Higgs particles

A. I. Vaĭnshteĭn, V. I. Zakharov, and M. A. Shifman

*Institute of Nuclear Physics, Siberian Branch, Academy of Sciences of the USSR, Novosibirsk
and Institute of Theoretical and Experimental Physics, Moscow*

Contemporary unified gauge models of the electroweak interaction contain not only quarks, leptons, and intermediate vector bosons, but also elementary spinless fields, whose inclusion is essential for renormalizability of the theory—the so-called Higgs bosons. A large class of problems related to the Higgs bosons is considered. The greatest attention is devoted to the Weinberg-Salam model, in which there is one such boson. The characteristics of this boson are discussed in detail: the coupling constants are fixed, and bounds on the mass are given. Theoretical estimates are given for its production cross sections, lifetime, and relative probabilities of decays into different channels. Various generalizations of the model with a greater number of Higgs particles are considered. A brief review is given of closely related questions: the hypothesis of "grand unification," classical solutions of the Polyakov-'t Hooft monopole type, and axions. A special section is devoted to an alternative variant in which no elementary Higgs fields are introduced (the so-called "Technicolor" scheme).

PACS numbers: 12.30.Ez

CONTENTS

1. INTRODUCTION

The experimental data on the weak interactions are described by a local four-fermion interaction. However, contemporary theories[1] are based on the fact that the primary interaction is of Yukawa type, and includes new particles: intermediate vector $(W^{\pm}, Z^{0})$ and scalar (H) bosons. The existence of vector bosons, while not proved by experiment, seems very natural: the weak interaction has a V–A structure, and the amplitudes of the known processes factorize. It is no accident that the hypothesis of vector bosons was put forward many years ago.

Unlike vector bosons, scalar (or Higgs) particles— which constitute the subject of our review—do not manifest themselves in any way in experiments at accessible energies. It is most probable that the interaction of fermions with Higgs bosons conserves parity, strangeness, and the other additive quantum numbers. Moreover, the coupling constants are suppressed, so that exchanges of scalar particles are unimportant in practice.

Therefore elementary scalar particles are at present nothing but the result of a theoretical fantasy. Nevertheless, it is difficult to dispense with them, and the observation of scalar bosons would confirm the theory in its least trivial constructions.

Why do we need Higgs bosons? Roughly speaking, for the following reasons: a) to construct renormalizable models of the weak interactions, and b) to account for the nonuniversality of the particle masses.

It turns out that a theory with a single vector boson is not closed. In such a theory, it is easy to describe the observable phenomena to first order in the Fermi constant G_F, but the higher-order graphs diverge.

Moreover, the coupling constants of the vector fields are "too" symmetric. The point is that only a theory of gauge vector fields is renormalizable.[2] This means that the coupling constants of the $W^{\pm}$ and Z^{0} are universal, just as the electric charges are universal. In addition, according to contemporary ideas, the strong interactions are also mediated by gauge fields, the gluons. The scalar particles remain as the only source of nonuniversality. The masses of the quarks and leptons are not universal, and it is assumed that the scalar particles are responsible for the origin of the masses. Mass then appears not as a result of emission and absorption of quanta of the scalar field, but as a result of the interaction with the classical part of the scalar field, which extends over all space.

Scalar particles are now experiencing a rejuvenation. Although the Weinberg–Salam model, which introduces these particles in a realistic context, was proposed over 11 years ago, the decisive experiments to test it have been performed in recent years.[3] Physicists have now become confident that the model correctly describes the low-energy experiments (energies at which no intermediate bosons are produced are considered to be low). The creators of this model—S. Glashow, S. Weinberg, and A.Salam—have recently been

awarded the Nobel Prize.

Of course, tremendous efforts will be expended in the near future to detect the intermediate bosons directly. The success of the hunt for vector bosons now seems predetermined. Moreover, their masses can be reliably predicted:

$$m_W \approx 80 \text{ GeV}, \quad m_Z \approx 90 \text{ GeV}.$$

Very little can be said about the mass of the scalar boson. If there is a single boson, then

$$m_H > 7\text{--}8 \text{ GeV}.$$

Note the absence of an upper limit on the mass. This theoretical uncertainty may also be covered by the richness of the spectrum of Higgs bosons: there may be relatively light and neutral, and also charged, scalar particles, and they may even exist undetected among the states that have already been observed.

For the theoretician, Higgs mesons are of special interest because they relate to that part of the theory which does not seem firmly established at the present time and which may change in some way in the future. Higgs bosons touch upon many mysteries of modern physics.

Therefore a detailed review of Higgs bosons would be very appropriate at the present time. Our objective is more modest. This review includes, as it were, two excursive routes: for those interested in the theory, and those interested in the experimental consequences. Both begin in Sec. 2, where we give an exposition of the "standard" Weinberg–Salam model. This section can be regarded as an extended introduction. Estimates of the production cross sections of Higgs bosons and the probabilities of their various decays are given later in Sec. 5 ("How to search for Higgs bosons"), which can be read immediately after Sec. 2. Sections 3, 6, and 7 ("Around the standard model," "The axion," and "An alternative to Higgs bosons") are aimed at recreating at least partially the atmosphere of tempting hypotheses, uncertainties, and problems that surrounds the Higgs bosons. Finally, in Sec. 4 we discuss the low-energy theorems, which demonstrate that the Higgs bosons, if they are observed, would make it possible to count the heavy states. This section refers mainly to the theoretical part, although it may also elucidate certain statements pertaining to the experimental consequences.

Of course, in the majority of cases we do not claim to give a complete exposition. A detailed discussion of the Weinberg–Salam model would already require a review larger in volume than the present one. The gaps can be filled from other sources. We list some of them.

The Weinberg–Salam model was considered in the reviews of Vaĭnshteĭn and Khriplovich, Abers and Lee, and Bernstein.[4] The possibilities of a search for Higgs bosons were first discussed in the detailed original papers of Bogomol'nyĭ and of Ellis, Gaillard, and Nanopoulos.[5] Cosmological consequences of theories with spontaneous symmetry breaking were discussed in a review by Linde.[6] A concise and lucid exposition

of many problems can be found in reports by Weinberg.[7] Of the more recent literature, we mention a review by Gaillard and lectures by Ellis.[8] Unification of the weak, electromagnetic, and strong interactions was discussed in a review by Matinyan.[9] As to individual theoretical problems (for example, the axion), we hope that the references to the original papers quoted in the text may serve as a starting point for the interested reader.

The material of the review is largely traditional. Thus, the fundamental possibilities of a search for Higgs bosons were already analyzed in the earliest original papers. Here we re-examine the numbers. There is greater scope in the choice of theoretical problems. In particular, the low-energy theorems were not discussed in the other reviews. The exposition of the problem of the axion is not entirely standard. We have also included a discussion of the criterion of "naturalness" of a given model, and we have referred to a possible alternative to the Higgs mechanism.

2. THE WEINBERG-SALAM MODEL

It is natural to begin the systematic exposition with the Weinberg–Salam model,[1] which we shall also call the standard model. The study of this model gives a general idea of the consequences of renormalizable theories of the weak interactions. Moreover, the model is in excellent agreement with experiment, and it is necessary to understand whether the data confirm the hypothesis of the existence of Higgs bosons and, if so, to what extent they restrict their properties.

This section consists of three subsections. We first give an account of the basic ideas. We then present the complete Lagrangian of the standard model and discuss the comparison of the model with experiment.

a) Basic ideas

1) *University of the coupling constants of the vector bosons.* It is assumed that the vector bosons, like the photon, are described by gauge fields and that their coupling constants are universal, just as the electric charges are universal.[10] For charged currents, it was verified long ago that the coupling constants are independent of the particle species in the case of β decays of the muon and the neutron. Nowadays, universality is also extended to the neutral currents. Moreover, a unified theory of the weak and electromagnetic interactions is constructed, i.e., it is assumed that the coupling constants for electromagnetic and semiweak interactions are of the same order and that the observed difference between the intensities of the transitions is due solely to the mass of the W boson:

$$G_F \sim \frac{\alpha}{m_W^2},$$

where G_F is the Fermi constant, $G_F = 10^{-5} m_p^{-2}$, and $\alpha = 1/137$.

However, the idea of universality cannot be developed completely, and two independent coupling constants are introduced: g for the triplet of vector fields, and g' for the singlet[11] (the reason for this is discussed

somewhat later). In the same way that the interaction of the photon is introduced as an interaction with the electric charge, the triplet of fields interacts with the weak isospin T_W, and the singlet field interacts with the weak hypercharge Y_W. If the photon is actually included in the theoretical scheme, the electric charge must be expressed in terms of the generators of the group, and the formula for the charge takes the form

$$Q = T_{3W} + \frac{Y_W}{2}. \tag{2.1}$$

To find the coupling constant for the interaction of the vector fields with a given particle, we must know to which multiplet of the weak-isospin group this particle belongs. The well-known distinguished role of the left-handed[1] particles in the weak interactions makes it natural to assume that only the left-handed components appear in the nontrivial representation of the weak-isospin group, while the right-handed components appear in singlets.

In more detail, one introduces the following doublets of left-handed fields:

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \quad \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \quad \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L, \quad \begin{pmatrix} u \\ d' \end{pmatrix}_L, \quad \begin{pmatrix} c \\ s' \end{pmatrix}_L, \quad \begin{pmatrix} t \\ b' \end{pmatrix}_L, \tag{2.2}$$

where e, μ, τ are charged leptons, ν_e, ν_μ, ν_τ are the corresponding neutrinos, u, c, t are quarks with charge 2/3, and d', s', b' are linear combinations of the quark fields with charge −1/3, for example, $d' = d \cos\theta_C + s \sin\theta_C$, where θ_C is the Cabibbo angle.

We also list the singlets of right-handed leptons and quarks:

$$e_R, \quad \mu_R, \quad \tau_R, \quad u_R, \quad d_R, \quad c_R, \quad s_R, \quad t_R, \quad b_R. \tag{2.3}$$

Using the relation (2.1), it is easy to find the values of Y_W for each of the particles. We see, incidentally, why it is necessary to introduce two coupling constants: the average charge of the leptons or quarks is not equal to zero. Of course, one might think that heavier, hitherto unobserved quarks and leptons enlarge (2.2) to representations with average charge zero (for example, triplets). Such models with a single coupling constant and a triplet of vector fields have been proposed,[12] but the experimental observation of weak neutral currents has removed them from the limelight.

The expression for the Lagrangian of the interaction with the gauge fields is, as always, obtained by replacing the ordinary derivative $\partial_\mu \psi$ by the covariant derivative $D_\mu \psi$:

$$D_\mu \psi = \left(\partial_\mu - ig \frac{\tau}{2} b_\mu + ig' \frac{Y_W}{2} a_\mu \right) \psi, \tag{2.4}$$

where b_μ is the triplet of vector fields, a_μ is the singlet field, and τ denotes the Pauli matrices (normalized by the condition $\mathrm{Tr}\,\tau_i \tau_k = 2\delta_{ik}$), which act in the weak-isospin space, for example, τ^- transforms ν_{eL} into e_L^-, etc.

Equations (2) and (4) completely fix the form of the interaction in terms of the fields b_μ and a_μ. However, the physical states are those with definite mass—the

photon A_μ and the Z^0 boson. Clearly, the photon field can be written as

$$A_\mu = \frac{1}{\sqrt{g^2 + g'^2}} (-g' b_\mu^3 + g a_\mu). \tag{2.5}$$

In fact, the coupling of the field b_μ^3 with T_{3W} is proportional to g, and the coupling of a_μ with the hypercharge is proportional to $-g'/2$. As a result, A_μ interacts identically with T_{3W} and $Y_W/2$, i.e., with the electric charge [see (2.1)]. The overall factor is fixed by the normalization condition.

The orthogonal combination

$$Z_\mu = \frac{1}{\sqrt{g^2 + g'^2}} (g b_\mu^3 + g' a_\mu) \tag{2.6}$$

describes the Z^0 boson.

2) *Renormalizability. Higgs bosons.* Renormalizability can be understood as the assertion that the amplitudes for processes calculated in perturbation theory do not grow too fast at high energies.[13] The attempt to satisfy this requirement leads to Higgs bosons. Let us elucidate this by means of a simple example.

A well-known renormalizable theory is quantum electrodynamics. Estimates of cross sections in quantum electrodynamics can therefore be used to understand what behavior of the cross sections is admissible. Consider, for example, the annihilation $e^+ e^- \to \gamma\gamma$ at high energy and large momentum transfer. From dimensional considerations, it is obvious that

$$\frac{d\sigma}{dt} (e^+ e^- \to \gamma\gamma) \sim \frac{\alpha^2}{s^2} \qquad (s \sim t). \tag{2.7}$$

The coupling constants of the vector bosons are also dimensionless, and one might think that the same estimate of the cross section applies to the production of intermediate bosons:

$$e^+ e^- \to Z^0 Z^0 \tag{2.8}$$

(we recall that $s \gg m_Z^2$).

However, there is another source of growth of the cross sections for the production of massive particles as a function of energy. The point is that the average over the spins of the vector boson e_μ is given by the formula

$$\overline{e_\mu e_\nu} = -\frac{1}{3} \left(g_{\mu\nu} - \frac{k_\mu k_\nu}{m_Z^2} \right), \tag{2.9}$$

where m_Z is the mass of the Z boson, and k_μ is its 4-momentum [the relation (2.9) is particularly obvious in the rest system: $k_0 = m_Z$, $\overline{e_i e_k} = (1/3)\delta_{ik}$ for $i, k = 1, 2, 3$].

The factors k_μ/m_Z explicitly contain the ratio E/m_Z, and they may lead to an additional growth of the cross section and require special investigation.[14] Note that their appearance is due to the longitudinally polarized states, i.e., states in which the polarization and momentum three-vectors are parallel.

We denote by M_μ the matrix element for emission of the Z boson, so that $M_\mu e_\mu$ is its production amplitude. Then the quantity $k_\mu M_\mu$ is equal to the matrix element of the divergence of the source of the vector bosons.

---

[1] By the left-handed (right-handed) components of the fermion fields, we mean $\psi_{L,R} = \frac{1}{2}(1 \pm \gamma_5)\psi$.
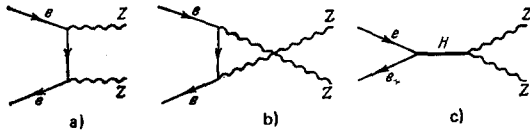
FIG. 1. Diagrams describing the process $e^+e^- \to ZZ$.

If the source of the Z bosons contains an axial current, it is not conserved:

$$\partial_\mu \left(\bar{e}\gamma_\mu\gamma_5 e\right) = 2im_e\bar{e}\gamma_5 e. \qquad (2.10)$$

As a result, the cross section corresponding to the diagrams of Figs. 1a and 1b grows inadmissibly fast with energy, the coefficient being proportional to the mass of the lepton.

In somewhat more detail, the matrix element corresponding to the diagrams of Figs. 1a and 1b takes the form

$$M^{(\text{axial})} = e_\mu^{(1)} e_\nu^{(2)} M_{\mu\nu},$$
$$M_{\mu\nu} = -\frac{g^2+g'^2}{16} \bar{e}\left(\gamma_\nu\gamma_5 \frac{1}{\hat{e}_1 - \hat{k}_1 - m_e}\gamma_\mu\gamma_5 + \gamma_\mu\gamma_5 \frac{1}{\hat{l}_1 - \hat{k}_2 - m_e}\gamma_\nu\gamma_5\right)e, \qquad (2.11)$$

where $e_{1,2}$ and $k_{1,2}$ are the momenta of the leptons and the Z bosons, and we have retained only the axial coupling of the leptons with the Z bosons.

The amplitude for production of longitudinally polarized bosons is (for $k_{1,2} \gg m_Z$) proportional to

$$k_{1\mu}k_{2\nu}M_{\mu\nu} = -\frac{g^2+g'^2}{4} m_e\bar{e}e,$$

and by squaring the matrix element we obtain for the cross section at high energies the expression

$$\frac{d\sigma^{(a+b)}}{dt} = \frac{1}{512\pi} \frac{(g^2+g'^2)^2}{s^2} \frac{m_e^2 s}{m_Z^4} \qquad (s \sim t \gg m_Z^2), \qquad (2.12)$$

which is inadmissible in a renormalizable theory [cf. (2.8)].

Thus, with no additional particles the theory is unsatisfactory. A possible way out is to introduce scalar particles which interact with the leptons in proportion to their mass. If we write their interaction Lagrangian as

$$\mathcal{L}_h = c_e H\bar{e}e + c_Z HZ_\mu Z_\mu,$$

we must have

$$c_e c_Z = -\frac{1}{2}(g^2 + g'^2) m_e. \qquad (2.13)$$

Then at high energies the diagram of Fig. 1c cancels the growth of the cross section for production of longitudinal bosons which we found above.

By considering other processes, it is possible to reconstruct all the coupling constants of the Higgs boson.[15] We shall not dwell on this in greater detail. What is important for us, however, is the conclusion drawn above: the coupling constant of the Higgs boson is proportional to the mass of the particle.

3) *Spontaneous symmetry breaking*. The existence of a particle whose interaction with other particles is proportional to their masses seems a rather exotic hypothesis at first sight. However, there is another method of constructing renormalizable theories, which is more perspicuous from a physical point of view and

which renders our conclusion more natural.

Let us try to reverse the assertion: it is not the interaction of the Higgs boson that is proportional to the mass, but conversely, the mass of the particles arises only from this interaction and is proportional to the coupling constant. The realization of this proposal is simple. Suppose that the field H has a constant part over all space (as we say, a vacuum expectation value $\langle H \rangle_0$). Then the interaction with the lepton

$$c_e H\bar{e}e$$

leads to a nonzero mass $m_e = c_e\langle H\rangle_0$.

The field H does indeed have a constant part if this is energetically favorable. For example, the potential energy

$$U(h) = \text{const}\cdot(H^2 - H_0^2)^2 \qquad (2.14)$$

has the consequence that $H \neq 0$ at the position of the minimum. In fact, the example (2.14) is not simply academic but is practically the only one if allowance is made for the fact that a renormalizable interaction cannot contain higher powers of the field than H^4.

An even more interesting result follows from the application of the same idea to gauge fields. In the first place, a gauge field acquires mass. Secondly, the appearance of mass is necessarily accompanied by a rearrangement of the degrees of freedom. Indeed, a massless vector field has two independent polarizations, whereas a massive one has three.

Let us elucidate this for the simple example[16] of a single massless vector field interacting with a charged scalar field φ (we retain the notation H for the Higgs boson of the standard model). We begin with a system having four degrees of freedom: two for a massless vector field b_μ, and two for the charged scalar field φ. If $\langle\varphi\rangle_0 \neq 0$, the system is equivalent to a single massive vector field (three degrees of freedom) and a single neutral scalar field (one degree of freedom).

The proof of this assertion is very simple. The Lagrangian has the form

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F_{\mu\nu} + (\partial_\mu\varphi^+ - ieb_\mu\varphi^+)(\partial_\mu\varphi + ieb_\mu\varphi) - U(\rho), \qquad (2.15)$$

where the potential energy $U(\rho)$ depends only on the modulus of $\rho^2 = \varphi^+\varphi$ [see (2.14)] and ensures that $\langle\varphi\rangle_0 \neq 0$. Clearly, the field b_μ then requires mass. Thus, in perturbation theory the propagator of the vector particle is replaced by

$$\frac{g_{\mu\nu}}{q^2} \to \frac{g_{\mu\nu}}{q^2} + \left(g_{\mu\nu} - \frac{q_\mu q_\nu}{q^2}\right)\frac{2e^2|\langle\varphi\rangle_0|^2}{q^4}, \qquad (2.16)$$

where the second term corresponds to the diagrams of Fig. 2. The expression (2.16) can be represented as the first term of the expansion in the mass of the propagator

$$\frac{g_{\mu\nu} - (q_\mu q_\nu/q^2)}{q^2 - m^2} + \text{longitudinal terms}, \qquad m^2 = 2e^2|\langle\varphi\rangle_0|^2.$$

The longitudinal terms $q_\mu q_\nu$ are actually unimportant because of the conservation of the current, the source of the field b_μ.
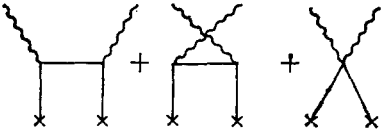
FIG. 2. Diagrams describing the polarization operator of an Abelian vector field b_μ (wavy line) interacting with a charged scalar field φ (solid line). The lines with crosses correspond to the vacuum expectation value $\langle\varphi\rangle_0$.

It is easy to rewrite the Lagrangian (2.16) in terms of new fields and in a general form without resorting to perturbation theory. Introducing the notation

$$c_\mu = b_\mu - \frac{i}{e}\,\partial_\mu \ln\varphi, \qquad (2.17)$$

we have

$$\mathscr{L} = -\frac{1}{4}\,F_{\mu\nu}F_{\mu\nu} + \frac{1}{2}\cdot 2e^2\rho^2 c_\mu^2 + (\partial_\mu\rho)^2 - U(\rho), \qquad (2.18)$$

where the intensity $F_{\mu\nu}$ is expressed in terms of c_μ in the usual way, since b_μ and c_μ differ by a gauge transformation.

It can be seen that the Lagrangian (2.18) describes a field c_μ with mass $\sqrt{2}e|\langle\varphi\rangle_0|$ and a scalar neutral field ρ.

In the foregoing, we have summarized the papers of Englert and Brout and of Higgs[16] in 1964 (we note that the first of these appeared one month earlier). It is these papers that paved the way to the introduction into the theory of weak interactions of scalar bosons (Weinberg and Salam,[1] 1967), which are now known as Higgs bosons.

The example considered above refers to an Abelian field ("photons"). However, the calculations can be generalized at once to the case of a non-Abelian field[17] (a triplet of vector fields of the isotopic group, etc.). Clearly, the scalar fields must possess the appropriate charge, for example, they must be doublets of the isospin group.

It is very important that the concept of spontaneous symmetry breaking gives a perspicuous meaning to the assertion that theories of the Weinberg–Salam type are renormalizable. Indeed, the behavior of the amplitudes at high energy or at large virtuality is important for renormalizability. Under these conditions, we can neglect the constant part of the scalar field, and the Feynman graphs for the massive theory actually reduce to Feynman graphs for a system of massless vector fields and scalar fields. It is well known that such a theory is renormalizable.

The *conclusion* common to subsections 2 and 3 is that gauge models with spontaneous symmetry breaking are renormalizable. These models involve not only vector bosons, but also scalar bosons.

b) The Lagrangian of the standard model

After the preliminary discussion of the preceding subsection, there should be no mystery about the Lagrangian of the standard model:

$$\mathscr{L} = -\frac{1}{4}\,b_{\mu\nu}b_{\mu\nu} - \frac{1}{4}\,a_{\mu\nu}a_{\mu\nu} + (D_\mu\varphi)^+ D_\mu\varphi - U(\varphi)$$
$$+ \sum_k (\bar{L}_k i\hat{D}L_k + \bar{R}_k i\hat{D}R_k) - \sum_{i,k} h_{ik}(\bar{L}_i\varphi R_k + \bar{R}_k\varphi^+ L_i)$$
$$- \sum_{i,k} h'_{ik}(\bar{L}_i\varphi_c R_k + \bar{R}_k\varphi_c^+ L_i), \qquad (2.19)$$
$$b_{\mu\nu} = \partial_\mu b_\nu - \partial_\nu b_\mu + g\,[b_\mu, b_\nu], \quad a_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu;$$

here the first two terms describe the triplet and singlet of vector fields, and $b_{\mu\nu}$ and $a_{\mu\nu}$ are the corresponding intensities. In addition, φ denotes the doublet of scalar fields,

$$\varphi = \begin{pmatrix} \varphi^+ \\ \varphi^0 \end{pmatrix}, \quad \varphi_c = i\tau_2\varphi^* = \begin{pmatrix} \varphi^{0*} \\ -\varphi^{+*} \end{pmatrix},$$

and L_i and R_i are a doublet and singlet of fermion fields, for example,

$$L_1 \equiv \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \quad R_1 \equiv e_R, \quad L_2 \equiv \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \quad R_2 \equiv \mu_R \quad \left(\psi_{L,R} = \frac{1}{2}(1\pm\gamma_5)\,\psi \right).$$

The covariant derivative was already introduced earlier [see (2.4)]. The constants h_{ik} and h'_{ik} are unknown *a priori*. Unlike g and g', they vary from doublet to doublet.

The Lagrangian (2.19) is invariant with respect to the group $SU(2)\times U(1)$. This invariance ensures renormalizability of the model. We note that the bare masses would break the invariance, and they are assumed to be equal to zero. Masses arise as a result of spontaneous symmetry breaking.

In particular, it is assumed that the potential energy is

$$U(\varphi) = -\mu^2\varphi^+\varphi + f^2(\varphi^+\varphi)^2,$$

so that the solution $\varphi = 0$ is unstable. On the contrary, at the minimum we have

$$\varphi_{\min} = \begin{pmatrix} 0 \\ \frac{\eta}{\sqrt{2}} \end{pmatrix}, \qquad \eta = \frac{\mu}{f} \qquad (2.20)$$

and the true field, which is subject to quantization, is the deviation of φ from the vacuum expectation value $\varphi_{\min}$.

The Lagrangian (2.19) can be rewritten in terms of massive vector fields. To do this, we introduce the notation

$$\varphi(x) = \varphi_{\min} + \frac{1}{\sqrt{2}}(H + i\tau\psi(x))\begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad (2.21)$$

and exploit gauge invariance to choose $\psi(x) \equiv 0$ (this gauge is called the unitary gauge). Then

$$\mathscr{L} = -\frac{1}{4}\,b_{\mu\nu}b_{\mu\nu} - \frac{1}{4}\,a_{\mu\nu}a_{\mu\nu} + \frac{1}{2}\left[\frac{g^2}{4}(\eta+H)^2 b_\mu b_\mu + \frac{g'^2}{4}(\eta+H)^2 a_\mu a_\mu \right.$$
$$\left. + \frac{gg'}{2}(\eta+H)^2 a_\mu b_\mu^3\right] + \frac{1}{2}\,\partial_\mu H\partial_\mu H - \mu^2 H^2 - \mu f H^3 - \frac{f^2}{4}\,H^4$$
$$+ \text{fermion part.} \qquad (2.22)$$

The Lagrangian (2.22) describes two charged vector fields $W_\mu^\pm = (b_\mu^1 \mp ib_\mu^2)/\sqrt{2}$ with mass $m_W^2 = g^2\eta^2/4$, two neutral vector fields A_μ and Z_μ [see (2.5) and (2.6)] with masses

$$m_A^2 = 0, \quad m_Z^2 = \frac{g^2+g'^2}{4}\,\eta^2,$$

and a scalar field with mass $m_H^2 = 2\mu^2 = 2f^2\eta^2$.

The Lagrangian (2.22) permits an explicit calculation of the tree graphs, which is sufficient for our purposes. The reader may acquaint himself with the quan-

tization procedure, for example, in the review of Abers and Lee.[4]

c) Comparison with experiment

In the standard model, all observable cross sections are expressed in terms of the constant $G_F = g^2/4\sqrt{2}m_W^2$ and the so-called Weinberg angle θ_W:

$$\text{tg}\,\theta_W = \frac{g'}{g}. \tag{2.23}$$

All the data on the manifold neutral reactions and on nonconservation of parity in the eN interaction (atoms and eN scattering)[5] agree with the theory for

$$\sin^2\theta_W = 0.22 \pm 0.02. \tag{2.24}$$

Of course, exchanges of Higgs bosons give here a small contribution $\sim m_f^2/m_H^2$, so that the existing agreement between theory and experiment cannot serve as a proof of their existence. However, there is one indirect piece of evidence for the reality of the Higgs mechanism of mass generation. For the discussion, it is convenient to divide all the existing tests of the theory into two classes: tests of the universality of the coupling constants and of the relations between the masses.

1) *Universality of the coupling constants*. Universality of the charged currents was verified many years ago. For the neutral currents, new predictions arise. It follows from the definition of the field Z [see (2.6)] that its source is

$$T_{3W} - \sin^2\theta_W \cdot Q, \tag{2.25}$$

and this prediction is independent of the mass of the Z boson, which in the general case would have to be regarded as a free parameter. We see, in particular, that the neutral axial current is a component of an isovector. The overwhelming majority of tests of the Weinberg–Salam model reduce to a test of (2.25).

2) Relation between the masses of the vector bosons. In addition to universality of the coupling constants, the model predicts a definite relation between the masses of the W and Z bosons:

$$\frac{m_Z^2\cos^2\theta_W}{m_W^2} = 1, \tag{2.26}$$

while analysis of the data leads to 0.98 ± 0.05.

Of course the relation (2.26) follows directly from the Lagrangian (2.22). However, in view of the importance of the relation (2.26), we shall give an elementary explanation of why this relation occurs.

Renormalizability requires that the coupling constant of the Higgs boson be proportional to the mass of the particle [see (2.18)]. By choosing the constants, this requirement can easily be fulfilled in all cases except the gauge fields: their coupling constants are universal and cannot be varied. Therefore the only way of satisfying the requirement of universality of the coupling constants of the vector bosons and also their proportionality to the mass is to require a definite relation between the masses. Thus we arrive at (2.26).

However, we must return once again to the statement that "the coupling constants of the gauge fields are fixed." They are fixed if one indicates the representation according to which a given field transforms. So far, we have assumed that the scalar fields belong to a doublet. We then find the relation (2.26). If, for example, we introduce a quartet of scalar particles, we would find a perfectly definite but different relation between the masses [it would differ from (2.26) by a coefficient]. On the other hand, it is equally obvious that nothing changes if we introduce several Higgs doublets, since the coupling constant for the interaction of the gauge field with a particle depends only on the representation to which the particle belongs.

Thus we arrive at the following important *conclusion*: the Higgs particles belong to one or several doublets[2)] of the group $SU(2)_W \times U(1)_W$.

To conclude this entire section, it should be noted that experiments at ultralow energies (in comparison with the masses of the W and Z) have made it possible to test a surprisingly large part of the theory: not only the universality of the coupling constants, but also, in part, the mechanism of mass generation. Of course, it cannot be excluded that all the agreement between theory and experiment is fortuitous, and the model will in any case remain a hypothesis until intermediate vector and scalar bosons are observed.

3. AROUND THE STANDARD MODEL

In this section, we discuss individual theoretical problems: whether the Weinberg–Salam model is unique; if not, how to choose the correct model; what can be said about the number and masses of Higgs bosons; whether we can recognize that there is a condensate of scalar fields in the vacuum—such is the list of questions which we have chosen from among those that inevitably occur to anyone who has "learned" the Weinberg–Salam model. These questions have no exhaustive answers. Therefore our goal will be to discuss the fundamental possibilities rather than the numbers. As we have already mentioned in the Introduction, the experimental consequences proper have been relegated to Sec. 5.

a) Grand unification

The Weinberg–Salam model is by no means the only variant of a renormalizable theory of the weak interactions. Using the same basic principles, it is easy to mass-produce new models. We can change the initial invariance group of the Lagrangian and the multiplets to which the fermions and scalar fields belong (for an attempt at classification of models, see Ref. 18). The main selection criterion here is agreement with experiment. A basic difficulty in choosing the correct model is the fact that in theories with spontaneous symmetry breaking a single multiplet combines particles with completely different masses.

[2)] The possible existence of several different multiplets of Higgs particles, so that the relation (2.26) is satisfied approximately and fortuitously, is not discussed for "esthetic" reasons.

Of course, the greatest interest lies in models which attempt to solve at least some of the fundamental problems which have so far been brushed aside. In particular, we cannot fail to mention so-called "grand unification," which includes the strong, electromagnetic, and weak interactions[19,20] (see also the review of Ref. 9).

The basic idea is that at small distances all interactions are described by a single coupling constant and that the Lagrangian possesses an invariance which unifies all the currently known symmetries. For example, in the "grand unification" model of Ref. 20 the $SU(5)$ symmetry group of the Lagrangian incorporates both the $SU(3)_{color}$ subgroup of the strong interactions and the $SU(2)_W \times U(1)_W$ group of the Weinberg–Salam model.

To account for the experimentally observed differences between the coupling constants, we must assume that for some scale of masses M characterizing the masses of superheavy vector bosons there occurs a first spontaneous symmetry breaking, so that at distances greater than $1/M$ we can neglect only the masses of the gauge fields corresponding to $SU(3)_{color} \times SU(2)_W \times U(1)_W$. Then, for $r \sim 1/300$ GeV, there occurs a second spontaneous symmetry breaking, which is described by the Weinberg–Salam model.

It is easy to estimate the value of M at which splitting of the weak and strong interactions occurs[21]:

$$\frac{\alpha_s(\mu)}{1 + 7\frac{\alpha_s(\mu)}{2\pi}\ln\frac{M}{\mu}} = \frac{8}{3}\frac{\alpha(\mu)}{1 - \frac{10}{3}\frac{\alpha(\mu)}{2\pi}\ln\frac{M}{\mu}}, \qquad (3.1)$$

where the left-hand side represents the strong-interaction coupling constant, extrapolated to small distances $r \sim 1/M$, and the right-hand side is the electromagnetic-interaction coupling constant α, multiplied by 8/3, at these same distances. The factor 8/3 represents $1/\sin^2\theta_W$ in the $SU(5)$ scheme at distances $r \sim 1/M$.

Thus the theory involves a new (enormous) mass scale:

$$M \sim 10^{16} \text{ GeV}. \qquad (3.2)$$

Of course, enlargement of the group also leads to an increase in the number of Higgs bosons. In a model with $SU(5)$ symmetry, the first spontaneous symmetry breaking is associated with a 24-plet, and the second with a 45-plet of scalar particles. Thus the number of elementary scalar particles may be large (in some variants, as much as 1000).

However, the new particles are very heavy, and in what follows we shall consider only the scale of masses $\leqslant 100$ GeV, where we can use the classification according to the group $SU(2) \times U(1)$.

b) "Natural" and "unnatural" models

Of the theoretical criteria which have been proposed for the choice of a correct theory of the weak interactions, the most interesting one seems to be the requirement of "naturalness."[22] This means that the fundamental experimental facts, such as conservation

of electric charge and the absence of neutral strangeness-changing currents, are reproduced by the model for an arbitrary, and not for some special, choice of the parameters. In Ref. 22 a derivation is given of conditions which ensure the absence of neutral currents which change strangeness or other flavors for an arbitrary mass matrix of the quarks.

It must be borne in mind that effective neutral currents arise not only as a result of Z-boson exchange, but also as a result of higher orders (exchange of a $W^+ W^-$ pair) as well as exchange of Higgs bosons. Here we shall discuss the latter mechanism, since we are interested in restrictions on the spectrum and interactions of Higgs particles.

In the standard model, the requirement that there is no change of strangeness when a Higgs boson is exchanged is fulfilled in a natural way. To see this, we return to the example discussed in subsection 2.b2—the annihilation of fermions into two Z bosons. Specifically, we consider the transition $\bar{s}d \rightarrow ZZ$. Strangeness is conserved in the interaction of the Z boson determined by $T_{3W} - \sin^2\theta_W \times Q$:

$$\mathcal{L}_Z \propto Z_\mu \left[\frac{1}{2}\bar{d}_L^0\gamma_\mu d_L^0 + \frac{1}{2}\bar{s}_L^0\gamma_\mu s_L^0 - \sin^2\theta_W \left(\frac{1}{3}\bar{d}\gamma_\mu d + \frac{1}{3}\bar{s}\gamma_\mu s \right) \right]$$
$$= Z_\mu \left[\frac{1}{2}\bar{d}_L\gamma_\mu d_L - \frac{1}{2}\bar{s}_L\gamma_\mu s_L - \sin^2\theta_W \left(\frac{1}{3}\bar{d}\gamma_\mu d + \frac{1}{3}\bar{s}\gamma_\mu s \right) \right]$$
$$(d^0 = d\cos\theta_C + s\sin\theta_C, \quad s^0 = -d\sin\theta_C + s\cos\theta_C).$$

Therefore graphs of the type shown in Figs. 1a and 1b do not contribute to the process $\bar{s}d \rightarrow ZZ$. Then it follows from renormalizability that diagrams of the type shown in Fig. 1c are also absent. Consequently, there is no $\bar{s}dH$ vertex in the theory.

In other words, in the standard model diagonalization of the mass matrix leads simultaneously to a diagonal interaction of the Higgs bosons. This combination is obviously due to the fact that in the standard model there is only a single Higgs field and the fermion masses are due to the nonzero vacuum expectation value of this field. If we introduce a second isodoublet of scalar fields into the theory, the diagrams of Fig. 1c are, as before, forbidden by the condition of renormalizability at high energies. However, the absence of these graphs for $s \rightarrow \infty$ now implies only a single condition on the coupling constants of the three neutral Higgs bosons. At finite energies, the cancellation between the contributions of different Higgs bosons is "disturbed" and there is in general a mass difference between the K_L and K_S mesons to first order in the Fermi constant G_F, which is inadmissible from a phenomenological point of view.

Suppression of the transitions $s \rightarrow dH$ can be rendered natural in the case of two doublets φ_1 and φ_2 if it is assumed that the coupling of fermions with the scalar fields has the form[22]

$$c_1(\bar{u}, \bar{d}^0)_L \varphi_1 d_R + c_2(\bar{c}, \bar{s}^0)_L \varphi_1 s_R + c_3(\bar{u}, \bar{d}^0)_L \varphi_2^c u_R$$
$$+ c_4(\bar{c}, \bar{s}^0)_L \varphi_2^c c_R + \text{H. c.} + \ldots, \qquad (3.3)$$

i.e., that the various Higgs fields give mass to the quarks with charges 2/3 and −1/3. Then the situation in relation to the s and d quarks, for example, is the same as in the case of a single doublet, and strange-

ness is conserved in the interaction of Higgs bosons.

If the form (2.3) is not to be destroyed when allowance is made for higher orders, there must exist a symmetry which maintains it. Such a symmetry is invariance with respect to two discrete transformations of the form

$$
\begin{aligned}
&1) \quad \varphi_1 \to -\varphi_1, \quad s_R \to -s_R, \quad d_R \to -d_R, \\
&2) \quad \varphi_2 \to -\varphi_2, \quad u_R \to -u_R, \quad c_R \to -c_R.
\end{aligned} \tag{3.4}
$$

This discrete symmetry also imposes restrictions on the potential of the interaction between scalar particles. A potential which satisfies these restrictions has the form[23]

$$
\begin{aligned}
U(\varphi_1, \varphi_2) = {} &\mu_1 \varphi_1^\dagger \varphi_1 + \mu_2 \varphi_2^\dagger \varphi_2 + \lambda_1 (\varphi_1^\dagger \varphi_1)^2 + \lambda_2 (\varphi_2^\dagger \varphi_2)^2 \\
&+ \lambda_3 (\varphi_1^\dagger \varphi_1)(\varphi_2^\dagger \varphi_2) + \lambda_4 (\varphi_1^\dagger \varphi_2)(\varphi_2^\dagger \varphi_1) + \lambda_5 (\varphi_1^\dagger \varphi_2)^2 + \lambda_5^* (\varphi_2^\dagger \varphi_1)^2.
\end{aligned} \tag{3.5}
$$

We now give an example of a model in which suppression of the transition $s \to dH$ cannot in any way be natural. Suppose that, in addition to the doublets of left-handed particles (2.2), there also exist doublets of right-handed quarks:

$$
\binom{u_R}{b_R}, \quad \binom{t_R}{d_R}, \quad \binom{c_R}{s_R}. \tag{3.6}
$$

Such models, which restore the symmetry between the right- and left-handed components of the spinors, were discussed very intensively in 1975–1978 (see, for example, Ref. 24).

It is easy to see that in this model the amplitude of the process $sd \to W^+W^-$ has an unacceptable growth with energy unless one introduces diagrams of the type shown in Fig. 1c, so that renormalizability requires the coupling $\bar{s}(1 + \gamma_5)dH$. The problems which must be solved when conservation of strangeness is not natural were examined in detail for this case by the present authors in the review of Ref. 25, where it was not only stated that the model is unnatural, but it was also attempted to ascertain whether this leads to a direct contradiction with experiment. However, it is necessary to warn the reader that the model (3.6) is now ruled out by direct experimental data.[9]

We note that in general electric charge may not be conserved in a model with two doublets. Indeed, in the case of a single doublet, by means of a rotation in isotopic space the vacuum expectation value can be reduced to the form

$$
\langle \varphi \rangle_0 = \binom{0}{\sigma}, \tag{3.7}
$$

where σ is a real number. It is important that only a single component has a vacuum expectation value. We *define* electric charge so that with this choice of axes the upper component of the doublet corresponds to a neutral particle. Then the vacuum condensate is not charged, and electric charge is conserved in reactions between particles.

If there are two doublets, then in the general case only one of them can be reduced to the form (3.7):

$$
\langle \varphi_1 \rangle_0 = \binom{0}{\sigma}, \quad \langle \varphi_2 \rangle_0 = \binom{\delta}{\eta + i\chi}, \tag{3.8}
$$

where σ, δ, η, and χ are real numbers. If $\delta \neq 0$, conservation of electric charge is violated by exchanges

with the vacuum and the photon is massive—the model is unacceptable.

The matrix of vacuum expectation values is determined by the form of the Lagrangian. It is easy to see that the potential energy depends quadratically on the parameter δ. Therefore the solution $\delta = 0$ is ensured by fulfillment of a certain inequality for the original coupling constants in the Lagrangian of the scalar fields (3.5). It can be said that conservation of electric charge in a model with two doublets is obtained at a much lower price than the absence of neutral strangeness-changing currents. The latter requires not inequalities, but very specific relations between the coupling constants for the interaction of the scalar fields with the quarks, which can be guaranteed at the price of the discrete symmetry (3.4).

A detailed discussion of various models from the point of view of their naturalness (including models with "grand unification") is given in Ref. 26.

Conclusion: The standard model with one doublet of scalar fields is the most natural one. In a model with two doublets, it is necessary to assume the existence of an additional discrete symmetry. There are, of course, models in which the requirement of naturalness cannot in general be satisfied (if, for example, doublets of right-handed fermions are introduced).

c) Conservation laws in models with spontaneous symmetry breaking

From the material of the preceding subsection the reader could not fail to form the impression that the conservation laws in models with spontaneous symmetry breaking are rather accidental in character and are determined by the properties of the scalar fields.

This property of the models seems fairly general. Thus, in models with two doublets we can expect transitions with nonconservation of the muon charge (for example, $\mu A \to eA'$) with probability of order 10^{-10} of the usual weak processes.[27] Violation of CP invariance also occurs in a natural way.[28]

The reasons for both features are the same. The initial Lagrangian leads to conservation laws for such quantities as the electric and muon charges, and so forth. However, if these conservation laws are to manifest themselves in the form of conserved quantum numbers, there must be a corresponding invariance of the vacuum condensate. Thus, in the example involving electric charge in the case of one doublet considered above, the vacuum expectation value (3.7) is invariant with respect to multiplication of the upper component of the doublet by a phase factor; it is these transformations that are associated with the electric charge. An increase in the number of scalar fields naturally reduces the symmetry of the vacuum condensate.

But what happens to the original conservation laws when there is no corresponding invariance of the vacuum condensate? There are two variants: either there are massless Goldstone particles and the original symmetry manifests itself in low-energy theorems for

the interaction amplitudes of these particles, or the considered current is a source of a vector field. In the second variant, instead of Goldstone modes, the vector field acquires a longitudinally polarized state—it becomes massive. The symmetry of the Lagrangian manifests itself in the form of conserved quantum numbers at high momenta, where the masses and vacuum expectation values of the fields can be neglected.

d) Masses of the Higgs bosons

From a practical point of view, it is important to know how heavy the Higgs bosons are. Scalar particles were introduced in Sec. 2 on the basis of a treatment of the amplitudes in the infinite-energy limit $s \to \infty$. Clearly, such a treatment does not restrict the mass in any way and elucidates only the fundamental possibility of renormalizing the Lagrangian.

Further arguments are required to predict the mass. There are few convincing results. In the standard model[29] we have

$$m_H > 7\text{--}8 \text{ GeV}, \tag{3.9}$$

and models with several doublets can also contain lighter scalar particles.

1) *Lower limit on the mass.* We begin with a few words about the origin of the lower limit on the mass of the Higgs boson. In terms of the parameters appearing in the Lagrangian, the mass is

$$m_H^2 = 2f^2\eta^2 \tag{3.10}$$

[see (2.22)]. The vacuum expectation value η can be found from the mass of the W boson: $\eta = (2m_W/e)\sin\theta_W = (G_F\sqrt{2})^{-1/2}$. As to the self-interaction constant f^2 of the scalar fields, it cannot be determined directly from experiment and one might think that by reducing f^2 we could arrive at an arbitrary mass m_H.

However, in addition to the bare term $f^2\varphi^4$, an effective self-interaction arises as a result of the graphs containing virtual vector fields[30] (Fig. 3). Since the coupling constant for the interaction of the scalar field with the gauge vector field is fixed, the effective value f^2_{eff} cannot be smaller than $f^2_{\text{eff}} \sim e^4$, and this leads to the bound (3.9).

In a model with two doublets, knowledge of the mass m_W does not permit a determination of the two vacuum expectation values $\langle\varphi_1\rangle$ and $\langle\varphi_2\rangle$.

2) *Upper limit on the mass.* There is no reasonable upper limit on the mass m_H. However, a condition which is frequently cited in the literature is[31]

$$m_H < \sqrt{\frac{8\pi\sqrt{2}}{3G_F}} \approx 1 \text{ TeV}. \tag{3.11}$$

The bound (3.11) arises from the requirement that

FIG. 3. Diagram leading to an effective self-interaction of the scalar field (solid line) as a result of exchange of W and Z bosons (wavy lines). The unitary gauge is used.

the graphs of lowest order in the constant f^2 do not violate the unitarity condition. Indeed, a large coupling constant f^2 corresponds to large mass [see (3.10)], and for $m_H \sim 1$ TeV the interaction between Higgs bosons becomes strong. In this case, we must clearly sum the Feynman graphs.

It must be borne in mind, however, that there are certainly renormalizable theories in which we cannot confine ourselves to the first order of perturbation theory (a good example is quantum chromodynamics). Therefore it seems more consistent not to impose *a priori* the requirement of smallness of the constant f^2, but to consider whether we know from experimental data that the self-interaction of the Higgs fields is weak.

It turns out that the strong interaction in the Higgs sector has practically no influence on the effective *fermion* Lagrangian at low energies.[32] This is so because the contribution of high energies is suppressed by the factor m_f^2/m_W^2 (where m_f is the fermion mass).

Roughly speaking, the expansion in the coupling constant at low energies has the form

$$G_F \to G_F + G_F(G_F\Lambda^2)\frac{m_f^2}{m_W^2},$$

where either Λ is of order m_H if m_H is below the unitarity limit, or $\Lambda \sim G_F^{-1/2}$ if the mass of the Higgs boson is very large. In either case, the higher-order correction is numerically small and cannot be used for an upper bound on the mass of the Higgs boson.

Since the mass of the W boson exceeds the masses of the known fermions, the intermediate vector bosons might be the best tool for detecting the strong interaction in the Higgs sector. In particular, a special investigation is required to determine whether the relation between the masses m_Z and m_W [see (2.26)] is preserved if there is no perturbation theory in the Higgs sector. We refer to the opinion of Weinberg[33] that the strong self-interaction of the Higgs bosons destroys the relation $m_W = m_Z\cos^2\theta_W$.[3] However, we could not devise a proof of this assertion. Moreover, a counterexample is mentioned briefly in Sec. 7.

3) *Hierarchy of masses.* New problems involving the masses arise in models with "grand unification." As we have already mentioned, the idea of unification of the weak and strong interactions presupposes the existence of two scales of masses, which differ from one another by many orders of magnitude:

$$\langle\varphi_1\rangle \approx 10^{14} \langle\varphi_2\rangle. \tag{3.12}$$

Even if we assume that the parameters in the Lagrangian are chosen appropriately, the radiative corrections in general destroy the relation (12). A detailed discussion of this problem can be found in Ref. 34.

[3] In a recent paper, S. Weinberg [Phys. Rev. D 19, 1277 (1979)] abandoned his view and argued that even with a strong interaction in the Higgs sector the relation $m_W^2 = m_Z^2\cos^2\theta_W$ is not modified.

437   Sov. Phys. Usp. 23(8), Aug. 1980

Vaĭnshteĭn *et al.*   437

Conclusion: In the standard model, it is unlikely that the Higgs boson is lighter than 7–8 GeV. No reasonable upper limit on the mass exists. In models with "grand unification," it is difficult to maintain the conjectured hierarchy of masses.

e) Classical excitations of the vacuum condensate

We have so far discussed spontaneous symmetry breaking only in connection with the renormalizability of the weak interactions. Is the concept of spontaneous symmetry breaking simply a prescription for writing down renormalizable Lagrangians, or does it have deeper physical content?

The problem of renormalizability is solved at the quantum level with the treatment of loop graphs. However, the clearest manifestation of spontaneous symmetry breaking would be associated with classical excitations of the vacuum: it is possible to have solutions of the classical equations in which a scalar field is different from $\langle\varphi\rangle_0$ and depends on the coordinate. The observation of the corresponding effects would be of tremendous interest. It is no accident that the section on spontaneous symmetry breaking in the review of Abers and Lee[4] was introduced with the following epigraph due to Nambu: "If my view is correct, the Universe may have a kind of domain structure. In one part of the Universe you may have one preferred direction of the axis; in another part, the direction of the axis may be different."[4)]

A detailed discussion of such questions lies outside the scope of the present review. We shall confine ourselves to two examples of nontrivial classical solutions, both of which refer to models which are simpler than the standard one.

1) *Cosmological consequences.* An example of cosmology in a theory with spontaneous (discrete) symmetry breaking was analyzed for the first time in Ref. 35. Suppose that there is a single scalar field φ described by the Lagrangian

$$\mathcal{L}(\varphi) = \frac{1}{2}(\partial_\mu\varphi)^2 - f^2(\varphi^2 - \eta^2)^2. \tag{3.13}$$

The minimum of the energy then corresponds to either $\langle\varphi\rangle = \eta$ or $\langle\varphi\rangle = -\eta$.

We can imagine, however, that in one part of space $\langle\varphi\rangle = +\eta$, and in another part $\langle\varphi\rangle = -\eta$. In particular, the classical solution satisfying the boundary conditions $\varphi(x) \to \pm\eta$ for $x \to \pm\infty$ has the form

$$\varphi(x) = \eta\, \text{th}\,[\sqrt{2}\,f\eta\,(x - x_0)]. \tag{3.14}$$

We can say that at $x = x_0$ there is a wall, i.e., a region of transition from $\langle\varphi\rangle = \eta$ to $\langle\varphi\rangle = -\eta$. The solution (3.14) describes a distribution of the field φ with energy greater than that of the vacuum: the energy density per unit area of the wall is $4\sqrt{2}f\eta^3/3$.

In this sense, the production of walls is energetically unfavorable. From the point of view of cosmological applications, it is therefore a crucial observation that

---

[4)]It is also no accident, however, that cosmology is not discussed in the text of the review.

walls would nevertheless necessarily be produced. The point is that in the hot-Universe model different regions of space are not interconnected by causal signals and the choice between the possibilities $\langle\varphi\rangle = \pm\eta$ would be made independently in different regions of space. One can find the minimum density of walls per unit volume.

Since very considerable energy is accumulated in the walls, their existence would have a radical effect on the entire evolution of the Universe, which can be excluded by the observational data.

As to the realistic standard model, there are no walls in this model (the change of sign in $\langle\varphi\rangle$ can be eliminated by a gauge transformation and therefore cannot lead to observable consequences). The cosmological consequences of the standard model relate primarily to the following observation.[36] At high temperature, as at high energy or strong virtuality, the initial symmetry of the Lagrangian is restored. Allowance for this fact may alter previous ideas about the development of the Universe at the initial moments of time.

The interested reader should refer to the review of Linde.[6]

2) *Magnetic monopole of Polyakov and 't Hooft.*[37] The classical excitations of the Higgs field can manifest themselves in the form of elementary particles of a new type, whose stability is related to the topology of the scalar field. The existence of such particles is essential for models with a simple non-Abelian symmetry group, i.e., models with one independent coupling constant of the gauge interaction.

The simplest model of this type contains the photon and charged vector bosons $W^\pm$, i.e., it is the electrodynamics of the W boson, whose mass is introduced by the Higgs procedure. The Lagrangian of the model can be written as[12]

$$\mathcal{L} = -\frac{1}{4}G^a_{\mu\nu}G^a_{\mu\nu} + \frac{1}{2}(D_\mu\varphi^a)D_\mu\varphi^a + \frac{1}{2}\mu^2\varphi^a\varphi^a - \frac{1}{4}f^2(\varphi^a\varphi^a)^2,$$

$$G^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu - e\varepsilon^{abc}W^b_\mu W^c_\nu, \tag{3.15}$$

$$D_\mu\varphi^a = \partial_\mu\varphi^a - e\varepsilon^{abc}W^b_\mu\varphi^c \qquad (a = 1, 2, 3).$$

The vacuum state corresponds to the classical part of a field φ^a of the form $(\varphi^+, \varphi^0, \varphi^-)_{cl} = (0, F, 0)$; the field W^0_μ describes a massless photon, $W^\pm_\mu$ is a massive boson, $m_W = eF$, $(\varphi^0 - F)$ is a Higgs boson, and $m_H = \sqrt{2}fF$ (in the unitary gauge, where $\varphi^\pm \equiv 0$).

It turns out that in addition to the solution with φ^a_{cl}, which is independent of the coordinates, there are other stable classical solutions. We assume that the direction of the isovector φ^a in isotopic space depends on the spatial coordinates in the following way:

$$\varphi^a = \varphi(r)\,r^a. \tag{3.16}$$

Gauge invariance makes it possible, by a choice of the gauge, to orient the field φ^a at any point in a given direction. However, the angle of rotation must be a continuous function of the point of space. It is easy to see that the field (3.16) cannot be "combed" in a single direction over all space by means of a continuous transformation. Polyakov designated such an object by the

picturesque term "hedgehog."

The distinction between the ordinary vacuum and the field (3.16) is topological in character and can be related to the so-called topological charge. Clearly, the solution with minimum energy and lowest nonzero topological charge is stable. An example of such a solution was in fact given in the preceding subsection in the problem of a one-dimensional wall.

By examining the covariant derivative of the field φ^a given by the expression (3.16), it can be seen that the corresponding vector field must be sought in the form

$$W_k^a = \varepsilon_{akm} r_m W(r). \qquad (3.17)$$

The expression for the energy E of the system can be represented in the following form:

$$E - E_0 = \frac{m_W}{\alpha} \int_0^\infty dx \left[\omega'^2 + \frac{\omega^2}{2x^2}(\omega - 2)^2 + \frac{1}{2} x^2 s'^2 \right.$$
$$\left. + s^2(1 - \omega)^2 + \frac{1}{8}\beta x^2 (1 - s^2)^2 \right], \qquad (3.18)$$

where E_0 is the energy of the vacuum, and we have made use of the following dimensionless quantities: $x = eFr$, $\beta = 2f^2/e^2$, $s(x) = x\varphi/eF^2$, and $\omega = -Wx^2/eF^2$. The corresponding differential equations can be written down without difficulty.

The monopole mass is a quantity of order m_W/α:

$$M_{max} = \frac{m_W}{\alpha} \varepsilon(\beta), \qquad (3.19)$$

and the value of the coefficient $\varepsilon(\beta)$ was determined by numerical calculation in Ref. 38: $\varepsilon(\beta)$ varies from 1 to 1.8 when β varies from 0 to $\beta = \infty$.

There are no monopoles in the Weinberg–Salam model: the additional $U(1)$ symmetry has the consequence that any distribution can be "combed."

However, as has been emphasized by many authors, it is desirable to have a simple group as the initial symmetry group. Apart from the esthetic advantage due to the presence of a single coupling constant, this automatically leads to quantization of electric charge, since the charge operator is one of the generators of the group. Thus, in the example of the $SU(2)$ group discussed above, we have $Q = T_3$, where T_3 is the third component of isospin and obviously takes on only discrete values. For the Weinberg–Salam $SU(2) \times U(1)$ model, $Q = T_3 + \frac{1}{2}Y$, where the hypercharge Y is related to the $U(1)$ group and can be arbitrary.

In models of grand unification, the group $SU(2)_W \times U(1)_W \times SU(3)_{color}$ is enlarged to a simple group, for example, $SU(5)$, so that a monopole occurs. However, in estimating the monopole mass we must interpret m_W as the intermediate-boson mass which arises in the first spontaneous symmetry breaking, $M_{mon} \sim 4\pi g^{-2} \times 10^{16}$ GeV. In other words, the monopole mass in such models is enormous.

We note that heavy monopoles may appear in the early stages of the evolution of a hot Universe. In Ref. 39 a calculation was made of the density of "relict" monopoles, which was found to be very large, much greater than the existing experimental limits. It is not clear, however, whether such calculations can be used for the grand-unification monopoles with mass of the order of the Planck mass ($\sim G^{-1/2}$, where G is the gravitational constant). Masses of order 10 TeV were discussed in Ref. 39.

4. LOW-ENERGY THEOREMS

This short section is devoted to the low-energy theorems. Their treatment is, as it were, the link between the purely theoretical problems discussed in the preceding section and the practical estimates which constitute the subject of the next section. By low energy, we mean here the case in which the mass of the intermediate state, for example, the W boson, is much greater than the mass of the Higgs boson whose decay or production is being considered.

The low-energy theorems demonstrate a unique property of the Higgs bosons: if they can be observed experimentally, they make it possible to examine even smaller distances and to count the number of states with mass exceeding the mass of the scalar boson. This really is a unique property, since the requirement of a certain isolation of the region of low energies from the region of ultrahigh energies can normally be regarded as one of the formulations of renormalizability (in a renormalizable theory, the cross sections do not grow rapidly at high energy, and dispersion relations demonstrate that the high-energy contribution is unimportant at low energies).

The breakdown of the usual ideas about the unimportance of heavy states is due to the fact that the interaction of the Higgs bosons is itself proportional to the mass. We shall list all the cases in which heavy intermediate states are important. They amount to processes involving two gluons or photons and an arbitrary number of Higgs bosons.

a) Decay of the Higgs boson into two gluons

The decay of the Higgs boson into two gluons[40,41] occurs in the single-loop approximation: the boson is converted into a pair of quarks, which then annihilate into gluons (Fig. 4).

The result for this diagram is simple and leads to the following expression for the effective interaction of the Higgs boson with gluons:

$$\mathscr{L}^{eff} = \sum_{m_q > m_H} \frac{c_{Hq}}{m_q} \frac{\alpha_s}{12\pi} G_{\mu\nu}^a G_{\mu\nu}^a H, \qquad (4.1)$$

where c_{Hq} is the coupling constant of the Higgs boson, m_q is the quark mass, $G_{\mu\nu}^a$ is the intensity of the gluon field, and α_s is the strong-interaction coupling constant. The relation (4.1) is valid for $m_q > m_H$.

In the standard theory, $c_{Hq} = -\sqrt{G_F\sqrt{2}}\, m_q$, and each heavy quark gives the same contribution, regardless
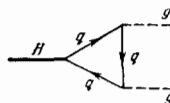

FIG. 4. Transition of the H boson into gluons via a quark loop.

of its mass. It might appear that a logarithmic dependence on the mass nevertheless remains in the factor α_s. However, this is not so: the higher-order corrections actually cancel this apparent dependence [the anomalous dimensionality of $\alpha_s (G_{\mu\nu}^a)^2$ is zero].

For the probability of the decay $H \to 2g$, we obtain

$$\Gamma (H \to 2g) = \frac{G_F m_H^3}{4\sqrt{2}\,\pi} \left(\frac{\alpha_s (m_H)}{3\pi} n_h \right)^2, \tag{4.2}$$

where n_h is the number of heavy quarks. The condition of "heaviness" of a quark is necessary for the validity of the initial expression (4.1). If $m_q \ll m_H$, there is a form factor, which in the final analysis reduces the contribution of the given quark. Actually, the formula can be applied if $m_q \gtrsim 0.2 m_H$ (the contribution of lighter quarks can, of course, be taken into account explicitly).

The coupling constant for the interaction of the Higgs boson with two gluons can be measured not only in its decays into hadrons, but also in the process of production of the Higgs boson in hadron–hadron collisions.

b) Coupling constant of the Higgs boson with the nucleon

As an application of the low-energy theorem of the preceding subsection, we shall determine the coupling constant for the interaction of the Higgs boson with the nucleon (at zero momentum transfer).[40] The direct coupling of the Higgs boson with the u and d quarks

$$-\sqrt{G_F \sqrt{2}}\, H\, (m_u \bar{u}u + m_d \bar{d}d)$$

gives a relatively small contribution

$$-\langle N|\sqrt{G_F \sqrt{2}}\,(m_u \bar{u}u + m_d \bar{d}d)|N\rangle \approx -\sqrt{G_F \sqrt{2}}\,(15\,\text{MeV})\,\bar{u}_N u_N \tag{4.3}$$

(where u_N is a spinor describing the nucleon), since the u and d quarks are light:

$$m_u \approx 4\ \text{MeV}, \quad m_d \approx 7\ \text{MeV}. \tag{4.4}$$

The heavy intermediate states turn out to be more important. To determine their contribution, we must calculate the matrix element

$$\langle N|\sqrt{G_F \sqrt{2}}\,\frac{\alpha_s}{12\pi}\, n_h G_{\mu\nu}^a G_{\mu\nu}^a|N\rangle, \tag{4.5}$$

where $G_{\mu\nu}^a$ is the operator of the gluon field intensity [see (4.1)]. This problem of calculating the matrix element, which is at first sight very complicated, can be solved almost exactly.

The point is that the following expression can be obtained for the trace of the energy–momentum tensor in quantum chromodynamics:

$$\theta_{\mu\mu} \to -9\,\frac{\alpha_s}{8\pi}\, G_{\mu\nu}^a G_{\mu\nu}^a + \sum_{q=u,\,d,\,s} m_q \bar{q}q, \tag{4.6}$$

where the first term is the so-called anomaly in the trace of the energy–momentum tensor. This expression for $\theta_{\mu\mu}$ holds for the matrix elements between states with small momenta even when heavy quarks are introduced in the theory, since the latter appear only through the loops and the corresponding contribution is proportional to $1/m_q$.

When the operator (4.6) is averaged with respect to the nucleon, the quark terms can be neglected: the

light quarks are very light. On the other hand, the matrix element of the trace of the energy–momentum tensor with respect to the nucleon at zero momentum transfer is known and is simply equal to the nucleon mass. Thus, we obtain the final result

$$\langle N|\sqrt{G_F \sqrt{2}}\,\frac{\alpha_s}{12\pi}\, n_h G_{\mu\nu}^a G_{\mu\nu}^a|N\rangle = -\sqrt{G_F \sqrt{2}}\,\frac{2n_h}{27}\, m_N \bar{u}_N u_N. \tag{4.7}$$

Taking into account only the known heavy quarks c and b, this already exceeds (4.3).

We note that in the literature one quite frequently encounters incorrect estimates of the interaction of H particles with the ordinary hadrons, owing to neglect of the anomaly in $\theta_{\mu\mu}$.

c) Decay of the Higgs boson into two photons[5b, 42]

Clearly, this decay is in many respects similar to the decay $H \to 2g$ and is also sensitive to heavy intermediate states. The difference is that, in addition to quarks, allowance must be made for other heavy charged particles: vector bosons and leptons (charged scalar bosons can also occur in generalizations of the standard model, but we do not consider them here).

From the computational point of view, diagrams involving W bosons present the greatest difficulties. In general, they cannot be calculated without a consistent treatment of the quantization of the Weinberg–Salam Lagrangian—a problem which we have avoided in the present review. However, we can obtain an explicit result without performing any new calculations, but applying only the well-known results for the renormalization of the charge (the Gell-Mann–Low function) in a theory with vector bosons.

The trick reduces to the following. Since we are considering the low-energy theorems, the Higgs field can be regarded as an external field which is independent of the coordinates. Then inclusion of the interaction with this field is equivalent to replacement of the bare masses of the fermions and W bosons by the effective masses:

$$\begin{aligned} m_f &\to m_f (1 + \sqrt{G_F \sqrt{2}}\,H), \\ m_W &\to m_W (1 + \sqrt{G_F \sqrt{2}}\,H), \end{aligned} \tag{4.8}$$

this being obvious, for example, from the form of the Lagrangian [see Eq. (2.22) of Sec. 2].

Suppose, further, that we know the amplitude M_{AB} of some process $A \to B$ without Higgs bosons. Then the amplitude of the process involving the production of an extra H boson at zero momentum of the external particles can be determined by differentiating M_{AB} with respect to the mass:

$$M (A \to B + H) = \sqrt{G_F \sqrt{2}} \left(\sum_f m_f \frac{\partial}{\partial m_f} + m_W \frac{\partial}{\partial m_W} \right) M_{AB}. \tag{4.9}$$

The validity of this assertion is immediately apparent from (4.8).

Further, let $A \to B$ denote the transition $\gamma \to \gamma$ or $g \to g$, i.e., we are considering the polarization operator of the photon or gluon in the single-loop approximation. It is well known that the result diverges and contains $\ln \Lambda$. It is obvious from dimensional argu-

440   Sov. Phys. Usp. 23(8), Aug. 1980

Vaĭnshteĭn et al.   440

ments that for small photon momentum the cutoff parameter Λ can appear only in the form of the ratio Λ/m, where m is the mass of the intermediate particles.

The result is paradoxical: we can keep track of only the divergent part in the diagram, determine the coefficient of $\ln\Lambda$, apply the relation (4.9), and find the amplitude for emission of a Higgs scalar particle at ultralow energies!

Specifically, the effective Lagrangian of the $\gamma \to \gamma$ transition in the single-loop approximation has the form

$$\mathcal{L}^{\text{eff}}(\gamma \to \gamma) = -\frac{1}{4} F_{\mu\nu} F_{\mu\nu} \left(1 - 7\frac{\alpha}{4\pi} \ln\frac{\Lambda^2}{m_W^2} + \frac{4}{3} \sum_f Q_f^2 \frac{\alpha}{4\pi} \ln\frac{\Lambda^2}{m_f^2} \right),$$
(4.10)

where Q_f and m_f are the fermion charge and mass (for the quarks, allowance must be made for the three color varieties), and Λ is the cutoff parameter. Applying the relation (4.9), we obtain the effective Lagrangian for the decay $H \to 2\gamma$:

$$\mathcal{L}^{\text{eff}}(H \to 2\gamma) = \frac{\alpha}{8\pi} V\overline{G}\sqrt{2} \left(-7 + \frac{4}{3} \sum Q_F^2 \right) F_{\mu\nu} F_{\mu\nu}.$$
(4.11)

The quantity $(-7 + 4\sum_f Q_f^2/3)$ is the coefficient in the Gell-Mann–Low function. The sign of the W-boson contribution, which is opposite to that of the fermion contribution, corresponds to asymptotic freedom of a theory with non-Abelian gauge fields: the behavior at ultrahigh energies can be tested in the low-energy region.

For the $H \to 2\gamma$ decay width, we find from (4.11) that

$$\Gamma(H \to 2\gamma) = \left(7 - \frac{4}{3} \sum_f Q_f^2 \right)^2 \left(\frac{\alpha}{4\pi} \right)^2 \frac{G_F m_H^3}{8\pi\sqrt{2}}.$$
(4.12)

The relation (4.12) "counts" the heavy charged particles.

d) Production of several scalar particles

If the threshold for production of Higgs bosons is attained, it will be possible to study the reaction

$$g + g \to H + H$$
(4.13)

or, in the more general case, the reaction $gg \to nH$ with several scalar particles. The initial gluons correspond to hadrons, so that the reaction (4.13) should be understood in the context of the parton model.

The amplitude for the reaction (4.13) is also determined by the heavy quarks, and each quark gives the same contribution, regardless of its mass. An explicit expression for the amplitude is readily obtained by applying the operation (4.9) several times to the $g \to g$ transition amplitude.[42]

Conversely, if we increase the number of gluons or photons and consider, for example, the decay $H \to 3g$, then the dominant contribution here comes from states with relatively low mass.

5. HOW TO SEARCH FOR HIGGS BOSONS

In this section, we consider processes of production and decay of Higgs bosons. Much is uncertain in the estimates. First, practically nothing can be said about
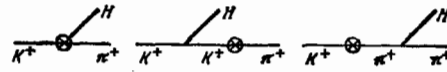

FIG. 5. The decay $K^+ \to \pi^+ H$. The crosses indicate transitions determined by the weak interaction (W-boson exchange).

the mass m_H. Secondly, the coupling constants of the H boson also vary from model to model. In the standard model, at any rate, the second problem does not arise, and in most of this section we take the interaction of Higgs bosons as it occurs in the standard model:

$$\mathcal{L}_{\text{int}} = -\sqrt{G_F V \sqrt{2}}\, H \left(\sum_q m_q \bar{q}q + \sum_l m_l \bar{l}l + 2m_W W_\mu^+ W_\mu^- + m_Z^2 Z_\mu^2 \right).$$
(5.1)

A basic feature of this interaction is the growth of the coupling constant with the particle mass. The general picture for neutral Higgs particles is evidently the same also in models with several doublets. A specific feature of these models is the presence of charged scalar particles, which are discussed in subsection 5f.

a) Higgs bosons in decays of other particles

1) *The K and η mesons*. A sufficiently light Higgs boson might be found in the decay $K^+ \to \pi^+ H$[5]. To determine the probability of this decay, we can, as was done in the preceding section, apply the low-energy theorem, which permits a determination of the interaction amplitude of the Higgs field for small 4-momentum. In fact, the introduction of a coordinate-independent Higgs field is equivalent [see (5.1)] to multiplication of all the bare masses by the factor $(1 + \sqrt{G_F\sqrt{2}H})$.

The process $K^+ \to \pi^+ H$ is described by the diagrams of Fig. 5, where the pole contributions occur because of the dependence on the bare masses of the kaon and pion propagators.

Since the squares of the masses of the Goldstone particles, m_K^2 and m_π^2, are linear in the quark masses, the HKK and $H\pi\pi$ vertices are equal to $-(G_F\sqrt{2})^{1/2}m_K^2$ and $-(G_F\sqrt{2})^{1/2}m_\pi^2$. As to the $K-\pi$ transition, it is proportional to m_q/m_W^2. The origin of m_W^{-2} is obvious, and the proportionality to m_q follows from the vanishing at $m_q = 0$. In fact, if we make use of partial conservation of the axial current, the $K-\pi$ transition is related (in the limit $m_{u,d} = 0$) to the $K \to 2\pi$ decay amplitude by the equation

$$\langle \pi^+ | \mathcal{L}_W | K^+ \rangle = -if_\pi \sqrt{2} \langle \pi^+ \pi^- | \mathcal{L}_W | K_S \rangle,$$
(5.2)

this decay being forbidden in the $SU(3)$ limit $(m_u = m_d = m_s)$.

As a result, the amplitude for direct H emission is obtained from the $K-\pi$ transition by multiplication by $-\sqrt{G_F\sqrt{2}}$. Inclusion of the pole graphs doubles the result, and by using the relation (5.2) we find that the relative probability is

$$\frac{\Gamma(K^+ \to \pi^+ H)}{\Gamma(K^+ \to \text{all})} = 2.1 \cdot 10^{-4} \sqrt{1 - 0.18 \frac{m_H^2 - m_\pi^2}{m_\pi^2}} \sqrt{1 - 0.05 \frac{m_H^2 - m_\pi^2}{m_\pi^2}}.$$
(5.3)

Such a large value is apparently ruled out experimentally. We note that the H boson with small mass decays mainly into an e^+e^- pair (or $\mu^+\mu^-$ if $m_H > 2m_\mu$),

while the decay $K^+ \to \pi^+ e^+ e^-$ has been observed at the level of 2.6×10^{-7}.

The absence of the decay $K^+ \to \pi^+ H$ gives a bound on the mass of the H boson:

$$m_H > m_K - m_\pi = 350 \text{ MeV}. \tag{5.4}$$

This bound is perhaps the only consequence of the existing data.

An analogous treatment of the decay $\eta \to \pi^0 H$ leads to the conclusion that there is a strong suppression.[5b] The $\eta - \pi$ transition amplitude is proportional to the quark masses, and the pole diagrams cancel with the nonpole diagrams.

2) *Heavy quarkonium.* Two heavy quarks are now already known: the c quark ($m_c \approx 1.25$ GeV) and the b quark (mass 4.5 GeV). Few physicists doubt that there must be at least one further heavy t quark. From experiment, $m_t > 15$ GeV.

The heavy quarks form bound states—quarkonium, the first example of which was the J/ψ meson. In the decays of quarkonium, there must be a monochromatic photon line corresponding to the decay

$$(Q\bar{Q}) \to \gamma + H,$$

provided that the decay is energetically allowed.

Quarkonium is simple for a theoretical analysis, and the decay probability can be estimated reliably.[41] Thus, for the Υ meson with mass 10 GeV, we have

$$\frac{\Gamma(\Upsilon \to H\gamma)}{\Gamma(\Upsilon \to \mu^+\mu^-)} = \frac{G_F M_\Upsilon^2}{4\sqrt{2}\pi\alpha}\left(1 - \frac{m_H^2}{M_\Upsilon^2}\right) \approx 0.9 \cdot 10^{-2}, \tag{5.5}$$

where we have assumed that $1 - (m_H^2/M^2) \approx 1$. We note also that $B(\Upsilon \to \mu^+\mu^-) \sim 1/20$. In addition, we can estimate the decay into a Higgs boson and hadrons:

$$\frac{\Gamma(\Upsilon \to H + X)}{\Gamma(\Upsilon \to \mu^+\mu^-)} \sim \frac{9}{64} \cdot \frac{\pi^2 - 9}{\pi^2}\left(\frac{\alpha_s}{\alpha}\right)^2 (G_F\sqrt{2}M_\Upsilon^2) \approx 0.8 \cdot 10^{-2}. \tag{5.6}$$

It is interesting that the probabilities for production of the H boson in conjunction with a photon and hadrons are comparable with one another. The reason for this lies in the phase-space suppression of the decay $\Upsilon \to H + 2g$.

Conclusion: The search for the Higgs boson in the decays of heavy quarkonium is evidently a perfectly realistic experimental task.

b) Direct production of H bosons in e^+e^- collisions

A rather large number of estimates of the cross section for the production of Higgs particles can be found in the literature. They are all full of pessimism. The phrase "elusive particle" has become established in relation to H bosons.

Apparently, the most promising possibility is associated production of H bosons in conjunction with W bosons, Z bosons, etc. (Provided, of course, that the required energy is available for the production of the W, Z, etc.).

In Refs. 5b, 43, and 44 the following reaction was considered:

$$e^+e^- \to Z + H.$$

The cross section for this reaction is

$$\sigma(e^+e^- \to ZH) = \frac{G_F^2 m_Z^4}{48\pi} \cdot \frac{1}{2}[1 + (1 - 4\sin^2\theta_W)^2]f(s, m_Z), \tag{5.7}$$

where

$$f(s, m) = \frac{1}{s}\sqrt{1 - \frac{(m+m_H)^2}{s}}\sqrt{1 - \frac{(m-m_H)^2}{s}}$$
$$\times \left\{\left[1 - \frac{(m+m_H)^2}{s}\right]\left[1 - \frac{(m-m_H)^2}{s}\right] + 12\frac{m^2}{s}\right\}\left(1 - \frac{m^2}{s}\right)^{-2}. \tag{5.8}$$

The value of the cross section is an appreciable fraction of, or even greater than, the value of the "standard" electrodynamic cross section $\sigma(e^+e^- \to \mu^+\mu^-) = 4\pi\alpha^2/3s$. In Fig. 6 we show the ratio

$$\frac{\sigma(e^+e^- \to ZH)}{\sigma(e^+e^- \to \mu^+\mu^-)} \tag{5.9}$$

as a function of m_H for various values of s. It can be seen from this figure that the annihilation into ZH has a probability of the same order of magnitude as for annihilation into $\mu^+\mu^-$, this probability being practically independent of m_H almost up to the production threshold.

It is probably worth noting that although the ratio (5.9) is not small, the "standard" cross section is itself very small in the studied energy region,

$$\sigma(e^+e^- \to \mu^+\mu^-) = \frac{4\pi\alpha^2}{3s} \sim 10^{-35} \text{ cm}^2$$

$$\text{for } \sqrt{s} \sim 100\text{GeV},$$

so that for work at such energies accelerators with high luminosity are required.

At the more modest energies of PEP and PETRA, H particles might be produced in association with c, b, and t quarks or τ leptons. However, estimates show that their yield is small here and comprises from 10^{-6} to 10^{-4} of the total cross section.

c) Quark mechanism of H-boson production $p\bar{p}$ collisions

The results of the preceding subsection can be generalized directly to the case of the reaction[44]

$$p + \bar{p} \to W(Z) + H + X,$$

which at the elementary level looks like $q\bar{q} \to WH$ and is therefore described by practically the same formulas as the reaction $e^+e^- \to ZH$.

In more detail, we have for the corresponding cross

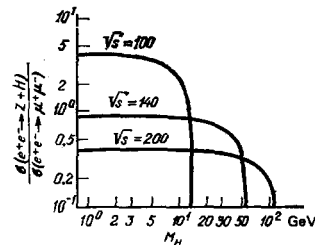FIG. 6. Cross section for the reaction $e^+e^- \to ZH$ as a function of the H-boson mass at various energies in units of the cross section $\sigma(e^+e^- \to \mu^+\mu^-)$. For the Weinberg angle, we have taken the value $\sin^2\theta_W = 0.25$.

sections

$$\sigma(p\bar{p} \to W^{\pm}(Z) + H + X)$$

$$= \frac{G_F^2 m_{W(Z)}^4}{24\pi} \int_{\tau_0}^{1} dx_1 \int_{\tau_0/x_1}^{1} dx_2 F_{W(Z)}(x_1, x_2) f(x_1 x_2 s, m_{W(Z)}),$$

where $\tau_0 = (m_W + m_H)^2/s$, and the functions $F_{W(Z)}(x_1, x_2)$ have the form

$$F_W = \frac{1}{3} u(x_1) d(x_2),$$

$$F_Z = \frac{1}{12} u(x_1) u(x_2) \left[1 + \left(1 - \frac{8}{3} \sin^2 \theta_W \right)^2 \right] \qquad (5.10)$$

$$+ \frac{1}{12} d(x_1) d(x_2) \left[1 + \left(1 - \frac{4}{3} \sin^2 \theta_W \right)^2 \right];$$

here $u(x)$ and $d(x)$ are the distribution functions of valence u and d quarks in the proton. The function $f(s, m)$ is defined in (5.8).

The cross section depends on the values of the masses and energies, as well as on the assumptions about the quark distribution functions. However, the uncertainty is relatively small. According to the estimate of Ref. 44, for the H-boson mass $m_H \sim 10$ GeV and $\sqrt{s} > 300$ GeV the yield of Higgs particles comprises approximately 10^{-3} of the yield of W and Z bosons:

$$\frac{\sigma(p\bar{p} \to W(Z) + H + X)}{\sigma(p\bar{p} \to W(Z) + X)} \sim (0.5 - 1.0) \cdot 10^{-3}. \qquad (5.11)$$

Again, it should be added that the expected value of the cross section for the reaction $p\bar{p} \to W^{\pm} + X$ is of order 10^{-34} cm².

Bremsstrahlung of the H particle in the process of production of W or Z bosons is also convenient for investigation because detection of the W(Z) may be an excellent trigger for the H particle, whose identification is a complex experimental task.

We note that Eq. (10) is also appropriate for pp collisions with the obvious replacement of $F_{W(Z)}$, for example, $u(x_1) d(x_2) \to u(x_1) \bar{d}(x_2) + u(x_2) \bar{d}(x_1)$.

Of course, quark–antiquark collisions can lead to annihilation into a single H boson—a process which can take place even below the threshold for production of the W and Z bosons. However, such a mechanism of single production is so strongly suppressed by the small masses of the light quarks that it does not seem accessible to experimental observation.

Evidently, the principal mechanism of single production of H particles in hadronic collisions will be gluon annihilation into the H particle. We now turn to the discussion of this mechanism.

d) Gluon mechanism of H-boson production

The cross section in this case is determined by the amplitude for the transition

$$g + g \to H. \qquad (5.12)$$

The gluon mechanism is of special interest, since the amplitude for the process (5.12) depends on the number n_h of heavy quarks (see the discussion in subsection 4a). On the other hand, the distribution function $D_G(x, Q^2)$ of the gluons in the proton is as yet poorly known, and this makes the estimate uncertain. It is to be hoped that D_G will be determined experimentally
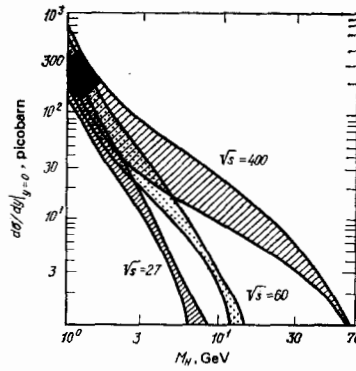

FIG. 7. Cross section for the process $pp \to HX$, related to the transition $gg \to H$. It is assumed that there exist three quarks with mass $m_q > 0.2 m_H$.

from the production cross sections of heavy particles (quarkonium).

For the cross section of the reaction $pp \to H$ +hadrons, we have[45]

$$\frac{d\sigma}{dy}\Big|_{y=0} = \frac{\alpha_s^2}{32\pi} \frac{G_F^2 n_h^2}{9\sqrt{2}}$$

$$\times \tau D_G^2(\sqrt{\tau}, m_H^2), \qquad (5.13)$$

where y (the rapidity of the H particle, $\tau = m_H^2/s$, and the gluon distribution function $D_G(x, Q^2)$ appear for $Q^2 = m_H^2$.

In Fig. 7 we show numerical estimates of the cross section taken from Ref. 45, where distributions of the following form were assumed:

$$D_G = 0.5 \sum_{n=4}^{8} c_n (n+1) x^{-1} (1-x)^n, \text{ where } c_n > 0.$$

The sum of the c_n was normalized to 1, which corresponds to half of the proton momentum going into gluons. The uncertainty in the c_n leads to the corridors indicated in Fig. 7.

It can be seen that for mass $m_H = 10$ GeV and $\sqrt{s} \gtrsim 400$ GeV the production cross section may be of order 10^{-35} cm², which is approximately two orders of magnitude greater than the cross section for associated production of WJ or ZH in $p\bar{p}$ collisions at the same energies. However, the gain is partially lost because of the difficulty in identification of the H boson. For example, the decay $H \to \mu^+ \mu^-$, which is convenient for observation, is very improbable if $m_H > 4$ GeV (see the following subsection for further details).

It seems justified to draw the following general *conclusion*: The production cross sections of Higgs particles are in general very small, and there is good reason for calling them "elusive." Associated production of ZH pairs in $e^+ e^-$ annihilation seems to offer the best prospects for experimental investigation. It is worth searching for H bosons in the decays of heavy quarkonium (this is once again a privilege of the physics of $e^+ e^-$ collisions). In hadronic collisions, H particles are produced either in association with the W(Z) in qq annihilation, or singly through the gluon mechanism.

e) Lifetime and decay modes

Both the total lifetime and the relative decay probabilities depend very strongly on the mass of the H particle. A general rule says that the dominant decays are those into particles with the maximally heavy mass (for hadronic decays, the quark mass is important).

Thus, in the mass interval 0.5–1 GeV the dominant decay is $H \to \mu^+\mu^-$, whose width is

$$\Gamma(H \to \mu^+\mu^-) = \frac{G_F m_H m_\mu^2}{4\sqrt{2}\pi}\left(1 - \frac{4m_\mu^2}{m_H^2}\right)^{3/2}, \qquad (5.14)$$

which corresponds to a lifetime of the order of 10^{16} sec.

Another appreciable mode in this mass region is the decay into 2π:

$$\frac{\Gamma(H \to \pi^+\pi^-)}{\Gamma(H \to \mu^+\mu^-)} \approx \frac{m_\pi^4}{2m_\mu m_H^2}|f(m_H^2)|^2. \qquad (5.15)$$

The form factor $f(m_H^2)$ can be estimated from the resonance formula $f = m_c^2/(m_c^2 - m_H^2 - im_c\Gamma_c)$, where m_c and Γ_c are the mass and width of the resonance in the $\pi\pi$ channel. If $\Gamma_c > 200$ MeV, the ratio (5.15) is always less than unity.

When $m_H > 1$ GeV, the dominant decays are those involving the production of particles containing strange quarks, i.e., modes such as $K\bar{K}$ +pions, φ +pions, or η +pions. The total width can be estimated as the decay $H \to s\bar{s}$:

$$\frac{\Gamma(H \to s\bar{s})}{\Gamma(H \to \mu^+\mu^-)} = \frac{3m_s^2}{m_\mu^2}, \quad \Gamma(H \to s\bar{s}) \approx 40\,\text{eV}\left(\frac{m_H}{1\,\text{GeV}}\right), \qquad (5.16)$$

where the factor 3 is due to color, and for the mass of the strange quark we have taken the value $m_s = 150$ MeV.

The next threshold correspond to the possibility of decay into a $\tau^+\tau^-$ lepton pair and into charmed particles. Formulas analogous to (5.14) and (5.16) give

$$\Gamma(H \to \tau^+\tau^-) \approx 20\,\text{keV}\left(\frac{m_H}{10\,\text{GeV}}\right),$$

$$\Gamma(H \to c\bar{c}) \approx 32\,\text{keV}\left(\frac{m_H}{10\,\text{GeV}}\right).$$

Then we find b quarks, and so forth.

We mention also the decays into two photons and into two gluons, which were discussed in Sec. 4:

$$\Gamma(H \to 2\gamma) = 0.1\,\text{eV}\left(7 - \frac{4}{3}\sum_{m_f > m_H} Q_f^2\right)^2\left(\frac{m_H}{10\,\text{GeV}}\right)^3,$$

$$\Gamma(H \to 2g) = 1.6\,\text{keV}\left(\frac{\alpha_s(m_H)\,n_h}{0.15}\right)^2\left(\frac{m_H}{10\,\text{GeV}}\right)^3.$$

here n_h is the number of quarks with mass greater than m_H, and $\alpha_s(m_H)$ is the strong-interaction coupling constant (approximately 0.15 at $m_H = 10$ GeV).

Conclusion: The total width of the Higgs boson varies from several electron volts for $m_H < 1$ GeV to several MeV for $m_H \lesssim 100$ GeV. The dominant decays are those into the heaviest of the energetically accessible particles.

f) Charged scalar particles

In the standard model, there are no elementary charged scalar bosons. Such particles occur in models with two or more doublets of scalar fields (see sub-section 3d). Their masses are not necessarily large. A charged Higgs boson with relatively light mass (of the order of several GeV) would be a real gift of nature: its identification would be not at all as difficult as that of a neutral boson.

The clearest effect would be a semiweak decay of heavy stable particles into such a Higgs boson.[46] We have in mind heavy particles containing the b quark or, if this "sally" turns out to be unsuccessful, then particles containing the t quark.

The point is that with the emission of a charged H boson, unlike a neutral boson, there is a change of strangeness, charm, beauty, etc. The corresponding coupling constants are of order $\sqrt{G_F\sqrt{2}}m_q$, so that for the ratio of the probabilities of the decay $b \to H^-c$ and of the ordinary decay $b \to c\bar{u}d$ (or $b \to c\bar{c}s$) we obtain

$$\frac{\Gamma(b \to H^-c)}{\Gamma(b \to c\bar{u}d)} \sim \frac{6\pi^2}{G_F m_b^2} \sim 10^5.$$

Thus, the decay into the H boson certainly dominates, provided that it is energetically allowed. This difference between the probabilities of Higgs and ordinary decays is due to the fact that the probability of the decay $b \to H^-c$ is proportional to the first power of G_F.

The fact that the decays of charmed particles are described by the ordinary theory leads to the following bound on the possible $H^\pm$ mass:

$$m_{H^+} > 1.5\,\text{GeV}.$$

It will evidently become known in the near future whether there are any anomalies in the decays of particles containing b quarks.

As to the $H^\pm$ decays, the dominant ones are

$$H^- \to \tau\bar{\nu}_\tau, \quad \text{if} \quad 1.8\,\text{GeV} < m_H < 2.5\,\text{GeV},$$

or

$$H \to c\bar{s}, \quad \text{if} \quad m_H > 2.5\,\text{GeV}.$$

The last case is characterized by an appreciable number of strange particles in the final state.

6. THE AXION

In the Weinberg–Salam model, there is a distinction between the roles of the strong and weak interactions: the weak interactions give mass to the quarks, and the strong interactions are responsible for the production of bound states of these quarks. The weak and strong interactions act on different degrees of freedom of the quarks: the weak interactions in the space of flavors (isospin, strangeness, charm, and so forth), and the strong interaction in color space. In our review, we have discussed the Higgs mechanism of mass generation, and the strong interactions were mentioned only incidentally. One might think that this is justified. In reality, it is not: the combination of the Higgs mechanism with quantum chromodynamics leads in general to a strong violation of *CP* invariance. More precisely, *CP* invariance is not a "natural" symmetry of the strong interactions. To avoid this unpleasant conclusion, we must introduce a new light pseudoscalar particle—the

444   Sov. Phys. Usp. 23(8), Aug. 1980

Vainshteĭn *et al.*   444

The axion has been intensively discussed for the past one and a half years. To all appearances, there are no particles with the properties predicted in the original papers. Therefore we shall not dwell in detail on the experimental predictions which have been made. The purpose of the present section is to show how the idea of introducing the new particle arises. At the same time, we shall see that its properties are not rigidly fixed, and the existing data therefore say nothing about the fundamental possibility that the axion exists.

a) CP invariance of the strong interactions (naive approach)

For several years, it has been believed[50] that CP invariance arises in a natural way in quantum chromodynamics. We shall give the corresponding arguments, which were in fact known long before the creation of QCD.[51] In the next subsection, we shall see why these arguments do not guarantee CP conservation in quantum chromodynamics.

For simplicity, we shall consider only a single massive quark Q^i ($i=1,2,3$ is the color index). The renormalizable Lagrangian contains operators with dimension $d \leqslant 4$. Using only relativistic invariance and conservation of color, we can write down the most general form of the operators with $d=4$ constructed from the field Q^i (the kinetic part of the Lagrangian):

$$\mathscr{L}_{kin}^{(0)} = \overline{Q}_i \, (a + b\gamma_5) \, i\gamma_\mu \, \partial_\mu Q^i; \tag{6.1}$$

here a and b are arbitrary numbers. The quark–gluon interaction is obtained uniquely—we must replace the ordinary derivative $\partial_\mu Q$ by the covariant derivative $D_\mu Q$.

To reduce $\mathscr{L}_{kin}$ to standard form, we make the change of variables Q^i and $\overline{Q}^i$ (we do not consider the degenerate case $a^2 = b^2$):

$$Q'^i = Q^i, \quad \overline{Q}'_i = \overline{Q}_i (a + b\gamma_5). \tag{6.2}$$

Then $\mathscr{L}_{kin}$ has the form

$$\mathscr{L}_{kin} = \overline{Q} i\gamma_\mu D_\mu Q, \tag{6.3}$$

where we have omitted the prime and matrix notation with respect to color is understood.

Thus the kinetic part of the Lagrangian always leads to a CP-invariant form. For what follows, it is important that the reduction of $\mathscr{L}_{kin}$ to standard form does not definitively fix the choice of the variables Q and $\overline{Q}$ and, in particular, $\mathscr{L}_{kin}$ remains invariant with respect to the substitution

$$Q' = e^{i\alpha\gamma_5}Q, \quad \overline{Q}' = \overline{Q}e^{i\alpha\gamma_5}. \tag{6.4}$$

The mass term, which is an operator with dimension $d=3$, can be written in general form as

$$\mathscr{L}_m = -\overline{Q} \, (m_1 + im_2\gamma_5) \, Q, \tag{6.5}$$

where $m_{1,2}$ are parameters, which are real by virtue of the Hermiticity of the Lagrangian. The term proportional to m_2 looks like a CP-noninvariant term. However, transforming to a new basis by means of the substitution (6.4), it is easy to reduce $\mathscr{L}_m$ to the usual

form by a choice of α:

$$\mathscr{L}_m = -\sqrt{m_1^2 + m_2^2}\,\overline{Q}'Q', \quad \alpha = \text{arctg}\,\frac{m_2}{m_1}.$$

The logic presented above (which can be easily generalized to the case of several quarks) leads to the conclusion that CP invariance holds regardless of the choice of the initial parameters.

b) Role of the regulator fields

With a more careful treatment, it turns out that the γ_5 invariance of the kinetic energy which we used in the preceding subsection is a formal symmetry of the Lagrangian which cannot be maintained when we take into account the need to eliminate the ultraviolet divergences. To calculate the Q-quark loops, we must regularize the theory. It is convenient to do this by the Pauli–Villars procedure, by introducing a field R with mass m_R, which is Bose-quantized. The Lagrangian of the fields Q and R has the form

$$\mathscr{L}^{Q,R} = \overline{Q} i\gamma_\mu D_\mu Q - m_Q \overline{Q} (\cos\alpha + i\gamma_5 \sin\alpha) Q$$
$$+ \overline{R} i\gamma_\mu D_\mu R - m_R \overline{R} (\cos\alpha_R + i\gamma_5 \sin\alpha_R) R. \tag{6.6}$$

For a γ_5 rotation of the field Q, we must rotate the field R through the same angle if the Feynman integrals are to remain the same; this is in contrast with the discussion of the preceding subsection. If we rotate Q through an angle α, the mass term of the field Q is reduced to $-m_Q\overline{Q}Q$, and for the field R the parameter α_R is replaced by $\alpha_R - \alpha$ (note that the value $\alpha_R = 0$ is usually understood in the literature).

Thus, CP noninvariance is localized in the term $-m_R \sin(\alpha_R - \alpha)\overline{R} i\gamma_5 R$, and the question is whether any contribution from it remains in the limit $m_R \to \infty$. It turns out that this is the case. Thus, to first order in $\alpha_R - \alpha$, the triangle diagram of Fig. 8 is finite for $m_R \to \infty$. Since we are considering noninvariance under a γ_5 rotation, it is natural that this same diagram determines the anomaly in the divergence of the axial current.[52]

For arbitrary $\alpha_R - \alpha$, the CP-noninvariant term discussed above reduces for $m_R \to \infty$ to the following addition to the Lagrangian:

$$\Delta\mathscr{L} = (\alpha - \alpha_R)\frac{g_s^2}{32\pi^2} G_{\mu\nu}^a \widetilde{G}_{\mu\nu}^a, \quad \widetilde{G}_{\mu\nu}^a = \frac{1}{2}\varepsilon_{\mu\nu\gamma\delta}G_{\gamma\delta}^a, \tag{6.7}$$

where g_s is the strong-interaction coupling constant, and $G_{\mu\nu}^a$ is the gluon field intensity.

It is obvious that from the outset we could add to the Lagrangian a term of this form—the so-called θ term:

$$\mathscr{L}_\theta = \theta\frac{g_s^2}{32\pi^2} G_{\mu\nu}^a \widetilde{G}_{\mu\nu}^a. \tag{6.8}$$

Since the sum of $\Delta\mathscr{L}$ and $\mathscr{L}_\theta$ appears, it is not the
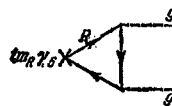

FIG. 8. Triangle diagram whose contribution is finite in the limit $m_R \to \infty$, where m_R is the mass of the regulator field. The dashed lines correspond to gluon loops.

individual values of θ, α, and α_R that are important, but only the quantity

$$\tilde{\theta} = \theta + \alpha - \alpha_R. \qquad (6.9)$$

If there are several quarks, the corresponding α_i and α_{R_i} are added.

c) The θ term and violation of CP invariance

The Lagrangian (6.8) is P- and CP-odd, so that in the general case CP invariance is not conserved. However, the θ term has one peculiarity, which explains why it was not discussed previously. The point is that the addition to the Lagrangian which we have found can be represented in the form of a total derivative (of a gauge-noninvariant quantity):

$$G^a_{\mu\nu}\tilde{G}^a_{\mu\nu} = \partial_\mu K_\mu,$$

$$K_\mu = 2\varepsilon_{\mu\nu\lambda\sigma}\left(A^a_\nu \partial_\lambda A^a_\sigma + \frac{g_s}{3} f^{abc} A^a_\nu A^b_\lambda A^c_\sigma\right).$$

Naively, one might think that a total derivative is unimportant.

However, in 1975 the discovery was made of classical solutions to the Yang–Mills equations—instantons,[34] for which the θ term is manifestly important:

$$\left(\int d^4x\, G^a_{\mu\nu}\tilde{G}^a_{\mu\nu}\right)_{\text{instanton}} = \frac{32\pi^2}{g_s^2}.$$

The general statement that total derivatives are unimportant breaks down because the instanton field does not fall off sufficiently fast at infinity.

Without relying on the single-instanton approximation, it can be seen that if $\tilde{\theta} \neq 0$, then CP invariance is actually violated in physical processes,[53] the characteristic scale of violation being of order

$$\tilde{\theta}\,\frac{m_q}{\mu},$$

where m_q is the mass of the lightest quark (~ 4 MeV), and μ is a characteristic mass (~ 200 MeV).

Thus, the naive arguments of subsection 6a are invalid, and CP invariance is not a "natural" symmetry of the strong interactions.

d) $U(1)$ symmetry and the axion

If there is a massless quark in the theory, there is no CP violation from the θ term. This can already be seen from the fact that the angle α which appears in the expression (6.9) for $\tilde{\theta}$ is undetermined and can always be chosen so that $\tilde{\theta}$ vanishes. We recall that the angle α was introduced in (6.4) as the angle of γ_5 rotation and was fixed by the CP-invariant form of the mass term.

The introduction of the new axion field[47-49] makes it possible to have $U(1)$ symmetry with respect to a γ_5 rotation even for a massive quark. The effective parameter θ^{eff} becomes dependent on the vacuum expectation value of the axion field. This vacuum expectation value is chosen by the condition of minimality of the energy of the vacuum, which corresponds precisely to $\theta^{\text{eff}} = 0$, i.e., to CP conservation.

We shall analyze the introduction of the axion for the example of a single quark Q.[47,53] Suppose that the mass of this quark arises spontaneously from the interaction with a complex field φ:

$$\mathcal{L} = \bar{Q}i\gamma_\mu D_\mu Q - h\left(\varphi \bar{Q}_R Q_L + \varphi^+ \bar{Q}_L Q_R\right)$$
$$+ \partial_\mu \varphi^+ \partial_\mu \varphi + m^2\varphi^+\varphi - f^2(\varphi^+\varphi)^2 - \frac{1}{4}G^a_{\mu\nu}G^a_{\mu\nu}.$$

Formally, this Lagrangian is invariant with respect to $U(1)$ transformations:

$$Q \to e^{i\alpha\gamma_5}Q, \qquad \varphi \to e^{-2i\alpha}\varphi.$$

The field φ has a nonzero vacuum expectation value of the form

$$\varphi = \frac{m}{f\sqrt{2}}e^{i\beta}.$$

The phase β is not determined in the classical approximation. It appears in the mass term of the quark, and when allowance is made for quark loops, as we have seen in the preceding subsections, it redefines the parameter θ: $\theta^{\text{eff}} = \theta + \beta$. The difference is that now this is not a fixed parameter of the theory, but a field which is chosen by the condition of minimality of the energy of the vacuum.

It is easy to verify that the energy of the vacuum is quadratic in θ^{eff} for small θ^{eff}. In fact, the term of first order is proportional to $\langle 0|G^a_{\mu\nu}\tilde{G}^a_{\mu\nu}|0\rangle$, where an average is taken over the CP-even vacuum, and reduces to zero. Therefore β is such that $\theta^{\text{eff}} = \theta + \beta = 0$, and there is automatically no CP violation.

What particles does such a scheme describe? In addition to the quark with mass $m_Q = h\varphi_0$, there is a scalar particle corresponding to oscillations of the modulus of φ, with mass $m\sqrt{2}$, and a pseudoscalar particle associated with oscillations of the phase. The mass of this particle—the axion—arises only as a result of the loops and is proportional to $\alpha_s/\langle\varphi_0\rangle$.

If the parameter m tends to infinity (for fixed h and f), the masses of the quark and scalar field also tend to infinity. But the mass of the axion tends to zero, as does its interaction with gluons. Thus, in this limit the axion becomes a sterile particle and does not interact with our world. There are no other experimental predictions (apart from the existence in principle of a practically massless particle), and everything looks like a theoretical phantom. To some extent, such a solution is justifiable: the whole problem of CP invariance arose from distances of the order of the inverse cutoff momentum and can be solved at the price of introducing a very heavy quark. However, there remains a feeling that the entire construction is clumsy.

In their original papers, Weinberg[48] and Wilczek[49] proposed a more optimistic model of the axion. These authors introduced not a new quark, but only a new scalar field which interacts with the old quarks, i.e., they made the Weinberg–Salam model somewhat more complicated. This leads to definite experimental consequences, which are apparently already inconsistent with experiment. However, it is important to understand that the solution of Weinberg and Wilczek is not obligatory.

7. A NEW STRONG INTERACTION IN THE 1 TeV REGION OF ENERGY—AN ALTERNATIVE TO HIGGS BOSONS

To enable the reader to judge better whether Higgs bosons are obligatory, we shall briefly discuss a possible variant of the theory in which there are no elementary scalar fields. But we must make the reservation that such variants are much less well developed.

The idea is to make Higgs particles composite.[33,55] To do this, one introduces a new strong interaction with properties similar to those of the ordinary interaction, but having a characteristic scale of masses and energies of order 1 TeV instead of 1 GeV. The growth of the weak-interaction amplitudes comes to an end at energies of the order of a TeV, i.e., in the region where the weak interactions have an effective coupling constant of order unity.

a) Dynamical mass

To understand how mass can arise in a theory without elementary scalar particles, we turn to an example of low-energy pion physics, or chiral invariance.[56]

If the quarks are massless, the axial current is conserved, and this would seem to require that the nucleon mass is also equal to zero. However, there is another way out—the existence of a massless particle.[57] The matrix elements of the axial currents then have the form

$$\langle A \mid j_\mu^5 \mid B \rangle = \left(g_{\mu\nu} - \frac{q_\mu q_\nu}{q^2} \right) M_\nu^{AB}(q), \quad q = p_B - p_A, \qquad (7.1)$$

where the quantity M_ν^{AB} is finite and nonzero at $q = 0$.

The pole at $q^2 = 0$ in the expression (7.1) is identified with the pion, which becomes massless in the limit of exact chiral symmetry. We shall now digress from real symmetry breaking and regard the pion as massless.

A necessary condition for such a rearrangement of the states—the appearance of a fermion mass and the emergence of a massless pseudoscalar particle—is the formation of a vacuum condensate:

$$\langle 0 \mid \bar{q}q \mid 0 \rangle \neq 0. \qquad (7.2)$$

The Goldstone (massless) particle is an excitation of this condensate. Whether or not a condensate is formed is a question of dynamics. For example, some materials are ferromagnetic, and domains—regions with aligned spins—are formed in them, while others are not. If a condensate is formed [Eq. (7.2)], the theory involves a mass parameter and the particle masses can be expressed in terms of it. In quantum chromodynamics, it can be seen directly to some extent how this occurs.[58]

The dynamical character of the nucleon mass manifests itself in the appearance of a form factor. The scale of masses Λ on which the form factor varies is determined by the distances at which the strong interactions have a coupling constant of order 1: $\alpha_s(\Lambda) \sim 1$. Even if there were no means of establishing directly that the nucleon is not point-like, it would be possible

to infer that the πN interaction is not fundamental on the basis of the nonrenormalizability of the phenomenological chiral Lagrangian describing this interaction. The fact that the axial coupling constant g_A of the nucleon differs from 1 might also suggest that the nucleon is not elementary.

At the present time, there is no evidence that history is repeating itself in the case of the quarks and (or) leptons. However, we cannot exclude the possibility that such indications will appear when we go to higher energies.

b) A new strong interaction

It is remarkable that, even in the absence of any experimental evidence for the dynamical character of the masses of the quarks or W bosons, we can indicate the scale of masses on which new phenomena should occur if the masses are dynamical.[33,55]

For this purpose, we turn to the question of the mass of the intermediate vector boson. In the standard model, the massive vector boson arose as a result of "unification" of the massless vector and scalar particles.

It is quite obvious that the role of the elementary massless scalar particle can also be played by a composite Goldstone boson.[59] Indeed, suppose that there is a massless vector boson (W) interacting with some (hypothetical) π^T meson. We shall find the mass operator of the vector boson.

The propagator $D_{\mu\nu}(q)$ of a vector particle can be written in general form as

$$D_{\mu\nu} = -i \left(g_{\mu\nu} - \frac{q_\mu q_\nu}{q^2} \right) D(q^2), \qquad D(q^2) = \frac{1}{q^2 - q^2 \Pi(q^2)},$$

$$(g_{\mu\nu}q^2 - q_\mu q_\nu) \Pi(q^2) = ig^2 \int d^4x e^{iqx} \langle 0 \mid T j_\mu(x) j_\nu(0) \mid 0 \rangle. \qquad (7.3)$$

If there is a massless π^T meson, the function $\Pi(q^2)$ has a pole

$$\Pi(q^2) = g^2 \frac{(F_\pi^T)^2}{q^2},$$

corresponding to this pion, where F_π^T is the corresponding residue. Substituting in (7.3), we see that the vector particle has acquired a mass

$$m_W^2 = g^2 (F_\pi^T)^2$$

(and the Goldstone particle has become unobservable).

Substituting the empirical value of the mass m_W, we obtain

$$F_\pi^T \sim 250 \, \text{GeV}. \qquad (7.4)$$

It is now completely obvious that the ordinary pion cannot account for the mass of the W: its coupling constant f_π is 2000 times smaller.

Consequently, we must assume that there exists a world of new strong interactions in which the chiral symmetry of the Lagrangian is broken. This world may be similar in many respects to the known strong interactions. However, there is certainly an important difference: the scale of masses in the new world is approximately 1000 times greater, i.e., new phenomena occur in the tera-electron-volt region.

From this TeV region, at the energies available to us only Goldstone particles appear; the remaining states are too heavy for their excitation to be appreciable.

The first test of the proposed hypothesis for the origin of the vector-boson masses is to examine whether the relation between m_Z and m_W [Eq. (2.26) of Sec. 2] is satisfied. It turns out that it is:

$$\frac{m_Z^2}{m_W^2} = \frac{|\langle 0 | j_\mu^Z | \pi_T^0 \rangle|^2}{|\langle 0 | j_\mu^W | \pi_T^+ \rangle|^2} = \frac{g^2 + g'^2}{g^2} = \frac{1}{\cos^2 \theta_W}, \qquad (7.5)$$

where we have made use of the ordinary isotopic relations to connect the coupling constants of the Z boson with π_T^0 and the W^+ with π_T^+.

The correction to the relation (7.5) due to the other states (for example, ρ_T) is small. It can be estimated as

$$g^2 \left[\frac{F_\pi}{(g_\rho^2/4\pi) m_\rho} \right]_T ,$$

i.e., it is very small if the ratio of the coupling constants in the new world (i.e., with the index T) is the same as in ours.

c) Pseudo-Goldstone mesons

Attempts at realistic constructions involve repeated application of the trick of unification of massless vector bosons and Goldstone particles into massive vector particles. A discussion of these attempts would lie beyond the scope of the present review, particularly because they are evidently far from a definitive variant. However, it is important to note that models with spontaneous symmetry breaking lead to new particles: the so-called pseudo-Goldstone mesons with mass of the order of the W-boson mass.[33,55] The point is that with spontaneous symmetry breaking of the new strong interactions there are in general many massless pseudoscalar particles. Their fate varies. Only three of them are not observable, owing to mixing with the $W^\pm$ and Z bosons.

Some of them acquire mass when the electromagnetic and weak interactions are taken into account. Such particles are called pseudo-Goldstone particles: the corresponding current is conserved only in a certain limit, when no allowance is made for electromagnetic interactions. Their mass would obviously be of order

$$m^2 \sim (1 \text{ TeV})^2 e^2$$

and comparable with the mass $m_{W,Z}$.

Finally, the most difficult problem is presented by the massless particles associated with spontaneous breaking of the strict symmetry of the Lagrangian, for which there are no corresponding vector fields.

d) The problem of massless particles

If there are no scalar fields, the well-known $SU(2) \times U(1)$ Lagrangian is invariant with respect to rotations of the right-handed components of the s and d quarks:

$$s_R \rightarrow s_R \cos \theta + d_R \sin \theta, \quad d_R \rightarrow d_R \cos \theta - s_R \sin \theta.$$

In fact, both s_R and d_R are singlets of the group $SU(2) \times U(1)$, and the requirement of renormalizability prevents us from giving the s or d quark a bare mass,

so that s_R and d_R are indistinguishable (in the standard model, the s and d quarks interact differently with the scalar particles).

If the observed difference between the masses of the ordinary and strange particles arises as a result of spontaneous symmetry breaking, there must be a Goldstone boson φ^0 which is coupled to s_R and d_R.

The vertex for the decay $K^+ \rightarrow \pi^+ + \varphi^0$ would look like

$$M(K^+ \rightarrow \pi^+ \varphi^0) = \frac{1}{F} \langle \pi^+ \varphi^0 | \bar{s}_R \gamma_\mu d_R \partial_\mu \varphi^0 | K^+ \rangle,$$

and even for $F = 3 \times 10^3$ TeV the relative probability of this decay would be of order unity.

Thus, in a theory without Higgs bosons it is necessary to introduce some new interaction which distinguishes s_R and d_R.

8. CONCLUSIONS

The Higgs mechanism of mass generation which is incorporated in the Weinberg–Salam model has already played a major part in the development of elementary-particle physics in recent years. The construction of renormalizable models has stimulated experimental investigations. The experimental confirmation of the predictions has in turn strengthened the faith in Lagrangian field theory and the conviction that renormalizable theories are really distinguished. It may also be recalled that it is in Higgs models that the first discoveries were made of applications of topologically nontrivial solutions to elementary particles, which were then taken over to quantum chromodynamics.

Higgs bosons have become such familiar objects that it is difficult to imagine that there are as yet no direct or convincing indirect experimental proofs of their existence. The search for Higgs bosons seems one of the most deserving tasks for contemporary experimental physics.

It is rather paradoxical, but models with Higgs bosons are now becoming the victim of their own success. The belief in the omnipotence of the theory is making different demands on the theoretical constructions which have already been created. For example, unification of not only the electromagnetic and weak interactions, but also the strong interactions, seems inevitable.

When studying such questions, the theoretician cannot help remembering that the well-known models have not solved many fundamental problems, such as the following:

1) the problem of zero charge when allowance is made for the Higgs sector of the models;

2) the problem of "natural" conservation of CP parity in strong transitions;

3) the problem of calculating the mass spectrum of the quarks and leptons;

4) the problem of the different mass scales in the "grand synthesis," etc.

It is not clear at the present time whether the theory

will be limited to the known ideas in the solution of such problems, or whether there will arise new radical solutions that may make also Higgs bosons unnecessary.

In conclusion, we would like to express gratitude to L. B. Okun' for helpful discussions and a kind invitation to write the present review for Uspekhi Fizicheskikh Nauk, and also to A. A. Ansel'm for discussions and remarks.

[1]S. Weinberg, Phys. Rev. Lett. 19, 1264 (1967); A. Salam, in: Elementary Particle Theory (ed. N. Svartholm), Almquist and Forlag, Stockholm (1968).

[2]C. N. Yang and R. Mills, Phys. Rev. 96, 191 (1954); G. 't Hooft, Nucl. Phys. B33, 173 (1971); B35, 167 (1971).

[3]C. Baltay, in: Proc. of the 19th Intern. Conf. on High Energy Physics, Tokyo, (1978), p. 882; L. M. Barkov and M. S. Zolotarev, Pis'ma Zh. Eksp. Teor. Fiz. 27, 379 (1978); 28, 544 (1978) [JETP Lett. 27, 357 (1978); 28, 503 (1978)]; Phys. Lett. B85, 308 (1979); C. Y. Prescott et al., Phys. Lett. B77, 347 (1978); B84, 524 (1978).

[4]A. I. Vaĭnshteĭn and I. B. Khriplovich, Usp. Fiz. Nauk 112, 685 (1974) [Sov. Phys. Usp. 17, 263 (1975)]; E. S. Abers and B. W. Lee, Phys. Rept. C9, 1 (1973); J. Bernstein, Rev. Mod. Phys. 46, 7 (1974) [Russian translation of the last two papers, in: Kvantovaya teoriya kalibrovochnykh poleĭ (Quantum Theory of Gauge Fields), Mir, Moscow (1977)].

[5]a) E. B. Bogomol'nyĭ, Yad. Fiz. 18, 574 (1973); 20, 984 (1974) [Sov. J. Nucl. Phys. 18, 297 (1973); 20, 522 (1975)]. b) J. Ellis, M. K. Gaillard, and D. V. Nanopoulos, Nucl. Phys. B106, 292 (1976).

[6]A. Linde, Rep. Prog. Phys. 42, 389 (1979).

[7]S. Weinberg, in: A Festschrift for I. I. Rabi, Academy of Sciences, New York (1978); Proc. of the 19th Intern. Conf. on High Energy Physics, Tokyo (1978), p. 907; Phys. Lett. B82, 387 (1979).

[8]M. K. Gaillard, Commun. Nucl. Part. Phys. 8, 31 (1978); J. Ellis, Weak Interactions at High Energies: Proc. of Summer Institute on Particle Physics, SLAC Report No. 215 (1978).

[9]S. G. Matinyan, Usp. Fiz. Nauk 130, 3 (1980) [Sov. Phys. Usp. 23, 1 (1980)].

[10]J. Schwinger, Ann. Phys. (N.Y.) 2, 407 (1957).

[11]S. L. Glashow, Nucl. Phys. 22, 579 (1961).

[12]H. Georgi and S. L. Glashow, Phys. Rev. Lett. 28, 1494 (1972).

[13]L. D. Landau, Zh. Eksp. Teor. Fiz. 10, 718 (1940); Sobranie trudov (Collected Works), Vol. 1, Nauka, Moscow (1969), p. 322.

[14]A. I. Vaĭnshteĭn and I. B. Khriplovich, Yad. Fiz. 13, 198 (1971) [Sov. J. Nucl. Phys. 13, 111 (1971)]; A. I. Vainshtein, Lett. Nuovo Cimento 5, 680 (1972).

[15]E. B. Bogomol'nyĭ, V. I. Zakharov, and L. B. Okun', in: Élementarnye chastitsy (Elementary Particles), No. 1, Atomizdat, Moscow (1973), p. 49.

[16]F. Englert and R. Brout, Phys. Rev. Lett. 13, 321 (1964); P. W. Higgs, Phys. Lett. 12, 132 (1964); Phys. Rev. 145, 1156 (1966).

[17]T. W. B. Kibble, Phys. Rev. 155, 1554 (1967).

[18]J. D. Bjorken and C. H. Llewellyn Smith, Phys. Rev. D 7, 887 (1973).

[19]J. C. Pati and A. Salam, Phys. Rev. D 8, 1240 (1973).

[20]H. Georgi and S. L. Glashow, Phys. Rev. Lett. 32, 438 (1974).

[21]H. Georgi, H. R. Quinn, and S. Weinberg, Phys. Rev. Lett. 33, 451 (1974).

[22]S. L. Glashow and S. Weinberg, Phys. Rev. D 15, 1958 (1977).

[23]N. G. Deshpande and E. Ma, Phys. Rev. D 18, 2574 (1978).

[24]H. Fritzsch, M. Gell-Mann, and P. Minkowski, Phys. Lett. B59, 256 (1975).

[25]M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, Phys.

[26]M. S. Chanowitz, J. Ellis, and M. K. Gaillard, Nucl. Phys. B128, 506 (1977).

[27]J. D. Bjorken and S. Weinberg, Phys. Rev. Lett. 38, 622 (1977); A. A. Anselm and D. I. D'yakonov, Nucl. Phys. B145, 271 (1978).

[28]S. Weinberg, Nucl. Phys. B37, 657 (1976); A. A. Anselm and D. I. D'yakonov, Nucl. Phys. B145, 271 (1978).

[29]A. Linde, Pis'ma Zh. Eksp. Teor. Fiz. 23, 73 (1976) [JETP Lett. 23, 64 (1976)]; S. Weinberg, Phys. Rev. Lett. 36, 294 (1976).

[30]S. Coleman and E. Weinberg, Phys. Rev. D 7, 1888 (1973).

[31]D. A. Dicus and V. S. Mathur, Phys. Rev. D 7, 3111 (1973); B. W. Lee, C. Quigg, and H. G. Thacker, Phys. Rev. D 16, 1519 (1977).

[32]M. Veltman, Acta Phys. Pol. B8, 475 (1977); Phys. Lett. B70, 253 (1977).

[33]S. Weinberg, Phys. Rev. D 13, 974 (1976).

[34]E. Gildener, Phys. Rev. D 14, 1667 (1976); K. T. Mahanthappa and K. T. Unger, Phys. Lett. B78, 604 (1978).

[35]Ya. B. Zel'dovich, I. Yu. Kobzarev, and L. B. Okun', Zh. Eksp. Teor. Fiz. 67, 36 (1974) [Sov. Phys. JETP 40, 14 (1975)].

[36]D. A. Kirzhnits, Pis'ma Zh. Eksp. Teor. Fiz. 15, 745 (1972) [JETP Lett. 15, 529 (1972)]; D. A. Kirzhnitz and A. D. Linde, Phys. Lett. B42, 471 (1972).

[37]A. M. Polyakov, Pis'ma Zh. Eksp. Teor. Fiz. 20, 430 (1974) [JETP Lett. 20, 194 (1974)]; G. 't Hooft, Nucl. Phys. B79, 276 (1974).

[38]E. B. Bogomol'nyĭ and M. S. Marinov, Yad. Fiz. 23, 676 (1976) [Sov. J. Nucl. Phys. 23, 355 (1976)].

[39]Ya. B. Zeldovich and M. Yu. Khlopov, Phys. Lett. B79, 237 (1978).

[40]M. A. Shifman, A. I. Vainshtein, and V. I. Zhakharov, Phys. Lett. B78, 443 (1978).

[41]F. Wilczek, Phys. Rev. Lett. 39, 1304 (1977).

[42]A. I. Vaĭnshteĭn et al., Yad. Fiz. 30, 1368 (1979) [Sov. J. Nucl. Phys. 30, 711 (1979)].

[43]B. L. Ioffe and V. A. Khoze, Preprint LIYaF-274, Leningrad (1976).

[44]S. L. Glashow, D. V. Nanopoulos, and A. Yildiz, Phys. Rev. D 18, 1724 (1978).

[45]H. Georgi, S. Glashow, M. Manacek, and D. V. Nanopoulos, Phys. Rev. Lett. 40, 692 (1978).

[46]E. Golowich and T. C. Yang, Phys. Lett. B80, 245 (1979).

[47]R. D. Peccei and H. R. Quinn, Phys. Rev. Lett. 38, 1440 (1977); Phys. Rev. D 16, 1791 (1977).

[48]S. Weinberg, Phys. Rev. Lett. 40, 233 (1978).

[49]F. Wilczek, Phys. Rev. Lett. 40, 279 (1978).

[50]S. Weinberg, Phys. Rev. Lett. 31, 494 (1973); Phys. Rev. D 8, 4482 (1973).

[51]B. L. Ioffe, Zh. Eksp. Teor. Fiz. 38, 1263 (1960) [Sov. Phys. JETP 11, 911 (1960)].

[52]S. L. Adler, Phys. Rev. 177, 2426 (1969).

[53]M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, Preprint ITEP-64 (1979).

[54]A. Belavin, A. Polyakov, A. Shwartz, and Yu. Tyupkin, Phys. Lett. B59, 85 (1975).

[55]L. Susskind, Stanford University Preprint (1979); S. Dimopoulos and L. Susskind, Stanford University Preprint ITP-626 (1979).

[56]Y. Nambu, Phys. Rev. Lett. 4, 380 (1960); Y. Nambu and G. Jona-Lasinio, Phys. Rev. 122, 485 (1961).

[57]J. Goldstone, Nuovo Cimento 19, 15 (1961).

[58]M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, Nucl. Phys. B147, 385, 445 (1979).

[59]J. Schwinger, Phys. Rev. 125, 397 (1962); A. A. Migdal and A. M. Polyakov, Zh. Eksp. Teor. Fiz. 51, 135 (1966) [Sov. Phys. JETP 24, 91 (1967)]; R. Jackiw and K. Johnson, Phys. Rev. D 8, 2386 (1973); J. M. Cornwall and R. F. Norton, Phys. Rev. D 8, 3338 (1973).

Rev. D 18, 2583 (1978); 19, 2815 (errata) (1979).

449   Sov. Phys. Usp. 23(8), Aug. 1980

Vainshteĭn et al.   449