

$$\mathscr{E}_{RN} = -\sum_{n=1}^{N} J_{Rn}\,T_c\left[P_{Rn}\right] \tag{VI.18}$$

(using the definition $\mathscr{E}_{R0} = 0$). The dominant terms in this sum are for large n since $J_{Rn} \sim \Lambda^n$ and $T_c \sim 1$ (Eq. (V.30)). For large n, $P_{Rn} \simeq P_c$. Let

$$\gamma = T_c\left[P_c\right] \tag{VI.19}$$

Then for large N (using Eq. (VI.13)),

$$\mathscr{E}_{RN} \simeq -\Lambda^N \gamma a^{-1} \beta^{-N} \Lambda(\Lambda-\beta)^{-1} \tag{VI.20}$$

A more careful calculation gives the first correction to Eq. (VI.20) to be

$$\mathscr{E}_{RN} \simeq -\Lambda^N \beta^{-N} \gamma a^{-1} \Lambda(\Lambda-\beta)^{-1} + \Lambda^N \gamma_1 \Lambda(\Lambda-1)^{-1} \tag{VI.21}$$

where γ_1 is a constant; also

$$J_{RN} \simeq a^{-1} \beta^{-N} \Lambda^N + \gamma_2 \Lambda^N \tag{VI.22}$$

where γ_2 is a constant.

With the above approximations for P_{RN}, J_{RN}, and $\mathscr{E}_{RN}$ one can write

$$H_{RN} = a^{-1} \Lambda^N \beta^{-N} \mathscr{H}_{cN} + \Lambda^N \mathscr{H}_{dN} \tag{VI.23}$$

where $\mathscr{H}_{cN}$ is a Hamiltonian with the decomposition $(J, \mathscr{E}, N, P) = (1, -\gamma\Lambda(\Lambda-\beta)^{-1}, N, P_c)$, and $\mathscr{H}_{dN}$ is a Hamiltonian with the decomposition $(1, \gamma_1 \Lambda(\Lambda-1)^{-1}, N, \gamma_2 P_c + P_d)$. The only N dependence of $\mathscr{H}_{cN}$ and $\mathscr{H}_{dN}$ is in terms of how many degrees of freedom are kept in Eq. (V.2), since neither the J, $\mathscr{E}$, nor P components of $\mathscr{H}_{cN}$ or $\mathscr{H}_{dN}$ depend on N.

Now compare H_{RN} with H_{RN+1}. The difference between $\mathscr{H}_{cN}$ and $\mathscr{H}_{cN+1}$ is only in terms containing a_{N+1}, b_{N+1}, etc., and such terms are of order Λ^{-N-1} or less. So the energy levels of $\mathscr{H}_{cN}$ are approximately the same as the energy levels of $\mathscr{H}_{cN+1}$ only each level of $\mathscr{H}_{cN}$ corresponds to four of $\mathscr{H}_{cN+1}$ due to the extra degrees of freedom in $\mathscr{H}_{cN+1}$. In H_{RN}, the $\mathscr{H}_{cN}$ term dominates the $\mathscr{H}_{dN}$ term by a factor β^N; neglecting the $\mathscr{H}_{dN}$ term, the energy levels of $\beta^{-1}\Lambda H_{RN}$ and H_{RN+1} are approximately equal. This establishes the basic claim of this section. The energy levels of H_{RN} and H_{RN+1} are both subsets of the energy levels of H_R. So the energy levels of $\Lambda\beta^{-1}H_R$ are approximately equal to the energy levels of H_R. In scaling the Hamiltonian an extra factor β has appeared.

Let us consider the errors in approximate scale invariance. One is comparing $\Lambda\beta^{-1} H_{RN}$ with H_{RN+1}. The basic energy scale for these Hamiltonians is $\Lambda^{N+1}\beta^{-N-1}$. Neglecting the $\mathscr{H}_{dN}$ and $\mathscr{H}_{dN+1}$ terms in H_{RN} and H_{RN+1} means one has an error of order Λ^N. This is small by a factor β^{N+1} from the basic energy scale but huge on an absolute scale (remember N must be large for all our approximations to hold). There is also an error which is of order Λ^{-N-1} in $\mathscr{H}_{cN+1}$ when one neglects the a_{N+1} terms; this becomes an error of order β^{-N-1} in H_{RN+1} which is negligible compared to Λ^N ($\beta \sim 1/2$ while $\Lambda > 4 \times 10^6$). Due to the error of order Λ^N the matching between H_{RN+1} and $\beta^{-1}\Lambda H_{RN}$ is close only for energy levels with energies large compared to Λ^N, i.e., only highly excited states.

One can now get a scaling law for the leading correction to scale invariance. Namely one can take $\mathscr{H}_{dN}$ into account but still neglect the difference between $\mathscr{H}_{cN}$ and $\mathscr{H}_{cN+1}$ and the difference between $\mathscr{H}_{dN}$ and $\mathscr{H}_{dN+1}$. In this case one can write

$$H_{RN} = H_{cN} + H_{dN} \tag{VI.24}$$

with $H_{cN} = a^{-1}\Lambda^N\beta^{-N}\mathscr{H}_{cN}$, $H_{dN} = \Lambda^N\mathscr{H}_{dN}$; then

$$H_{RN+1} \simeq \Lambda\beta^{-1} H_{cN} + \Lambda H_{dN} \tag{VI.25}$$

Since H_{dN} is small compared to H_{cN}, the energies of $H_{cN} + H_{dN}$ consist, to a first approximation, of energies of H_{cN} plus expectation values of H_{dN}. The correction therefore scales by a factor Λ when $N \to N+1$ while the dominant term in the energy scales by $\Lambda\beta^{-1}$.

The unrenormalized Hamiltonian had two parts, the free meson energy term and the interaction term, but both parts scaled by Λ when $N \to N+1$. The renormalized Hamiltonian also has two parts to a first approximation but the two parts scale differently, the dominant term scaling by $\Lambda\beta^{-1}$ while the leading correction scales by Λ.

It was crucial for the proof of scale invariance that the constants θ_{RN} approach a limit ($\pi/2$) for $N \to \infty$. As long as θ_{RN} changes with N the energy level structure of H_R on the scale Λ^N will differ by more than a scale factor from the structure on the scale Λ^{N+1}. This is due to the nontrivial dependence of the energy levels of H_{RN} on θ_{RN}. In particular, in perturbation theory, when θ_{RN} is small, the change from θ_{RN} to θ_{RN+1} is nonnegligible in order θ_{R0}^3 (see Section VII for details). Hence in third order or higher in θ_{R0}, H_R does not show scale invariance. It is only when N is so large that $\theta_{RN} \simeq \pi/2$ that scale invariance becomes apparent; but for these values of N an expansion in θ_{R0} is silly even if θ_{R0} is small: the true expansion parameter turns out to be $\sqrt{N}\,\theta_{R0}$ which is huge, instead of θ_{R0}.

VII. RENORMALIZATION AND THE ROLE OF THE TRANSFORMATION T

The renormalization program carried out in this paper followed the conventional pattern, in that a renormalized coupling constant was defined and held fixed in the limit of infinite cutoff. The transformations T and T_A were introduced as part of the technique of solving the cutoff Hamiltonians; their properties were useful in proving the existence of the renormalized Hamiltonian. An analysis of the renormalization program of Section V shows that the transformations T and T_A play a more fundamental role in the renormalization than one might think. In part A of this section it is shown that the renormalized Hamiltonian is determined more by the properties of the transformation T_A than by properties of the original unrenormalized Hamiltonian of Section II. In part B, the problem of "why renormalization?" is considered; it is shown that three features of the model Hamiltonian cause the renormalization program to be nontrivial. These three features are: first, the model has an infinite number of degrees of freedom; second,the m^{th} degree of freedom with m large dominates the degrees of freedom with m small; third, scale invariance makes the behavior of the degrees of freedom for large m similar for different m. In part C, the renormalization theory of this paper is compared with the theory of Gell-Mann and Low for quantum electrodynamics.[8]

A. Renormalization and the Transformation T_A

The analysis of the renormalization program to be given here concerns very basic questions; to set the stage for these questions it is worth reviewing the role of the Hamiltonian in ordinary quantum mechanics. In nonrelativistic quantum mechanics, a system is well defined once the Hamiltonian is specified. Any hermitian (self-adjoint) Hamiltonian defines a unique and acceptable quantum mechanics. To specify the Hamiltonian, one must first define the basic observables of the system (e.g., position, momentum, or spin operators). Then one specifies the Hamiltonian as a function of these observables. In principle one could define the Hamiltonian in a different way, by giving a list of its eigenvalues and eigenvectors. This is rarely done in practice because the eigenvalues and eigenvectors are generally very complicated expressions, often not expressible in closed form. In contrast, the Hamiltonian is often a simple function of the observables (for example, compare the Coulomb Hamiltonian of the helium atom with its eigenvalues and eigenvectors).

In Section II of this paper we defined a model quantum theory in an entirely conventional manner. The "observables" a_n, a_n^+, b_n, b_n^+, and $\tau^{\pm}$ were defined, and the Hamiltonian written as a simple function of these observables, with one free parameter g_0. Then in Sections IV and V the techniques for solving the model were defined, and it was shown that after renormalization the theory had finite eigenvalues. The finite theory again depended on one free parameter, which however was the renormalized constant θ_{R0} instead of g_0.

The construction of the renormalized Hamiltonian in Section V was a complicated process. In summary, one chose a renormalized coupling constant θ_{R0}. One constructed a sequence of Hamiltonians H_{RN} by starting with the point $P_{R0} = R(\theta_{R0})$ and constructing the sequence P_{RN} through the relation $P_{RN} = T_A(P_{RN+1})$. The full renormalized Hamiltonian consisted of a limit of H_{RN} for $N \to \infty$ suitably defined. This construction leaves unclarified some fundamental questions. Does the renormalized theory solve the unrenormalized Hamiltonian of Section II? If not, what problem does it solve? Is the renormalized coupling constant a fundamental parameter in the theory? If not, can it be replaced by one that is? Is the unrenormalized Hamiltonian the simple expression which underlies and defines the rather complicated spectrum of renormalized energy levels; if not, where do we look for simplicity?

It is difficult to answer these questions conclusively because there are problems of interpretation. For example, one must decide what is a "fundamental" parameter, and what is "simple." However, in trying to answer the questions of the previous paragraph, two results become clear. The first is that the relation of the unrenormalized, uncutoff Hamiltonian to the renormalized energy levels is fundamentally different than the relation of a simple Coulomb Hamiltonian to its eigenvalues. How to characterize the new relationship can be debated, but certainly it is not the old and comfortable relationship of elementary quantum mechanics. The second result is this: there is a key fact which must figure in any discussion of the new relationship of Hamiltonian to energy levels, a key idea which must be used to obtain any fundamental understanding of why we must introduce an essentially phenomenological parameter (the renormalized coupling constant) in defining the renormalized theory. The crucial fact is the existence of a fixed point of the transformation T_A, namely, the point $P_c = R(\pi/2)$. The point P_c has already been encountered in Section VI: it is the limit of the points P_{RN} (involved in the definition of H_{RN}) as $N \to \infty$. The role of the fixed point cannot be summarized in a few words; a detailed analysis of its function will be given later in this section.

The relation of the unrenormalized uncutoff Hamiltonian to the renormalized theory can be summarized in terms of the following two results which will be proven later in this section.

1. If $\left\{g_{0M}\right\}$ is a sequence of coupling constants which approach a finite limit g_0 as $M \to \infty$, then the energy levels of the unrenormalized cutoff Hamiltonians $H_M(g_{0M})$ approach the energy levels of the uncutoff free Hamiltonian (Eq. (II.1) with $g_0 = 0$) as $M \to \infty$, except for an additive constant.

2. If $\{g_{0M}\}$ is a sequence of coupling constants which approach ∞ as $M \to \infty$ the energy levels of $H_M(g_{0M})$ may or may not approach a limit as $M \to \infty$. For any θ_{R0} with $0 < \theta_{R0} < \pi/2$, there exists a sequence $\{g_{0M}\}$ with $g_{0M} \to \infty$ as $M \to \infty$, such that the energy levels of $H_M(g_{0M})$ approach the energy levels of the renormalized Hamiltonian $H_R(\theta_{R0})$ as $M \to \infty$ (apart from an additive constant).

The first result means that if the uncutoff unrenormalized Hamiltonian with finite g_0 is defined as a limit of cutoff Hamiltonians, then its solution is the same as the solution of the free uncutoff Hamiltonian, and in particular is not related to any of the renormalized theories with interaction. The second result means that a single uncutoff unrenormalized Hamiltonian, the one with $g_0 = \infty$, has an infinite number of possible solutions depending on what sequence $\{g_{0M}\}$ is used in the cutoff Hamiltonians. So instead of each renormalized Hamiltonian corresponding to a separate unrenormalized Hamiltonian, one finds that all the renormalized Hamiltonians solve a single unrenormalized Hamiltonian. The nonuniqueness of the solution of the unrenormalized Hamiltonian with $g_0 = \infty$ is discussed further below.

Now the results quoted above will be proven. It is helpful to prove the following. If $\theta < \theta_1$ and both lie between 0 and $\pi/2$, then

$$f_L(\theta) < f_L(\theta_1) \qquad (\text{for } \theta < \theta_1) \tag{VII.1}$$

The proof is based on Theorem 5. From 5b, $f_L(\theta_1) - f_L(0)$ is positive for $\theta_1 > 0$. From 5d (Eq. (V.54))

$$|f_L(\theta_1) - f_L(\theta)| > .025\ |\theta_1 - \theta| \tag{VII.2}$$

From 5b, $F_L(\theta)$ is continuous in θ. Hence $f_L(\theta_1) - f_L(\theta)$ cannot change sign anywhere in the range $0 \leq \theta < \theta_1$. Hence Eq. (VII.1) holds. To prove the first result, consider a sequence $\{g_{0M}\}$ with a finite limit g_0 as $M \to \infty$. Consider the unrenormalized Hamiltonians $H_M(g_{0M})$. Using the transformation T one can generate effective Hamiltonians $H_N(M)$ with coupling constants $\theta_N(M)$ having the same energy levels as $H_M(g_{0M})$. The constants $\theta_N(M)$ satisfy Eqs. (V.92) and (V.93) and

$$\tan \theta_M(M) = \sqrt{2}\, g_{0M} \tag{VII.3}$$

Let θ be an upper bound to $\theta_M(M)$; since g_{0M} has a finite limit, one can choose θ to be less than $\pi/2$. Define a sequence $\left\{\theta_L\right\}$ to be: $\theta_0 = \theta$, $\theta_L = f_L(\theta_{L-1})$. Due to Eq. (VII.1), θ_{M-N} is an upper bound for $\theta_N(M)$. The sequence $\left\{\theta_L\right\}$ is a decreasing sequence with limit 0 as $L \to \infty$; this follows from the inequality (V.31). Hence $\theta_N(M) \to 0$ as $M \to \infty$ for fixed N. Hence in the limit $M \to \infty$, $H_N(M)$ becomes a free Hamiltonian, which is result 1. To prove the second result, consider the sequence $\left\{g_{0M}\right\}$ defined in Section V following Eq. (V.93) corresponding to a given nonzero renormalized constant θ_{R0}. Again one has constants $\theta_N(M)$ satisfying Eqs. (V.92), (V.93), and (VII.3), but now $\theta_0(M)$ is fixed to be θ_{R0}. From Eq. (V.31), $\theta_1(M+1) > \theta_0(M+1) = \theta_0(M)$; using Eq. (VII.1) repeatedly one gets $\theta_{N+1}(M+1) > \theta_N(M)$ for all N, and hence $g_{0M+1} > g_{0M}$. So $\left\{g_{0M}\right\}$ is an increasing sequence. It cannot have a finite upper bound, for if so then θ_{R0} would have to be zero. Hence $g_{0M} \to \infty$ as $M \to \infty$. By the analysis of Section V the Hamiltonians $H_N(M)$ have well defined limits as $M \to \infty$. (In Section V the unrenormalized Hamiltonians H_M have a ground state energy subtraction; if this subtraction is not made then only the energy differences of levels of $H_N(M)$ have a limit as $M \to \infty$.) Such a sequence $\left\{g_{0M}\right\}$ exists for any θ_{R0}, so result 2 is proved.

The fact that the uncutoff Hamiltonian with $g_0 = \infty$ has an infinite number of solutions can be blamed on the fixed point P_c of T_A. This result can be seen by studying the behavior of the double sequence $P_N(M)$ of points in S_A defined in

Section V as part of the renormalization analysis. The points $P_N(M)$ have the following properties:

a. $P_M(M)$ has the decomposition $(\theta_M(M), 0, 0)$, i.e., the components $\underset{\sim}{A}_k$ and C_k are all zero. The point $P_M(M)$ corresponds to the unrenormalized Hamiltonian $H_M(g_{0M})$ with finite cutoff M and g_{0M} given by Eq. (VII.3).

b. $P_0(M)$ has θ coordinate θ_{R0}, by definition.

c. $P_{N-1}(M) = T_A\left[P_N(M)\right]$

When $M \rightarrow \infty$, $\theta_M(M) \rightarrow \pi/2$, so $P_M(M)$ has a limit $(\pi/2, 0, 0)$ when $M \rightarrow \infty$. Denote this point by P_U. The point P_U corresponds to the unrenormalized, uncutoff Hamiltonian with $g_0 = \infty$.

The point $P_c = R(\pi/2)$ (the fixed point of T_A) also has θ coordinate $\pi/2$, but it is easily seen that the components $\underset{\sim}{A}_k$ and C_k of P_c cannot vanish. Hence P_c is distinct from P_U.

One can think of the points $P_N(M)$, for fixed M, as defining a trajectory C(M). If one takes the limit of the trajectories C(M) for $M \rightarrow \infty$, one gets a double trajectory $C_A \oplus C_B$. The trajectory C_A goes from P_U to P_c, i.e., it connects the point P_U representing the unrenormalized Hamiltonian to the fixed point P_c. The trajectory C_B connects the renormalized point P_{R0} to the fixed point P_c. The first trajectory is an infinite sequence of points $(P_U, P_{U1}, P_{U2}, \ldots)$ all with $\theta = \pi/2$, satisfying $P_{UN} = T_A(P_{UN-1})$ and with the limit P_c as $N \rightarrow \infty$. The trajectory C_B consists of the renormalized points P_{RN} lying on the curve R, again with limit P_c as $N \rightarrow \infty$. The trajectories C(M) with M large lie close to the limiting trajectories: the first few points on C(M) (e.g., $P_M(M)$, $P_{M-1}(M)$, etc.), lie close to the first few points on C_A. The last few points on C(M) (e.g., $P_1(M)$, $P_2(M)$, etc.) lie close to the first few points on C_B. The points near the middle of the trajectory C(M) (e.g., $P_{M/2}(M)$) all lie close to P_c.

The trajectories $C(M)$, C_A, and C_B are illustrated in Fig. 1. Figure 1 is an artist's conception of what these trajectories might look like if the space S_A was a two-dimensional space instead of an infinite dimensional space. The two dimensions are θ and a coordinate x replacing the infinite dimensional space defined by the sequences $\{\underset{\sim}{A}_k\}$ and $\{C_k\}$. One can see explicitly in Fig. 1 that the points $P_N(M) \rightarrow P_{RN}$ as $M \rightarrow \infty$ and $P_{M-N}(M) \rightarrow P_{UN}$ as $M \rightarrow \infty$. One can also see the clustering of points about P_c.

Now return to the problem of the infinite number of solutions of the unrenormalized Hamiltonian. The nonuniqueness is connected with the fixed point P_c, because the limiting trajectory $C_A \oplus C_B$ is nonunique only on the section C_B. The trajectory C_A connecting P_U to P_c is uniquely determined by P_U and the recursion formula $P_{UN} = T_A(P_{UN-1})$. The trajectory C_B connecting P_c to P_{R0} is nonunique; it is a different trajectory for each different value of θ_{R0}. So the nonuniqueness arises at the point P_c.

The next question is, how is the nonuniqueness related to the properties of the fixed point P_c? In order to discuss this question it is necessary to know the behavior of the transformation T_A in the neighborhood of P_c; this behavior will now be investigated.

Assume that the transformation T_A is differentiable in the vicinity of P_c, so that if P is any point near P_c, one can write

$$T_A(P) = P_c + U_A(P-P_c) + \text{order } (P-P_c)^2 \tag{VII.4}$$

where U_A is a <u>linear</u> transformation. Now consider a trajectory of points P_N, namely a sequence of points satisfying

$$P_{N+1} = T_A(P_N) \tag{VII.5}$$

and suppose that the trajectory lies in the vicinity of P_c. Then approximately

$$P_{N+1} - P_c = U_A(P_N - P_c) \tag{VII.6}$$

Consider therefore the trajectories defined by U_A, that is, sequences of points Q_N satisfying

$$Q_{N+1} = U_A(Q_N) \tag{VII.7}$$

Since this is a linear equation, an arbitrary solution can be written as a linear combination of a set of linearly independent "basic" solutions $Q_{N\alpha}$ ($\alpha = 1, 2, 3 \ldots$ labels different linearly independent trajectories). The simplest type of solution is of the form

$$Q_{N\alpha} = Q_{0\alpha} (r_\alpha)^N \tag{VII.8}$$

where $Q_{0\alpha}$ is a point (determined up to a scale factor) and r_α is a constant. $Q_{0\alpha}$ is an eigenvector of the transformation U_A:

$$r_\alpha Q_{0\alpha} = U_A(Q_{0\alpha}) \tag{VII.9}$$

and r_α is an eigenvalue. Since U_A does not have to be a self-adjoint transformation, the eigenvalues r_α need not be real; also there may be trajectories $Q_{N\alpha}$ behaving as $N(r_\alpha)^N$, $N^2(r_\alpha)^N$, etc., under special circumstances. Since U_A is a transformation on a space with an infinite number of dimensions, there will be an infinite set of basic solutions $Q_{N\alpha}$. These solutions divide into three possible categories. Those with $|r_\alpha| > 1$ are called "unstable" trajectories; these trajectories move away from P_c as one keeps applying the transformation T_A. Those with $|r_\alpha| < 1$ are stable trajectories; the stable trajectories approach P_c as one keeps applying T_A. For example, the trajectory C_A connecting P_U with P_c is a stable trajectory; the trajectory C_B is an unstable trajectory. There can also be "neutral" trajectories with $|r_\alpha| = 1$, in special cases.

A crucial question is: how many linearly independent unstable trajectories does U_A have? The answer is one; the proof is as follows. There must be at least one basic unstable trajectory, for if all the basic trajectories were stable then all linear combinations of the basic trajectories would also be stable, i.e., all solutions of Eq. (VII.6) would be stable. But we know there are unstable solutions, namely, the trajectories C_B for any θ_{R0} (to be precise, the parts of these trajectories lying near P_c). On the other hand, there cannot be more than one basic unstable trajectory. For if there were two linearly independent unstable trajectories, say Q_{N1} and Q_{N2}, then one could form a linear combination of these, say $\beta_1 Q_{N1} + \beta_2 Q_{N2}$, such that the θ coordinate of $\beta_1 Q_{11} + \beta_2 Q_{21}$ is 0. This means the θ coordinate of $(P_c + \beta_1 Q_{11} + \beta_2 Q_{21})$ is $\pi/2$. But now the θ coordinate of $P_c + \beta_1 Q_{N1} + \beta_2 Q_{N2}$ will be $\pi/2$ for all N because T_A does not change θ if $\theta = \pi/2$. But then the sequence of points $P_c + \beta_1 Q_{N1} + \beta_2 Q_{N2}$ must approach P_c as $N \to \infty$ using Theorems 8 - 10 of Section V. This means $\beta_1 Q_{N1} + \beta_2 Q_{N2}$ is a stable trajectory. Then we could use $\beta_1 Q_{N1} + \beta_2 Q_{N2}$ as a basic trajectory instead of Q_{N2} say, which leaves only one unstable trajectory. The trajectories C_B for different θ_{R0} must all be multiples of the single unstable trajectory. This result has already been demonstrated in Section VI; cf., Eqs. (VI.12) and (VI.17).

It will now be shown that the number of linearly independent unstable trajectories of U_A determines the number of free parameters in the renormalized Hamiltonian. In other words, the degree of nonuniqueness of the solution of the unrenormalized Hamiltonian is determined by the number of unstable solutions of the linearized transformation U_A.

To show this we must discuss what would have happened if U_A had two or more linearly independent unstable trajectories. It will be shown that in this case the nonuniqueness of the solution of the unrenormalized Hamiltonian involves two or more free parameters. To be precise, we show that one can construct sequences $P_N(M)$ such that

1. $\lim_{M \to \infty} P_M(M) = P_U$

2. $\lim_{M \to \infty} P_N(M) = P_{RN}(a_1 \ldots a_k)$

3. $P_{N-1}(M) = T_A(P_N(M))$

where the point P_{RN} depends on k phenomenological parameters $a_1 \ldots a_k$, k being the number of linearly independent unstable solutions of U_A. Having shown that such sequences exist for any choice of the parameters $a_1 \ldots a_k$, it is clear that there is a k-parameter family of renormalized Hamiltonians, defined by the points $P_{RN}(a_1 \ldots a_k)$ for all N, all of which can be considered solutions of the single unrenormalized Hamiltonian P_U.

To prove the existence of the sequences $P_N(M)$, it is sufficient to consider the part of the sequence lying near P_c, say the points $P_N(M)$ with

$$L < N < M - L$$

where L is large but held fixed as $M \to \infty$. So long as

1a. $\lim_{M \to \infty} P_{M-L}(M) = P_{UL}$ (P_{UL} is the L^{th} point on the trajectory C_A)

2a. $\lim_{M \to \infty} P_L(M) = P_{RL}(a_1 \ldots a_k)$

one can reconstruct the remainders of the sequences using T_A or T_A^{-1} and satisfy the original requirements. If L is large enough, P_{UL} and P_{RL} will be near P_c and we can assume that

3a. $P_{N-1}(M) = P_c + U_A(P_N(M) - P_c)$

Since $P_N(M) - P_c$ satisfies the linearized equation, it must be a linear combination of the basic solutions for each M:

$$P_N(M) - P_c = \sum_\alpha \beta_\alpha(M)\, Q_{M-N\alpha} \tag{VII.10}$$

(Q depends on M-N rather than N, so that the index of Q increases as one applies U_A.) The sequence P_{UN} must also be a linear combination of the basic solutions:

$$P_{UN} = \sum_\alpha \gamma_\alpha\, Q_{N\alpha} + P_c \tag{VII.11}$$

Furthermore, since $P_{UN} \rightarrow P_c$ as $N \rightarrow \infty$ the coefficients γ_α must be zero for all unstable trajectories. Suppose, to be specific, that the unstable trajectories correspond to $1 \leq \alpha \leq k$ and that the trajectories for $\alpha > k$ are stable. Then $\gamma_\alpha = 0$ for $\alpha \leq k$. The requirement that $P_{M-L}(M) \rightarrow P_{UL}$ as $M \rightarrow \infty$ means that $\beta_\alpha(M)$ must satisfy

$$\lim_{M \rightarrow \infty} \beta_\alpha(M) = \gamma_\alpha \tag{VII.12}$$

The requirement that $P_L(M)$ have a limit as $M \rightarrow \infty$ means that $\sum_\alpha \beta_\alpha(M)\, Q_{M-L\alpha}$ must have a limit for $M \rightarrow \infty$. For the stable trajectories $Q_{M-L\alpha} \rightarrow 0$ as $M \rightarrow \infty$ and since $\beta_\alpha(M) \rightarrow \gamma_\alpha$ which is finite, the stable trajectories drop out in this limit. Assume that the unstable trajectories have pure exponential form (Eq. (VII.8); the author has not examined alternative forms in detail). Then the limit is $\sum_{\alpha=1}^{k} \beta_\alpha(M)\, (r_\alpha)^{M-L}\, Q_{0\alpha}$. For this to have a limit it is sufficient to have

$$\beta_\alpha(M) = a_\alpha (r_\alpha)^{-M} \qquad (1 \leq \alpha \leq k) \tag{VII.13}$$

where the constants a_α are arbitrary. Since $|r_\alpha| > 1$ for $\alpha \leq k$, the constants $\beta_\alpha(M)$ for $\alpha \leq k$ have the limit 0 as $M \rightarrow \infty$ as required by Eq. (VII.12). To complete the specification of $\beta_\alpha(M)$, put

$$\beta_\alpha(M) = \gamma_\alpha \qquad (\alpha > k)\,. \tag{VII.14}$$

With this specification of $\beta_\alpha(M)$, the points $P_N(M)$ satisfy the requirements 1a to 3a. The limit P_{RL} has the form

$$P_{RL} = P_c + \sum_{\alpha=1}^{k} a_\alpha \, Q_{0\alpha}(r_\alpha)^{-L} \tag{VII.15}$$

which has k arbitrary constants, as was stated at the beginning. In fact the renormalized points P_{RL} (for sufficiently large L) are just a linear combination of the k unstable trajectories of U_A, with the coefficients representing free parameters in the renormalized Hamiltonian.

In fact the transformation U_A has only one unstable trajectory, the renormalized Hamiltonian has only one free parameter and Eq. (VII.15) reduces to Eq. (VI.17) where the free parameter is a (which depends on θ_{R0}). It was also shown in Section VI that the eigenvalue of U_A (r_1 in Eq. (VII.15) or β^{-1} in Eq. (VI.17)) determines the scaling properties of the renormalized Hamiltonian at small distances.

As a final comment one notes that the unrenormalized Hamiltonian could be chosen to be any point P with $\theta = \pi/2$; the renormalized Hamiltonians are independent of the choice of the unrenormalized Hamiltonian since the sequences $P_N(M)$ will in the limit of large M go from the unrenormalized point to P_c and then along the unstable trajectory to a renormalized point P_{R0}.

In summary, the renormalized Hamiltonian is determined by properties of the fixed point P_c rather than of a particular unrenormalized Hamiltonian. The sequence of renormalized Hamiltonians P_{RN} approaches P_c as $N \to \infty$; for large N, $P_{RN} - P_c$ must be a linear combination of the unstable trajectories leaving P_c, and the different renormalized theories can be labelled by the coefficients a_α relating $P_{RN} - P_c$ to unstable trajectories. I think it is this relation of the renormalized theory to unstable trajectories leaving a fixed point which is simple, to answer the question raised earlier. The coefficients a_α are, I think, as close as one can get to being fundamental parameters in the theory.

B. Why Renormalization?

In this part we shall try to understand what features of the model Hamiltonian make renormalization necessary. The first step in the analysis will be to show that the transformation T is divergence free. Then the reason for the appearance of divergences in perturbation theory will be examined.

The statement that the transformation T is divergence free means the following. Let H be a Hamiltonian in S. Let H' be T(H). Let H have a decomposition $(J, \mathscr{E}, N, \theta, \underset{\sim}{A}_k, C_k)$ and H' have a decomposition $(J', \mathscr{E}', N-1, \theta', \underset{\sim}{A}_k', C_k')$. Then as discussed in Section V, if J, $\mathscr{E}$, θ, $\underset{\sim}{A}_k$ and C_k are held fixed while N varies, the quantities J', $\mathscr{E}'$, θ', $\underset{\sim}{A}_k'$, and C_k' are independent of N and cannot diverge for $N \to \infty$. Furthermore the transformation is continuous, that is if H and H'' are two Hamiltonians with transforms H' and H''', then $H' \to H'''$ when $H \to H''$. This continuity is uniform in N.

To understand the significance of T being divergence-free, one can study the divergences that appear in ordinary perturbation theory and see that they arise despite the finiteness of T. Consider the unrenormalized cutoff Hamiltonian H_M with a small bare coupling constant g_0 and large cutoff M. Consider also the effective Hamiltonian $H_0(M)$ which describes the ground state and first few excited states of H_M. That g_0 is small means the angle $\theta_M(M)$ (also called θ_{0M}, as in Eq. (V.87)) is small, and an expansion in g_0 can easily be converted into an expansion in $\theta_M(M)$. The effective Hamiltonian $H_0(M)$ is known if one knows three parameters $J_0(M)$, $\mathscr{E}_0(M)$, and $\theta_0(M)$ and the curve $Q_M(t)$ in S_A. The curve $Q_M(t)$ is well behaved for large M: as $M \to \infty$ it approaches the limit curve R(t). From Eq. (V.104), $J_0(M)$ is a simple function of $\theta_0(M)$. So any divergences in the low lying energy levels of H_M as $M \to \infty$ must be due to divergences in $\mathscr{E}_0(M)$ or $\theta_0(M)$ as $M \to \infty$. A divergence in $\mathscr{E}_0(M)$ affects only the ground state energy but not energy differences between the ground state and excited states. A divergence in

$\theta_0(M)$ means a divergence in differences of energy levels at least through the scale factor $J_0(M)$. The divergence in $\theta_0(M)$ can be identified as a coupling constant divergence while a divergence in $\mathscr{E}_0(M)$ is a ground state energy divergence.

To study the divergences in $\mathscr{E}_0(M)$ and $\theta_0(M)$ one uses Eqs. (V.91) and (V.92) of Section V. Let $\theta_M(M)$ be denoted θ_M; $\mathscr{E}_M(M)$ is zero (we do not make an energy subtraction in H_M). From the inequality (V.31) one finds that for θ small

$$f_L(\theta) = \theta - \eta_L \theta^3 \tag{VII.16}$$

with $\eta_L \simeq 1/2$. For $L \to \infty$, η_L approaches a limit η since $f_L(\theta)$ has a limit. To a first approximation one neglects the θ^3 term in Eq. (VII.16); then one gets $\theta_0(M) \simeq \theta_M$. To a second approximation one replaces θ^3 by θ_M^3; then Eq. (V.92) becomes:

$$\theta_N(M) = \theta_{N+1}(M) - \eta_{M-N}\theta_M^3 \tag{VII.17}$$

which gives

$$\theta_0(M) = \theta_M - \left\{\sum_{n=1}^{M} \eta_n \right\} \theta_M^3 \tag{VII.18}$$

For large M this becomes

$$\theta_0(M) \simeq \theta_M - M\eta\, \theta_M^3 \tag{VII.19}$$

and one has a divergence linear in M. This corresponds to a logarithmic divergence in the cutoff momentum (since the cutoff momentum is Λ^M). The energy $\mathscr{E}_0(M)$ is dominated by a contribution from $J_M(M)$:

$$\mathscr{E}_0(M) \simeq \Lambda^M (\cos \theta_M)^{-1}\ T_c\left[P_M(M)\right] \tag{VII.20}$$

Since $T_c \simeq -1$ for any argument, $\mathscr{E}_0(M)$ is linearly divergent in the cutoff momentum. These are the divergences one expects.

The divergence in $\mathscr{E}_0(M)$ is easy to understand. The ground state energy of H_M gets contributions from each meson degree of freedom represented in H_M. The degree of freedom m contributes an energy of order Λ^m for that is the energy scale for mesons in state ψ_m. The dominant energy is Λ^M associated with mesons having the cutoff momentum. So $\mathscr{E}_0(M)$ is of order Λ^M. In any case the divergence in $\mathscr{E}_0(M)$ as $M \to \infty$ arises because the scale factor $J_M(M) \to \infty$ as $M \to \infty$. This type of divergence occurs also in relativistic theories as mass renormalization. In some field theories the mass is linearly divergent. The cause of this is that when the cutoff is large the natural energy scale for self mass effects is the cutoff. Then one must let the bare mass in the Lagrangian be of order the cutoff and chosen very carefully so that all cutoff-dependent self masses cancel and the physical mass is much smaller than the cutoff.

The coupling constant divergence in $\theta_0(M)$ is more subtle. There is no question of a cutoff dependent scale here; θ is a dimensionless variable. The divergence is proportional to the number of degrees of freedom. It arises because the transformation T must be iterated M times to give $H_0(M)$ starting from H_M. These iterations define a sequence of constants $\theta_N(M)$. The difference between $\theta_N(M)$ and $\theta_{N+1}(M)$ is finite for all N and small in perturbation theory. However, these differences add in going from θ_M to $\theta_0(M)$, hence the divergence.

One sees from the above discussion that the divergences of perturbation theory derive from two causes. The linear divergence is due to the energy scale of the cutoff Hamiltonian H_M being Λ^M instead of the pion mass. The logarithmic divergence is due to the transformation T being iterated M times in going from θ_M to $\theta_0(M)$. The cause of the logarithmic divergence must be pursued further. Why was it necessary to compute $\theta_0(M)$ by an iterative process? Will an iterative method in which $\theta_0(M)$ is calculated in M steps always make $\theta_0(M)$ divergent when $M \to \infty$?

To set up the discussion pretend that the details of the analysis of the model had been different from what was reported in Section V. Suppose that the cutoff energy Λ^M had not been crucial for the discussion of the model, but that still one defined a sequence of constants $\theta_N(M)$ in going from θ_M to $\theta_0(M)$. What might one expect in this case? Then when M and N are large one would expect that there could be no appreciable difference between $\theta_N(M)$ and $\theta_{N+1}(M)$ for in both cases the effective cutoff (Λ^N or Λ^{N+1}) is large compared to the only important length. Most of the difference between $\theta_0(M)$ and θ_M would be due to the difference $\theta_0(M) - \theta_1(M)$ or $\theta_1(M) - \theta_2(M)$; the differences $\theta_N(M) - \theta_{N+1}(M)$ for large N would go to zero and could not accumulate to make $\theta_0(M)$ diverge for $M \to \infty$.

So the essential question is why the difference $\theta_N(M) - \theta_{N+1}(M)$ does not go to zero for large N, at least in perturbation theory. The answer lies in two features of the cutoff Hamiltonian H_M and the effective Hamiltonians $H_N(M)$. The first is that meson degrees of freedom of order N dominate the Hamiltonian $H_N(M)$ rather than meson degrees of freedom of order 1. As a result, the change from $H_N(M)$ to $H_{N-1}(M)$, which means eliminating the N^{th} degree of freedom, is a nontrivial change. Thus one can hardly expect $\theta_{N-1}(M)$ to be the same as $\theta_N(M)$ no matter how large N is. If by contrast the meson degrees of freedom of order 1 had been the dominant degrees of freedom in $H_N(M)$ for large N, then dropping the N^{th} degree of freedom would have been a negligible change and $\theta_{N-1}(M)$ would probably have been equal to $\theta_N(M)$. The second important feature is scale invariance. Scale invariance means that if the degrees of freedom of order 1 can be neglected (which is true for large N) then the process of going from $H_N(M)$ to $H_{N-1}(M)$ is indistinguishable from the process of going from $H_{N-1}(M)$ to $H_{N-2}(M)$. In particular if $H_{N-1}(M)$ differs from $H_N(M)$ only by a scale factor and an additive constant then $H_{N-2}(M)$ differs from $H_{N-1}(M)$ only by the same scale factor and another additive constant. Now if $\theta_N(M)$ is small, N is large and $M \gg N$, $H_{N-1}(M)$ does differ from $H_N(M)$ by little more than a scale factor and an additive constant. This is because $H_N(M)$ is defined by the constants $J_N(M)$, $\mathscr{E}_N(M)$, $\theta_N(M)$ and the point $Q_{M-N}(\theta_N(M))$ while $H_{N-1}(M)$ is defined by $J_{N-1}(M)$, $\mathscr{E}_{N-1}(M)$, $\theta_{N-1}(M)$, and $Q_{M-N+1}(\theta_{N-1}(M))$. If $\theta_N(M)$ is small then $\theta_{N-1}(M) \simeq \theta_N(M)$; since $Q_L(t) \approx R(t)$ when L is large, $Q_{M-N}(\theta_N(M)) \simeq Q_{M-N+1}(\theta_{N-1}(M))$ so only the scale factor $J_N(M)$ and constant $\mathscr{E}_N(M)$ can differ appreciably from $J_{N-1}(M)$ and $\mathscr{E}_{N-1}(M)$. But under these circumstances the effect of the transformation T on $H_N(M)$ and $H_{N-1}(M)$ is essentially the same, except for the effect on the scale factors J and the constants $\mathscr{E}$. This is scale invariance, and it means in particular that the difference $\theta_{N-2}(M) - \theta_{N-1}(M)$ is the same as the difference $\theta_{N-1}(M) - \theta_N(M)$ when $\theta_N(M)$ is small; hence the divergence in $\theta_0(M)$ in perturbation theory is proportional to M rather than some other function of M.

In conclusion, the fact that meson degrees of freedom of order the cutoff dominate the cutoff Hamiltonians makes renormalization inevitable. The divergence problem is not just an artifact of perturbation theory. Since the dominance of the degrees of freedom of order the cutoff is due to the energy of a meson increasing as its momentum increases, which is also true in relativistic theories, one expects that renormalization will be inevitable for strongly coupled relativistic theories too. We note also that not only does the transformation T determine basic properties of the renormalized theory, as shown in part A; it is also divergence free. Clearly one will want to try to define an analogous transformation for relativistic theories.

C. Analogy to the Renormalization Theory of Gell-Mann and Low

Gell-Mann and Low, in 1954, presented an analysis of the renormalization of Quantum Electrodynamics, and predicted that there would be an "eigenvalue condition" for the bare coupling constant.[8] That is, the bare coupling constant e_0 would have to have a fixed value independent of the value of the renormalized coupling constant. To be precise, they predicted that there would be a function $\psi(x)$ with the property that if e_0 is finite then e_0 is a root of the equation $\psi(e_0^2) = 0$. To show this, Gell-Mann and Low of necessity had to obtain ideas from perturbation theory and then extrapolate to the region of strong bare coupling constant. This involves several speculations, some of which will be criticized below. Nevertheless, the analysis of Gell-Mann and Low remains after 16 years the most sensible discussion in the literature of nonperturbative renormalization theory for relativistic field theory.

Here is a brief review of the Gell-Mann-Low theory: Let e be the physical (renormalized) electron charge and let m be the physical electron mass. Let $d_c(k^2/m^2, e^2)$ be the renormalized photon propagator apart from a factor k^{-2}. The customary normalization requirement for d_c is assumed:

$$d_c(0, e^2) = 1 \tag{VII.21}$$

Gell-Mann and Low define a generalization of the usual renormalization procedure for electrodynamics, with a different definition of the renormalized charge. In the Gell-Mann-Low program, the renormalized charge is a quantity e_λ depending on a subtraction point λ. The photon propagator is (apart from the factor k^{-2}) a function $d(k^2/\lambda^2, m^2/\lambda^2, e_\lambda^2)$ with the normalization condition

$$d(1, m^2/\lambda^2, e_\lambda^2) = 1 \tag{VII.22}$$

The propagator d is related to the usual propagator d_c through the relation

$$e^2 d_c(k^2/m^2, e^2) = e_\lambda^2\, d(k^2/\lambda^2, m^2/\lambda^2, e_\lambda^2) \tag{VII.23}$$

In particular, putting $k^2 = \lambda^2$ gives

$$e_\lambda^2 = e^2 \, d_c(\lambda^2/m^2, \, e^2) \tag{VII.24}$$

which gives the definition of e_λ in terms of e. In the Gell-Mann-Low program, all other amplitudes (electron propagator, vertex function, etc.) are functions of e_λ, and all depend on the reference momentum λ as well as m and various momenta. The subtraction procedure of Gell-Mann and Low is defined so that the bare coupling constant e_0 is the limit of e_λ as $\lambda \to \infty$.

Gell-Mann and Low then argue that the function $d(k^2/\lambda^2, \, m^2/\lambda^2, \, e_\lambda^2)$ has a finite limit when $m \to 0$ holding k^2, λ^2, and e_λ^2 fixed. This should also be true of other amplitudes. They give an example of this from fourth order perturbation theory, and then argue that it is true in general because the momentum k provides an infrared cutoff. Whether the finiteness assumption is true is still an open question; the author knows of no reason to doubt it, and it will be assumed to be correct in the following.

If $d_c(k^2/m^2, \, e^2)$ is expanded in powers of e^2 for k^2 large, the coefficients involve logarithms of k^2/m^2, so that the effective expansion parameter is $e^2 \ln(k^2/m^2)$ not e^2; this means that radiative corrections become important when $\ln(k^2/m^2)$ is sufficiently large, no matter how small e is. In contrast, as Gell-Mann and Low make clear, the fact that d is independent of m^2/λ^2 when m^2/λ^2 is small means that the expansion of $d(k^2/\lambda^2, \, m^2/\lambda^2, \, e_\lambda^2)$ in powers of e_λ^2 involves no large logarithms if k and λ are simultaneously large so that k^2/λ^2 is of order 1. In fact, in this case the coefficients of e_λ^2, e_λ^4, etc. are of order 1 no matter how large k and λ are.

To compute e_0 from Eq. (VII.24) directly would be difficult since for any e the radiative corrections to $d_c(\lambda^2/m^2, \, e^2)$ are infinite in the limit $\lambda \to \infty$. So

Gell-Mann and Low develop an indirect procedure which requires knowing only $d(k^2/\lambda^2, 0, e_\lambda^2)$ for k^2 near λ^2. The radiative corrections to d will be important because, as will be seen, one will have to consider coupling constants e_λ of order 1. But unless one must consider the limit $e_\lambda \to \infty$, the radiative corrections will be finite. The trick of Gell-Mann and Low is to observe that one can use the function d to set up an equation for $de_\lambda/d\lambda$. From Eq. (VII.23) one finds that, for any λ and λ'

$$e_\lambda^2\, d(k^2/\lambda^2, m^2/\lambda^2, e_\lambda^2) = e_{\lambda'}^2\, d(k^2/\lambda'^2, m^2/\lambda'^2, e_{\lambda'}^2) \tag{VII.25}$$

Putting $k=\lambda'$ gives

$$e_{\lambda'}^2 = e_\lambda^2\, d(\lambda'^2/\lambda^2, m^2/\lambda^2, e_\lambda^2) \tag{VII.26}$$

If λ and λ' are both much larger than m one can neglect the m dependence. Differentiating with respect to λ' and then putting $\lambda'=\lambda$ and approximating m/λ by 0 gives

$$2e_\lambda(de_\lambda/d\lambda) = 2\,\psi(e_\lambda^2)/\lambda \tag{VII.27}$$

where

$$\psi(x) = yx \left.\frac{\partial d(y, 0, x)}{\partial y}\right|_{y=1} \tag{VII.28}$$

The function $\psi(x)$ has a power series expansion in x for small x with finite coefficients; Gell-Mann and Low assume it has a well-defined extrapolation to values of x of order 1. To compute the limit of e_λ for $\lambda \to \infty$ one must solve the differential equation (VII.27). If $de_\lambda/d\lambda$ does not go to zero for finite e_λ, then necessarily an infinite increase in λ will give an infinite increase in e_λ. So the only way e_λ can stay finite as $\lambda \to \infty$ is for $\psi(e_\lambda^2)$ to have a zero. If $\psi(x)$ has a zero at $x=x_0$ and is positive for $x < x_0$ (ψ is positive for small x from perturbation theory) then the solution e_λ^2 of Eq. (VII.27) will be an increasing function of λ with the limit x_0 as $\lambda \to \infty$ (assuming $e_\lambda^2 < x_0$ when λ is of order m, as it will be if e is small).

If $\psi(x)$ has a zero at x_0 then the function e_λ^2 will have the limit x_0 as $\lambda \to \infty$ for any value of e sufficiently small. This demonstrates the main result of Gell-Mann and Low: the bare coupling constant e_0 is <u>independent</u> of the physical coupling constant e, at least over some finite range for e. Even if $\psi(x)$ does not have a zero, the solution e_λ will have the limit ∞ for $\lambda \to \infty$ independently of the value of e: the bare coupling constant is again independent of the physical coupling constant. (This is true only for certain forms of the function $\psi(x)$. If the integral $\int_1^\infty dx/\psi(x)$ is finite then $e_\lambda \to \infty$ for some finite value of λ and the theory becomes nonsense for larger values of λ. This leads to contradictions discussed below.)

Thus Gell-Mann and Low predicted for electrodynamics the result that <u>one</u> unrenormalized Lagrangian would have an infinite number of solutions. This is exactly the result that was proved for the model in part A of this section.

The differential equation (VII.27) can be regarded as analogous to the transformation equations

$$P_{RN} = T_A(P_{RN+1}) \qquad \text{(VII.29)}$$

which is involved in the definition of the renormalized Hamiltonian of the model. Equation (VII.27) tells how a coupling constant e_λ changes as λ changes, while Eq. (VII.29) tells how an infinite set of coupling constants change as N changes. One can think of the function ψ as defining an infinitesimal transformation on a one dimensional coupling constant space. In the limit $\lambda \to \infty$, e_λ goes to a fixed point of the transformation defined by ψ (if $\psi(e_0^2) = 0$, then for $e_\lambda = e_0$, $de_\lambda/d\lambda = 0$: thus e_0 is a fixed point). This is analogous to the result that the limit of P_{RN} as $N \to \infty$ is a fixed point of T_A. Thus Gell-Mann and Low discovered the idea that a fixed point of a transformation is important in renormalization. There are differences between Gell-Mann and Low's fixed point e_0 and the fixed point P_c, these differences will be emphasized below. These differences do not alter the fact that Gell-Mann and Low discovered the essential idea of a fixed point. Since they discovered the idea in the context of relativistic field theory, this is encouragement to believe that the analysis of the fixed point in the model is relevant to relativistic field theory and not just a consequence of the many simplifications which were made in defining the model.

There are two basic differences between the transformation T_A defined for the model and the transformation ψ of electrodynamics. First, the function ψ can only be computed after electrodynamics has been solved, whether by a perturbation expansion or whatever. This is because ψ is defined in terms of the renormalized propagator which is itself part of the solution of electrodynamics. In particular if electrodynamics does not have a solution except as a perturbation expansion then the ψ function will not exist for strong coupling. In contrast the transformation T_A is defined before one knows whether the model has a solution. In the model of this paper the renormalized theory exists; but there are other models for which there is no renormalized theory (except one with no coupling). A particular example is a derivative of the Lee model constructed by analogy with the model of this paper. An earlier version of such a model was described in a previous paper[7] and from the analysis given there it is easy to see what happens in the truncated Lee model. One defines a transformation analogous to T_A, and uses it to construct curves analogous to $Q_L(t)$. However these curves do not exist over the full range $0 \leq t \leq \pi/2$ but rather over a range $0 \leq t \leq t_L$ where the constants t_L form a decreasing sequence with the limit 0 as $L \rightarrow \infty$. The reason for this is that if a Hamiltonian has component θ, the Lee model transformation takes θ into θ' where $\theta' < \theta$ for any $\theta > 0$ including $\theta = \pi/2$. This means also that the Lee model T_A has no fixed point analogous to P_c. This analysis assumes that one does not permit complex coupling constants, as would be necessary if one wants to obtain nontrivial renormalized solutions. Since every time one considers a new theory the existence of a fixed point of the corresponding transformation T_A is in doubt, and since renormalizability depends on there being such a fixed point (at least for the two examples considered; a general analysis of renormalization theory indicates renormalization could be possible for some types of transformations without fixed points) it is important that T_A be defined without reference to the renormalized theory.

The second difference between ψ and T_A is that ψ acts on a one-dimensional space, while T_A acts on an infinite dimensional space. In order to formulate the transformation ψ as a transformation on one variable one has to know that the renormalized theory depends on only one phenomenological parameter. For example, in pseudoscalar meson theory where there are two phenomenological parameters, one must replace ψ by a transformation on a two-dimensional space. But the lesson of the model of this paper is that the number of phenomenological parameters is not known until one has found the fixed point of T_A and determined the number of unstable solutions of T_A near the fixed point. The fact that T_A is a transformation on an infinite set of coupling constants means one is not committed in advance to a particular number of phenomenological constants. Furthermore one is not restricted to theories with interactions which are renormalizable. As long as T_A is a transformation on the space of all possible couplings, renormalizable or not, the customary reason for considering only renormalizable interactions becomes irrelevant. The customary reason is that nonrenormalizable interactions require an infinite set of counter terms to be renormalized; but now these counter terms are all present anyways in the phenomenological Hamiltonians (or Lagrangians, perhaps). So if the renormalization theory of the model can be generalized to relativistic field theory there is hope to study pure quark models or the Fermi interaction, although there is no guarantee that the corresponding transformations will have fixed points.

ACKNOWLEDGEMENTS

I wish to thank Professor S. Drell and members of the SLAC theory group for their hospitality during the author's stay at SLAC.

REFERENCES AND FOOTNOTES

1. This idea persists despite the results of S. Adler and Wu-Ki Tung, Phys. Rev. Letters 22, 978 (1969) and R. Jackiw and G. Preparata, Phys. Rev. Letters 22, 975, 1162 (1969) who show that equal time commutators are affected in perturbation theory by renormalization.
2. T. D. Lee, Phys. Rev. 95, 1329 (1954).
3. W. Thirring, Ann. Phys. (N.Y.) 3, 91 (1958); K. Johnson, Nuovo Cimento 20, 773 (1961).
4. The anomalous scaling properties of the Thirring model are implicit in Johnson's solution (Ref. 3); see also Ref. 6.
5. The two-dimensional ϕ^4 theory analyzed by Jaffe and Glimm is in this class. For references see A. Jaffe, Rev. Mod. Phys. 41, 576 (1969).
6. K. Wilson, Phys. Rev. 179, 1499 (1969).
7. K. Wilson, Phys. Rev. 140, B445 (1965).
8. M. Gell-Mann and F. E. Low, Phys. Rev. 95, 1300 (1954). See also M. Baker and K. Johnson, Phys. Rev. 183, 1292 (1969).
9. The full Hamiltonian is given by Eq. (1) of Ref. 7.
10. Cf. the paper of A. Wightman in High Energy Electromagnetic Interactions and Field Theory, M. Levy, Ed. (Gordon and Breach Science Publishers, New York, 1967), especially pp. 245-262 and references cited therein.
11. See also Section III of Ref. 7.
12. See Section VII, part C.
13. Throughout this paper $\leq$ means "not greater than"; there is no implication that equality can be realized.

14. Equation (V.104) means that the coefficient of $(a_0^+ a_0 + b_0^+ b_0 - 1)$ in $H_N(M)$ is simply Λ^N independently of the value of $\theta_N(M)$. This is also an immediate consequence of Eq. (V.33) of Theorem 2.

15. J. Wess, Nuovo Cimento 18, 1086 (1960).

16. N. N. Bogoliubov and V. Shirkov, Introduction to the Theory of Quantized Fields (Interscience Publishers, Inc., New York, 1959), pp. 528-529.

17. $\bar{Q}$ is a symbol completely independent of Q.

18. The symbol g is moonlighting here. Sorry.

19. The function f is not the function defined in Section V. Sorry.

APPENDIX A

It is proven here that an iterative solution to Eq. (IV.9) exists and that $1 + R^+ - R$ has an inverse provided that

$$\|H_I\| < .2\,\Delta E \tag{A.1}$$

where ΔE is the energy difference between the ground states and first excited state of H_0. Define a sequence of operators $\{R_n\}$ by

$$R_0 = 0 \tag{A.2}$$

$$R_n = (1-P)(E_0 - H_0)^{-1}(1 - P - R_{n-1})\,H_I(P + R_{n-1}) \qquad (n > 0) \tag{A.3}$$

Then

$$R = \lim_{n \to \infty} R_n \tag{A.4}$$

To prove the existence of the limit the following equation is useful

$$R_{n+1} - R_n = (1-P)(E_0 - H_0)^{-1}\left\{(1-P-R_n)\,H_I\,(R_n - R_{n-1}) - (R_n - R_{n-1})\,H_I\,(P + R_{n-1})\right\} \tag{A.5}$$

Now it is shown that

$$\|R_n\| < .4 \tag{A.6}$$

Proof: This is true for n=0. Suppose it is true for n-1. Now

$$\|(1-P)(E_0 - H_0)^{-1}(1-P)\| = \Delta E^{-1} \tag{A.7}$$

$$\|1-P\| = 1 \tag{A.8}$$

$$\|P\| = 1 \tag{A.9}$$

and $R_{n-1} = (1-P)\,R_{n-1}$ from Eq. (A.3). So

$$\|R_n\| \le \Delta E^{-1}(1.4)(.2\,\Delta E)(1.4) < .4 \tag{A.10}$$

Q.E.D.

Likewise one can show that

$$||R_{n+1} - R_n|| \leq .4 \times (.56)^n \tag{A.11}$$

Hence from the Cauchy criterion, R exists. It is easily shown that R satisfies Eq. (IV.9). The bound (A.6) implies that

$$||R^+ - R|| < .8 \tag{A.12}$$

which means the inverse of $1 + R^+ - R$ exists as a power series in $R^+ - R$.

APPENDIX B

In this appendix the transformation T will be defined in detail. It will be shown that T has the form of Eqs. (V.16) - (V.19). Then Theorems 1 - 4 of Section V will be proven. The only assumption made in this appendix is that $\Lambda > 4 \times 10^6$.

The first problem is to define T. Let H be a Hamiltonian in S. Let H have the decomposition $(J, \mathscr{E}, N, P_A)$ where $P_A \in S_A$. Let P_A have the decomposition $(\theta, \underset{\sim}{A}_k, C_k)$. Let $H = H_0 + H_I$ with H_0 given by Eq. (V.14). Define H_{eff} using Eq. (IV.29). To define T we must specify the decomposition of H_{eff}. The decomposition of H_{eff} must be defined because it is not unique, as was pointed out in Section V. This nonuniqueness means that one must often prove properties for the decomposition of an operator which are obvious or already established for the operator itself. To define this decomposition we will write out in detail the steps leading to H_{eff}, and specify the decomposition of each of the operators arising in the calculation. The operator H_I has the form $H_I = J\mathscr{H}_I$ with

$$\mathscr{H}_I = \left\{ \sum_{k=1}^{N} \underset{\sim}{V}_k \cdot \underset{\sim}{B}_{k-1} + \sum_{k=0}^{N} C_k \right\} \tag{B.1}$$

where

$$\underset{\sim}{B}_0 = (m, \sqrt{2}\, g\, \tau^+, \sqrt{2}\, g\, \tau^-) + \underset{\sim}{A}_0 \tag{B.2}$$

$$\underset{\sim}{B}_k = \underset{\sim}{A}_k \qquad (k > 0) \tag{B.3}$$

and $m = \cos\theta$, $g = (1/\sqrt{2}) \sin\theta$. The equations which define H_{eff} are as follows (including the iterative definitions of R and $(1 + R^+R)^{\pm 1/2}$):

$$R_0 = 0 \tag{B.4}$$

$$R_n = (E_0 - H_0)^{-1} (1 - P - R_{n-1})\, H_I (P + R_{n-1}) \tag{B.5}$$

$$R = \lim_{n \to \infty} R_n \tag{B.6}$$

$$Q_0 = 0 \tag{B.7}$$

$$Q_n = 1/2 \left(R^+ R - Q_{n-1}^2 \right) \tag{B.8}$$

$$Q = \lim_{n \to \infty} Q_n \tag{B.9}$$

$$\bar{Q}_0 = 0 \tag{B.10}$$

$$\bar{Q}_n = -Q - Q\bar{Q}_{n-1} \tag{B.11}$$

$$\bar{Q} = \lim_{n \to \infty} \bar{Q}_n \tag{B.12}$$

$$H_{eff} = (P+Q)\, H_I\, (P+R)\, (P+\bar{Q}) + PE_0 \tag{B.13}$$

In these formulae P is the projection operator onto the two ground states of H_0, $(P+Q)$ is $(1+R^+R)^{1/2}$ P and[17] $(P+\bar{Q})$ is $P(1+R^+R)^{-1/2}$, and E_0 is the ground state energy of H_0.

A particular form for $\mathscr{H}_I$ has been given in Eq. (B.1). The operators $\underset{\sim}{B}_k$ and C_k will be called the decomposition of $\mathscr{H}_I$. Analogous decompositions will now be defined for R_n, etc. The equations (B.4) - (B.13) involve three basic operations: multiplication of H_I, R_n, etc. with themselves, multiplication with P or with $(E_0 - H_0)^{-1}$. So it is sufficient to define the decomposition of any of these products. Let X be an operator with decomposition $(\underset{\sim}{D}_k, F_k)$ say. Then PX has the obvious decomposition $(P\underset{\sim}{D}_k, PF_k)$, and analogously for $(E_0 - H_0)^{-1}$ X. This is a legitimate decomposition since the only requirement on a decomposition $(\underset{\sim}{D}_k, F_k)$ is that $\underset{\sim}{D}_k$ and F_k do not involve meson operators numbered above k (no upper bounds on $\underset{\sim}{D}_k$ and F_k will be imposed now). Since P and $(E_0 - H_0)^{-1}$ act in the space of nucleons and 0-mesons (meson operators numbered 0), this restriction is satisfied by $P\underset{\sim}{D}_k$ and PF_k, or $(E_0 - H_0)^{-1} \underset{\sim}{D}_k$ and $(E_0 - H_0)^{-1} F_k$. Now let Y be another operator with decomposition $(\underset{\sim}{A}_k, C_k)$. One must define a decomposition $(\underset{\sim}{G}_k, L_k)$ for the product XY. The decomposition is as follows:

$$\underset{\sim}{G}_0 = F_0 \underset{\sim}{A}_0 + \underset{\sim}{D}_0 C_0 \tag{B.14}$$

$$L_0 = F_0 C_0 \tag{B.15}$$

For $k \geq 1$:

$$\begin{aligned} \underset{\sim}{G}_k = & \sum_{m=0}^{k} \left\{ F_m \underset{\sim}{A}_k + \underset{\sim}{D}_k C_m \right\} + \sum_{m=0}^{k-1} \left\{ F_k \underset{\sim}{A}_m + \underset{\sim}{D}_m C_k \right\} \\ & + \sum_{n=1}^{k} \sum_{m=0}^{n-1} \left\{ \underset{\sim}{D}_k (\underset{\sim}{T}_n \cdot \underset{\sim}{A}_m) + (\underset{\sim}{T}_n \cdot \underset{\sim}{D}_m) \underset{\sim}{A}_k \right\} \\ & + \sum_{n=0}^{k-1} \sum_{m=0}^{k-1} \left\{ \underset{\sim}{D}_n (\underset{\sim}{T}_k \cdot \underset{\sim}{A}_m) + (\underset{\sim}{T}_k \cdot \underset{\sim}{D}_n) \underset{\sim}{A}_m \right\} \end{aligned} \tag{B.16}$$

$$\begin{aligned} L_k = & \sum_{n=0}^{k-1} \sum_{m=0}^{k-1} (\underset{\sim}{T}_k \cdot \underset{\sim}{D}_n)(\underset{\sim}{T}_k \cdot \underset{\sim}{A}_m) \\ & + \sum_{n=1}^{k} \sum_{m=0}^{n-1} \left\{ (\underset{\sim}{T}_n \cdot \underset{\sim}{D}_m) C_k + F_k (\underset{\sim}{T}_n \cdot \underset{\sim}{A}_m) \right\} \\ & + \sum_{m=0}^{k} F_k C_m + \sum_{m=0}^{k-1} F_m C_k \end{aligned} \tag{B.17}$$

$\underset{\sim}{T}_n$ is defined by Eqs. (V.4) - (V.6). It is clear from these formulae that $\underset{\sim}{G}_k$ and L_k do not involve meson operators numbered above k. With some straightforward algebra one can verify that the operator product XY is given by

$$XY = \sum_{k=1}^{N} \underset{\sim}{V}_k \cdot \underset{\sim}{G}_{k-1} + \sum_{k=0}^{N} L_k \tag{B.18}$$

It can be shown that the decomposition is associative, i.e., a triple product (XY)Z has the same decomposition as X(YZ).

With the rules specified above and Eqs. (B.1) - (B.13), the decomposition of ($H_{eff} - E_0 P$) is uniquely defined. Note that the number of degrees of freedom N nowhere enters into the calculation of $\underset{\sim}{G}_k$ and L_k. So if the operators $\underset{\sim}{B}_k$ and C_k in the decomposition of $\mathscr{H}_I$ are defined for <u>all</u> k and are independent of N, then the decompositions of R_n, etc. (including H_{eff}) will also be defined for all k and independent of N. It will be presumed from now on that decompositions are defined and computed for all k. Note also that H_I, H_0, and E_0 are all proportional to J. This makes R_n, Q_n, etc. independent of J and H_{eff} proportional to J. To be specific, H_{eff} has the form

$$H_{eff} = PE_0 + J\left\{\sum_{k=1}^{N} \underset{\sim}{V}_k \cdot \underset{\sim}{G}_{k-1} + \sum_{k=0}^{N} L_k\right\} \tag{B.19}$$

where $\underset{\sim}{G}_k$ and L_k depend on $\tau^{\pm}$ and meson operators numbered 0 to k.

Since H_{eff} acts within the subspace projected by P, the dependence of H_{eff} on the 0-meson operators (a_0, etc.) and $\tau^{\pm}$ can be reduced to a dependence on $\tau_R^{\pm}$, the raising and lowering operators for the ground states of H_0. When this is done, $\underset{\sim}{G}_k$ and L_k depend only on $\tau_R^{\pm}$ and meson operators numbered 1 to k. To put H_{eff} in a form in which it can be contained in the space S, one must renumber the meson operators 1 - N to run from 0 to N - 1, e.g., $a_1 \rightarrow a_0$, $a_2 \rightarrow a_1$, etc. Also one replaces $\tau_R^{\pm}$ by $\tau^{\pm}$. Under this renumbering, $\underset{\sim}{V}_k$ becomes $\Lambda^{-1}\, \underset{\sim}{V}_{k-1}$; H_{eff} is

$$H_{eff} = E_0 + J\,\Lambda^{-1}\left\{\sum_{k=0}^{N-1} \underset{\sim}{V}_k \cdot \underset{\sim}{G}_k + \Lambda \sum_{k=0}^{N} L_k\right\} \tag{B.20}$$

where $\underset{\sim}{G}_k$ and L_k depend on $\tau^{\pm}$ and meson operators numbered 0 to k-1. PE_0 is replaced by E_0 because there is no longer any possible reference to states outside the subspace projected by P.

Now consider $\underset{\sim}{G}_0$ and L_0. They involve no meson operators; they can be expressed purely in terms of $\tau^{\pm}$. Furthermore $\underset{\sim}{G}_0$ and L_0 satisfy the appropriate hermiticity, charge conservation, charge conjugation, and time reversal requirements, because these requirements are preserved by the equations defining the decomposition of H_{eff}. These requirements force L_0 to be a real constant and $\underset{\sim}{G}_0$ to have the form

$$\underset{\sim}{G}_0 = (m'', \sqrt{2}\, g''\, \tau^+, \sqrt{2}\, g''\, \tau^-) \tag{B.21}$$

where m'' and g'' are real constants.

It is now easy to define a decomposition of H_{eff} in the space S. Denote the decomposition $(J', \mathscr{E}', N', P'_A)$ with P'_A having the decomposition $(\theta', \underset{\sim}{A}'_k, C'_k)$. Comparison of Eq. (B.20) with Eqs. (V.1), (V.2), (V.8), and (V.9) leads to the following formulae

$$J' = \Lambda^{-1} J\, (m''^2 + 2g''^2)^{1/2} \tag{B.22}$$

$$\mathscr{E}' = E_0 + JL_0 \tag{B.23}$$

$$N' = N-1 \tag{B.24}$$

$$\theta' = \tan^{-1}(\sqrt{2}\, g''/m'') \tag{B.25}$$

$$\underset{\sim}{A}'_k = (m''^2 + 2g''^2)^{-1/2}\, \underset{\sim}{G}_{k+1} \tag{B.26}$$

$$C'_k = \Lambda(m''^2 + 2g''^2)^{-1/2}\, L_{k+1} \tag{B.27}$$

The quantities m'', g'', L_0, $\underset{\sim}{G}_k$, and L_k depend only on P_A, not on J, $\mathscr{E}$, or N. Hence, Eq. (B.22) has the form of Eq. (V.17) with $T_B(P_A) = (m''^2 + 2g''^2)^{1/2}$. Also Eqs. (B.25) - (B.27) define the transformation $T_A(P_A)$ of Eq. (V.19). Finally, the ground state energy of H_0 (defined by Eq. (V.14)) is

$$E_0 = \mathscr{E} - J \tag{B.28}$$

(of Table I). Hence $\mathscr{E}'$ has the form of Eq. (V.18) with

$$T_c(P_A) = -1 + L_0 \tag{B.29}$$

The next problem is to prove Theorems 1 - 4 of Section V. The proofs involve a very large number of upper bounds, and are quite complex. To guard against subtle errors, all bounds have been obtained as explicit numbers multiplying powers of Λ. In principle, it would have been sufficient to know that bounds existed in the form of unknown sufficiently large numbers multiplying known powers of Λ. In addition the use of numbers saves symbols. In the following $\leq$ means only $\ngtr$ (the equality need not be realized). These proofs are crucial to the renormalization of the model of this paper; they are condemned to an appendix because they are special to the model whereas the analysis of Section V is of more general interest.

To start with, one needs an upper bound for the decomposition of the product XY given bounds on X and Y. Let X and Y have decompositions $(\underset{\sim}{D}_k, F_k)$ and $(\underset{\sim}{A}_k, C_k)$ as before. It is convenient to define an abstract bound for X. This bound will consist of three numbers (d, e, f). By definition, X has a bound (d, e, f) if

$$\begin{aligned}
&\|\underset{\sim}{D}_0\| \leq \underset{\sim}{m} d && \|\underset{\sim}{D}_k\| \leq \underset{\sim}{m} e\, \Lambda^{-k} && (k \geq 1) \\
&\|F_0\| \leq f\, \Lambda^{-1} && \|F_k\| \leq e\, \Lambda^{-2k} && (k \geq 1)
\end{aligned} \tag{B.30}$$

where $\|\underset{\sim}{D}_0\|$ is a vector with components $\|D_{01}\|$, $\|D_{02}\|$, $\|D_{03}\|$, $\|D_{01}\|$ being the ordinary operator bound. Also $\underset{\sim}{m}$ is the vector $(m, \sqrt{2}\, g, \sqrt{2}\, g)$, and m and g are as defined before.

Suppose X has a bound (d, e, f) and Y has a bound (a, b, c). Then it can be shown that XY has a bound (g, h, ℓ),[18] with

$$g = \Lambda^{-1} (af + dc) \tag{B.31}$$

$$h = 5\, ad + \Lambda^{-1} (bf + ec + \sqrt{70}\, (ae + bd)) + \Lambda^{-2} (14be) \tag{B.32}$$

$$\ell = cf\, \Lambda^{-1} \tag{B.33}$$

(These bounds were computed assuming only that $\Lambda > 21$.) A brief summary of the proof of these bounds is as follows. The operators T_{ki} have bounds

$$\|T_{ki}\| = \Lambda^{-k} \tag{B.34}$$

(This is proved by a straightforward calculation.) Next one puts bounds on the sums $\sum_{n=1}^{\infty} \|\underset{\sim}{T}_n\|$, $\sum_{k=0}^{\infty} \|\underset{\sim}{A}_n\|$, etc. (which are also bounds for finite sums such as $\sum_{n=0}^{k} \|\underset{\sim}{A}_n\|$). One gets

$$\sum_{n=1}^{\infty} \|\underset{\sim}{T}_n\| = \underset{\sim}{1}\, \Lambda^{-1} (1 - \Lambda^{-1})^{-1} \leq 1.05\, \Lambda^{-1}\, \underset{\sim}{1} \tag{B.35}$$

where $\underset{\sim}{1}$ is the vector (1, 1, 1). Also

$$\sum_{n=0}^{\infty} \|\underset{\sim}{A}_n\| \leq \underset{\sim}{m}a + \underset{\sim}{m}b\, \Lambda^{-1} (1 - \Lambda^{-1})^{-1} \leq \underset{\sim}{m} (a + 1.05\, \Lambda^{-1} b) \tag{B.36}$$

using the definition of the bound (a, b, c). Similar formulae can be obtained for sums of $\|C_n\|$, $\|\underset{\sim}{D}_n\|$, and $\|F_n\|$. Now one constructs upper bounds for all the terms in Eqs. (B.14) - (B.17) for $\underset{\sim}{G}_0$, L_0, $\underset{\sim}{G}_k$ and L_k. For example, one term in $\underset{\sim}{G}_k$ is

$$\left\| \sum_{n=1}^{k} \sum_{m=0}^{n-1} \underset{\sim}{D}_k\, \underset{\sim}{T}_n \cdot \underset{\sim}{A}_m \right\| \leq \|\underset{\sim}{D}_k\| \sum_{n=1}^{\infty} \|\underset{\sim}{T}_n\| \cdot \sum_{m=0}^{\infty} \|\underset{\sim}{A}_m\|$$

$$\leq \underset{\sim}{m}e\, \Lambda^{-k} (1.05\, \Lambda^{-1})\, \underset{\sim}{1} \cdot \underset{\sim}{m}\, (a + 1.05\, \Lambda^{-1}\, b) \tag{B.37}$$

Now $\underset{\sim}{1} \cdot \underset{\sim}{m} = m + 2\sqrt{2}g$. An upper bound on $m + 2\sqrt{2}g$ results from

$$(m + 2\sqrt{2}g)^2 - 5(m^2 + 2g^2) = -(2m - \sqrt{2}g)^2 < 0 \qquad \text{(B.38)}$$

Because $m^2 + 2g^2 = 1$ (normalization condition), one gets the bound

$$\underset{\sim}{1} \cdot \underset{\sim}{m} \leq \sqrt{5} \qquad \text{(B.39)}$$

Hence

$$\left\| \sum_{n=1}^{k} \sum_{m=0}^{n-1} \underset{\sim}{D}_k \underset{\sim}{T}_n \cdot \underset{\sim}{A}_m \right\| \leq \underset{\sim}{m} \Lambda^{-k-1} (1.05\sqrt{5})(ae + 1.05\Lambda^{-1} be) \qquad \text{(B.40)}$$

Similarly one finds bounds for all terms in Eqs. (B.14) - (B.17); the result is that (g, h, ℓ) given by Eqs. (B.31) - (B.33) is an upper bound for the product XY. It is convenient to introduce a shorthand for Eqs. (B.31) - (B.33): we define the "product" (a, b, c) (d, e, f) to be the quantities (g, h, ℓ) given by Eqs. (B.31) - (B.33). This product can be shown to be associative and commutative so algebraic expressions involving these products can be manipulated using ordinary algebra. This simplifies the calculations.

Using the bound quoted above for products, one can construct a set of upper bounds for the operators R_n, Q_n, etc. These bounds are listed in Table IV. They are not least upper bounds. The operator $\bar{\bar{H}}$ in Table IV is defined in terms of H_{eff} by

$$H_{eff} = PE_0 + JP\, \underset{\sim}{M} \cdot \underset{\sim}{V}_1 P + J\bar{\bar{H}} \qquad \text{(B.41)}$$

where H_{eff} is the effective Hamiltonian as of Eqs. (B.13) and (B.19) before re-numbering the meson operators, and $\underset{\sim}{M}$ is

$$\underset{\sim}{M} = (m, \sqrt{2}g\,\tau^+, \sqrt{2}g\,\tau^-) \qquad \text{(B.42)}$$

The operator $\mathscr{H}_J$ in Table IV is defined by

$$\mathscr{H}_J = \sum_{k=1}^{N} \underset{\sim}{V}_k \cdot \underset{\sim}{A}_{k-1} + \sum_{k=0}^{N} C_k \tag{B.43}$$

so that

$$\mathscr{H}_I = \underset{\sim}{M} \cdot \underset{\sim}{V}_1 + \mathscr{H}_J \tag{B.44}$$

The proofs of the bounds of Table IV are mostly straightforward and only examples of the proofs will be given here. In some cases the bounds of Table IV are gross overestimates of the bounds one calculates in the proofs quoted below.

The bounds on $\mathscr{H}_J$ and $\mathscr{H}_I$ are simple consequences of the definition of the space S, in particular the bounds (V.10) - (V.13). In computing the bound on $\mathscr{H}_I$ one also uses the inequality $g^2 < 1/2$ which follows from the definition (V.9) of g (also one uses $\Lambda > 200$).

In proving the bound on R_n, it is convenient to eliminate the factor J by defining

$$H_0 = E_0 + J\mathscr{H}_0 , \tag{B.45}$$

Now write the equation for R_n as

$$\begin{aligned} R_n &= (-\mathscr{H}_0)^{-1} (1-P) \underset{\sim}{M} \cdot \underset{\sim}{V}_1 P \\ &+ (-\mathscr{H}_0)^{-1} (1-P) \left\{ \mathscr{H}_J P - R_{n-1} \mathscr{H}_I P + \mathscr{H}_I R_{n-1} - R_{n-1} \mathscr{H}_I R_{n-1} \right\} \end{aligned} \tag{B.46}$$

The proof of the bound on R_n is by induction. It is true of R_0. Assume it is true of R_{n-1}. To bound the first term of Eq. (B.46) one needs the following bounds:

$$\| (-\mathscr{H}_0)^{-1} (1-P) \| = 1 \tag{B.47}$$

$$\| (1-P) \tau^+ P \| = \| (1-P) \tau^- P \| = g \tag{B.48}$$

These bounds can be obtained by explicit calculation using Table I. With this information one finds that $(-\mathscr{H}_0)^{-1}(1-P)\,\underset{\sim}{M}\cdot\underset{\sim}{V}_1\,P$ has a bound (g, 0, 0). The second term in Eq. (B.46) can be bounded using the bound $\|P\| = 1$, Eq. (B.47), and the bounds of Table IV for $\mathscr{H}_J$, R_{n-1}, and $\mathscr{H}_I$. Schematically one has

$$|R_n| \leq (g,\ 0,\ 0) + |\mathscr{H}_J| + 2|R_{n-1}||\mathscr{H}_I| + |\mathscr{H}_I|\ |R_{n-1}|^2 \tag{B.49}$$

where $|\mathscr{H}_J|$ means $200\ g^2\ (\Lambda^{-1},\ \Lambda^{-1},\ 1)$, etc. After calculating the products explicitly using Eqs. (B.31) - (B.33), one finds that this expression is less than the bound of Table IV for $|R_n|$. Hence the bound of Table IV for R_n holds for all n, by induction. The same bound holds for R because R is the limit of R_n for $n \to \infty$.

The bound on $R_n - R_{n-1}$ is also proven by induction. The bound on $R_1 - R_0$ is true because it is larger than the bound of Table IV for R_1. Then one computes a bound on $R_{n+1} - R_n$, given the bound for $R_n - R_{n-1}$ and using Eq. (A.5) of Appendix A. Since the bound for $|R_n - R_{n-1}|$ goes to zero as $n \to \infty$, the decomposition of R_n approaches a limit for $n \to \infty$; the limit defines a decomposition for R.

To get a bound for $\bar{H}$, one writes

$$\bar{H} = P\underset{\sim}{M}\cdot\underset{\sim}{V}_1\,(1-P)\,R + P\mathscr{H}_J(P+R) + Q\mathscr{H}_I(P+R)(P+\bar{Q}) + P\mathscr{H}_I(P+R)\,\bar{Q} \tag{B.50}$$

and sums the bounds of each term.

Now one can get bounds on m'' and g''. Let $\bar{H}$ have a decomposition $(\underset{\sim}{D}_k,\ F_k)$. Comparing Eq. (B.41) with Eq. (B.19), one gets

$$\underset{\sim}{G}_0 = P\underset{\sim}{M}P + \underset{\sim}{D}_0 \tag{B.51}$$

$$\underset{\sim}{G}_k = \underset{\sim}{D}_k \qquad (k > 0) \tag{B.52}$$

$$L_k = F_k \qquad \text{(all k)} \tag{B.53}$$

Explicit calculation using Table I gives

$$\underset{\approx}{\mathrm{PMP}} = \left(m, \sqrt{2}\, g\,(1-g^2)\, \tau_R^+, \sqrt{2}\, g\,(1-g^2)\, \tau_R^-\right) \tag{B.54}$$

From Table IV, $\underset{\approx}{D}_0$ has the bound (remember $\Lambda > 4 \times 10^6$)

$$\|\underset{\approx}{D}_0\| \le \underset{\approx}{m}\, g^2 \times 25000\ \Lambda^{-1} \le .01\, \underset{\approx}{m}\, g^2 \tag{B.55}$$

The bound on $\underset{\approx}{D}_0$ is a bound on the difference $\underset{\approx}{G}_0$ - $\underset{\approx}{\mathrm{PMP}}$. $\underset{\approx}{G}_0$ can be expressed in terms of m'' and g'' (Eq. (B.21); for $\tau^\pm$ in Eq. (B.21) read $\tau_R^\pm$ since in the present analysis we have not yet substituted $\tau^\pm$ for $\tau_R^\pm$). Using Eqs. (B.21), (B.54), (B.51), and (B.55), one gets the bounds

$$|m'' - m| \le .01\, m\, g^2 \tag{B.56}$$

$$|g'' - g\,(1-g^2)| \le .01\, g^3 \tag{B.57}$$

From these bounds one gets bounds on $\tan\theta'$ (Eq. (B.25))

$$(\sqrt{2}g/m)\,(1-1.01g^2)\,(1+.01g^2)^{-1} \le \tan\theta' \le (\sqrt{2}g/m)\,(1-.99g^2)\,(1-.01g^2)^{-1} \tag{B.58}$$

Using the bound $g^2 \le 1/2$, one can simplify these bounds; inserting $\sqrt{2}\, g = \sin\theta$, and $m = \cos\theta$, one gets

$$\tan\theta\,(1-.51\sin^2\theta) \le \tan\theta' \le \tan\theta\,(1-.48\sin^2\theta) \tag{B.59}$$

To complete the proof of Theorem 2, one notes that (cf. Eq. (B.29))

$$T_c(P_A) = -1 + L_0 = -1 + F_0 \tag{B.60}$$

From the bound on $\overline{\overline{H}}$, F_0 is less than $210\, g^2\, \Lambda^{-1}$ which is less than .01. Hence one obtains Eq. (V.30).

To prove Theorem 1, one must have bounds for $\|\underset{\sim}{A}'_k\|$ and C'_k in terms of m' and g'. One has bounds for $\|\underset{\sim}{G}_k\|$ and $\|L_k\|$ in terms of m and g (from Table IV and Eqs. (B.52) and (B.53)):

$$\|\underset{\sim}{G}_k\| \le \underset{\sim}{m}\, g^2 \times 40\, \Lambda^{-k} \qquad (k \ge 1) \tag{B.61}$$

$$\|L_k\| \le g^2 \times 40\, \Lambda^{-2k} \qquad (k \ge 1) \tag{B.62}$$

From Eq. (B.59) one has $\theta' < \theta$, and therefore m < m'. To get a bound on g in terms of g', one uses Eq. (B.59). Let

$$(1 - .51 \sin^2 \theta)^2 = 1 - \beta \tag{B.63}$$

Then

$$\begin{aligned} g^2/g'^2 &= \sin^2 \theta/\sin^2 \theta' = \sin^2 \theta \left\{1 + (\tan^2 \theta')^{-1}\right\} \\ &\le \sin^2 \theta \left\{1 + \cos^2 \theta\, (1-\beta)^{-1} (\sin^2 \theta)^{-1}\right\} \\ &= 1 + (1 - \sin^2 \theta)\, \beta (1-\beta)^{-1} \end{aligned} \tag{B.64}$$

The maximum value of β occurs for $\sin \theta = 1$ and is less than .8. Except for very small θ, $1-\beta$ is larger than $1 - \sin^2 \theta$ making g^2/g'^2 less than $1+\beta$. Hence

$$g^2 \le 1.8\, g'^2 \tag{B.65}$$

Hence

$$\underset{\sim}{m} \le (1.8)^{1/2}\, \underset{\sim}{m}' \tag{B.66}$$

(the inequality is true for each component of the two vectors). Also, from Eqs. (B.56) and (B.57) and $m^2 + 2g^2 = 1$ one gets a minimum value for $m''^2 + 2g''^2$, which in turn gives a bound

$$(m''^2 + 2 g''^2)^{-1/2} < 2.03 \tag{B.67}$$

The bounds (B.61), (B.62), (B.66), and (B.67), substituted in Eqs. (B.26) and (B.27), lead to the bounds

$$|\underset{\sim}{A}'_k| \leq 200\ \underset{\sim}{m}'\ g'^2\ \Lambda^{-k-1} \tag{B.68}$$

$$|C'_k| \leq 200\ g'^2\ \Lambda^{-2k-1} \tag{B.69}$$

To complete the proof of Theorem 1, one must show that $\underset{\sim}{A}'_k$ and C'_k satisfy hermiticity requirements and symmetry requirements with respect to charge conservation, charge conjugation and time reversal. The symmetry requirements are easily established since all the intermediate operators R_n, etc., have the same symmetries as H. One easily verifies that if X and Y are operators whose decompositions obey the symmetry requirements then the product XY has a decomposition obeying the symmetry requirements. The rest of the proof of symmetry is omitted. Hermiticity is more complicated because R_n and R are not hermitian, and one must use Eq. (IV.3) instead of Eq. (IV.19) to show that H_{eff} is hermitian. However a proof can still be constructed that A'_k and C'_k satisfy the hermiticity requirements of S_A. The basic result needed for the proof is that if XY has a decomposition $(\underset{\sim}{G}_k, L_k)$ then Y^+X^+ has the hermitian conjugate decomposition $(G_{k1} \rightarrow G^+_{k1}, G_{k2} \rightarrow G^+_{k3}, G_{k3} \rightarrow G^+_{k2}, L_k \rightarrow L^+_k)$. The proof is omitted.

Now Theorem 3 will be proven. If an operator X has a decomposition $(\underset{\sim}{A}_k, C_k)$, we will call the A_{k1} the "1-components" of X.

Note the following. Let operators X and Y have the decompositions $(\underset{\sim}{D}_k, F_k)$ and $(\underset{\sim}{A}_k, C_k)$ respectively. Let A_{k1} vanish for all k and D_{k1} vanish for $k > 0$, and let D_{01} be a c-number. Let the product XY have decomposition $(\underset{\sim}{G}_k, L_k)$. Then from Eqs. (B.14) and (B.16),

$$G_{01} = D_{01}\ C_0 \tag{B.70}$$

$$G_{k1} = D_{01} \left\{ C_k + \sum_{m=0}^{k-1} \underset{\sim}{T}_k \cdot \underset{\sim}{A}_m \right\} \qquad (k > 0) \tag{B.71}$$

Exactly the same formulae for G_{01} and G_{k1} result from decomposing the commuted product YX. This means that the commutator $[X, Y]$ has no 1-components in its decomposition.

If D_{01} is zero also then XY has no 1-components. Now consider Theorem 3. Let the 1-components of H_J vanish. We prove by induction that R_n has no 1-components. This is obviously true of R_0. Assume it is true of R_{n-1}. Consider Eq. (B.5). The operator $(1-P)\, H_I P$ has no 1-components because

$$(1-P)\, H_I P = (1-P)\, \underset{\sim}{M} \cdot \underset{\sim}{V}_1 P + (1-P)\, H_J P \quad ; \tag{B.72}$$

H_J has no 1-components by assumption and $(1-P)\, M_1 P = (1-P)\, mP$ vanishes. The product $R_{n-1}\, H_I R_{n-1}$ can be written $R_{n-1}\, PH_I\, (1-P)\, R_{n-1}$ so is a product of terms none of which contain 1-component. So this product has no 1-components. The remaining term in R_n can be written $(E_0 - H_0)^{-1}\, (1-P) \left[H_I, R_{n-1}\right] P$. The operators $X = H_I$, $Y = R_{n-1}$ satisfy the conditions noted above so the commutator $\left[H_I, R_{n-1}\right]$ has no 1-components. Hence R_n has no 1-components. It follows that R has no 1-components, nor do Q and $\bar{Q}$. Now consider H_{eff} (Eq. (B.13)). Using the fact that $Q = QP$, $\bar{Q} = P\bar{Q}$, $R = (1-P)\, R$, and that $PH_I\, (1-P)$ has no 1-components, one sees that the 1-components of H_{eff} are contained in $(P+Q)\, PH_I\, (P+\bar{Q})$. This can be written

$$PH_I\, (P+\bar{Q}) + PH_I Q\, (P+\bar{Q}) + \left[Q,\ PH_I\right] (P+\bar{Q})$$

The commutator has no 1-components by the argument noted above. The other terms can be written

$$PH_I\, (P+Q)\, (P+\bar{Q}) = PH_I\, P \tag{B.73}$$

since $(P+Q)\, (P+\bar{Q})$ is P. This means that the only 1-component in H_{eff} comes from $P\underset{\sim}{M} \cdot V_1\, P$. This means that in Eq. (B.20), G_{01} is m and G_{k1} vanishes for $k \geq 1$. This means that $m'' = m$ and $A'_{k1} = 0$, which is Theorem 3.

Finally Theorem 4 will be proven. This requires that two Hamiltonians, say H and H_A, be compared. Let H and H_A both be elements of S. Then for each of the operators $\mathscr{H}_I$, $\mathscr{H}_J$, R_n, etc., associated with H, there is a corresponding operator ($\mathscr{H}_{IA}$, $\mathscr{H}_{JA}$, R_{An}, etc.) associated with H_A. The decomposition of $\mathscr{H}_J$ is ($\underset{\sim}{A}_k$, C_k); the decomposition of $\mathscr{H}_{JA}$ is ($\underset{\sim}{A}_{kA}$, C_{kA}). One has

$$\mathscr{H}_I = \underset{\sim}{M} \cdot \underset{\sim}{V}_1 + \mathscr{H}_J \tag{B.74}$$

$$\mathscr{H}_{IA} = \underset{\sim}{M}_A \cdot \underset{\sim}{V}_1 + \mathscr{H}_{JA} \tag{B.75}$$

where

$$\underset{\sim}{M} = (m, \ \sqrt{2}\, g\, \tau^+, \ \sqrt{2}\, g\, \tau^-) \tag{B.76}$$

$$\underset{\sim}{M}_A = (m_A, \sqrt{2}\, g_A\, \tau^+, \ \sqrt{2}\, g_A\, \tau^-) \tag{B.77}$$

and

$$m^2 + 2g^2 = m_A^2 + 2g_A^2 = 1 \tag{B.78}$$

The assumptions of Theorem 4 are that

$$(m - m_A)^2 + 2\,(g - g_A)^2 \leq d_1^2 \tag{B.79}$$

$$\|\underset{\sim}{A}_{kA} - \underset{\sim}{A}_k\| \leq d_2\, \underset{\sim}{u}\, \Lambda^{-k-1} \qquad \text{(all k)} \tag{B.80}$$

$$\|C_{kA} - C_k\| \leq d_2\, \Lambda^{-2k-1} \qquad \text{(all k)} \tag{B.81}$$

where $\underset{\sim}{u}$ is the vector $(1/\sqrt{2})\ (1,\ 1,\ 1)$. The objective is to obtain bounds on $m_A' - m'$, $g_A' - g'$, $\underset{\sim}{A}_{kA}' - \underset{\sim}{A}_k'$ and $C_{kA}' - C_k'$, all in terms of d_1 and d_2. The bounds will be computed by the same techniques as in the proofs of Theorems 1 and 2. One change is that in defining the bound (a, b, c) of an operator X the vector $\underset{\sim}{u}$ is substituted in Eq. (B.30) for $\underset{\sim}{m}$. From Eq. (B.78) it follows that $\underset{\sim}{m} \leq \sqrt{2}\, \underset{\sim}{u}$, $\underset{\sim}{m}_A \leq \sqrt{2}\, \underset{\sim}{u}$, which means a bound (a, b, c) from Table IV (which implies the use of the vector $\underset{\sim}{m}$) can be changed into a bound using the vector $\underset{\sim}{u}$ simply by the replacement $a \rightarrow \sqrt{2}\, a$, $b \rightarrow \sqrt{2}\, b$, $c \rightarrow c$. The bounds for $\mathscr{H}_I$, etc., expressed in terms of $\underset{\sim}{u}$, also bound $\mathscr{H}_{IA}$, etc.

A problem arises in comparing A'_{kA} with A'_k. As part of the definition of $\underset{\sim}{A}'_k$ one took $\underset{\sim}{G}_{k+1}$ and replaced τ_R^+ by τ^+. The operator τ_R^+ is an operator in the full Hilbert space of mesons labelled 0 - k and the nucleons, although it is non-zero only in the subspace of the ground states $|P\rangle$ and $|N\rangle$ of H_0 plus mesons labelled 1 to k. The operator τ^+ acts in a separate space isomorphic to this subspace. Now when the operators $\underset{\sim}{A}_{kA}$ are calculated one starts from $\underset{\sim}{G}_{k+1A}$ expressed in terms of operators $\tau_{RA}^{\pm}$ which are <u>different</u> from $\tau_R^{\pm}$. This is because $\tau_{RA}^{\pm}$ are raising and lowering operators for a different pair of states $|PA\rangle$ and $|NA\rangle$ namely the ground states of H_{0A}. However in A'_k and A'_{kA} the same operators $\tau^{\pm}$ appear. Thus it will simplify matters to make a unitary transformation on $\underset{\sim}{G}_{k+1A}$ which takes $\tau_{RA}^{\pm}$ into $\tau_R^{\pm}$; after this has been done one can make comparisons in the full space of 0 - N mesons plus nucleons instead of the separate space involving $\tau^{\pm}$ plus 1 - N mesons. Let the unitary transformation be U_A. One wants U_A^+ to take eigenstates of H_{0A} into eigenstates of H_A. In particular if P_A projects the ground states of H_{0A}, one wants

$$U_A^+ P_A U_A = P \tag{B.82}$$

Then one replaces $\underset{\sim}{G}_{kA}$ by $U_A^+ \underset{\sim}{G}_{kA} U_A$ before comparing with $\underset{\sim}{G}_k$, and likewise for L_{kA}.

One can take Eqs. (B.4) to (B.13), replace R_n by R_{An}, etc., and then transform them all by $U_A^+ \cdot\cdot\, U_A$. Note that $\mathscr{H}_{0A}$ and $\mathscr{H}_0$ (cf. Eq. (B.45) have the same eigenvalues (0, 1, and 2) (cf. Table I) so $U_A^+ \mathscr{H}_{0A} U_A = \mathscr{H}_0$. From now on let R_{An} stand for what was $U_A^+ R_{An} U_A$, and likewise for R_A, Q_{An}, Q_A, $\bar{Q}_{An}$, $\bar{Q}_A$, $\bar{\bar{H}}_A$

(cf. Eq. (B.41)) and H_{Aeff}. However, $\mathscr{H}_{IA}$ and $\mathscr{H}_{JA}$ will still be the untransformed operators; denote $U_A^+ \mathscr{H}_{IA} U_A$ by $\mathscr{H}_{IA}'''$ and $U_A^+ \mathscr{H}_{JA} U_A$ by $\mathscr{H}_{JA}'''$. The equations for R_{An}, Q_{An}, etc. are now obtained from Eqs. (B.4) - (B.13) by replacing H_I by $\mathscr{H}_{IA}'''$ and by inserting an overall scale factor J_A in the formula for H_{Aeff}.

Now define the following differences

$$\mathscr{H}_{Ia} = \mathscr{H}_{IA}''' - \mathscr{H}_I$$

$$\mathscr{H}_{Ja} = \mathscr{H}_{JA}''' - \mathscr{H}_J$$

$$\underset{\sim}{M}_a = \underset{\sim}{M}_A''' - \underset{\sim}{M}$$

$$R_{an} = R_{An} - R_n \text{, etc.} \tag{B.83}$$

where

$$\underset{\sim}{M}_A''' = U_A^+ \underset{\sim}{M}_A U_A \tag{B.84}$$

One can write equations for the differences R_{an}, etc., as follows:

$$R_{a0} = 0 \tag{B.85}$$

$$R_{an} = (-\mathscr{H}_0)^{-1} (1-P) \Big\{ (1-R_{An-1}) \mathscr{H}_{Ia} (P+R_{An-1})$$

$$- R_{an-1} \mathscr{H}_I (P+R_{An-1}) + (1-R_{n-1}) \mathscr{H}_I R_{an-1} \Big\} \qquad (n > 0) \tag{B.86}$$

$$Q_{an} = (1/2) \Big\{ R_A^+ R_a + R_a^+ R - Q_{An-1} Q_{an-1} - Q_{an-1} Q_{n-1} \Big\} \qquad (n > 0) \tag{B.87}$$

$$\bar{Q}_{an} = - Q_a - Q_A \bar{Q}_{an-1} - Q_a \bar{Q}_{n-1} \tag{B.88}$$

$$\bar{H}_a = P\,\mathscr{H}_{Ja}\,P + Q_A\,\mathscr{H}_{Ia}\,(P+R_A)\,(P+\bar{Q}_A) + Q_a\,\mathscr{H}_I\,(P+R_A)\,(P+\bar{Q}_A)$$

$$+ Q\,\mathscr{H}_I\,R_a\,(P+\bar{Q}_A) + Q\,\mathscr{H}_I\,(P+R)\,\bar{Q}_a + P\,\mathscr{H}_{Ia}\,(R_A+\bar{Q}_A+R_A\bar{Q}_A)$$

$$+ P\,\mathscr{H}_I\,(R_a+\bar{Q}_a+R_a\,\bar{Q}_A+R\,\bar{Q}_a) \tag{B.89}$$

Knowing upper bounds for $\bar{H}_a$, one easily obtains upper bounds for $\left\{U_A^+\,\underset{\sim}{G}_{kA}\,U_A - \underset{\sim}{G}_k\right\}$ and $\left\{U_A^+\,L_{kA}\,U_A - L_k\right\}$.

The first step in deriving upper bounds is to get upper bounds for $\mathscr{H}_{Ja}$ and $\underset{\sim}{M}_a$. One has

$$\underset{\sim}{M}_a = \left(m_A - m,\ \sqrt{2}\,g_A\,U_A^+\,\tau^+\,U_A - \sqrt{2}\,g\,\tau^+,\ \sqrt{2}\,g_A\,U_A^+\,\tau^-\,U_A - \sqrt{2}\,g\,\tau^-\right) \tag{B.90}$$

Write $U_A = 1 + V_A$; then

$$g_A\,U_A^+\,\tau^+\,U_A - g\,\tau^+ = (g_A - g)\,U_A^+\,\tau^+\,U_A + g\,V_A^+\,\tau^+\,U_A + g\,\tau^+\,V_A \tag{B.91}$$

So one needs a bound for V_A. The operator U_A is

$$U_A = \sum_{n=1}^{8} |n\rangle_A\,\langle n| \tag{B.92}$$

where $|n\rangle$ $(1 \le n \le 8)$ are the eigenstates of H_0 and $|n\rangle_A$ are the eigenstates of H_{0A}. These are known explicitly from Table I. An upper bound for $\|V_A\|^2$ is obtained by computing the trace of $V_A^+ V_A$. The trace is

$$\mathrm{Tr}\ V_A^+ V_A = \mathrm{Tr}\left\{2 - U_A - U_A^+\right\} = \sum_{n=1}^{8}\left\{2 - \langle n|n_A\rangle - \langle n_A|n\rangle\right\} \tag{B.93}$$

In fact one can compute the trace separately for states of a given charge; the maximum of these traces is still greater than $\|V_A\|^2$. The traces are

$$\text{charge} = 2 \text{ or } -1: \qquad \mathrm{Tr}\ V_A^+ V_A = 0 \tag{B.94}$$

$$\text{charge} = 0 \text{ or } 1: \qquad \mathrm{Tr}\ V_A^+ V_A = 2\left\{(m - m_A)^2 + 2\,(g - g_A)^2\right\} \tag{B.95}$$

The latter result was obtained using Table I and Eq. (B.78). From this it follows that

$$\|V_A\| \leq \sqrt{2}\, d_1 \tag{B.96}$$

Also $|m - m_A|$ and $\sqrt{2}|g - g_A|$ are less than d_1 (from Eq. (B.79)); $\sqrt{2}\, g$ and $\sqrt{2}\, g_A$ are less than 1; and $\|U_A\| = 1$. Using these results in Eq. (B.90) gives

$$\|\underset{\sim}{M}_a\| < 5.5\, d_1\, \underset{\sim}{u} \tag{B.97}$$

A similar calculation for $\mathscr{H}_{Ja}$, using the bound of Table IV for $\mathscr{H}_J$ and $\mathscr{H}_{JA}$, plus the bounds $\underset{\sim}{m} < \sqrt{2}\, \underset{\sim}{u}$, $\underset{\sim}{m}_A < \sqrt{2}\, \underset{\sim}{u}$, gives the bound shown in Table V. One can now obtain bounds on R_{an}, R_a, etc., using Eqs. (B.85) - (B.89). One uses Eqs. (B.31) - (B.33) to obtain bounds on products (with $\underset{\sim}{u}$ replacing $\underset{\sim}{m}$ in the definition of the bound (a, b, c)). The results are shown in Table V.

Write the decomposition of $\overline{H}_a$ as

$$\overline{H}_a = \sum_{k=1}^{N} \underset{\sim}{V}_k \cdot D_{ak-1} + \sum_{k=0}^{N} F_{ak} \tag{B.98}$$

The bound of Table V for $\overline{H}_a$ gives the following bounds:

$$\begin{aligned} \|\underset{\sim}{D}_{a0}\| &\leq \left(7200\, \Lambda^{-1} d_1 + 16\, \Lambda^{-1} d_2\right) \underset{\sim}{u} \\ \|\underset{\sim}{D}_{ak}\| &\leq \left(230\, d_1 + 27000\, \Lambda^{-1} d_2\right) \underset{\sim}{u}\, \Lambda^{-k} \qquad (k > 0) \\ \|F_{ak}\| &\leq \left(230\, d_1 + 27000\, \Lambda^{-1} d_2\right) \Lambda^{-2k} \qquad (k > 0) \end{aligned} \tag{B.99}$$

Consider the significance of $\underset{\sim}{D}_{a0}$. It is a difference $\underset{\sim}{D}_{A0} - \underset{\sim}{D}_0$. From Eqs. (B.51), (B.54), and (B.21)(one must substitute $\tau_R^{\pm}$ for $\tau^{\pm}$ in Eq. (B.21)), $\underset{\sim}{D}_0$ itself is

$$\underset{\sim}{D}_0 = \left(m'' - m,\ \sqrt{2}\left[g'' - g\ \ (1 - g^2)\right] \tau_R^{+},\ \sqrt{2}\left[g'' - g\,(1 - g^2)\right] \tau_R^{-}\right) \tag{B.100}$$

Correspondingly

$$\underset{\sim}{D}_{A0} = \left(m''_A - m_A,\ \sqrt{2}\left[g''_A - g_A\left(1 - g_A^2\right)\right]\tau_R^+,\ \sqrt{2}\left[g''_A - g_A\left(1 - g_A^2\right)\right]\tau_R^-\right) \tag{B.101}$$

So $\underset{\sim}{D}_{a0}$ involves differences such as $(m''_A - m_A) - (m'' - m)$.

We can use the bound on $\underset{\sim}{D}_{a0}$ to prove the inequalities of Eq. (V.37) (the first inequality of Theorem 4). In the notation of this appendix the quantity d'_1 is defined as

$$(d'_1)^2 = (m'_A - m')^2 + 2\,(g'_A - g')^2 \tag{B.102}$$

where

$$m' = \cos\theta' = m''/(m''^2 + 2g''^2)^{1/2} \tag{B.103}$$

$$g' = \sin\theta' = \sqrt{2}\,g''/(m''^2 + 2g''^2)^{1/2} \tag{B.104}$$

and analogous formulae hold for m'_A and g'_A; θ' is the angle in the decomposition of T(H) (cf. Eq. (B.25)).

To get bounds on d'_1 requires some further manipulations which are most conveniently done with another set of vectors. Define the following two-dimensional vectors:

$$\underset{\sim}{x} = \left(m,\ \sqrt{2}\,g\left[1 - g^2\right]\right) \tag{B.105}$$

$$\underset{\sim}{x}'' = (m'',\ \sqrt{2}\,g'') \tag{B.106}$$

$$\underset{\sim}{x}' = (m',\ \sqrt{2}\,g') \tag{B.107}$$

and analogously $\underset{\sim}{x}_A$, $\underset{\sim}{x}''_A$, and $\underset{\sim}{x}'_A$. Define

$$\hat{\underset{\sim}{x}} = |\underset{\sim}{x}|^{-1}\,\underset{\sim}{x}, \quad \text{etc.} \tag{B.108}$$

Then

$$\underset{\sim}{x}' = \hat{x}'', \qquad \underset{\sim}{x}'_A = \hat{x}''_A \tag{B.109}$$

Now one has

$$d_1' = |\underset{\sim}{x}_A' - \underset{\sim}{x}'| = |\hat{\underset{\sim}{x}}_A'' - \hat{\underset{\sim}{x}}''| \tag{B.110}$$

The bound d_1' will be computed in two parts, first relating $\hat{x}_A'' - \hat{x}''$ to $\hat{x}_A - \hat{x}$ and then bounding $\hat{x}_A - \hat{x}$. Write

$$\phi = |\hat{x}_A'' - \hat{x}'' - \hat{x}_A + \hat{x}| \tag{B.111}$$

$$\psi = |\hat{x}_A - \hat{x}| \tag{B.112}$$

Then

$$\psi - \phi \le d_1' \le \psi + \phi \tag{B.113}$$

To compute ψ, it is convenient to let

$$\begin{aligned} \hat{x} &= (\cos\ \omega,\ \sin \omega) \\ \hat{x}_A &= (\cos\ \omega_A,\ \sin\ \omega_A) \end{aligned} \tag{B.114}$$

and $m = \cos\ \theta$, $\sqrt{2}\, g = \sin\ \theta$, $m_A = \cos\ \theta_A$, $\sqrt{2}\, g = \sin\ \theta_A$. Then[19]

$$\omega = \tan^{-1}\left[\tan\ \theta\ (1 - 1/2 \sin^2\ \theta)\right] = f(\theta) \tag{B.115}$$

and ω_A is $f(\theta_A)$. The derivative $f'(\theta)$ has the form

$$f'(\theta) = N(y)/D(y) \tag{B.116}$$

where

$$y = \sin^2\ \theta \tag{B.117}$$

$$N(y) = (1 - 1/2\ y - y\ (1-y)) \tag{B.118}$$

$$D(y) = (1 - y^2 + 1/4\ y^3) \tag{B.119}$$

Analyzing the form for $f'(\theta)$ one sees that the numerator N decreases for $0 < y < 3/4$ and increases for $3/4 < y$, the denominator D decreases over the whole range $0 < y < 1$. So one has the following bounds:

$$N(.75)/D(0) < f'(\theta) < \mathrm{Max}\left\{N(0)/D(.75),\ N(1)/D(1)\right\} \tag{B.120}$$

Evaluated, this gives

$$.4375 < f'(\theta) < 2 \tag{B.121}$$

Hence by the mean value theorem

$$.436\,|\theta_A - \theta| \le |\omega_A - \omega| \le 2\,|\theta_A - \theta| \tag{B.122}$$

Now the definitions of d_1 and ψ are equivalent to

$$d_1 = |2 \sin 1/2\,(\theta_A - \theta)| \tag{B.123}$$

$$\psi = |2 \sin 1/2\,(\omega_A - \omega)| \tag{B.124}$$

One can show that

$$(\sin az) > a \sin z \tag{B.125}$$

when $0 < z < \pi/2$ and $0 < a < 1$. The result of Eqs. (B.124), (B.122), (B.125) and then (B.123), is

$$\psi \ge |2 \sin .218\,(\theta_A - \theta)| \ge .436\, d_1 \qquad (z = .5\,(\theta_A - \theta)) \tag{B.126}$$

$$\psi \le |2 \sin (\theta_A - \theta)| \le 2\, d_1 \qquad (az = .5\,(\theta_A - \theta)) \tag{B.127}$$

The next step is to bound ϕ. It is convenient to define

$$\underset{\sim}{g}(\underset{\sim}{x}) = \hat{\underset{\sim}{x}} \tag{B.128}$$

Then

$$\phi \le \phi_1 + \phi_2 \tag{B.129}$$

where

$$\phi_1 = |\underset{\sim}{g}(\underset{\sim}{x}''_A) - \underset{\sim}{g}(\underset{\sim}{x}_A + \delta\underset{\sim}{x})| \tag{B.130}$$

$$\phi_2 = |\underset{\sim}{g}(\underset{\sim}{x}_A + \delta\underset{\sim}{x}) - \underset{\sim}{g}(\underset{\sim}{x} + \delta\underset{\sim}{x}) - \underset{\sim}{g}(\underset{\sim}{x}_A) + \underset{\sim}{g}(\underset{\sim}{x})| \tag{B.131}$$

and

$$\delta\underset{\sim}{x} = \underset{\sim}{x}'' - \underset{\sim}{x} \tag{B.132}$$

Now by the mean value theorem

$$\phi_1 \leq \text{Max}\ (0 \leq \lambda \leq 1) \left| \delta \underset{\sim}{x}_a \cdot \nabla \underset{\sim}{g} \left(\underset{\sim}{x}''_A - \lambda\, \delta \underset{\sim}{x}_a \right) \right| \tag{B.133}$$

where

$$\delta \underset{\sim}{x}_a = \underset{\sim}{x}''_A - \underset{\sim}{x}'' - \underset{\sim}{x}_A + \underset{\sim}{x} \tag{B.134}$$

and by a second order mean value theorem

$$\phi_2 \leq \text{Max}\ (0 \leq \lambda \leq 1,\ 0 \leq \mu \leq 1) \left| (\delta \underset{\sim}{x} \cdot \nabla)(\underset{\sim}{x}_a \cdot \nabla)\, \underset{\sim}{g}(\underset{\sim}{x} + \lambda \delta \underset{\sim}{x} + \mu \underset{\sim}{x}_a) \right| \tag{B.135}$$

where

$$\underset{\sim}{x}_a = \underset{\sim}{x}_A - \underset{\sim}{x} \tag{B.136}$$

(and $\delta \underset{\sim}{x} \cdot \nabla$ acts on $\underset{\sim}{g}$, not on $\underset{\sim}{x}_a \cdot \nabla$). From Eqs. (B.56), (B.57), (B.78), (B.105) and (B.106),

$$|\delta \underset{\sim}{x}| \leq .005 \tag{B.137}$$

and from Eq. (B.99)

$$|\delta \underset{\sim}{x}_a| \leq \left(7200\, \Lambda^{-1} d_1 + 16\, \Lambda^{-1} d_2 \right) \tag{B.138}$$

Furthermore

$$|\underset{\sim}{x}_a|^2 = (m_A - m)^2 + 2 (g_A - g)^2 (1 - g_A^2 - g_A g - g^2)^2 \tag{B.139}$$

and since g_A and g are less than $1/\sqrt{2}$ one has

$$-1/2 < \left(1 - g_A^2 - g_A g - g^2 \right) < 1 \tag{B.140}$$

So

$$|\underset{\sim}{x}_a|^2 < d_1^2 \tag{B.141}$$

Now let $\underset{\sim}{y}$ be an arbitrary vector; one can most easily compute $\delta \underset{\sim}{x}_a \cdot \nabla \underset{\sim}{g}(\underset{\sim}{y})$ and $(\delta \underset{\sim}{x} \cdot \nabla)(\underset{\sim}{x}_a \cdot \nabla)\, \underset{\sim}{g}(\underset{\sim}{y})$ using a coordinate system with the first axis parallel to $\underset{\sim}{y}$. If $\delta x_\parallel$ and $\delta x_\perp$ are the components of $\delta \underset{\sim}{x}$ parallel and perpendicular to y (and likewise for $x_{a\parallel}$, etc.), one has

$$\delta \underset{\sim}{x}_a \cdot \nabla\, \underset{\sim}{g}(\underset{\sim}{y}) = \left(0,\ \delta x_{a\perp} |\underset{\sim}{y}|^{-1} \right) \tag{B.142}$$

and

$$(\delta \underset{\sim}{x} \cdot \nabla)(\underset{\sim}{x}_a \cdot \nabla)\, \underset{\sim}{g}(\underset{\sim}{y}) = \left(-\delta x_\perp\ x_{a\perp} |\underset{\sim}{y}|^{-2},\ -\left(\delta x_\parallel\ x_{a\perp} + \delta x_\perp\ x_{a\parallel}\right) |\underset{\sim}{y}|^{-2}\right) \tag{B.143}$$

In absolute value

$$|(\delta \underset{\sim}{x}_a \cdot \nabla)\ \underset{\sim}{g}(\underset{\sim}{y})| \leq |\delta \underset{\sim}{x}_a|\ |\underset{\sim}{y}|^{-1} \tag{B.144}$$

$$|(\delta \underset{\sim}{x} \cdot \nabla)(\underset{\sim}{x}_a \cdot \nabla)\ \underset{\sim}{g}(\underset{\sim}{y})| \leq (2/\sqrt{3})\ |\delta \underset{\sim}{x}|\,|\underset{\sim}{x}_a|\,|\underset{\sim}{y}|^{-2} \tag{B.145}$$

The second inequality is proved using the relation

$$2\, x_{a\parallel}\ x_{a\perp}\, \delta x_\parallel\, \delta x_\perp < (4/3)\,(\delta x_\parallel\ x_{a\parallel})^2 + (1/3)\left[(\delta x_\parallel\ x_{a\perp})^2 + (\delta x_\perp\, x_{a\parallel})^2 + (\delta x_\perp\, x_{a\perp})^2\right] \tag{B.146}$$

To use the bound (B.144) to obtain a bound for ϕ_1 (cf. Eq. (B.133)), one puts $\underset{\sim}{y} = \underset{\sim}{x}''_A - \lambda\, \delta \underset{\sim}{x}_a$; hence

$$|\underset{\sim}{y}|^{-1} \leq \left\{|\underset{\sim}{x}''_A| - |\delta \underset{\sim}{x}_a|\right\}^{-1} \tag{B.147}$$

Now (from Eq. (B.137) and the analogous bound for $|\underset{\sim}{x}''_A - \underset{\sim}{x}_A|$

$$|\delta \underset{\sim}{x}_a| \leq |\delta \underset{\sim}{x}| + |\underset{\sim}{x}''_A - \underset{\sim}{x}_A| \leq .01 \tag{B.148}$$

and $|\underset{\sim}{x}''_A| > .49$ from Eq. (B.67) (which holds for $|\underset{\sim}{x}''_A|$ as well as $|\underset{\sim}{x}''|$). Thus (remember that $\Lambda > 4 \times 10^6$)

$$\phi_1 \leq 2.1\left(7200\, \Lambda^{-1}\, d_1 + 16\, \Lambda^{-1}\, d_2\right) \leq .004\, d_1 + 10^{-5}\, d_2 \tag{B.149}$$

To get a bound for ϕ_2, one uses Eq. (B.145) with $\underset{\sim}{y}$ being $(1-\mu)\, \underset{\sim}{x} + \mu\, \underset{\sim}{x}_A + \lambda \delta \underset{\sim}{x}$. Since $\underset{\sim}{x} \cdot \underset{\sim}{x}_A \geq 0$ and since μ and $1-\mu$ are non-negative,

$$|(1-\mu)\, \underset{\sim}{x} + \mu\, \underset{\sim}{x}_A|^2 \geq (1-\mu)^2\, |\underset{\sim}{x}|^2 + \mu^2 |\underset{\sim}{x}_A|^2 \tag{B.150}$$

But $|\underset{\sim}{x}|$ and $|\underset{\sim}{x}_A|$, and $(1-\mu)^2 + \mu^2$, are all larger than or equal to 1/2. So

$$|(1-\mu)\, \underset{\sim}{x} + \mu\, \underset{\sim}{x}_A| \geq 1/\sqrt{8} \tag{B.151}$$

Hence

$$|\underset{\sim}{y}| \geq (1/\sqrt{8}) - .005 > 1/3 \tag{B.152}$$

Hence from Eqs. (B.135), (B.145), (B.152), (B.137), and (B.141),

$$\phi_2 \le 6\sqrt{3}\,|\delta \underset{\sim}{x}|\,|\underset{\sim}{x}_a| \le .052\, d_1 \tag{B.153}$$

From Eqs. (B.113), (B.126), (B.127), (B.129), (B.149), and (B.153),

$$.38\, d_1 - 10^{-5}\, d_2 < d_1' < 20\, d_1 + 10^{-5}\, d_2 \tag{B.154}$$

which is the first inequality of Theorem 4.

To obtain the second inequality of Theorem 4, one starts from Eqs. (B.26), (B.27), (B.52), (B.53), (B.106), and the corresponding equations for A_{kA}', etc., from which one can obtain

$$A_{kA}' - A_k' = \left\{ |\underset{\sim}{x}_A''|^{-1} - |\underset{\sim}{x}''|^{-1} \right\} D_{Ak+1} + |\underset{\sim}{x}''|^{-1}\, D_{ak+1} \tag{B.155}$$

$$C_{kA}' - C_k' = \Lambda \left\{ |\underset{\sim}{x}_A''|^{-1} - |\underset{\sim}{x}''|^{-1} \right\} F_{Ak+1} - \Lambda |\underset{\sim}{x}''|^{-1}\, F_{ak+1} \tag{B.156}$$

Now, from Eqs. (B.67), (B.134), (B.136), (B.138), and (B.141)

$$|\underset{\sim}{x}_A''|^{-1} - |\underset{\sim}{x}''|^{-1} \le |\underset{\sim}{x}_A'' - \underset{\sim}{x}''|\; |\underset{\sim}{x}_A''|^{-1}\, |\underset{\sim}{x}''|^{-1} \le 4.07\, |\delta \underset{\sim}{x}_a + \underset{\sim}{x}_a|$$

$$\le 4.07 \left(d_1 + 7200\, \Lambda^{-1}\, d_1 + 16\, \Lambda^{-1}\, d_2 \right) \tag{B.157}$$

From Table V, $\underset{\sim}{D}_{Ak+1}$ and F_{Ak+1} have bounds

$$\|\underset{\sim}{D}_{Ak+1}\| \le 40\, \underset{\sim}{u}\, \Lambda^{-k-1} \tag{B.158}$$

$$\|F_{Ak+1}\| \le 40\, \Lambda^{-2k-2} \tag{B.159}$$

From Eqs. (B.155) - (B.159), (B.67), and (B.99) one gets

$$\|A_{kA}' - A_k'\| < \left\{ 1100\, d_1 + .06\, d_2 \right\} \underset{\sim}{u}\, \Lambda^{-k-1} \tag{B.160}$$

$$\|C_{kA}' - C_k'\| \le \left\{ 1100\, d_1 + .06\, d_2 \right\} \Lambda^{-2k-1} \tag{B.161}$$

which proves the second inequality of Theorem 4.

TABLE I

Eigenstates of the Hamiltonian $m(a^+a + b^+b - 1) + g(a+b^+)\tau^+ + g(a^+ + b)\tau^-$, where a^+ creates π^+, b^+ creates π^-, $|p\rangle$ and $|n\rangle$ are nucleon states, and $\mu = m(m^2 + 2g^2)^{-1/2}$, $\gamma = g(m^2 + 2g^2)^{-1/2}$. The other four eigenstates are obtained by charge conjugation ($p \leftrightarrow n$, $\pi^+ \leftrightarrow \pi^-$).

Eigenvalue	Eigenstate
$-(m^2 + 2g^2)^{1/2}$	$1/2\,(1+\mu)\,|p\rangle - \gamma\,|n\pi^+\rangle + 1/2(1-\mu)\,|p\pi^+\pi^-\rangle$
0	$|p\pi^+\rangle$
0	$\gamma\,|p\rangle + \mu\,|n\pi^+\rangle - \gamma\,|p\pi^+\pi^-\rangle$
$(m^2 + 2g^2)^{1/2}$	$1/2\,(1-\mu)\,|p\rangle + \gamma\,|n\pi^+\rangle + 1/2\,(1+\mu)\,|p\pi^+\pi^-\rangle$

TABLE II

Breakdown of H_{eff} by type of operator for each order in Λ. The symbols x_m, τ_R, and $(x_m)^2$ are explained in the text. Any operator listed for a power Λ^m can occur for lower powers of Λ also.

Order in Λ	Types of Operators
Λ^M	constant
Λ^{M-1}	$x_{M-1}\,\tau_R$
Λ^{M-2}	$x_{M-2}\,\tau_R$, $(x_{M-1})^2\,\tau_R$
Λ^{M-3}	$x_{M-3}\,\tau_R$, $x_{M-2}\,x_{M-1}\,\tau_R$
Λ^{M-4}	$x_{M-4}\,\tau_R$, $x_{M-3}\,x_{M-1}\,\tau_R$, $(x_{M-2})^2\,\tau_R$, $x_{M-2}(x_{M-1})^2\,\tau_R$

TABLE III

Breakdown of H'_{eff} by type of operator for each order in Λ. Cf. Table II.

Order in Λ	Types of Operators
Λ^M, Λ^{M-1}	constant
Λ^{M-2}	$x_{M-2}\,\tau'_R$
Λ^{M-3}	$x_{M-3}\,\tau'_R$, $(x_{M-2})^2\,\tau'_R$
Λ^{M-4}	$x_{M-4}\,\tau'_R$, $x_{M-3}\,x_{M-2}\,\tau'_R$

TABLE IV

Operator bounds obtained in the proofs of Theorems 1 and 2, assuming $\Lambda > 4 \times 10^6$

Operator	Bound	Operator	Bound
$\mathscr{H}_I$	$(1.5, .5, 100)$	$Q_n - Q_{n-1}$	$13\, g^2 \times 10^{4-n}(\Lambda^{-1}, 1, \Lambda^{-1})$
$\mathscr{H}_J$	$200\, g^2 (\Lambda^{-1}, \Lambda^{-1}, 1)$	$\overline{Q}_n, \overline{Q}$	$14\, g^2 \times 10^3(\Lambda^{-1}, 1, \Lambda^{-1})$
R_n, R	$g\,(2, 65, 160)$	$\overline{Q}_n - \overline{Q}_{n-1}$	$14\, g^2 \times 10^{4-n}(\Lambda^{-1}, 1, \Lambda^{-1})$
$R_n - R_{n-1}$	$16\, g \times 10^{2-n}(1, \Lambda, 1)$	$\overline{H}$	$g^2 (25 \times 10^3 \Lambda^{-1}, 40, 210)$
Q_n, Q	$13\, g^2 \times 10^3(\Lambda^{-1}, 1, \Lambda^{-1})$		

TABLE V

Bounds on operators needed in the proof of Theorem 4. The bound (d, e, f) is defined by Eq. (B. 30) except that the vector $\underset{\sim}{u} = 1/\sqrt{2}\ (1, 1, 1)$ replaces the vector $\underset{\sim}{m}$.

Operator	Bound
$\underset{\sim}{M}_a \cdot \underset{\sim}{V}_1$	$d_1 (5.5, 0, 0)$
V_A	$\sqrt{2}\, d_1$
$\mathscr{H}_{Ja}$	$d_1 (450\Lambda^{-1}, 450\Lambda^{-1}, 300) + d_2 (\Lambda^{-1}, \Lambda^{-1}, 1)$
R_{an}, R_a	$d_1 (8, 800, 1600) + d_2 (30\Lambda^{-1}, 2200\Lambda^{-1}, 1.3)$
Q_{an}, Q_a	$d_1 (3.25 \times 10^5 \Lambda^{-1}, 245, 3.25 \times 10^5 \Lambda^{-1}) + d_2 (7.9\Lambda^{-1}, 1100\Lambda^{-1}, 2200\Lambda^{-1})$
$\overline{Q}_{an}, \overline{Q}_a$	$d_1 (3.3 \times 10^5 \Lambda^{-1}, 250, 3.3 \times 10^5 \Lambda^{-1}) + d_2 (8\Lambda^{-1}, 1200\Lambda^{-1}, 2400\Lambda^{-1})$
$\overline{H}_a$	$d_1 (7200\Lambda^{-1}, 230, 310) + d_2 (16\Lambda^{-1}, 27000\Lambda^{-1}, 1.1)$
$\overline{H}, \overline{H}_A$	$(20000\Lambda^{-1}, 40, 120)$

FIGURE CAPTION

1. Artist's conception of the trajectories C(3), C(7), C_A, and C_B projected on a two-dimensional space. The renormalized coupling constant is $\pi/4$. The curve R is also shown. The first few points on C_A, C_B, C(3) and C(7) are labelled explicitly: P_U is the first point on C_A, P_{R0} the first point on C_B.

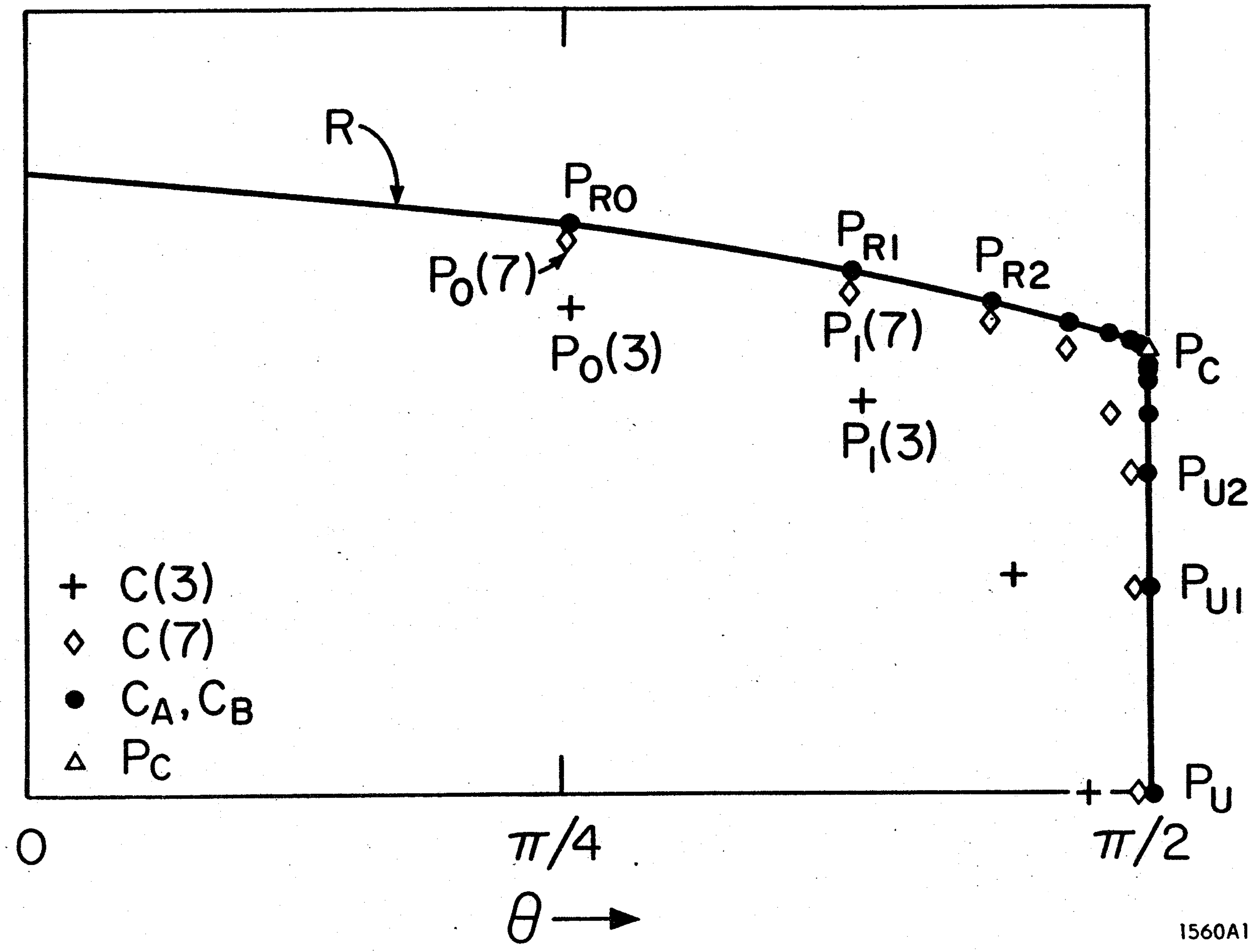
R
PR0
PR1
PR2
P0(7)
P0(3)
P1(7)
P1(3)
PC
PU2
PU1
PU
C(3)
C(7)
CA, CB
PC
0
π/4
π/2
θ
1560A1

Fig. 1