Commun. Math. Phys. 104, 177–206 (1986)

Communications in Mathematical Physics

Volume Dependence of the Energy Spectrum in Massive Quantum Field Theories

I. Stable Particle States

M. Lüscher

Theory Division, Deutsches Elektronen-Synchrotron DESY, D-2000 Hamburg 52
Federal Republic of Germany

Abstract. Due to polarization effects, the mass M of a stable particle in a quantum field theory enclosed in a large (space-like) box of size L and periodic boundary conditions in general differs from its infinite volume value m. As L increases, the finite size mass shift $\Delta m = M - m$ goes to zero exponentially with a rate, which depends on the particle considered and on the spectrum of light particles in the theory. This behaviour follows from an apparently universal asymptotic formula, already presented earlier, which relates Δm to certain forward elastic scattering amplitudes. A detailed proof of this basic relation is given here to all orders of perturbation theory in arbitrary massive quantum field theories.

1. Introduction

From experience with large scale numerical simulations of lattice gauge theories over the past few years, it has become plausible that with this method a reliable calculation of the hadron masses and other low energy parameters in QCD will ultimately be possible. Because of the limited capacity of today's computer systems, the lattices one can simulate are however rather small so that, for example, a lattice of size $L = 5$ fermi and spacing $a = (2\ \mathrm{GeV})^{-1}$ would already be considered huge by present standards. Hadrons contained in such small volumes occupy a significant fraction of the available space and one therefore expects that the calculated masses show some dependence on L. Thus, for the correct interpretation of the data obtained from Monte Carlo simulations, a theoretical understanding of these finite size effects is needed and studies with variable L must be made to check the theoretical expectations.

Finite volume effects are also interesting in their own right and their investigation may prove useful for purposes other than merely controlling a systematic error source. The reason for this is that they probe the system at distances large compared to the lattice spacing. In general, they are therefore universal (i.e. independent of the form and magnitude of the ultra-violet cutoff) and often contain useful information on the infinite volume system. In statistical

mechanics this observation has long been converted into a powerful tool for the numerical calculation of critical exponents at second order phase transitions, for example (see ref. [1] for a review and ref. [2] for a recent paper in this field). More recently, the finite size method has also been applied to asymptotically free field theories in an attempt to calculate the spectrum of the low-lying stable particles analytically [3, 4, 5]. When combined with data obtained from Monte Carlo simulations, this approach may result in a significant test of scaling and a determination of the Λ-parameter, which is free of extrapolation ambiguities [6, 7, 8]. Finally, in the present work the volume dependence of energy values is related to scattering amplitudes, which makes it possible, in certain cases, to compute 3-particle on-shell coupling constants [9, 10] and scattering lengths [11] by numerical simulation.

In this and the following paper [11], massive quantum field theories enclosed in an $L \times L \times L$ box with periodic boundary conditions are considered (time remains unrestricted). Due to the finite volume, the spectrum of the Hamilton operator (i.e. of the transfer matrix in lattice theories) is then discrete and the corresponding energy values depend on L in a way, which apart from some gross features, is dynamically determined. For large volumes one expects that there are distinguished zero momentum eigenstates of the Hamilton operator, which can be interpreted as states of a single stable particle at rest. The corresponding energy values $M_i(L)$ (i labels the different particles) are close to the rest masses m_i of these particles as defined at $L = \infty$. This paper is devoted to the question of how exactly $M_i(L)$ approaches m_i in the limit where $L \to \infty$ and all other parameters of the theory are kept fixed. Note that this limit is different from the usual finite size scaling limit, where L is made large but the parameters in the Lagrangian are tuned in such a way that L/ξ is fixed, where ξ is a correlation length.

The physical origin of the size dependence of the mass M_i of a "pointlike" stable particle is that such particles polarize the vacuum around them, i.e. they are accompanied by a cloud of virtual particles. In the simplest case, the diameter of the cloud is roughly equal to the Compton wave length λ_0 of the lightest particle in the theory. When enclosed in a box, the energy of the particle starts to deviate from its infinite volume mass m_i as soon as the cloud is squeezed by the box, i.e. for $\lambda_0 \simeq L$. This physical picture can be translated into an exact asymptotic formula relating the size dependence of the masses M_i to certain forward elastic scattering amplitudes (cf. Eq. (2.22) below). I have already presented this formula some time ago together with a number of applications [12]. In this paper, a detailed proof of the formula is given within the framework of Feynman diagrams.

A seemingly different physical situation occurs when the particle considered is a bound state of two other "pointlike" stable particles with a binding energy small compared to its mass. The wave function of the bound particles then falls off exponentially with a characteristic length λ, which may be substantially larger than λ_0. One therefore expects that finite size effects on the binding energy are large up to sizes $L \simeq \lambda$ and only then go to zero exponentially as $L \to \infty$. An exactly soluble 2-dimensional example displaying this behaviour has recently been discussed in ref. [13]. Here it will be shown (Sect. 3) that the bound state situation is actually not so different from a squeezed polarization cloud as discussed above, since in both cases finite size effects arise from particle exchange "around the

world" and are described by similar relativistic amplitudes, the main differences being of a kinematical nature.

The stable mesons and baryons in QCD are bound states of quarks, but the situation here is quite different from the one just discussed, because quarks are confined. As suggested by simple models, this presumably implies that the wave functions of the valence quarks inside a hadron are going to zero more rapidly than exponentially as the distance between them increases [1]. One therefore expects that the probability for a single quark to separate from its partner(s) and walk around the periodic box is quickly going to zero for growing L so that, in unquenched QCD, the leading finite size effect on the hadron masses at large L is not due to this process, but arises from the squeezing of the virtual pion cloud around these particles as discussed above. Only this latter mechanism was taken into account for the estimation of the size dependence of the pion and nucleon masses presented in ref. [12].

As already indicated above, the proof of the basic relation between finite size mass shifts and elastic scattering amplitudes will be given to all orders in perturbation theory, i.e. I shall assume that the dynamics of the particles considered can be described by a Lagrangian quantum field theory, where all fields are massive and the couplings are small. The interaction Lagrangian can be arbitrarily complicated and the theory may also have a fixed ultra-violet cutoff. In all cases, the resulting formulae are independent of these details and refer only to the physical masses and scattering amplitudes of the particles. In view of this universality, I believe that the result is in fact true beyond perturbation theory.

Another important question is whether the relations so derived are also valid in pure non-Abelian gauge theories and QCD, where perturbation theory in the gauge coupling constant involves massless fields and the arguments given in this paper cannot immediately be applied. However, one can always describe the low energy properties of these theories by effective Lagrangians [2], which, for an accurate description, are perhaps very complicated, but are of exactly the type tractable by the Feynman diagram technique of Sect. 2. Note that because of the universality of the final result, the precise form of the effective Lagrangian is never needed, i.e. the effective Lagrangian only catalyses the proof.

These arguments suggest that the mass shift formulae proved in this paper are of a basically kinematical nature and that they are valid in arbitrary massive quantum field theories, a conclusion, which is also supported by exactly soluble models [12] and a recent numerical study of finite size effects in the 0(3) non-linear σ-model in two dimensions [7].

Although this paper is self-contained, the reader is advised to first consult ref. [12] for an overview and illustrations (no concrete applications will be

---

1 Velikson and Weingarten [14] have recently calculated Coulomb gauge quark wave functions in quenched lattice QCD and find that they are decaying rapidly, although, in the limited range of distances available, a deviation from an exponential law is not seen

2 At least in principle, such effective Lagrangians could be constructed by "integrating out" the high frequency modes in the functional integral. Alternatively, one may adopt Weinberg's point of view [15] that the class of all effective Lagrangians reproducing the global symmetries and the spectrum of low lying particles of an underlying field theory contains no more information than would be implied by basic principles (locality, analyticity, etc.) anyway

discussed here). To keep the presentation as simple as possible, proofs will only be given for spinless particles and the dimensionality of space-time is set equal to 4 throughout the paper. Also, I shall assume that the ultra-violet cutoff (if any) does not break Lorentz invariance. All these restrictions are in no way crucial to the argumentation and can easily be relaxed, in particular, with appropriate modifications the results also hold in lattice theories [9].

The bulk of this paper is devoted to the proof of the finite size mass shift formula alluded to above (Sect. 2). Although the details are worked out for simple scalar theories only, the method can easily be generalized to more complicated situations. In particular, the volume dependence of bound state masses can be calculated and one finds, in the non-relativistic limit, that the leading finite size effect on the binding energy is correlated with the fall off properties of the bound state wave function in the expected way (Sect. 3). A few selected remarks are included in the final Sect. 4.

2. Volume Dependence of the Mass Gap in Simple Scalar Theories

2.1. Basic Definitions. We here discuss theories of a real scalar field $\phi(x)$, which, in infinite volume, describes the physics of a single self-interacting particle ("meson") of mass m and spin 0. For the study of finite size effects, it is convenient to work with the connected euclidean correlation functions $\langle\phi(x_1)\dots\phi(x_n)\rangle$ of ϕ rather than time-ordered vacuum expectation values. The normalization of ϕ is chosen such that the meson pole in the euclidean propagator has unit residue, i.e. we have[3]

$$\langle\phi(x)\phi(0)\rangle = \int \frac{d^4p}{(2\pi)^4} e^{ipx} G(p), \tag{2.1}$$

$$G(p)^{-1} = m^2 + p^2 - \Sigma(p), \tag{2.2}$$

$$\Sigma(p) = \frac{\partial}{\partial p^\mu}\Sigma(p) = 0 \quad \text{for} \quad p^2 = -m^2. \tag{2.3}$$

Since it is assumed that there are no bound states or other additional stable particles, the meson pole at $p^2 = -m^2$ is the only singularity of $G(p)$ below the 2-particle threshold at $p^2 = -4m^2$.

In a finite volume of size L and periodic boundary conditions, the field ϕ satisfies

$$\phi(x^0, \mathbf{x} + L\mathbf{n}) = \phi(x^0, \mathbf{x}) \quad \text{for all} \quad \mathbf{n} \in \mathbb{Z}^3. \tag{2.4}$$

Denoting the connected euclidean correlation functions at $L < \infty$ by $\langle\phi(x_1)\dots\phi(x_n)\rangle_L$, we have

$$\langle\phi(x)\phi(0)\rangle_L = L^{-3}\sum_{\mathbf{p}} \int \frac{dp^0}{2\pi} e^{ipx} G_L(p), \tag{2.5}$$

$$G_L(p)^{-1} = m^2 + p^2 - \Sigma_L(p), \tag{2.6}$$

---

3 Euclidean 4-vectors are written as $p_\mu = p^\mu = (p^0, \mathbf{p})$, $\mathbf{p} = (p^1, p^2, p^3)$, and the euclidean scalar product is $p \cdot q = p^0 q^0 + \mathbf{p} \cdot \mathbf{q}$

where the momenta $\mathbf{p}$ take values

$$\mathbf{p}=\frac{2\pi}{L}\mathbf{n}, \quad \mathbf{n}\in\mathbb{Z}^3. \tag{2.7}$$

For large L, we expect (and shall later show) that $\sum_L(p)$ is close to $\sum(p)$, in particular $\sum_L(p)$ and $\frac{\partial}{\partial p^0}\sum_L(p)$ are nearly vanishing along the mass shell $p^2=-m^2$. Thus, for every fixed $\mathbf{p}$ of the form (2.7), $G_L(p)$ has a pair of poles in the complex energy plane at

$$p^0=\pm i\omega_L(\mathbf{p}), \tag{2.8}$$

$$\omega_L(\mathbf{p}) \underset{L\to\infty}{\sim} \omega(\mathbf{p})=\sqrt{m^2+\mathbf{p}^2}>0. \tag{2.9}$$

The meson mass M in finite volume is now defined by

$$M=\omega_L(\mathbf{0}), \tag{2.10}$$

or, equivalently, through the leading exponential decay of the 2-point function at large times:

$$\langle\phi(x)\phi(0)\rangle_L \underset{x^0\to\infty}{\propto} e^{-Mx^0}. \tag{2.11}$$

The asymptotic formula to be proved in the following subsections relates the finite size mass shift

$$\Delta m=M-m \tag{2.12}$$

to the (infinite volume) elastic meson scattering amplitude T. To write it down explicitly, some further preparation is needed. First of all we note that the scattering amplitude T can be expressed through the euclidean 4-point function in the following way. Define full propagator amputated correlation functions $G(p_1,\ldots,p_n)$ by

$$\langle\phi(x_1)\ldots\phi(x_n)\rangle=\int\frac{d^4p_1}{(2\pi)^4}\ldots\frac{d^4p_n}{(2\pi)^4}\,e^{i(p_1x_1+\ldots+p_nx_n)}$$
$$\cdot(2\pi)^4\,\delta(p_1+\ldots+p_n)G(p_1)\ldots G(p_n)G(p_1,\ldots,p_n). \tag{2.13}$$

For $z<0$, set

$$G_z(p_1,\ldots,p_n)=G(\tilde{p}_1,\ldots,\tilde{p}_n), \quad \tilde{p}=(zp^0,\mathbf{p}). \tag{2.14}$$

Then, using the spectral condition, one may show that for every fixed, real momentum configuration $p_1,\ldots,p_n$, $G_z(p_1,\ldots,p_n)$ extends to an analytic function of z in the half-plane $\mathrm{Re}\,z<0$. Furthermore, the elastic scattering amplitude is given by

$$T(\mathbf{p}',\mathbf{q}'|\mathbf{p},\mathbf{q})=\lim_{\varepsilon\to 0}G_{i-\varepsilon}(p',q',-p,-q), \tag{2.15}$$

where $\mathbf{p},\mathbf{q}$ are the momenta of the incoming mesons and $\mathbf{p}',\mathbf{q}'$ those of the outgoing particles. The energy components of the 4-momenta in Eq. (2.15) are $p^0=\omega(\mathbf{p})$,

$p'^0 = \omega(\mathbf{p}')$, etc. The normalization of the scattering amplitude so defined is such that the optical theorem reads

$$\operatorname{Im} T(\mathbf{p}, \mathbf{q}|\mathbf{p}, \mathbf{q}) = \sqrt{s(s-4m^2)}\,\sigma_{\text{tot}}(s)\,, \tag{2.16}$$

where s denotes the centre of mass energy squared and $\sigma_{\text{tot}}(s)$ is the total cross section.

The relation between Δm and T involves the forward amplitude

$$F = T(\mathbf{p}, \mathbf{q}|\mathbf{p}, \mathbf{q})\,. \tag{2.17}$$

F depends on a single Lorentz invariant, which is conveniently taken to be the crossing variable

$$\nu = (\omega(\mathbf{p})\omega(\mathbf{q}) - \mathbf{p}\cdot\mathbf{q})/m\,. \tag{2.18}$$

It follows from general principles that $F(\nu)$ is a boundary value of a function, also denoted by F, which is analytic in the cut plane shown in Fig. 1. Crossing symmetry implies that this analytic function is even in ν. Besides the physical cuts, F has no singularities in the simple theories considered here except perhaps for a pair of poles at $\nu = \pm\frac{1}{2}m$. These arise from 1-particle exchange reactions as follows. Let $\Gamma(p_1, \ldots, p_n)$ be the 1-particle irreducible part of the amputated n-point function $G(p_1, \ldots, p_n)$. Using the graphical notation of Fig. 2, the 4-point function (and hence the scattering amplitude T) can be decomposed into 1-particle irreducible parts as shown in Fig. 3. The poles of the forward amplitude F at $\nu = \pm\frac{1}{2}m$ stem from the first two 1-particle reducible diagrams in Fig. 3, because for these values of ν the momentum flowing through the middle propagators is just on the mass shell. For the residue of the pole, we thus have

$$\lim_{\nu \to \pm\frac{1}{2}m} (\nu^2 - \tfrac{1}{4}m^2)F(\nu) = \tfrac{1}{2}\lambda^2\,, \tag{2.19}$$

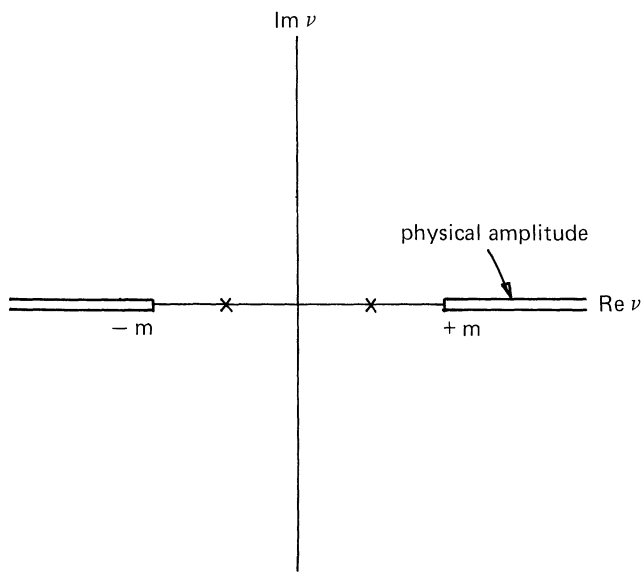

Fig. 1. Analyticity domain of the forward amplitude $F(\nu)$. There are cuts along the real line from $-\infty$ to $-m$ and from m to $+\infty$. Simple poles may occur at $\nu = \pm\frac{1}{2}m$

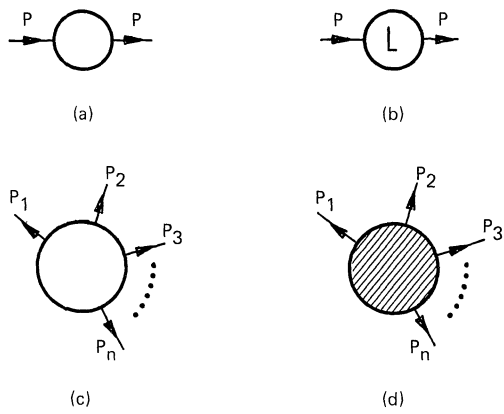

Fig. 2a–d. Graphical symbols used for (**a**) the full propagator $G(p)$, (**b**) the modified full propagator $\left(2\sum_{j=1}^{3}\cos p_j L\right)\cdot G(p)$, (**c**) the connected, full propagator amputated n-point function $G(p_1, \ldots, p_n)$, and (**d**) the n-point vertex function $\Gamma(p_1, \ldots, p_n)$

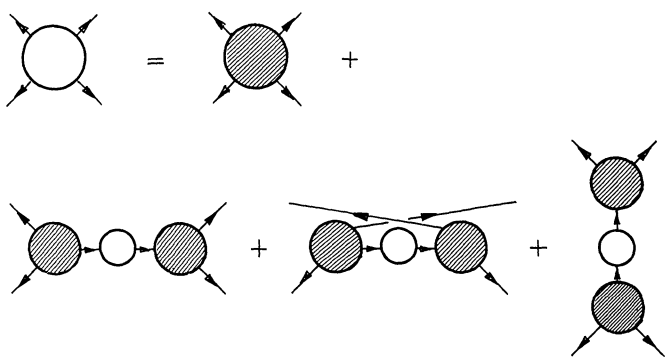

Fig. 3. Decomposition of the 4-point function into 1-particle irreducible parts

where the 3-meson coupling constant λ is given by

$$\lambda = \Gamma(p, q, k), \tag{2.20}$$

$$p+q+k=0, \quad p^2=q^2=k^2=-m^2 \tag{2.21}$$

(the complex point (2.21) is away from the singularities of the 3-point function, cp. Subsect. 2.4).

2.2. Statement of Result and Outline of Proof. We are now in a position to write down the long heralded asymptotic formula for the finite size mass shift:

$$\Delta m = -\frac{3}{16\pi m^2 L}\left\{\lambda^2 e^{-\frac{\sqrt{3}}{2}mL} + \frac{m}{\pi}\int_{-\infty}^{\infty} dy e^{-\sqrt{m^2+y^2}L} F(iy) + O(e^{-\tilde{m}L})\right\}. \tag{2.22}$$

Here, $\bar{m}$ is some mass larger than m, i.e. the error term in Eq. (2.22) is exponentially small compared to the first two terms. In perturbation theory, we shall find

$$\bar{m} \geqq \sqrt{\tfrac{3}{2}}\, m\,. \tag{2.23}$$

If $\lambda \neq 0$, this bound is actually saturated, but in other cases like the ϕ^4-theory with an ultra-violet cutoff, $\bar{m}$ is expected to be larger.

The implications of Eq. (2.22) have already been discussed at length in ref. [12] so that here we directly proceed to the proof of this formula. Deferring details to the following subsections, the argumentation is as follows. As explained in the introduction, the basic assumption is that the correlation functions of ϕ can be expanded in a series of Feynman diagrams with momentum space propagators

$$\tilde{\Delta}(p;m) = (m^2 + p^2)^{-1} \tag{2.24}$$

and arbitrary local vertices (the set of vertices must include mass and wave function renormalization counterterms to insure the validity of the normalization condition (2.3) to all orders of the expansion). If desired, the free propagator (2.24) may also be replaced by a propagator with an ultra-violet cut-off, for example

$$\tilde{\Delta}^{\Lambda}(p;m^2) = (m^2 + p^2)^{-1} - (\Lambda^2 + p^2)^{-1}\,, \qquad \Lambda \gg m\,. \tag{2.25}$$

In the course of the discussion, it will become clear that the validity of Eq. (2.22) is not affected by such a modification independently of how large Λ is.

The Feynman rules for the finite L correlation functions are exactly the same as in infinite volume except, of course, that the space-like components of the loop and external momenta are restricted to the discrete values (2.7). This immediately implies that, as asserted above, $\Sigma_L(p)$ converges to $\Sigma(p)$ as $L \to \infty$, because in this limit, the sums over loop momenta can be replaced by integrals. The finite size mass shift Δm is therefore small for large L and the pole equation

$$G_L(iM, \mathbf{0})^{-1} = 0 \tag{2.26}$$

can be solved by expanding in powers of Δm, which leads to

$$\begin{gathered} \Delta m = -\Sigma_L(\hat{p}) \Big/ \left(2m + i\frac{\partial}{\partial p^0}\Sigma_L(\hat{p})\right) + O((\Delta m)^2)\,, \\ \hat{p} = (\mathrm{im}, 0, 0, 0)\,. \end{gathered} \tag{2.27}$$

With the help of some abstract graph theory summarized in Subsect. 2.3, it may now be shown (Subsects. 2.5–2.7) that diagram by diagram one has

$$\Sigma_L(\hat{p}) - \Sigma(\hat{p}) = O\left(e^{-\frac{\sqrt{3}}{2}mL}\right), \tag{2.28}$$

$$\frac{\partial}{\partial p^0}\Sigma_L(\hat{p}) - \frac{\partial}{\partial p^0}\Sigma(\hat{p}) = O\left(e^{-\frac{\sqrt{3}}{2}mL}\right). \tag{2.29}$$

In the sum of all diagrams of a given order, $\Sigma(\hat{p})$ and $\dfrac{\partial}{\partial p^0}\Sigma(\hat{p})$ vanish because of the normalization condition (2.3), and it follows that

$$\Delta m = -\frac{1}{2m}\Sigma_L(\hat{p}) + O(e^{-\bar{m}L}) = O\left(e^{-\frac{\sqrt{3}}{2}mL}\right) \tag{2.30}$$

to all orders in perturbation theory (here and below, $\bar{m}$ denotes some mass satisfying the bound (2.23)).

The analysis of the L-dependence of the self-energy diagrams which leads to Eqs. (2.28), (2.29) also allows us to identify the class of graphs, which contribute to the leading exponential decays of $\Sigma_L(\hat{p})$ at large L. These graphs can be summed up in closed form and one obtains

$$\Sigma_L(\hat{p}) = \tfrac{1}{2}(I_1 + I_2 + I_3) + O(e^{-\bar{m}L}), \tag{2.31}$$

$$I_1 = \int \frac{d^4q}{(2\pi)^4} \left(2 \sum_{j=1}^{3} \cos q_j L\right) G(q + \tfrac{1}{2}\hat{p}) G(-q + \tfrac{1}{2}\hat{p}) \cdot \Gamma(-\hat{p}, q + \tfrac{1}{2}\hat{p}, -q + \tfrac{1}{2}\hat{p}) \Gamma(\hat{p}, -q - \tfrac{1}{2}\hat{p}, q - \tfrac{1}{2}\hat{p}), \tag{2.32}$$

$$I_2 = \int \frac{d^4q}{(2\pi)^4} \left(2 \sum_{j=1}^{3} \cos q_j L\right) G(q) G(0) \Gamma(-q, q, 0) \Gamma(-\hat{p}, \hat{p}, 0), \tag{2.33}$$

$$I_3 = \int \frac{d^4q}{(2\pi)^4} \left(2 \sum_{j=1}^{3} \cos q_j L\right) G(q) \Gamma(\hat{p}, q, -\hat{p}, -q). \tag{2.34}$$

The graphical representation of these integrals is displayed in Fig. 4. The proof of the mass shift formula (2.22) is now easily completed (Subsect. 2.7) by using complex contour integration to extract the asymptotic behaviour of I_1, I_2, and I_3 at large L. The analytic properties of the vertex functions, as far as they are needed for this last step, are established to all orders in perturbation theory in Subsect. 2.4 (the discussion there also serves as a simple illustration of the abstract graph theory developed in the following subsection).

2.3. Some Abstract Graph Theory. The proof of the statements made above requires some control over the topology of an arbitrary Feynman diagram. To facilitate this task, some notions and results from abstract graph theory are summarized here (a fuller account can be found in Nakanishi's book [16]).

(a) *Abstract Graphs.* An (abstract) graph $\mathscr{G}$ consists of a set of lines $\mathscr{L}$, a non-empty set of vertices $\mathscr{V}$ and two mappings i and f from $\mathscr{L}$ into $\mathscr{V}$ called incidence

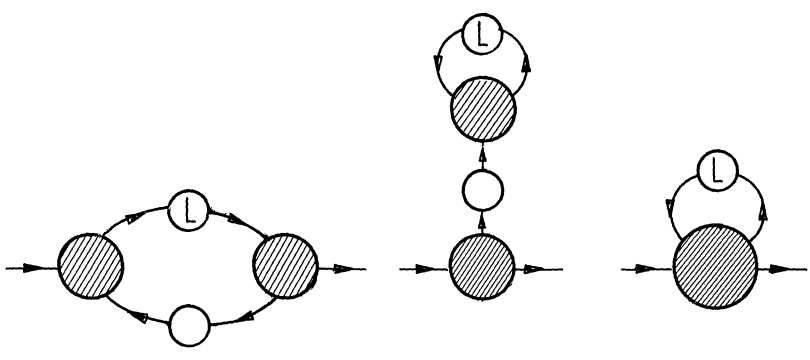

Fig. 4a–c. Graphical representation of the integrals I_1, I_2, and I_3 (from left to right). The notation is as in Fig. 2 and the momentum flowing into the diagrams is $\hat{p}$

relations[4]. For every line $\ell \in \mathscr{L}$, $i(\ell)$ is called its initial vertex and $f(\ell)$ its final vertex. $i(\ell)$ and $f(\ell)$ are also referred to as the endpoints of ℓ. It is possible that $i(\ell)$ coincides with $f(\ell)$ in which case ℓ is called a loop line.

(b) *Paths.* A path $\mathscr{P}$ in a graph $\mathscr{G}$ connecting the vertices $a \neq b$ is a subset of $\mathscr{L}$ with the property that there exists a sequence $a = v_1, v_2, \ldots, v_N = b$ of pairwise different vertices v_k and a labelling $\ell_1, \ell_2, \ldots, \ell_{N-1}$ of the lines in $\mathscr{P}$ such that v_k, v_{k+1} are the endpoints of ℓ_k. Note that $\mathscr{P}$ is a set of lines and does not contain the vertices v_k. In particular, two paths intersect if and only if they share a common line.

A graph $\mathscr{G}$ is connected, if for any pair of vertices $a \neq b$ there exists a path $\mathscr{P}$ in $\mathscr{G}$ connecting a and b. In a connected graph, every vertex is an endpoint of some line except in the case when there are no lines at all and a single vertex only. A general graph always divides into a number of connected components in the obvious way.

(c) *Loops.* A loop $\mathscr{C}$ in a graph $\mathscr{G}$ is a non-empty subset of $\mathscr{L}$ with the property that there exists a sequence $v_1, \ldots, v_N$ of pairwise different vertices and a labelling $\ell_1, \ldots, \ell_N$ of the lines in $\mathscr{C}$ such that v_k, v_{k+1} are the endpoints of $\ell_k (k = 1, \ldots, N-1)$ and v_N, v_1 are the endpoints of ℓ_N. In particular, $\mathscr{C} = \{\ell\}$ is a loop if ℓ is a loop line.

It is possible to define an orientation on a loop $\mathscr{C}$ in $\mathscr{G}$. This amounts to assigning a number, denoted $[\mathscr{C} : \ell]$, to every line $\ell \in \mathscr{C}$ such that

$$\begin{aligned} &[\mathscr{C} : \ell] \in \{1, -1\}, \\ &[\mathscr{C} : \ell] = [\mathscr{C} : \ell'] \quad \text{if} \quad i(\ell) = f(\ell'), \\ &[\mathscr{C} : \ell] = -[\mathscr{C} : \ell'] \quad \text{if} \quad i(\ell) = i(\ell') \quad \text{or} \quad f(\ell) = f(\ell') \; (\ell \neq \ell'). \end{aligned} \tag{2.35}$$

On every loop, there are two orientations, which differ by an overall sign.

A first result of abstract graph theory we shall rely on later is the following

Lemma 2.1. *Let $a \neq b$ be two vertices in a connected graph $\mathscr{G}$ and $\mathscr{C}_1, \ldots, \mathscr{C}_M$ a set of (pairwise) disjoint loops. Then there exists a path $\mathscr{P}$ in $\mathscr{G}$ connecting a and b such that $\mathscr{P} \cap \mathscr{C}_j$ is either empty or a path in $\mathscr{G}$ for all $j = 1, \ldots, M$.*

For a proof, see Appendix A.

(d) *Trees.* A tree T in a connected graph $\mathscr{G}$ is a maximal subset of $\mathscr{L}$ not containing any loop in $\mathscr{G}$. For every tree T, we define T^* to be the set of lines not belonging to T. It may be shown that trees always exist and that the number of lines in T^* is the same for all trees (and equal to the number of independent loops in $\mathscr{G}$). If T is a tree and $a \neq b$ are two vertices of $\mathscr{G}$, there exists a unique path $\mathscr{P} \subset T$ connecting a and b. Furthermore, for every line $\ell \in T^*$, there exists a unique loop in $\mathscr{G}$, which is contained in $T \cup \{\ell\}$. This loop necessarily passes through ℓ.

(e) *N-particle Irreducibility.* By deleting a line ℓ from a graph $\mathscr{G}$, a new graph denoted by $\mathscr{G} \backslash \{\ell\}$ is obtained. Thus, the set of lines of the new graph is $\mathscr{L} \backslash \{\ell\}$, its set of vertices is $\mathscr{V}$ and the incidence relations i and f are inherited in the obvious way. Similarly, a number $\ell_1, \ldots, \ell_N$ of lines can be deleted. If $\mathscr{G}$ is initially connected, the mutilated graph $\mathscr{G} \backslash \{\ell_1, \ldots, \ell_N\}$ in general decomposes into several connected components. A graph $\mathscr{G}$ is called N-particle irreducible between two vertices

4 $\mathscr{L}$ and $\mathscr{V}$ are assumed to be finite

$a, b \in \mathscr{V}$, if a and b always belong to the same connectivity component of $\mathscr{G} \setminus \{\ell_1, \ldots, \ell_N\}$ no matter which lines $\ell_1, \ldots, \ell_N$ are deleted.

N-particle irreducibility of a graph $\mathscr{G}$ implies a certain amount of analyticity of Feynman integrals associated to $\mathscr{G}$. To establish analyticity domains, the following result will be helpful.

Theorem 2.2. *Suppose $\mathscr{G}$ is a graph, which is N-particle irreducible between two of its vertices $a \neq b$. Then there exist $N+1$ disjoint paths $\mathscr{P}_1, \ldots, \mathscr{P}_{N+1}$ in $\mathscr{G}$ connecting a and b.*

For a proof see ref. [16, p. 37ff.].

(f) $\mathbb{Z}^3$ *Gauge Fields on a Graph.* As on regular lattices, it is possible to define gauge fields on an abstract graph $\mathscr{G}$. In particular, if the gauge group is $\mathbb{Z}^3$ (which will later turn out to be the relevant choice), a gauge field on $\mathscr{G}$ is an assignment of an integer vector $\mathbf{n}(\ell)$ to every line $\ell \in \mathscr{L}$. Another field $\mathbf{n}'(\ell)$ is then called gauge equivalent to $\mathbf{n}(\ell)$, if

$$\mathbf{n}'(\ell) = \mathbf{n}(\ell) + \boldsymbol{\lambda}(f(\ell)) - \boldsymbol{\lambda}(i(\ell)) \quad \text{for all} \quad \ell \in \mathscr{L}, \tag{2.36}$$

where $\boldsymbol{\lambda}(v)$, $v \in \mathscr{V}$, is some field of integer vectors. Equation (2.36) is also referred to as a gauge transformation. If $\mathscr{C}$ is an oriented loop in $\mathscr{G}$, one can define a gauge invariant quantity

$$\mathbf{W}(\mathscr{C}, \mathbf{n}) = \sum_{l \in \mathscr{C}} [\mathscr{C} : \ell] \, \mathbf{n}(\ell), \tag{2.37}$$

which is analogous to the Wilson loop in lattice gauge theories.

A useful way to label the gauge equivalence classes $[\mathbf{n}]$ of gauge fields on a connected graph $\mathscr{G}$ is the following. Choose some tree T in $\mathscr{G}$ and for every class $[\mathbf{n}]$ a representative field $\mathbf{n}(\ell)$ such that

$$\mathbf{n}(\ell) = 0 \quad \text{for all} \quad \ell \in T. \tag{2.38}$$

Fields satisfying (2.38) are said to be in the axial gauge (relative to T). It is trivial to show that in every class $[\mathbf{n}]$ there exists a unique member $\mathbf{n}(\ell)$, which is in the axial gauge. The class $[\mathbf{n}]$ can thus be characterized by the values assumed by this field along T^*.

(g) *Simple Gauge Fields.* As usual, gauge fields on a graph $\mathscr{G}$, which are gauge equivalent to $\mathbf{n}(\ell) = 0$, are referred to as pure gauge configurations. Another important class of gauge fields are those which are gauge equivalent to a configuration $\mathbf{n}(\ell)$ with $\mathbf{n}(\ell) = 0$ for all lines $\ell \in \mathscr{L}$ except for one line ℓ^*, which is contained in at least one loop in $\mathscr{G}$ and where one has $|\mathbf{n}(\ell^*)| = 1$. These fields are called simple.

A set of gauge independent simple fields can be constructed as follows. Define $\mathscr{L}_c$ to be the set of lines $\ell \in \mathscr{L}$, which are contained in at least one loop in $\mathscr{G}$. Two lines in $\mathscr{L}_c$ are called independent if there exists a loop in $\mathscr{G}$, which contains one of them but not the other. Now choose a maximal set $\{\ell_1, \ldots, \ell_N\}$ of pairwise independent lines in $\mathscr{L}_c$ and consider the $6N$ simple fields $\mathbf{n}(\ell; j, \mathbf{e})$ $(j = 1, \ldots, N,$

$\mathbf{e} \in \mathbb{Z}^3$, $|\mathbf{e}| = 1$) defined by

$$\mathbf{n}(\ell; j, \mathbf{e}) = \mathbf{e} \quad \text{if} \quad \ell = \ell_j,$$
$$= 0 \quad \text{otherwise}. \tag{2.39}$$

Then, it is easy to show that these configurations are a complete list of gauge independent simple fields.

(h) *Feynman Diagrams and Abstract Graphs.* Suppose $\mathscr{D}$ is a Feynman diagram contributing to an n-point vertex function. After assigning some arbitrary orientation to its lines, $\mathscr{D}$ defines, in a natural way, an abstract graph $\mathscr{G}$ consisting of a set of vertices $\mathscr{V}$, a set of lines $\mathscr{L}$ and incidence relations i and f. We distinguish between an abstract vertex $v \in \mathscr{V}$ and its coordinates in space-time, which will be denoted by $x(v)^\mu$. A vertex of $\mathscr{G}$ is called external, if one or more external momenta are leaving (entering) the corresponding vertex in $\mathscr{D}$. Apart from this qualification, all the vertices in $\mathscr{G}$ are treated on an equal footing independent of whether they correspond to vertices of different type in $\mathscr{D}$. Similarly, lines describing the propagation of different particles in $\mathscr{D}$ are not distinguished in $\mathscr{G}$.

2.4. Analyticity Properties of Vertex Functions. As already mentioned in Subsect. 2.2, some analyticity properties of the 3- and 4-point vertex functions are required for the proof of the mass shift formula (2.22). The analyticity domain established here derives from the 1-particle irreducibility of the vertex functions and will be sufficiently large for our purposes.

Define a complex domain

$$\mathbb{D} = \{(p, q) \in \mathbb{C}^4 \times \mathbb{C}^4 | (\mathrm{Im}\, p \pm \mathrm{Im}\, q)^2 < 4m^2\}. \tag{2.40}$$

Then we have

Theorem 2.3. *To all orders in perturbation theory, the vertex functions*

$$\Gamma(p, -q - \tfrac{1}{2}p, q - \tfrac{1}{2}p),$$
$$\Gamma(p, q, -p, -q),$$

which are initially defined for $(p, q) \in \mathbb{R}^4 \times \mathbb{R}^4$, *analytically extend to the whole domain* $\mathbb{D}$.

Proof. We consider only the 3-point function, the proof being similar for the 4-point function.

Let $\mathscr{D}$ be a Feynman diagram contributing to the 3-point vertex function and $\mathscr{G}$ the associated abstract graph. $\mathscr{G}$ is 1-particle irreducible between any two of its vertices and has 3 external vertices, denoted a, b, c, where the external momenta p, $-q - \frac{1}{2}p$, $q - \frac{1}{2}p$ leave the diagram $\mathscr{D}$.

Because $\mathscr{G}$ is 1-particle irreducible between b and c, Theorem 2.2 applies and it follows that there exist two disjoint paths $\mathscr{P}_1$, $\mathscr{P}_2$ connecting b and c (set $\mathscr{P}_1 = \mathscr{P}_2 = \emptyset$, if $b = c$). Suppose now we add an extra vertex z to $\mathscr{G}$ and two extra lines ℓ_b, ℓ_c connecting z with b and c. Then, this augmented graph is 1-particle irreducible between a and z, and hence there are disjoint paths $\mathscr{P}_3$, $\mathscr{P}_4$ connecting these vertices. Deleting ℓ_b and ℓ_c again, we are left with disjoint paths $\mathscr{P}_b$, $\mathscr{P}_c$ in $\mathscr{G}$, which connect a with b and c, respectively ($\mathscr{P}_b = \emptyset$, if $a = b$, and $\mathscr{P}_c = \emptyset$, if $a = c$).

The paths $\mathscr{P}_1, \mathscr{P}_2$ and $\mathscr{P}_b, \mathscr{P}_c$ can be used to define a flow of external momentum through the diagram $\mathscr{D}$ in such a way that the momenta $k(\ell)$ carried by the lines ℓ in $\mathscr{G}$ satisfy

$$\begin{aligned} k(\ell) &= \pm \tfrac{1}{2}p && \text{if} \quad \ell \in (\mathscr{P}_b \cup \mathscr{P}_c) \setminus (\mathscr{P}_1 \cup \mathscr{P}_2), \\ k(\ell) &= \pm \tfrac{1}{2}q && \text{if} \quad \ell \in (\mathscr{P}_1 \cup \mathscr{P}_2) \setminus (\mathscr{P}_b \cup \mathscr{P}_b), \\ k(\ell) &= \pm \tfrac{1}{2}p \pm \tfrac{1}{2}q && \text{if} \quad \ell \in (\mathscr{P}_b \cup \mathscr{P}_c) \cap (\mathscr{P}_1 \cup \mathscr{P}_2), \\ k(\ell) &= 0 && \text{otherwise}. \end{aligned} \tag{2.41}$$

In other words, the momentum carried by $\mathscr{P}_1, \mathscr{P}_2$ is $\frac{1}{2}q$ and the momentum flowing through $\mathscr{P}_b, \mathscr{P}_c$ is $-\frac{1}{2}p$.

When the diagram $\mathscr{D}$ is evaluated in momentum space, the total momentum flowing through a line ℓ is $k(\ell)+r(\ell)$, where $r(\ell)$ is a combination of loop momenta. Thus, as long as

$$\mathrm{Re}(k(\ell)+r(\ell))^2 > -m^2$$

for all lines ℓ and all loop momenta, the Feynman integral associated to $\mathscr{D}$ is not singular. Since $r(\ell)$ is real, the condition

$$(\mathrm{Im}\, k(\ell))^2 < m^2 \quad \text{for all} \quad \ell \in \mathscr{L},$$

is therefore sufficient to guarantee regularity. In view of Eq. (2.41), this criterion is satisfied for $(p, q) \in \mathbb{D}$, thus proving the theorem. □

2.5. Large L Behaviour of Self-Energy Diagrams. To study the L-dependence of Feynman diagrams at large L, it is useful to work in position space rather than momentum space. Thus, the infinite volume propagator is

$$\Delta(x; m) = \int \frac{d^4p}{(2\pi)^4} e^{ipx}(m^2+p^2)^{-1}, \tag{2.42}$$

and for finite L we have the well-known representation

$$\Delta_L(x; m) = \sum_{\mathbf{n} \in \mathbb{Z}^3} \Delta(x+nL; m), \qquad n = (0, \mathbf{n}). \tag{2.43}$$

This series converges rapidly, because $\Delta(x; m)$ decays exponentially at large x. In a position space Feynman integral, the vertex "factors" are homogeneous partial differential operators with constant coefficients acting on the arguments of the propagators in the diagram. An important point to note is that these differential operators follow directly from the Lagrange density and are hence independent of L.

Suppose now that $\mathscr{D}$ is a Feynman diagram contributing to $\Sigma_L(\hat{p})$ (the results obtained below also hold, with appropriate modifications, for $\frac{\partial}{\partial p^0}\Sigma_L(\hat{p})$ and arbitrary n-point vertex functions at momenta with purely imaginary energy components). The abstract graph $\mathscr{G}$ associated to $\mathscr{D}$ is 1-particle irreducible between any two of its vertices and has two external vertices, denoted a and b, where $\hat{p}$ flows in and out, respectively. It is possible that a and b coincide.

In position space, the contribution $\mathscr{J}_L(\mathscr{D})$ of the diagram $\mathscr{D}$ to $\sum_L(\hat{p})$ is an integral of the general form

$$\mathscr{J}_L(\mathscr{D}) = \prod_{v\in\mathscr{V}'} \int_{\mathbb{R}\times L^3} d^4x(v)\, \mathbb{V} \Big\{ e^{m(x(b)^0 - x(a)^0)} \prod_{\ell\in\mathscr{L}} \Delta_L(x(f(\ell)) - x(i(\ell)); m) \Big\}, \tag{2.44}$$

where $\mathscr{V}' = \mathscr{V}\backslash\{b\}$ and $\mathbb{V}$ is the product of the vertex "factors" as explained above. The integrand in Eq. (2.44) is a periodic function of the space-like coordinates of every vertex $v\in\mathscr{V}'$ and the integration over $\mathbf{x}(v)$ is accordingly restricted to a periodicity cell of volume L^3. The external momentum $\hat{p}$ flowing through the diagram is accounted for by the exponential factor in Eq. (2.44).

To obtain a more tractable expression, we now substitute the series (2.43) for the propagators Δ_L in Eq. (2.44). For every line $\ell\in\mathscr{L}$, we then have a summation variable $\mathbf{n}(\ell)\in\mathbb{Z}^3$ and, interchanging summations and integrations, $\mathscr{J}_L(\mathscr{D})$ becomes a sum of terms one for each $\mathbb{Z}^3$ gauge field configuration $\{\mathbf{n}(\ell)\}$ on $\mathscr{G}$ (cf. Subsect. 2.3). This summation can be split into two independent summations, one over the gauge equivalence classes $[\mathbf{n}]$ of gauge fields and the other over the gauge transformations $\boldsymbol{\lambda}(v)$, $v\in\mathscr{V}$, with $\boldsymbol{\lambda}(b)=0$. The latter can be combined with the integrations over $\mathbf{x}(v)$, $v\in\mathscr{V}'$, and one then obtains

$$\mathscr{J}_L(\mathscr{D}) = \sum_{[\mathbf{n}]} \mathscr{J}_L(\mathscr{D}, \mathbf{n}), \tag{2.45}$$

$$\mathscr{J}_L(\mathscr{D}, \mathbf{n}) = \prod_{v\in\mathscr{V}'} \int_{\mathbb{R}^4} d^4x(v)\, \mathbb{V} \Big\{ e^{m(x(b)^0 - x(a)^0)} \prod_{\ell\in\mathscr{L}} \Delta(x(f(\ell)) - x(i(\ell)) + n(\ell)L; m) \Big\}. \tag{2.46}$$

Note that now the vertices are integrated over all of $\mathbb{R}^4$. $\mathscr{J}_L(\mathscr{D}, \mathbf{n})$ is therefore gauge invariant and the summation over gauge equivalence classes in Eq. (2.45) is a well-defined operation.

For $\mathbf{n}=0$, $\mathscr{J}_L(\mathscr{D}, \mathbf{n})$ is independent of L and equal to $\mathscr{J}(\mathscr{D})$, the contribution of the diagram $\mathscr{D}$ to the infinite volume self-energy $\sum(\hat{p})$. Thus, for the contribution of $\mathscr{D}$ to the difference $\sum_L(\hat{p}) - \sum(\hat{p})$, we have

$$\mathscr{J}_L(\mathscr{D}) - \mathscr{J}(\mathscr{D}) = \sum_{[\mathbf{n}]\neq[\mathbf{0}]} \mathscr{J}_L(\mathscr{D}, \mathbf{n}). \tag{2.47}$$

The leading large L behaviour of the integrals $\mathscr{J}_L(\mathscr{D}, \mathbf{n})$ is described by the following theorem. It basically asserts that $\mathscr{J}_L(\mathscr{D}, \mathbf{n})$ falls off exponentially for non-trivial gauge fields $\mathbf{n}$, the rate $\varepsilon(\mathscr{G}, \mathbf{n})$ being determined by the topology of the diagram and the "strength" of the gauge field $\mathbf{n}$ (see Subsect. 2.6).

Theorem 2.4. *At large L, we have*

$$\ln \mathscr{J}_L(\mathscr{D}, \mathbf{n}) = -mL\varepsilon(\mathscr{G}, \mathbf{n}) + O(\ln L), \tag{2.48}$$

where $\varepsilon(\mathscr{G}, \mathbf{n})$ is given by

$$\varepsilon(\mathscr{G}, \mathbf{n}) = \min_x \Big\{ x(b)^0 - x(a)^0 + \sum_{\ell\in\mathscr{L}} |x(f(\ell)) - x(i(\ell)) + n(\ell)| \Big\}, \tag{2.49}$$

the minimum being taken over all possible positions $x(v)\in\mathbb{R}^4$ of the vertices $v\in\mathscr{V}$.

Proof. Using the heat kernel representation

$$\Delta(x; m) = \int_0^\infty dt (4\pi t)^{-2} \exp - \left(m^2 t + \frac{x^2}{4t} \right),$$

substituting $t \to tL/2m$, $x(v) \to Lx(v)$, and working out the vertex "factors," the integral (2.46) assumes the general form

$$\mathscr{J}_L(\mathscr{D}, \mathbf{n}) = \prod_{\ell \in \mathscr{L}} \int_0^\infty dt_\ell \prod_{v \in \mathscr{V}'} \int_{\mathbb{R}^4} d^4x(v) L^s \frac{P(t, x, \mathbf{n})}{Q(t)} e^{-mLR(t, x, \mathbf{n})}, \tag{2.50}$$

where s is some power, P and Q are polynomials and

$$R(t, x, \mathbf{n}) = x(b)^0 - x(a)^0 + \sum_{\ell \in \mathscr{L}} \tfrac{1}{2} \left\{ t_\ell + \frac{1}{t_\ell} [x(f(\ell)) - x(i(\ell)) + n(\ell)]^2 \right\}.$$

The integral (2.50) is of the saddle point type and can be evaluated, for large L, by expanding about the minima of R. This yields Eq. (2.48) with

$$\varepsilon(\mathscr{G}, \mathbf{n}) = \min_{t, x} R(t, x, \mathbf{n}).$$

Finally, Eq. (2.49) is obtained by performing the trivial minimization over the variables t_ℓ first. □

2.6. Properties of $\varepsilon(\mathscr{G}, \mathbf{n})$. The basic result is

Theorem 2.5. *Suppose* $\mathscr{C}_1, \ldots, \mathscr{C}_N$ *is a set of (pairwise) disjoint loops in* $\mathscr{G}$. *Then, we have*

$$\varepsilon(\mathscr{G}, \mathbf{n}) \geqq \frac{\sqrt{3}}{2} \sum_{j=1}^{N} |\mathbf{W}(\mathscr{C}_j, \mathbf{n})|, \tag{2.51}$$

where $\mathbf{W}(\mathscr{C}, \mathbf{n})$ *denotes the Wilson loop* (2.37).

Proof. According to Lemma 2.1, there exists a path $\mathscr{P}$ in $\mathscr{G}$ connecting a and b such that $\mathscr{P} \cap \mathscr{C}_j$ is either empty or a path in $\mathscr{G}$ for all $j = 1, \ldots, N$ (set $\mathscr{P} = \emptyset$ if $a = b$). To every line $\ell \in \mathscr{P}$ assign a number $[\mathscr{P} : \ell]$, which is 1 if the orientations of $\mathscr{P}$ and ℓ coincide, and -1 otherwise. Thus, $[\mathscr{P} : \ell]$ satisfies the same relations (2.35) as the loop orientation numbers $[\mathscr{C} : \ell]$ and, in addition,

$$[\mathscr{P} : \ell] = 1 \quad \text{if} \quad i(\ell) = a,$$
$$[\mathscr{P} : \ell] = -1 \quad \text{if} \quad f(\ell) = a.$$

Next, for every subset $\mathscr{S}$ of lines, define

$$\mathscr{E}(\mathscr{S}, \mathbf{n}) = \min_x \left\{ \sum_{\ell \in \mathscr{S} \cap \mathscr{P}} [\mathscr{P} : \ell] [x(f(\ell))^0 - x(i(\ell))^0] + \sum_{\ell \in \mathscr{S}} |x(f(\ell)) - x(i(\ell)) + n(\ell)| \right\}$$

($\mathscr{E}(\mathscr{S}, \mathbf{n}) = 0$ if $\mathscr{S} = \emptyset$). It is then easy to prove that

$$\varepsilon(\mathscr{G}, \mathbf{n}) = \mathscr{E}(\mathscr{L}, \mathbf{n}),$$
$$\mathscr{E}(\mathscr{S}_1, \mathbf{n}) \geqq \mathscr{E}(\mathscr{S}_2, \mathbf{n}) \quad \text{if} \quad \mathscr{S}_1 \supset \mathscr{S}_2,$$
$$\mathscr{E}(\mathscr{S}_1 \cup \mathscr{S}_2, \mathbf{n}) \geqq \mathscr{E}(\mathscr{S}_1, \mathbf{n}) + \mathscr{E}(\mathscr{S}_2, \mathbf{n}) \quad \text{if} \quad \mathscr{S}_1 \cap \mathscr{S}_2 = \emptyset.$$

In particular, we have

$$\varepsilon(\mathscr{G}, \mathbf{n}) \geqq \sum_{j=1}^{N} \mathscr{E}(\mathscr{C}_j, \mathbf{n}),$$

and (2.51) follows, if we can show that

$$\mathscr{E}(\mathscr{C}_j, \mathbf{n}) \geqq \frac{\sqrt{3}}{2} |\mathbf{W}(\mathscr{C}_j, \mathbf{n})|, \tag{2.52}$$

for all $j = 1, \ldots, N$.

Suppose first that $\mathscr{P} \cap \mathscr{C}_j = \emptyset$, and let $v_1, \ldots, v_M$ be consecutive vertices along $\mathscr{C}_j$. Then,

$$\mathscr{E}(\mathscr{C}_j, \mathbf{n}) = \min_x \left\{ \sum_{k=1}^{M} |x(v_{k+1}) - x(v_k) + [\mathscr{C}_j : \ell_k] n(\ell_k)| \right\},$$

where $v_{M+1} \equiv v_1$ and $\ell_k \in \mathscr{C}_j$ is the line with endpoints v_k, v_{k+1}. By repeated application of

$$|x - y + n| + |y - z + m| \geqq |x - z + n + m|, \tag{2.53}$$

one finds

$$\mathscr{E}(\mathscr{C}_j, \mathbf{n}) \geqq |\mathbf{W}(\mathscr{C}_j, \mathbf{n})|,$$

which is an even stronger inequality than (2.52).

Now consider the case $\mathscr{P} \cap \mathscr{C}_j \neq \emptyset$ and let $v_1, \ldots, v_M$ and $\ell_1, \ldots, \ell_M$ be as above. Since $\mathscr{P} \cap \mathscr{C}_j$ is a path contained in $\mathscr{C}_j$, it connects two vertices $v_r \neq v_s$. Thus, we have

$$\mathscr{E}(\mathscr{C}_j, \mathbf{n}) = \min_x \left\{ x(v_s)^0 - x(v_r)^0 + \sum_{k=1}^{M} |x(v_{k+1}) - x(v_k) + [\mathscr{C}_j : \ell_k] n(\ell_k)| \right\}.$$

Using (2.53), it follows that

$$\mathscr{E}(\mathscr{C}_j, \mathbf{n}) \geqq \min_x \{x^0 + |x| + |x + w|\},$$

where $w = (0, \mathbf{W}(\mathscr{C}_j, \mathbf{n}))$. Furthermore, applying the triangle inequality once more, the bound

$$\mathscr{E}(\mathscr{C}_j, \mathbf{n}) \geqq \min_{x^0} \{x^0 + \sqrt{(2x^0)^2 + \mathbf{w}^2}\},$$

is obtained, and the inequality (2.52) is now easily established by determining the minimum over x^0 by differentiation. □

An easy consequence of Theorem 2.5 is

Theorem 2.6. *If* $\mathbf{n}$ *is not a pure gauge configuration, we have* $\varepsilon(\mathscr{G}, \mathbf{n}) \geqq \sqrt{3}/2$.

Proof. Let T be a tree in $\mathscr{G}$ and choose the axial gauge for $\mathbf{n}$ (Subsect. 2.3). Because $\mathbf{n}$ is not a pure gauge configuration, there exists a line $\ell \in T^*$ such that $\mathbf{n}(\ell) \neq 0$. For the loop $\mathscr{C}$ with $\mathscr{C} \backslash \{\ell\} \subset T$, we therefore have $|\mathbf{W}(\mathscr{C}, \mathbf{n})| \geqq 1$, which by Theorem 2.5, implies $\varepsilon(\mathscr{G}, \mathbf{n}) \geqq \sqrt{3}/2$. □

The class of gauge fields $\mathbf{n}$, which make the leading contribution to the sum (2.47) at large L, is identified in the following

Theorem 2.7. *Suppose* $\mathbf{n}$ *is not a pure gauge configuration and* $\varepsilon(\mathscr{G}, \mathbf{n}) < \sqrt{3/2}$. *Then,* $\mathbf{n}$ *is a simple gauge field.*

The proof of this theorem is complicated and is therefore divided into digestible pieces. In what follows, we assume that $\mathbf{n}$ is a gauge field on $\mathscr{G}$, which is not a pure gauge configuration and which satisfies $\varepsilon(\mathscr{G}, \mathbf{n}) < \sqrt{3/2}$.

Lemma 2.8. *Let T be a tree in $\mathscr{G}$ and choose the axial gauge for $\mathbf{n}$. Then there exists an integer unit vector $\mathbf{e}$ and numbers $s(\ell) \in \{0, 1, -1\}$ such that $\mathbf{n}(\ell) = s(\ell)\mathbf{e}$ for all $\ell \in T^*$.*

Proof. Suppose $\ell \in T^*$ and let $\mathscr{C}$ be the loop in $\mathscr{G}$ with $\mathscr{C} \backslash \{\ell\} \subset T$. Then,

$$|\mathbf{W}(\mathscr{C}, \mathbf{n})| = |\mathbf{n}(\ell)|,$$

and since $\varepsilon(\mathscr{G}, \mathbf{n}) < \sqrt{3/2}$, Theorem 2.5 implies

$$|\mathbf{n}(\ell)| \in \{0, 1\}.$$

Let

$$T^* = \{\ell_1, \ldots, \ell_N\}, \tag{2.54}$$

be a labelling of the lines in T^* such that

$$\begin{aligned} |\mathbf{n}(\ell_j)| &= 1 \quad \text{for} \quad j = 1, \ldots, M, \\ &= 0 \quad \text{otherwise}. \end{aligned} \tag{2.55}$$

Because $\mathbf{n}$ is not a pure gauge configuration, we have $M \geqq 1$.

If $M = 1$, there is nothing left to prove. On the other hand, if $M \geqq 2$, we must show that

$$\mathbf{n}(\ell_i) = \pm \mathbf{n}(\ell_j) \quad \text{for all} \quad i, j \leqq M, i \neq j. \tag{2.56}$$

Let $\mathscr{C}_i$ be the loop passing through ℓ_i with $\mathscr{C}_i \backslash \{\ell_i\} \subset T$. Then, $\mathscr{C}_i \cap \mathscr{C}_j \neq \emptyset$, because Theorem 2.5 would otherwise imply $\varepsilon(\mathscr{G}, \mathbf{n}) \geqq \sqrt{3}$. $\mathscr{C}_i \backslash \{\ell_i\}$, $\mathscr{C}_j \backslash \{\ell_j\}$, and hence $\mathscr{C}_i \cap \mathscr{C}_j$ are paths contained in T. We can therefore construct a composed loop $\mathscr{C}_{ij}$ by

$$\mathscr{C}_{ij} = (\mathscr{C}_i \cup \mathscr{C}_j) \backslash (\mathscr{C}_i \cap \mathscr{C}_j). \tag{2.57}$$

$\mathscr{C}_{ij}$ passes through ℓ_i, ℓ_j and no other $\ell \in T^*$. Thus,

$$|\mathbf{W}(\mathscr{C}_{ij}, \mathbf{n})| = |\mathbf{n}(\ell_i) \pm \mathbf{n}(\ell_j)|,$$

where the sign depends on the relative orientation of $\mathscr{C}_i$ and $\mathscr{C}_j$ in $\mathscr{C}_{ij}$. Applying Theorem 2.5 once more, we have

$$|\mathbf{n}(\ell_i) \pm \mathbf{n}(\ell_j)| \in \{0, 1\},$$

and since $\mathbf{n}(\ell_i)$ and $\mathbf{n}(\ell_j)$ are integer unit vectors, (2.56) follows. □

In the following discussion, T always denotes a tree in $\mathscr{G}$ and $\mathbf{n}$ is assumed to be in the axial gauge. Furthermore, the elements of T^* are labelled as in Eqs. (2.54), (2.55). For $j = 1, \ldots, M$, define vertices u_j, v_j through

$$\begin{aligned} u_j &= i(\ell_j), \; v_j = f(\ell_j) \quad \text{if} \quad \mathbf{n}(\ell_j) = \mathbf{n}(\ell_1), \\ u_j &= f(\ell_j), \; v_j = i(\ell_j) \quad \text{if} \quad \mathbf{n}(\ell_j) = -\mathbf{n}(\ell_1). \end{aligned} \tag{2.58}$$

It is also helpful to introduce the reduced graph

$$\hat{\mathscr{G}} = \mathscr{G} \backslash \{\ell_1, \ldots, \ell_M\}, \tag{2.59}$$

which contains the tree T and is hence connected. By definition, $\mathbf{n}(\ell) = 0$ for all lines ℓ in $\hat{\mathscr{G}}$.

Lemma 2.9. *If $M \geqq 2$, the following statements hold.*

(1) *$u_i \neq v_j$ for all $i, j \leqq M$.*

(2) *Suppose $\mathscr{P}_1, \mathscr{P}_2$ are paths in $\hat{\mathscr{G}}$ connecting u_1, v_1 and u_2, v_2, respectively. Then, $\mathscr{P}_1 \cap \mathscr{P}_2 \neq \emptyset$.*

(3) *Suppose $\mathscr{P}_1, \mathscr{P}_2$ are paths in $\hat{\mathscr{G}}$ connecting u_1, v_2 and u_2, v_1, respectively. Then, $\mathscr{P}_1 \cap \mathscr{P}_2 \neq \emptyset$.*

Proof. (1) As above, let $\mathscr{C}_j$ be the loop passing through ℓ_j with $\mathscr{C}_j \backslash \{\ell_j\} \subset T$. Suppose $u_i = v_i$ for some i. Then, ℓ_i is a loop line and $\mathscr{C}_i \cap \mathscr{C}_j = \emptyset$ for all $j \neq i$. Since $M \geqq 2$, such j exist and Theorem 2.5 implies $\varepsilon(\mathscr{G}, \mathbf{n}) \geqq \sqrt{3}$, which contradicts our assumptions. Thus, $u_i \neq v_i$ for all i.

Suppose now that $u_i = v_j$ for some $i \neq j$. Then, the composed loop $\mathscr{C}_{ij}$ (Eq. (2.57)) passes through ℓ_i and ℓ_j in such a way that $|\mathbf{W}(\mathscr{C}_{ij}, \mathbf{n})| = 2$, which also leads to $\varepsilon(\mathscr{G}, \mathbf{n}) \geqq \sqrt{3}$. Thus, $u_i \neq v_j$ for all $i \neq j$.

(2) If $\mathscr{P}_1$ and $\mathscr{P}_2$ were disjoint, the loops $\mathscr{P}_1 \cup \{\ell_1\}$ and $\mathscr{P}_2 \cup \{\ell_2\}$ would also be disjoint, and hence $\varepsilon(\mathscr{G}, \mathbf{n}) \geqq \sqrt{3}$ by Theorem 2.5, which is a contradiction.

(3) We again assume $\mathscr{P}_1 \cap \mathscr{P}_2 = \emptyset$ and show that this leads to a contradiction. If $\mathscr{P}_1$ and $\mathscr{P}_2$ do not cross (i.e. if there is no vertex, which is an endpoint of a line in $\mathscr{P}_1$ and of another line in $\mathscr{P}_2$), the set $\mathscr{C} = \mathscr{P}_1 \cup \mathscr{P}_2 \cup \{\ell_1, \ell_2\}$ is a loop in $\mathscr{G}$. Furthermore, the orientations of ℓ_1 and ℓ_2 in $\mathscr{C}$ are such that $|\mathbf{W}(\mathscr{C}, \mathbf{n})| = 2$, and hence $\varepsilon(\mathscr{G}, \mathbf{n}) \geqq \sqrt{3}$, which is a contradiction.

If $\mathscr{P}_1$ and $\mathscr{P}_2$ cross, there exists a vertex z, which is an endpoint of some lines in $\mathscr{P}_1$ and $\mathscr{P}_2$. This vertex divides $\mathscr{P}_1$ into two paths $\mathscr{P}_1(u_1 : z)$ and $\mathscr{P}_1(z : v_2)$ connecting u_1, z and z, v_2, respectively (it is possible that z coincides with u_1, for example, in which case we set $\mathscr{P}_1(u_1 : z) = \emptyset$). Similarly, $\mathscr{P}_2$ divides into $\mathscr{P}_2(u_2 : z)$ and $\mathscr{P}_2(z : v_1)$. Because $\mathscr{P}_1$ and $\mathscr{P}_2$ are disjoint, so are the sets

$$\mathscr{L}_1 = \mathscr{P}_1(u_1 : z) \cup \mathscr{P}_2(z : v_1),$$
$$\mathscr{L}_2 = \mathscr{P}_2(u_2 : z) \cup \mathscr{P}_1(z : v_2).$$

$\mathscr{L}_1$ itself may not be a path, but it contains a path $\mathscr{Q}_1$ connecting u_1 and v_1 (note that $\mathscr{L}_1 \neq \emptyset$, because $u_1 \neq v_1$ by (1)). Similarly, there is a path $\mathscr{Q}_2$ contained in $\mathscr{L}_2$, which connects u_2 and v_2. Since $\mathscr{Q}_1 \cap \mathscr{Q}_2 = \emptyset$, there is a contradiction with (2), which has already been established above. □

Lemma 2.10. *If $M \geqq 2$, there exists a line $\hat{\ell}$ in $\hat{\mathscr{G}}$ such that $\hat{\mathscr{G}} \backslash \{\hat{\ell}\}$ decomposes into two disconnected components, one, denoted $\hat{\mathscr{G}}_u$, containing the vertices u_1, u_2 and the other, $\hat{\mathscr{G}}_v$, containing v_1, v_2.*

Proof. We assume that such a line $\hat{\ell}$ does not exist and derive a contradiction. Suppose we add 2 extra vertices U, V to $\hat{\mathscr{G}}$ and 4 extra lines connecting U with u_1, u_2 and V with v_1, v_2. The absence of $\hat{\ell}$ implies that this enlarged graph is 1-particle

irreducible between U and V. Thus, by Theorem 2.2 there exist two disjoint paths $\tilde{\mathscr{P}}_1, \tilde{\mathscr{P}}_2$ connecting U and V. Deleting the extra vertices and lines again, $\tilde{\mathscr{P}}_1$ and $\tilde{\mathscr{P}}_2$ are reduced to some disjoint paths $\mathscr{P}_1, \mathscr{P}_2$ in $\hat{\mathscr{G}}$ each of them connecting some u_i with some v_j with $i,j \in \{1,2\}$ (note that because of (1) in Lemma 2.9, the sets $\mathscr{P}_1, \mathscr{P}_2$ are not empty). Necessarily, the situation is then as in (2) or (3) of Lemma 2.9 so that $\mathscr{P}_1 \cap \mathscr{P}_2 = \emptyset$ is a contradiction. □

The proof of Theorem 2.7 is now easy to complete. Choosing the tree T and all the other notation as above, we need only consider the case $M \geqq 2$, because for $M=1$, Eq. (2.55) already implies that $\mathbf{n}$ is simple. Let $\hat{\ell}$ be the line in $\hat{\mathscr{G}}$, whose existence is guaranteed by Lemma 2.10. $\hat{\ell}$ is an element of T ($\hat{\mathscr{G}} \backslash \{\hat{\ell}\}$ would otherwise be connected). Thus, a new tree T' can be defined through

$$T' = (T \backslash \{\hat{\ell}\}) \cup \{\ell_1\}.$$

Relative to this tree, $\mathbf{n}$ is not in the axial gauge, because $\mathbf{n}(\ell_1) \neq 0$. Set

$$\lambda(w) = \mathbf{n}(\ell_1) \quad \text{if } w \text{ is a vertex of } \hat{\mathscr{G}}_u,$$
$$\lambda(w) = 0 \quad \text{if } w \text{ is a vertex of } \hat{\mathscr{G}}_v,$$

and let $\mathbf{n}'$ be the gauge transform of $\mathbf{n}$ by λ (cp. Eq. (2.36)). It is trivial to show that $\mathbf{n}'$ is in the axial gauge relative to T'. Furthermore, we also have $\mathbf{n}'(\ell_2) = 0$ and $\mathbf{n}'(\ell_j) = 0$ for all $j = M+1, \ldots, N$, because these latter lines belong either to $\hat{\mathscr{G}}_u$ or to $\hat{\mathscr{G}}_v$. The number M' of elements ℓ of T'^* with $\mathbf{n}'(\ell) \neq 0$ is therefore strictly smaller than M.

The procedure leading from T to T' can now be iterated and a series of trees is obtained with decreasing numbers of lines ℓ, where $\mathbf{n}(\ell) \neq 0$. After finitely many iterations, this number will have decreased to 1 and $\mathbf{n}$ is thus found to be a simple gauge field. □

2.7. Proof of the Mass Shift Formula (Final Steps). As already explained in Subsect. 2.2, the proof of the mass shift formula (2.22) proceeds via Eq. (2.31), a relation, which we can now prove to all orders of the Feynman diagram expansion using the results on the large L behaviour of self-energy Feynman integrals obtained above. Taken together, these results imply

$$\mathscr{J}_L(\mathscr{D}, \mathbf{n}) = O\left(e^{-\frac{\sqrt{3}}{2} mL}\right) \text{ if } \mathbf{n} \text{ is not pure gauge}, \tag{2.60}$$

$$\mathscr{J}_L(\mathscr{D}) - \mathscr{J}(\mathscr{D}) = \sum_{[\mathbf{n}] \text{ simple}} \mathscr{J}_L(\mathscr{D}, \mathbf{n}) + O(e^{-\bar{m}L}) \tag{2.61}$$

for all diagrams $\mathscr{D}$, in particular, Eq. (2.28) (and, similarly, Eq. (2.29)) follows immediately. After summing over all diagrams, Eq. (2.61) becomes

$$\Sigma_L(\hat{p}) = \sum_{\mathscr{D}} \sum_{[\mathbf{n}] \text{ simple}} \mathscr{J}_L(\mathscr{D}, \mathbf{n}) + O(e^{-\bar{m}L}), \tag{2.62}$$

where $\Sigma(\hat{p}) = 0$ has been used. Now recall that the simple classes $[\mathbf{n}]$ can be labelled by the set of fields (2.39) so that Eq. (2.62) may be rewritten in the form

$$\Sigma_L(\hat{p}) = \sum_{\mathscr{D}} \sum_{\ell \in \mathscr{D}} \hat{\mathscr{J}}_L(\mathscr{D}, \ell) + O(e^{-\bar{m}L}), \tag{2.63}$$

where ℓ runs over all lines in $\mathscr{D}$, which are contained in at least one loop and which are independent of each other (i.e. lines carrying the same momentum for all configurations of loop momenta are counted as one independent line). In momentum space, $\hat{\mathscr{J}}_L(\mathscr{D},\ell)$ is exactly equal to the infinite volume Feynman integral associated to $\mathscr{D}$, except that the integrand is multiplied by the extra factor

$$2\sum_{j=1}^{3}\cos(p_jL), \qquad p\text{: momentum flowing through } \ell. \tag{2.64}$$

These integrals are thus exactly of the type shown in Fig. 4 and, without great difficulties, one can prove that in fact the series (2.63) matches term by term with the Feynman diagram expansion of the right-hand side of Eq. (2.31).

Having established the basic relation (2.31), we now proceed to evaluate the integrals I_1, I_2, and I_3 for large L using complex contour integration. Consider first the simplest case, the integral I_3. Due to rotational invariance, it can be written as

$$I_3 = 6\int \frac{d^4q}{(2\pi)^4}\, e^{iq_1L} G(q)\Gamma(\hat{p},q,-\hat{p},-q). \tag{2.65}$$

By Theorem 2.3, $\Gamma(\hat{p},q,-\hat{p},-q)$ has no singularities in the complex q_1-plane for $0\leqq \mathrm{Im}\, q_1 < m\sqrt{3}$ (and real q_0, q_2, q_3). The propagator $G(q)$ is also analytic in this domain except for the meson pole at

$$q_1 = i\sqrt{m^2+q_\perp^2+q_0^2}, \qquad q_\perp = (q_2,q_3). \tag{2.66}$$

If we now shift the q_1 integration path from the real line to the line $\mathrm{Im}\, q_1 = m\sqrt{7}/2$, one obtains two terms, one from the meson pole (2.66) and the other from the integral along the new integration path. The latter contribution is more rapidly decaying at large L than the error term in Eq. (2.31) and is therefore negligible. Thus, we have

$$I_3 = 6\int_{\mathbb{B}} \frac{dq_0 d^2q_\perp}{(2\pi)^3 2|q_1|}\, e^{-|q_1|L}\Gamma(\hat{p},q,-\hat{p},-q) + O(e^{-\tilde{m}L}), \tag{2.61}$$

where q_1 is given by Eq. (2.66) and $\mathbb{B}$ is the ball

$$\mathbb{B} = \{(q_0,q_\perp)\in\mathbb{R}^3 \,|\, q_0^2+q_\perp^2 \leqq \tfrac{3}{4}m^2\} \tag{2.68}$$

(only when $(q_0,q_\perp)\in\mathbb{B}$ is the meson pole inside the strip $0\leqq \mathrm{Im}\, q_1 \leqq m\sqrt{7}/2$). Note that q is now on the meson mass shell and that, contrary to the oscillatory integral (2.65), the representation (2.67) immediately reveals what the large L behaviour of I_3 is.

The integral I_2 can be treated in exactly the same way as I_3 and one obtains

$$I_2 = 6\int_{\mathbb{B}} \frac{dq_0 d^2q_\perp}{(2\pi)^3 2|q_1|}\, e^{-|q_1|L} G(0)\Gamma(q,-q,0)\Gamma(\hat{p},-\hat{p},0) + O(e^{-\tilde{m}L}). \tag{2.69}$$

The integral I_1, on the other hand, is more complicated. There are two meson poles in this case located at

$$q_1^{\pm} = i\sqrt{m^2+q_\perp^2+\left(q_0\pm\frac{i}{2}m\right)^2}. \tag{2.70}$$

They are not purely imaginary and happen to coincide for $q_0=0$, which is a potential source of difficulty below. To avoid it, we first deform the q_0 integration path around $q_0=0$ into an infinitesimal half-circle in the complex lower half-plane. Then, noting

$$\begin{aligned} &0<\mathrm{Im}\, q_1^{\pm}<m\sqrt{7}/2 \quad \text{if} \quad (q_0, q_\perp)\in \mathbb{B}, \\ &\mathrm{Im}\, q_1^{\pm}>m\sqrt{3/2} \quad \text{otherwise}, \end{aligned} \tag{2.71}$$

the shift of the q_1 integration path can be performed as above and one obtains

$$I_1=I_1^+ + I_1^- + O(e^{-\bar{m}L}), \tag{2.72}$$

$$I_1^{\pm}=6\int_{\mathbb{B}} \frac{dq_0 d^2q_\perp}{(2\pi)^3}\left\{\frac{i}{2q_1} e^{iq_1 L} G(\tfrac{1}{2}\hat{p}\mp q)\right.$$

$$\left.\cdot\, \Gamma(-\hat{p}, q+\tfrac{1}{2}\hat{p}, -q+\tfrac{1}{2}\hat{p})\Gamma(\hat{p}, -q-\tfrac{1}{2}\hat{p}, q-\tfrac{1}{2}\hat{p})\right\}_{q_1=q_1^{\pm}}. \tag{2.73}$$

The integrals $I_1^{\pm}$ are not yet of the desired saddle point type (as Eq. (2.67), for example) and further contour shifting is needed. Consider first the integral I_1^+. At fixed $q_\perp$, $q_\perp^2<\frac{3}{4}m^2$, the q_0 integration path is along the real line from point A to point D in Fig. 5, A and D being characterized by

$$q_0=\pm\sqrt{\tfrac{3}{4}m^2-q_\perp^2}. \tag{2.74}$$

We now deform this integration path to the curve $ABCD$ shown in Fig. 5a, which is possible because q_1^+ and the other entries in the integrand are analytic inside the rectangle $ABCD$, as one may easily show using Theorem 2.3. Note that the meson pole of the propagator in the integrand is at $q_0=0$, which is outside the integration contour. After the deformation of the integration path, I_1^+ is a sum of 3 integrals corresponding to the straight lines AB, BC, and CD. The contribution of AB and CD is however negligible at large L, because $\mathrm{Im}\, q_1^+ \geqq m\sqrt{3/2}$ along these lines. In other words, the result is

$$I_1^+=6\int_{\mathbb{B}} \frac{dq_0 d^2q_\perp}{(2\pi)^3 2|q_1|} e^{-|q_1|L} G(\hat{p}-q)\cdot\Gamma(-\hat{p}, q, -q+\hat{p})\Gamma(\hat{p}, -q, q-\hat{p})+O(e^{-\bar{m}L}), \tag{2.75}$$

where q_1 is given by Eq.(2.66) $\left(\text{we have replaced } q_0 \text{ by } q_0-\frac{i}{2}m,\ q_0 \text{ real, along } BC\right)$.

The integral I_1^- can be treated similarly, the integration contour being displayed in Fig. 5b. The only difference is that now we also get a contribution from the meson pole at $q_0=0$ so that altogether we have

$$I_1^-=6\int_{\mathbb{B}'} \frac{d^2q_\perp}{(2\pi)^2 2|q_1'|} e^{-|q_1'|L}\frac{\lambda^2}{2m}$$

$$+6\int_{\mathbb{B}} \frac{dq_0 d^2q_\perp}{(2\pi)^3 2|q_1|} e^{-|q_1|L} G(\hat{p}+q)\Gamma(-\hat{p}, q+\hat{p}, -q)\Gamma(\hat{p}, -q-\hat{p}, q)+O(e^{-\bar{m}L}), \tag{2.76}$$

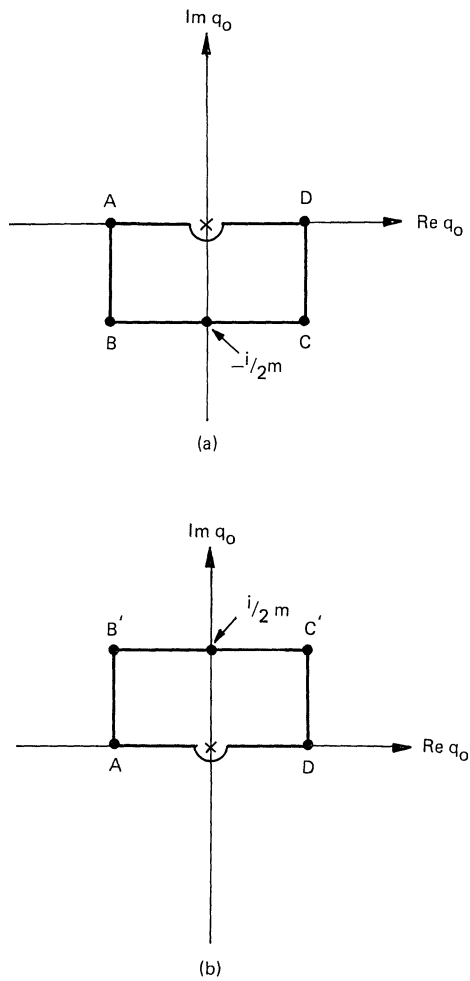

Fig. 5a and b. Integration contours in the complex q_0-plane, (**a**) for the calculation of I_1^+, (**b**) for I_1^-

where q_1 is given by Eq. (2.66), λ by Eq. (2.20) and

$$\begin{aligned} q_1' &= i\sqrt{\tfrac{3}{4}m^2 + q_\perp^2}\,, \\ \mathbb{B}' &= \{q_\perp \in \mathbb{R}^2 \mid q_\perp^2 \leqq \tfrac{3}{4}m^2\}\,. \end{aligned} \tag{2.77}$$

If we now add up all the integrals, the integrands combine to the forward scattering amplitude $F(v)$ defined in Subsect. 2.1, and we end up with

$$\begin{aligned} \Sigma_L(\hat{p}) = 3 \int_{\mathbb{B}'} \frac{d^2 q_\perp}{(2\pi)^2 2|q_1'|} e^{-|q_1'|L} \frac{\lambda^2}{2m} \\ + 3 \int_{\mathbb{B}} \frac{dq_0 d^2 q_\perp}{(2\pi)^3 2|q_1|} e^{-|q_1|L} F(-iq_0) + O(e^{-\tilde{m}L}). \end{aligned} \tag{2.78}$$

In this equation, we may just as well integrate over all real q_0 and $q_\perp$, the difference being of the same order as the error term at large L. Noting

$$\int \frac{d^2 q_\perp}{(2\pi)^2} \frac{1}{2\sqrt{\mu^2+q_\perp^2}} e^{-\sqrt{\mu^2+q_\perp^2}L} = \frac{1}{4\pi L} e^{-\mu L}, \tag{2.79}$$

we finally get

$$\Sigma_L(\hat{p}) = \frac{3\lambda^2}{8\pi m L} e^{-\frac{\sqrt{3}}{2}mL} + \frac{3}{8\pi^2 L} \int dy\, e^{-\sqrt{m^2+y^2}L} F(iy) + O(e^{-\tilde{m}L}), \tag{2.80}$$

which, in view of Eq. (2.30), agrees with the mass shift formula (2.22) and thus concludes the proof of this relation.

3. Volume Dependence of Bound State Masses

3.1. The Non-Relativistic Case. For the physical interpretation of the relativistic formulae derived later, it is useful to first consider the case of two non-relativistic bosons ("mesons") of mass m, which form a bound state of mass

$$m_B = 2m - E_B, \quad E_B > 0, \tag{3.1}$$

and spin 0 (the mesons are also assumed to be spinless). The hamiltonian $\mathbb{H}$ describing this system is an operator acting on scalar wave functions $\psi(\mathbf{x}, \mathbf{y})$, which are invariant under an interchange of the particle coordinates $\mathbf{x}$ and $\mathbf{y}$. Explicitly, we choose $\mathbb{H}$ to be of the form

$$\mathbb{H} = -\frac{1}{2m}(\Delta_x + \Delta_y) + V(\mathbf{x}-\mathbf{y}), \tag{3.2}$$

where Δ_x, Δ_y denote the Laplace operators with respect to $\mathbf{x}$ and $\mathbf{y}$. The potential V is assumed to be square integrable, rotationally symmetric and of finite range, i.e.

$$V(\mathbf{z}) = 0 \quad \text{for} \quad |\mathbf{z}| > R. \tag{3.3}$$

These assumptions are made for simplicity and could easily be relaxed without affecting the main results obtained below.

In infinite volume and at zero total momentum, the bound state wave function ψ_B depends only on the distance $r = |\mathbf{x}-\mathbf{y}|$ of the bound particles. It satisfies the Schrödinger equation

$$\mathbb{H}\psi_B = -E_B \psi_B \tag{3.4}$$

and falls off exponentially at large r:

$$\psi_B(r) = A \frac{e^{-\varkappa r}}{r} \quad (r > R), \tag{3.5}$$

$$\varkappa = \sqrt{m E_B}. \tag{3.6}$$

As is well-known, bound states give rise to poles in the analytically continued forward elastic meson scattering amplitude $F_{nr}(E)$, which is defined in terms of the

full (non-relativistic) amplitude $T_{nr}(\mathbf{p}', \mathbf{q}'|\mathbf{p}, \mathbf{q})$ through [5]

$$F_{nr}(E) = T_{nr}(\mathbf{p}, -\mathbf{p}|\mathbf{p}, -\mathbf{p}), \qquad E = \frac{\mathbf{p}^2}{m}. \tag{3.7}$$

If we choose the normalization

$$\int d^3z |\psi_B(|\mathbf{z}|)|^2 = 1, \tag{3.8}$$

the residue of $F_{nr}(E)$ at $E = -E_B$ is given by

$$\lim_{E \to -E_B} (E + E_B) F_{nr}(E) = \frac{32\pi^2}{m^2} |A|^2, \tag{3.9}$$

i.e. up to an irrelevant phase, the behaviour of ψ_B at large r is entirely determined by spectral data.

Suppose now that the mesons are enclosed in a box of size L and periodic boundary conditions. Then, the corresponding wave functions $\psi(\mathbf{x}, \mathbf{y})$ are periodic in both coordinates $\mathbf{x}, \mathbf{y}$ and the interaction potential V has to be replaced by

$$V_L(\mathbf{z}) = \sum_{\mathbf{n} \in \mathbb{Z}^3} V(\mathbf{z} + \mathbf{n}L) \tag{3.10}$$

to account for the interactions of a meson with the mirror images of its partner. Note that because V is of finite range, there are only a finite number of non-zero terms in the series (3.10) (at most one if $L > 2R$).

For large L, the finite volume Schrödinger equation

$$\mathbb{H}_L \psi = -E\psi, \tag{3.11}$$

$$\mathbb{H}_L = -\frac{1}{2m}(\Delta_x + \Delta_y) + V_L(\mathbf{x} - \mathbf{y}), \tag{3.12}$$

has a solution with $E \simeq E_B$ and $\psi \simeq \psi_B$. An asymptotic formula for the corresponding mass shift

$$\Delta m_B = E_B - E \tag{3.13}$$

can be derived as follows. Define

$$\psi_0(\mathbf{z}) = \sum_{\mathbf{n} \in \mathbb{Z}^3} \psi_B(|\mathbf{z} + \mathbf{n}L|), \qquad \mathbf{z} = \mathbf{x} - \mathbf{y}. \tag{3.14}$$

ψ_0 is periodic and hence an admissable finite volume wave function. Furthermore,

$$\mathbb{H}_L \psi_0 = -E_B \psi_0 + \eta, \tag{3.15}$$

$$\eta(\mathbf{z}) = \sum_{\mathbf{n} \neq \mathbf{n}'} V(\mathbf{z} + \mathbf{n}L) \psi_B(|\mathbf{z} + \mathbf{n}'L|) = O(e^{-\varkappa L}). \tag{3.16}$$

Thus, for large L, ψ_0 is almost a solution of the Schrödinger equation, and, with an appropriate normalization of the true solution ψ, one can show that

$$\psi = \psi_0 + O(e^{-\varkappa L}). \tag{3.17}$$

---

5 The sign and normalization of T_{nr} is chosen such that the non-relativistic optical theorem reads $\mathrm{Im} F_{nr}(E) = -\sqrt{E/m}\, \sigma_{\mathrm{tot}}(E)$. With this convention, the Born approximation is given by $T_{nr}(\mathbf{p}', -\mathbf{p}'|\mathbf{p}, -\mathbf{p}) = \tilde{V}(\mathbf{p}' - \mathbf{p}) + \tilde{V}(\mathbf{p}' + \mathbf{p})$, where $\tilde{V}(\mathbf{k}) = \int d^3z\, e^{-i\mathbf{k}\cdot\mathbf{z}} V(\mathbf{z})$

Taking the scalar product of ψ with Eq. (3.15) then leads to

$$\Delta m_B = \sum_{|\mathbf{n}|=1} \int d^3z \, \psi_B(|\mathbf{z}|)^* V(\mathbf{z}) \psi_B(|\mathbf{z}+\mathbf{n}L|) + O(e^{-\sqrt{2}\varkappa L}). \tag{3.18}$$

Finally, using Eqs. (3.3)–(3.6), the integral can be evaluated and the result

$$\Delta m_B = -24\pi |A|^2 \frac{e^{-\varkappa L}}{mL} + O(e^{-\sqrt{2}\varkappa L}) \tag{3.19}$$

is obtained.

In view of the decay properties of the bound state wave function ψ_B, it is no surprise that a formula like (3.19) holds. Still, it is remarkable that the detailed form of the potential V is irrelevant to the final result, which only refers to the particle masses m, m_B and the residue of the forward elastic scattering amplitude F_{nr} at the bound state pole (cf. Eqs. (3.6), (3.9)). Another interesting feature of Eq. (3.19) is that Δm_B is always negative, i.e. S-wave bound states get lighter when squeezed.

3.2. Calculation of Δm_B in Quantum Field Theory. The class of quantum field theories considered here is similar to the one discussed in Sect. 2, except that now we assume that the spectrum of stable particles contains an additional particle with mass $m_B < 2m$, which can decay virtually into two mesons. As far as the mesons are concerned, the notation of Subsect. 2.1 is taken over. For the bound state, we assume that there is a (euclidean) interpolating field χ normalized such that

$$\langle \chi(x) \chi(0) \rangle = \int \frac{d^4p}{(2\pi)^4} e^{ipx} G^\chi(p), \tag{3.20}$$

$$G^\chi(p)^{-1} = m_B^2 + p^2 - \Sigma^\chi(p), \tag{3.21}$$

$$\Sigma^\chi(p) = \frac{\partial}{\partial p^\mu} \Sigma^\chi(p) = 0 \quad \text{for} \quad p^2 = -m_B^2. \tag{3.22}$$

Vertex functions of r meson fields ϕ and s bound state fields χ are denoted by $\Gamma(p_1, \ldots, p_r; k_1, \ldots, k_s)$ (cp. Fig. 6). For simplicity, we shall furthermore assume that the transformation

$$\phi \to -\phi, \quad \chi \to \chi \tag{3.23}$$

is a symmetry of the theory so that

$$\Gamma(p_1, \ldots, p_r; k_1, \ldots, k_s) = 0 \quad \text{for } r \text{ odd}, \quad \langle \phi(x) \chi(0) \rangle = 0. \tag{3.24}$$

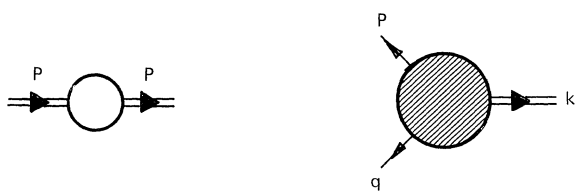

Fig. 6a and b. Graphical symbols used for (**a**) the full bound state propagator $G^\chi(p)$, and (**b**) the mixed 3-point vertex function $\Gamma(p, q; k)$

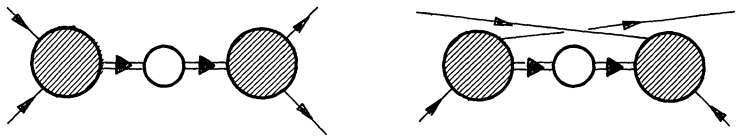

Fig. 7. Bound state exchange diagrams giving rise to poles in the meson scattering amplitude

In particular, the coupling constant λ vanishes and the bound state has no virtual decays into an odd number of mesons.

As in the non-relativistic case, the bound state gives rise to poles in the forward elastic meson scattering amplitude $F(v)$. They stem from the skeleton diagrams of Fig. 7 and are located at

$$v = \pm v_B, \qquad v_B = \frac{1}{2m}(m_B^2 - 2m^2). \tag{3.25}$$

The associated residue is

$$\lim_{v \to \pm v_B} (v^2 - v_B^2) F(v) = -\frac{v_B}{m} g^2, \tag{3.26}$$

where the $\phi\phi\chi$-coupling constant g is defined by

$$g = \Gamma(p, q; k), \quad p + q + k = 0, \quad p^2 = q^2 = -m^2, \quad k^2 = -m_B^2. \tag{3.27}$$

In finite volume, the meson mass shift Δm can be calculated as in the simple scalar theories considered in Sect. 2. The only difference is that now the Feynman diagrams also involve bound state propagators and the basic expression (2.31) for the self-energy $\Sigma_L(\hat{p})$ has to be modified accordingly. If we restrict ourselves to the case of small binding energies,

$$\sqrt{2}m < m_B < 2m, \tag{3.28}$$

the result of the Feynman diagram analysis is

$$\Sigma_L(\hat{p}) = I_1' + \tfrac{1}{2}(I_2' + I_3) + O(e^{-\bar{m}L}), \tag{3.29}$$

where $\bar{m}$ and I_3 are as before (Eqs. (2.23), (2.34)) and I_1', I_2' are the integrals graphically represented by Fig. 8a, b. Evaluating the integrals as in Subsect. 2.7, one finds

$$\Delta m = -\frac{3}{16\pi^2 m L} \int_{-\infty}^{\infty} dy\, e^{-\sqrt{m^2+y^2}L} F(iy) + O(e^{-\bar{m}L}), \tag{3.30}$$

which is identical with Eq. (2.22). In other words, the presence of the bound state has no influence on the leading finite size mass shift of the meson, in particular,

$$\Delta m = O(e^{-mL}) \tag{3.31}$$

for all m_B in the range (3.28).

We now proceed to calculate the bound state mass shift Δm_B. As for the mesons, one shows that

$$\Delta m_B = -\frac{1}{2m_B} \Sigma_L^{\chi}(\hat{p}_B) + O((\Delta m_B)^2), \tag{3.32}$$

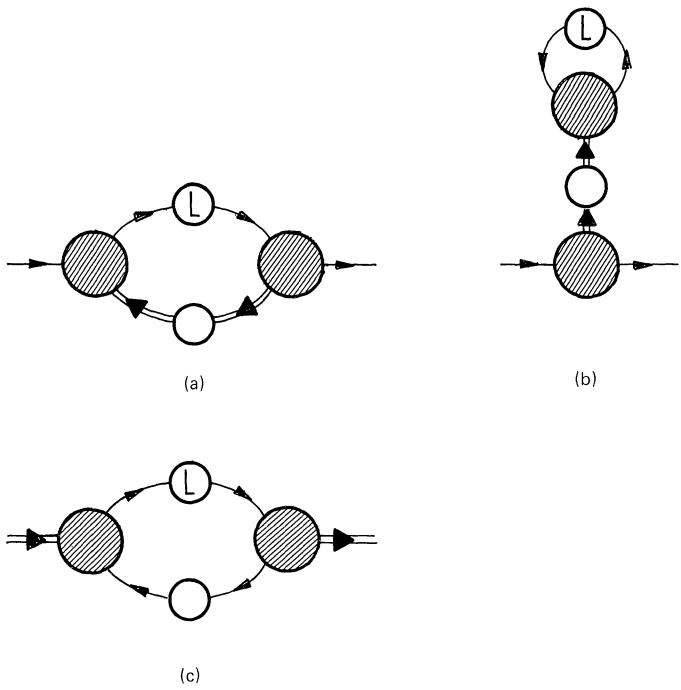

Fig. 8a–c. Graphical representation of the integrals I'_1, I'_2 contributing to $\sum_L(\hat{p})$ (diagrams (**a**) and (**b**)). Diagram (**c**) represents the integral I_1^χ, which is the leading term in the large L expansion of $\sum_L^\chi(\hat{p}_B)$

where $\sum_L^\chi$ is the finite volume self-energy of the χ-field and

$$\hat{p}_B=(\mathrm{i}m_B, 0, 0, 0)\,. \tag{3.33}$$

Summing up the leading Feynman diagrams contributing to $\sum_L^\chi(\hat{p}_B)$, we have

$$\sum_L^\chi(\hat{p}_B)=\tfrac{1}{2}I_1^\chi+O(e^{-\bar{m}_B L})\,, \tag{3.34}$$

the integral I_1^χ being graphically represented by the skeleton diagram shown in Fig. 8c. Provided m_B is in the interval (3.28), the error term in Eq. (3.34) is given by

$$\bar{m}_B \geqq \sqrt{2}\mu\,, \tag{3.35}$$

$$\mu=\sqrt{m^2-\tfrac{1}{4}m_B^2}\,. \tag{3.36}$$

Finally, evaluating I_1^χ for large L by complex contour integration, we end up with

$$\Delta m_B=-\frac{3g^2}{16\pi m_B^2 L}\,e^{-\mu L}+O(e^{-\bar{m}_B L}) \tag{3.37}$$

(the contribution of the $\phi\chi$ forward scattering amplitude is smaller than the error term). In quantum field theory, the calculation of Δm_B is thus very similar to the calculation of the meson mass shift Δm, and it is quite clear that the physical origin of the size dependence is the same in both cases.

As far as the dependence on L is concerned, Eq. (3.37) coincides with the non-relativistic formula (3.19). We are hence led to interpret the length μ^{-1} as the relativistic expression for the width of bound state wave function. An important point to note is that if m_B is very close to the elastic threshold, μ^{-1} is large and finite size effects on the binding energy are only slowly going to zero as L increases, in particular, for $m^{-1} < L < \mu^{-1}$, such a wide bound state is likely to be misinterpreted as a stationary scattering state (cf. ref. [11]).

In the limit of heavy, weakly bound mesons, one expects that the relativistic formula (3.37) reduces to the non-relativistic expression (3.19). This is in fact the case, because in this limit we have

$$\mu = \varkappa + O(E_B^{3/2}), \tag{3.38}$$

$$g^2 = 512\pi^2 m |A|^2 + O(E_B), \tag{3.39}$$

where the second relation has been obtained using Eqs. (3.9), (3.26) and $F = -4m^2 F_{nr}$ to account for the different normalizations of the relativistic and non-relativistic scattering amplitudes. Note that since the proof of the quantum mechanical formula did not rely on perturbation theory (and could in fact be made entirely rigorous using Green function methods), the correctness of the non-relativistic limit provides a non-trivial check on our quantum field theory calculations.

4. Concluding Remarks

The most important qualitative result of the analysis presented in this paper is that finite size effects on the stable particle masses fall off exponentially with a rate, which depends on the spectrum of light particles in the infinite volume theory. An asymptotically precise description of the size dependence of the masses is provided by the apparently universal formulae proved in Sects. 2, 3. These relations are obviously useful to control finite size effects in numerical studies of quantum field theories on a lattice (e.g. ref. [7]) and may also serve to estimate the strength of particle interactions at low energies [9, 10].

Compared to the simplicity of the finite size mass shift formulae (2.22) and (3.37), the proof given in this paper appears to be unduly complicated. However, it must be appreciated that a proof of Eq. (2.22) requires to control correlation functions at large times (to project on the mass shell) and simultaneously at large space-like distances. Simple transfer matrix methods therefore do not apply. An axiomatic approach does not seem promising either, because the condition that "the parameters in the Lagrangian are independent of the volume" is difficult to account for.

With little effort, the results obtained in this paper can be extended in several directions.

(a) To include particles with spin, one assumes the existence of appropriate (many component) interpolating fields and then follows the steps outlined in Subsect. 2.2. A subtle point to be observed is that the finite volume breaks the full rotational symmetry down to the cubic group $\mathcal{O}$. Since the irreducible representations of SU(2) with spin $s \geqq 2$ are reducible with respect to $\mathcal{O}$, the energy of a

particle at rest in general depends on the direction of its spin relative to the box. The corresponding set of finite size mass shifts is obtained by diagonalizing the finite L self-energy of the associated field, which is a non-trivial matrix in spin space in these cases.

(b) If there are several light particles in the theory, their contributions to the mass shifts must be added. In some cases, this leads to unexpectedly large finite size effects. For example, in the two-dimensional $O(n)$ non-linear σ-model, the theory converges to a massive free field theory as $n \to \infty$, but finite size effects survive in this limit, because there are n light particles.

(c) The methods of this paper can also be used to calculate the volume dependence of the masses of heavy particles, which are stabilized by conserved quantum numbers (such as baryon number). Unstable particles, on the other hand, cannot be treated this way, because the wave function of a resonance has an only slowly decaying scattering wave component and is hence expected to be more sensitive to the boundary conditions than a bound state wave function.

A calculation of the next to leading terms in the large L expansion of the stable particle masses would be a very non-trivial extension of the present work. Not only would one have to master the topology of Feynman diagrams to a higher degree than was needed to derive the leading terms, but a more complete knowledge of the analyticity properties of the vertex functions would also be required to be able to deform momentum integration paths sufficiently far away from the real axis (in the last step of the proof of the mass shift formula). Still, that the finite volume vertex functions could be written as an infinite series of skeleton diagrams of the type shown in Fig. 4, remains an attractive and logical possibility, which will perhaps be realized one day using more elegant methods.

Appendix A: Proof of Lemma 2.1

Let $\mathcal{Q}$ be an arbitrary path in $\mathcal{G}$. We say that $\mathcal{Q}$ is a good path, if $\mathcal{Q} \cap \mathcal{C}_j$ is either empty or a path in $\mathcal{G}$ for all $j = 1, \ldots, M$. Furthermore, two vertices $u, v \in \mathcal{V}$ are called nearest neighbors, if $u \neq v$ and if there exists a line ℓ, whose endpoints are u and v.

Define $\mathcal{V}'$ to be the set of all vertices, which can be connected to a by a good path in $\mathcal{G}$. It is obvious that all the nearest neighbors of a are in $\mathcal{V}'$. Our goal is to show that also $b \in \mathcal{V}'$. Since $\mathcal{G}$ is connected, it is sufficient to prove that $w \in \mathcal{V}'$ if w is a nearest neighbor of some $v \in \mathcal{V}'$ (and $w \neq a$).

Thus, let $\mathcal{Q}$ be a good path connecting a and v, w a nearest neighbor of v ($w \neq a$) and ℓ a line with endpoints v, w. Furthermore, let $a = v_1, v_2, \ldots, v_N = v$ be consecutive vertices along $\mathcal{Q}$ and $\ell_k \in \mathcal{Q}$ the line with endpoints v_k, v_{k+1} ($k = 1, \ldots, N-1$). Then the following cases can be distinguished:

a) $w = v_K$ for some $K \geqq 2$.

In this case, $\mathcal{Q}' = \{\ell_1, \ldots, \ell_{K-1}\}$ is a good path connecting a and w.

b) $w \neq v_k$ for all k and $\ell \notin \mathcal{C}_j$ for all j.

In this case, $\mathcal{Q}' = \mathcal{Q} \cup \{\ell\}$ is a good path connecting a and w.

c) $w \neq v_k$ for all k and $\ell \in \mathcal{C}_J$ for some J.

In this case, there exists a minimal K such that v_K is an endpoint of some line in $\mathcal{C}_J$. If $K = 1$, set $\mathcal{Q}' = \emptyset$ and otherwise $\mathcal{Q}' = \{\ell_1, \ldots, \ell_{K-1}\}$. Furthermore, let $\mathcal{Q}'' \subset \mathcal{C}_J$

be one of the paths connecting v_K and w. Then, $\mathscr{L}' \cup \mathscr{L}''$ is a path in $\mathscr{G}$ connecting a and w. It is good path, because $\mathscr{L}'$ is either empty or a good path, and because $\mathscr{C}_J \cap \mathscr{L}' = \emptyset$, $\mathscr{C}_j \cap \mathscr{L}'' = \emptyset$ for all $j \neq J$.

Thus, in all cases we have found a good path connecting a and w, and hence $w \in \mathscr{V}'$ as was to be shown.

References

1. Barber, M.N.: Finite-size scaling. In: Phase transitions and critical phenomena. Vol. 8, Domb, C., Lebowitz, J.L. (ed.). London: Academic Press 1983
2. Brézin, E., Zinn-Justin, J.: Finite size effects in phase transitions. Nucl. Phys. B**257** [FS14], 867 (1985)
3. Lüscher, M.: A new method to compute the spectrum of low-lying states in massless asymptotically free field theories. Phys. Lett. **118**B, 391 (1982)
4. Floratos, E.G., Petcher, D.: A two-loop calculation of the mass gap for the $O(N)$ model in finite volume. Nucl. Phys. B**252**, 689 (1985)
5. Lüscher, M.: Some analytic results concerning the mass spectrum of Yang-Mills gauge theories on a torus. Nucl. Phys. B**219**, 233 (1983)
   Lüscher, M., Münster, G.: Weak-coupling expansion of the low-lying energy values in the SU(2) gauge theory on a torus. Nucl. Phys. B**232**, 445 (1984)
6. Lüscher, M.: Project proposal to the EMC^2 collaboration (1983), unpublished
7. Bender, I., Berg, B., Wetzel, W.: Heidelberg preprint, HD-THEP-85-11 (1985)
8. Berg, B., Billoire, A.: DESY preprint 85-082 (1985)
9. Münster, G.: The size of finite size effects in lattice gauge theories. Nucl. Phys. B**249**, 659 (1985)
10. De Forcrand, Ph., Schierholz, G., Schneider, H., Teper, M.: The 0^{++} glueball mass in SU(3) lattice gauge theory: Towards definitive results. Phys. Lett. **152**B, 107 (1985)
11. Lüscher, M.: Volume dependence of the energy spectrum in massive quantum field theories, II. Scattering states (in preparation)
12. Lüscher, M.: On a relation between finite size effects and elastic scattering processes. In: Progress in gauge field theory. G. 't Hooft et al. (eds.) (Cargèse 1983). New York: Plenum 1984
13. Hochberg, D., Thacker, H.B.: Finite-volume effects on spectrum calculations: Monte Carlo study of an exactly solvable lattice field theory. Nucl. Phys. B**257** [FS14], 729 (1985)
14. Velikson, B., Weingarten, D.: Hadron wave functions in lattice QCD. Nucl. Phys. B**249**, 433 (1985)
15. Weinberg, S.: Phenomenological lagrangians. Physica **96**A, 327 (1979)
16. Nakanishi, N.: Graph theory and Feynman integrals. New York: Gordon and Breach 1971

Communicated by G. Mack

Received December 19, 1985