Toward the unification of weak, electromagnetic, and strong interactions: SU(5)

S. G. Matinyan

Physics Institute, Erevan
Usp. Fiz. Nauk **130**, 3–38 (January 1980)

The successes of the Weinberg-Salam model and quantum chromodynamics have made the attempt at the unification of weak, electromagnetic, and strong interactions of leptons and quarks in the framework of a single group into a serious proposition. The simplest and most economic unifying group is SU(5). The review is devoted to a detailed study of some of the questions associated with the unification of the interactions of the elementary particles in the example of the group SU(5). The questions considered include the part played by "natural" conservation of the quark flavor by neutral weak currents in searches for unifying groups, the "quantization" of charge (connection between the charges of quarks and leptons), the renormalization of unified groups to present-day energies, the instability of the nucleon, the finiteness of the number of flavors, the hierarchy of interactions unified in a single scheme, and the cosmological aspects of unification schemes. Other "grand unification" schemes are also considered briefly. Some possible ways of constructing unifying schemes are discussed in the conclusions.

PACS numbers: 11.30.Ly, 12.20.Hx, 12.40.Bb

CONTENTS

1. FAMILIES OF LEPTONS AND QUARKS

a) Experimental justification of the $SU(2)_L \otimes U(1)$ scheme. Doublets and singlets

The "standard" Weinberg–Salam model,[1] which combines weak and electromagnetic interactions on the basis of spontaneously broken $SU(2)_L \otimes U(1)$ symmetry, is being steadily confirmed by more and more experiments.

All the data relating to the neutral currents predicted by this model, including deep inelastic neutrino interactions, $\nu(\bar{\nu})p$ scattering, exclusive and inclusive pion production in neutrino or antineutrino beams, and elastic scattering of ν_μ or $\bar{\nu}_\mu$ by electrons are in good agreement with the predictions of the model. A serious argument against the standard model—the nonobservation of parity violation effects in experiments with bismuth atoms—has disappeared.[2] The scheme has received weighty confirmation through the SLAC experiment that measured the asymmetry in the scattering of longitudinally polarized electrons by deuterons and protons.[3] The complete set of these data can be well described by the single parameter that is not determined by the model—a value of the Weinberg angle with $\sin^2\theta_W$ in the region of 0.20 (Ref. 4), which fixes the structure of the neutral current. Also a problem of the past is the objection to the characteristic multiplet structure of the model [all SU(2) doublets are left-handed and all singlets right-handed] associated with neutrino processes in which charged currents participate: The so-called y anomaly, which gave birth from time to time to right-handed quark currents, has ceased to exist.

The discovery of parity violation in neutral currents in conjunction with experiments on the scattering of neutrinos and antineutrinos by electrons and nucleons confirmed the doublet structure of the left-handed leptons and quarks and the singlet structure of the right-handed fermions of the standard model and ruled out numerous gauge models proposed as alternatives to the Weinberg–Salam model [vector model, "symmetric"

 0038-5670/80/010001-20$01.10

SU(2)⊗U(1) model, "left-right" $SU(2)_L \otimes SU(2)_R \otimes U(1)$ model, etc.], which to some extent or other introduce right-handed doublets. The experiments have also confirmed an important feature of the Weinberg–Salam scheme, namely, the "minimality" of the spontaneous symmetry breaking mechanism, according to which the scalar Higgs meson H is placed in the lowest, doublet SU(2) multiplet, which leads to a simple connection between the masses of the $W^{\pm}$ and Z^0 bosons:

$$\beta^2 \equiv \frac{M_W^2}{M_Z^2 \cos^2 \theta_W} = 1.$$

In the general case, when there is only one Higgs multiplet with isospin I and the component with projection I_3 acquires vacuum expectation value, we obtain $\beta^2 = (I^2 - I_3^2 + I)/2I_3^2$. Then the cross sections for scattering of ν_μ or $\bar{\nu}_\mu$ by electrons, for example, must be multiplied (compared with the corresponding cross sections in the minimal model) by the factor β^2 (Ref. 5).

The neutral current experiments[6] (the ratio σ neut. curr/σ char. curr) give $\beta = 1.01 \pm 0.03$, which corresponds very well to the doublet structure of the Higgs meson. This, in its turn, means that the mass matrix of the gauge bosons (W_1, W_2, W_3, B) has the structure $(m_{W_1} = m_{W_2} = m_{W_3})$, reflecting an additional global symmetry of the Weinberg–Salam model,[7] the use of which may be of interest (see Ref. 8).

Taken together, these experimental confirmations of the predictions of the standard model already permit us to say (even though the $W^{\pm}$ and Z^0 bosons, which are essential elements of the scheme, have yet to be discovered) that this model is the theory of the weak and electromagnetic interactions, at least up to energies of the order of several tens of GeV.

The family of quarks are, from the point of view of their weak and electromagnetic interactions, excellently described by the $SU(2)_L \otimes U(1)$ scheme, and the convincing list of its experimental confirmations given above leads to the conclusion that, like the leptons, all left-handed quarks form doublets and right-handed quarks singlets.

The discovery of the c quark, predicted on the basis of this scheme to explain the suppression of strangeness-changing neutral currents[9] and capable (together with the then known electron, muon, and u, d, and s quarks) of canceling the Adler anomalies (needed for renormalizability),[10] the discovery in 1975 of the τ lepton and the detailed investigation of its properties[11] (V – A coupling to the corresponding neutrino ν_τ), the subsequent proof (1977) of the existence of the b quark with charge –1/3 (Ref. 12) and the high probability—because of the need once more to cancel the Adler anomalies (which again arise if a restriction is made to only the τ lepton and b quark)—that there exists a partner to the b quark with charge +2/3 (t quark) mean that there are at least three left-handed doublets of leptons,

$$\binom{\nu_e}{e^-}_L, \ \binom{\nu_\mu}{\mu^-}_L, \ \binom{\nu_\tau}{\tau^-}_L, \tag{1.1}$$

and right-handed singlets of charged leptons $e_{\bar{R}}, \mu_{\bar{R}}, \tau_{\bar{R}}$ (left-handed antileptons), three left-handed doublets of quarks with $Q = 2/3, -1/3$,

$$\binom{u}{d}_L, \ \binom{c}{s}_L, \ \binom{t}{b}_L, \tag{1.2}$$

and the right-handed singlets

$u_R, d_R, c_R, s_R, t_R, b_R$.

We shall assume that the masses of all the neutrinos are zero [the experimental upper limits are as follows[1)]: $m_{\nu_e} < 35$ eV (Ref. 13), $m_{\nu_\mu} < 570$ eV (Ref. 13), $m_{\nu_\tau} < 250$ MeV (Ref. 15)].

b) Violation of *CP* invariance

It is important to note that six quarks is the minimal number needed for a natural introduction of CP violation[16] [if one does not consider CP violation due to extension of the sector of the Higgs mesons (which means giving up "minimality"), through which one can introduce an arbitrary relative phase between the vertices of the interaction between the various scalar mesons and the fermions,[17] or CP violation associated with the possibility of letting some quarks participate in right-handed currents.[18] Both these possibilities are rather unaesthetic]. Let us briefly elucidate this point.

Suppose that for generality we have N doublets of left-handed quarks:

$$\binom{q_{Ai}}{q_{Ci}}_L \quad (i = 1, 2, \ldots, N) \quad \begin{matrix} (q_A \text{ is an upquark) with charge } +2/3, \\ q_C \text{ is a downquark) with charge } -1/3). \end{matrix}$$

In the standard scheme, the weak charged current can be written in the form

$$J_\mu = (\bar{q}_{A1}, \bar{q}_{A2}, \ldots, \bar{q}_{AN}) \gamma_\mu (1 + \gamma_5) T \begin{pmatrix} q_{C1} \\ q_{C2} \\ \vdots \\ q_{CN} \end{pmatrix},$$

where T is a unitary $N \times N$ matrix with, in general, N^2 real parameters. However, not all of them are physical, since we can always redefine the quark states $q_{Ai}, q_{Ci}: q_{Ai} \to q_{Ai} e^{i\varphi_{Ai}}$, etc., without changing the physical consequences. Since there are altogether $2N$ quarks, we have, bearing in mind that there is one common phase in the matrix element, $2N - 1$ phase parameters. Among the remaining $(N-1)^2$ real parameters there are $N(N-1)/2$ angles (generalized Cabibbo angles) [since an $N \times N$ orthogonal matrix admits $N(N-1)/2$ real parameters], and the remaining $(N-1)^2 - [N(N-1)/2] = (N-1)(N-2)/2$ parameters are phases in the T matrix, and it is therefore only for $N > 2$ that there arises the essential complexity in the current leading to CP violation. For one quark doublet it is, of course, not necessary to introduce even a Cabibbo angle, and for two doublets (as was the case in the standard model prior to the discovery of the Υ meson) there is only one angle and no essential phase in the current.

Unfortunately, it is impossible at present to determine exactly all three angles θ_i and the phase δ of CP violation in the T matrix for $N = 3$. The angle θ_1 is the Cabibbo angle θ_C, which mixes the d and s quarks and is determined from the β decay of the neutron and is equal to $13.2 \pm 0.5°$ (from the decay of K mesons and hyperons it is $13.2 \pm 0.2°$). Lepton decays with $|\Delta S| = 1$ give $\theta_3 < 16°$, and analysis of the mass difference of the

[1)]Cosmological arguments[14] give much lower neutrino masses: $m_{\nu_\mu} < 400$ eV, $m_{\nu_\tau} < 30$ eV.

K_S^0 and K_L^0 mesons gives $\theta_2 < 30°$. These upper limits can be somewhat reduced[19]: $\sin^2\theta_3 < 0.06$ and $\sin^2\theta_2 < 0.1$.

To determine θ_2 and θ_3, we require a detailed analysis of the weak processes of production and decay of particles containing b and t quarks, which is a matter for the future. However, on the basis of what we already know about θ_1, θ_2, and θ_3 we can say that the three doublets in (1.2) are mixed comparatively weakly, and the main weak transitions induced by charged currents are $u \leftrightarrow d$, $c \leftrightarrow s$, $t \leftrightarrow b$, these being proportional to c_1, $c_1c_2c_3$, and c_2c_3, respectively ($c_i \equiv \cos\theta_i$), while the transitions proportional to $s_i \equiv \sin\theta_i$ ($i = 1, 2, 3$) become important only when the main transitions are forbidden by a quantum number selection rule (as is the case, for example, in decays of strange particles and must apparently be the case in decays of particles containing b quarks).

The mixing of the quarks with $Q = -1/3$ (to accuracy s_i^2) ($s_1^2 = 0.05, s_2^2 < 0.1, s_3^2 < 0.06$) takes the form

$$\left.\begin{aligned} d' &= c_1 d - s_1 c_3 s, \\ s' &= s_1 c_2 d + c_1 c_2 c_3 s + (c_1 c_2 s_3 + s_2 c_3 e^{i\delta})\, b, \\ b' &= (c_1 s_2 c_3 + c_2 s_3 e^{i\delta})\, s - c_2 c_3 e^{i\delta} b. \end{aligned}\right\} \tag{1.3}$$

With regard to the phase parameter δ, which characterizes the CP violation, we can deduce that the decay $K_L^0 \to 2\pi$ tells us only that $\sin\delta > 5 \times 10^{-3}$ or $\delta > 0.3°$, since in this scheme the CP violating amplitudes with initial and final particles constructed only from the "ordinary" u, d, and s quarks are proportional to s_3 (one can show that they must also be proportional to $m_t^2 - m_c^2$).

If at least one neutrino has nonzero mass, there can also be weak mixing in the lepton sector.

c) "Natural" conservation of quark flavor by neutral weak currents

The experiments definitely indicate that neutral currents which change the flavor (strangeness, charm) are strongly suppressed. In the case of strangeness, the situation is well known—the suppression of neutral currents with $|\Delta S| = 1$ by many orders of magnitude compared with $G = 10^{-5} m_p^{-2}$ played a key role in the prediction of a new quantum number and is the basis of the GIM scheme.[9]

With regard to $|\Delta C| = 1$ neutral currents, if they exist, there must be appreciable D^0—$\bar{D}^0$ mixing, which, in its turn, would lead to an appreciable fraction of final states with $\Delta S = \pm 2$ in, for example, e^+e^- annihilation. In the case of complete mixing, the ratio

$$\varepsilon = \frac{N_{(\pm\mp)} - N_{(\pm\pm)}}{N_{(\pm\mp)} + N_{(\pm\pm)}},$$

where $N_{(\pm\mp)}$ is the number of events in e^+e^- anihilation with two K mesons with opposite charges and $N_{(\pm\pm)}$ is the number of events with a pair of K mesons of the same charge) must be close to zero; in the absence of mixing, $\varepsilon = 1$.

The experiment[20] gives $\varepsilon = 0.76 \pm 0.17$, which definitely rules out $|\Delta C| = 1$ neutral currents of order G.

Similarly, information on the suppression of neutral currents with $|\Delta C| = 1$ can be obtained by studying the production by a neutrino beam of charmed particles ($\nu + u \to \nu + c$) leading to a lepton of the "incorrect" sign in the final state [$c \to \mu^+(e^+) + X$]. The experimental data,[21] from which there follow upper limits on the cross sections for the production of charmed particles and the probabilities of their decays,

$$\frac{\sigma(\Delta C \neq 0,\ \text{n.c.})}{\sigma(\text{total, n.c.})} \leqslant 2.6\%, \quad \frac{w(\Delta C \neq 0,\ \text{n.c.})}{w(\text{total, n.c.})} \leqslant 2\%,$$

also rule out neutral currents with $|\Delta C| = 1$.

Thus, the experiments definitely indicate that in the lowest order in G neutral currents with both $|\Delta S| = 1$ and with $|\Delta C| = 1$ are suppressed. We can also speak of flavor conservation in the lepton sector, in which the muon and electron numbers are conserved with good experimental accuracy.

All that we have said above in this section suggests that there is a "natural" conservation by the neutral currents of all fermion flavors (the neutral currents are diagonal with respect to the flavor) (Weinberg, Glashow,[22] Paschos[23]).

By "natural" flavor conservation we mean that it is a consequence of the structure of the gauge group and its multiplet content irrespective of the choice of parameters of the theory such as the quark masses or mixing angles of Cabibbo type.

The requirement of "natural" flavor conservation imposes fairly stringent restrictions on the choice of the group and its multiplet structure, especially in the quark sector. Stated in general terms, the point is that in any gauge theory the vector bosons are the generators of the corresponding gauge algebra, and therefore they can couple only states belonging to one irreducible representation. It follows that the generators of the neutral currents cannot couple quarks of different flavors with equal electric charges if the quarks belong to two different representations of the gauge group.

The actual proof of this assertion for the group $SU(2) \otimes U(1)$ is very simple if we wish to forbid neutral flavor-changing transitions in the first order in G. The matrices coupling the left-handed (right-handed) fermions in the neutral weak current (Z^0 charges) have the form

$$Y_L = I_{3L} - \sin^2\theta_W Q,$$
$$Y_R = I_{3R} - \sin^2\theta_W Q,$$

where I_L and I_R are, respectively, the left- and right-handed isospin of the group SU(2), and Q is the charge [U(1) generator].

Flavor conservation means that Y_L and Y_R are diagonal in the quark basis (i.e., in the basis in which the mass matrix M of the quarks is diagonal).

"Natural" flavor conservation means that Y_L and Y_R must be diagonal *irrespective* of parameters such as the masses (and the Cabibbo angles), i.e., irrespective of M. Then the matrices Y_L and Y_R must act on any set of quarks with the same charge as matrices that are multiples of the unit matrix, i.e., they must depend only on the charge. It follows that quarks of the same charge must have the same I_{3L} and I_{3R}.

If now we wish to forbid flavor-changing neutral currents in the order αG, we must consider not only the operators of the neutral currents in the first order, corresponding to pole diagrams with exchange of the Z^0 boson, but also all operators (that do not change the charge) in the second order corresponding to the diagrams of Fig. 1 (for schemes that contain as gauge fields only $W^{\pm}$, Z^0, and γ).

In this case, because of the algebraic structure of the neutral and charged currents (participating here) and the relation $I_{\pm}I_{\mp}=I^2-I_3^2\pm I_3$, the effective matrix element of the neutral current in the second order will contain the operators $(I_{3L})^2$ and $(I_{3R})^2$ (for the diagrams in Figs. 1a–1d) and the operators $(I_L)^2$ and $(I_R)^2$ (for the diagrams in Figs. 1e and 1f) (this is true if terms $\sim m_q/m_W$ are ignored). Again requiring that the neutral current be diagonal with respect to the flavor irrespective of the mass matrix, we find that quarks with given charge must have the same isospins I_L and the same isospins I_R. In the case of $SU(2)\otimes U(1)$, this means that quarks of the same charge and the same helicity are either all singlets or all doublets.

We see that "natural" flavor conservation by the neutral currents imposes fairly stringent restrictions on the quark sector of the scheme. In particular, we cannot, if, for example, we restrict ourselves to charge $|Q_q|\leqslant 2/3$ of the quarks, have an unequal number of quarks with $Q=2/3$ and $Q=-1/3$, i.e., the number of quarks must definitely be even; this rules out the SU(3) scheme, in which there is a neutral current with the quantum numbers of the K^0 meson; if no other exactly conserved quantum numbers are introduced apart from the charge Q, then (if $|Q_q|\leqslant 2/3$) the only model permitted beside the standard model is the purely vector model $SU(2)_L\otimes SU(2)_R\otimes U(1)$ with degeneracy of the masses of quarks of the same charge.

There is one further important element of the theory which is strongly restricted by the requirement of natural flavor conservation; this is its Higgs structure.

It is, generally speaking, obvious that if the interaction of the Higgs scalar H with the quarks is nondiagonal with respect to the flavor, the $d+H\to s$ vertices can induce, for example, transitions with $|\Delta S|=2(\bar{s}+d\to H\to s+\bar{d})$ on an undesirable scale (the mass difference Δm of the K_S^0 and K_L^0 mesons requires that the corresponding effective coupling constant G_H of the four-fermion interaction induced by H exchange should not exceed $G\times 10^{-5}$).

The typical constant of the coupling of quarks to the H boson, $\Gamma[\bar{q}\Gamma_i(1+\gamma_5)qH_i-\bar{q}\Gamma_i^{+}(1-\gamma_5)qH_i^{+}]$, is of order

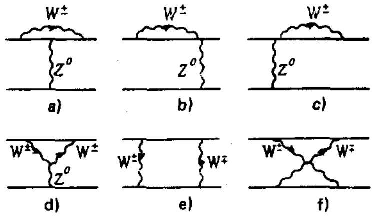

FIG. 1.

$m_q|\langle H\rangle$, where m_q is the characteristic mass of the quarks (or the difference of the masses), and $\langle H\rangle$ is the characteristic vacuum expectation value corresponding to the minimum of the potential $V(H)=\mu^2|H|^2+\lambda H$, $\langle H\rangle^2=-\mu^2/2\lambda$, $\mu^2<0$. The mass of the H meson is determined by the four-boson self-interaction and is of order $\sqrt{\lambda}\langle H\rangle$:

$$m_H^2=-2\mu^2\sim\lambda\langle H\rangle^2.$$

Thus, the effective four-fermion interaction induced by exchange of a Higgs boson has a coupling constant G_H of order

$$G_H\approx\frac{\Gamma^2}{m_H^2}\approx\frac{m_q^2}{\lambda\langle H\rangle^4}.$$

Taking $m_q\approx 1$ GeV and $\langle H\rangle\approx G^{-1/2}=300$ GeV $[m_{W^{\pm}}^2=(g^2/4)\langle H\rangle^2]$, we find that $G_H/G\approx 10^{-5}\lambda^{-1}$, and since $\lambda\ll 1(\lambda\sim\alpha)$, we see that if we do not eliminate the nondiagonal transitions $d\to H+s$ we can have an unacceptably large mass difference Δm.

Requiring thus that the coupling of each neutral Higgs boson to the fermions should conserve "naturally" all flavors [i.e., that the matrices Γ_i, applied to the quark flavor indices, be diagonal in the basis in which $M=\Gamma_i\langle H_i\rangle+M_0$ are diagonal, $M(Q)=\Gamma_Q\langle H_Q^0\rangle$], we can show that this requirement is equivalent to the requirement that the quarks of a given mass acquire their mass either through coupling *to one and only one* neutral Higgs boson[2)] or by means of M_0, an SU(2)-invariant mass term. The possibility that both mechanisms contribute simultaneously is in general ruled out (without the introduction of special tricks).

Our discussion hitherto has in some way or other been concerned with the doublet structure of the group $[SU(2)\otimes w]$. The above arguments can be generalized and strengthened,[27] by invoking additional (but now apparently necessary) arguments associated, for example, with the canceling of Adler anomalies. The generalization proceeds by considering an arbitrary group G_{QAD} that combines the weak and electromagnetic interactions (QAD stands for quantum asthenodynamics, the word deriving from the Greek word $\alpha\sigma\theta\epsilon\nu\eta'\zeta$, which means weak, or nonstrong) and includes neutral gauge

[2)]If there are n Higgs doublets $H_i=(H_i^+/H_i^0)$ $(i=1, 2, \ldots, n)$ with vacuum expectations $\langle H_i^0\rangle=\varkappa_i$, then, going over by means of a unitary transformation to the basis of χ^0 and $\xi_k^0(k=1, 2, \ldots, n-1)$, in which all $\langle\xi_k^0\rangle=0$ and $\langle\chi^0\rangle=\sqrt{\sum_i|\lambda_i|^2}$, we can obtain the "true" Higgs field χ^0 whose vacuum expectation value $\langle\chi^0\rangle$ is responsible for the mass of the fermions and the symmetry breaking $SU(2)\otimes U(1)\to U(1)$. The coupling of the χ^0 field to the fermions has the same structure as in the theory with one Higgs doublet, and χ^0 does not give flavor-changing interactions. Its mass cannot be arbitrary, since it is proportional to the constant $\sqrt{\lambda}$. In general, the remaining fields ξ_k^0 transmit flavor, but their masses can be chosen arbitrarily large without changing the structure of the theory, so that flavor-changing neutral currents can be suppressed arbitrarily strongly.[24] It is also appropriate to note that the electric dipole moment of the neutron in the case of one Higgs scalar in the $SU(2)_L\otimes U(1)$ scheme is of order 10^{-29} $e\cdot$cm, whereas it can be of order $10^{-24}-10^{-25}$ $e\cdot$cm in the case of several scalars.[25] In this last case, the relative probability of the decay $\mu\to e+\gamma$ is of order 10^{-8} (Ref. 26), contradicting the experiments ($\lesssim 2\times 10^{-10}$).

bosons Z^i corresponding to the generators of the group Y^i. Introducing as a generalization of Y_L and Y_R the matrices Y_L^i and Y_R^i of the coupling of the Z^i bosons to the L and R quarks of the theory, we can assert that if the neutral currents that change the quantum numbers of any of the flavors by unity in the order G are to be suppressed, than all the Y_L^i and Y_R^i must be mutually commuting, simultaneously diagonalizable linear operators of the QAD group. There then follows a generalization of the assertions made above concerning Y_L^i and Y_R^i at the tree and single-loop levels and concerning the minimality of the Higgs mechanism for generation of the mass of quarks with equal charge.

It should be noted that almost any gauge group can be taken as a "natural" group if one relaxes the restrictions on the charge composition of the quarks (allowing them to have charges $|Q_q| \geq 2/3$) and the vector bosons (by allowing them and, accordingly, the charged currents to have charges greater than or equal to two). For example, restricting ourselves to groups of rank $r=2$, we shall have for the maximal charges of the gauge bosons (and charged currents)[27] $Q_{max}=2$ in SU(3), $Q_{max}=3$ in SO(5), and $Q_{max}=5$ in G_2. Similarly, in SU(3), for example, it is necessary to introduce a quark with $Q=5/3$. The situation is the same with the other groups $r=2$. At higher ranks, the charge must increase considerably [in SU(4), for example, $Q_{max}=5$ and the charges of the quarks are (besides $Q=2/3$, $-1/3$)5/3, $-4/3$, and so forth].

On the basis of what we have said, it can be shown[27] that the requirement of "naturalness" and the decision not to introduce charged currents with $|\Delta Q| \geq 2$ and exotic charges of the quarks in conjunction with minimality in the Higgs sector leave as candidate for G_{QAD} only the group $SU(2)\otimes[U(1)]^N$, where $N \geq 1$. It is obvious that for arbitrary N this group is not free of Adler anomalies. To cancel the anomalies, it is necessary that each quark isodoublet cancel a corresponding lepton isodoublet (with allowance for the color of the quarks!). It is also necessary that there be equal numbers of right-handed A quarks ($Q=2/3$), right-handed C quarks ($Q=-1/3$), and right-handed C leptons ($Q=-1$). Of course, the number of right-handed massless neutral A leptons (l_{AR}) is not fixed by this condition.

As a result, there are allowed at least the following two possibilities for the *families* of leptons and quarks without anomalies and satisfying the requirement of natural conservation of the quark flavor in neutral currents:

$$\begin{pmatrix} q_A^i \\ q_C^i \end{pmatrix}_L, \quad q_{AR}^i, \quad q_{CR}^i, \begin{pmatrix} l_A \\ l_C \end{pmatrix}_L, \quad l_{CR}, \tag{I}$$

$$\begin{pmatrix} q_A^i \\ q_C^i \end{pmatrix}_L, \quad q_{AR}^i, \quad q_{CR}^i, \begin{pmatrix} l_A \\ l_C \end{pmatrix}_L, \quad l_{AR},\ l_{CR}, \tag{II}$$

$$q_A^i = u^i,\ c^i\ \ t^i, \ldots, q_C^i = d^i,\ s^i,\ b^i, \ldots,$$
$$l_A = \nu_e,\ \nu_\mu,\ \nu_\tau, \ldots, l_C = e^-,\ \mu^-,\ \tau^-, \ldots$$

Here, we have introduced effectively for the first time the color group $SU(3)_c$ in the form of the superscripts of the quark symbols, since the tripling of the number of quarks associated with color is, as we have just seen, responsible for canceling of the anomalies.

Thus, the further unification of the strong interactions with the "electroweak" interactions in a unified symmetry is to a considerable extent predetermined, but this is apparently difficult to avoid if we assume that $SU(3)_c$ is an absolutely exact symmetry of the strong interactions of quarks and gluons.

Are other possibilities besides (I) [which corresponds to the standard Weinberg-Salam scheme (1.1) and (1.2)] and (II) realized? In other words, is it possible to cancel anomalies for any N?

If we write down all the nontrivial conditions of cancellation of the anomalies,[27] (i.e., which do not follow from the concrete choice of the charges of the members of the family and the representations, in other words, include only diagonal matrices Y_L^i and Y_R^i)

$$\Sigma Y_L^i = 0,$$
$$\Sigma Q^2 (Y_L^i - Y_R^i) = 0,$$
$$\Sigma Q (Y_L^i Y_L^j - Y_R^i Y_R^j) = 0,$$
$$\Sigma (Y_L^i Y_L^j Y_L^k - Y_R^i Y_R^j Y_R^k) = 0$$

(summation over all members of the family), then, besides the usual solution for the hypercharge corresponding to (I) ($q_{A(C)L}$; $l_{A(C)L}$; q_{AR}, q_{CR}; l_{AR}, l_{CR}) with hypercharge $Y=(1/3, -1;\ 4/3, -2/3; 0, -2)(N=1)$, there is one further solution with the new "hypercharge"

$$Y' = (0,0;\ 1,\ -1,\ 1,\ -1),$$

so that in principle a group with $r=3$ is also possible:

$$G_{QAD} = SU(2) \otimes [U(1)]^2 \qquad (N=3).$$

Bearing in mind that $Y' \sim T_{3R}$, we can say that $SU(2)_L \otimes [U(1)]^2$ is isomorphic to $SU(2)_L \otimes SU(2)_R \otimes U(1)$, which, prior to the experimental discovery of parity nonconservation in neutral currents,[2,3] was a serious alternative to the Weinberg-Salam $SU(2)_L \otimes U(1)$ scheme and in its simple form contradicts these experiments.

There is a third solution to the above system, but completely unphysical:

$$Y'' = (0,\ 0;\ -7.5;\ 1,\ -\sqrt[3]{35}).$$

Thus, the rather tortuous (but full of brilliant experimental discoveries and theoretical ideas) path traversed in the last few years that we have tried to reconstuct briefly here suggests that the most natural group for the "electroweak" interaction is

$$G_{QAD} = SU(2)_L \otimes U(1), \qquad r=2,$$

with the families (I) (if there are no right-handed neutrinos in nature) or

$$G_{QAD} = SU(2)_L \otimes [U(1)]^2, \qquad r=3$$

with the families (II) (if there are right-handed neutrinos; then the neutrinos could be massive) as basic entity of 15 or 16 particles, respectively, in each family (with allowance for color).

At present, we know three families I (or II) (except for the as yet undiscovered t quark). Through the quark sector, these families are slightly mixed (we have no grounds for believing that in the lepton sector such a mixing occurs; if the neutrino masses are zero, there is certainly no mixing); the generalized Cabibbo angles and the phase δ of CP violation are not known theoretically and are external to the scheme, as is the Weinberg angle θ_W, which determines the connection between the groups $SU(2)_L$ and $U(1)$ ($\tan\theta_W = g'/g$).

In what follows, we shall consider in detail the scheme with the families (I), i.e., $G_{QAD} = SU(2)_L \otimes U(1)$, which in no way contradicts the experiments and proceeds from the assumption that all the neutrinos are left-handed.

Hitherto, the strong interactions of the quarks, which, we believe, are correctly described by quantum chromodynamics (QCD), have not entered our considerations except for the use of the necessary color degrees of freedom in the cancellation of the Adler anomalies and the counting of the members of the family.

In 1974, Georgi and Glashow[28] unified the group of the Weinberg–Salam theory to the QCD group $SU(3)_c$ in SU(5) symmetry.

At that time, this unification was regarded as one of many possiblities {the first unification of the weak, electromagnetic, and strong interactions on the basis of a single group was made by Pati and Salam[29]; then followed many other schemes [SO(10)],[30] the exceptional groups,[31] and so forth}. It was the same with the Weinberg–Salam model, which, although it was the simplest and most economic, had competitors worthy of more or less consideration.

Since the summer of 1978, especially after the Tokyo conference, at which the successful verification of the Weinberg–Salam model was summarized, interest in SU(5) symmetry as the most direct and economic unification of the "correct" group of the electroweak $SU(2)_L \otimes U(1)$ interaction with the QCD group $SU(3)_c$ in the framework of a single non-Abelian group revived with new vigor.

We shall now turn to the exposition of the scheme of such a unification (the corresponding models and constructions are now referred to by rather fine sounding names such as Grand Unification Models, Grand Synthesis, and so forth).

2. WHY IS UNIFICATION OF $SU(2)_L \otimes U(1)$ AND $SU(3)_c$ NEEDED?

The Weinberg angle θ_W, which determines the ratio of the weak and electromagnetic coupling constants [or rather g' and g in the Weinberg–Salam theory, $\tan\theta_W = g'/g$ ($\sin\theta_W = e/g$)], is an external parameter for the theory and there is no way in which it can be calculated; in other words, in the $SU(2)_L \otimes U(1)$ scheme the weak and and electromagnetic interactions are not unified to the end and the unified theory is not in fact perfect, since it contains, not one constant (the electric charge), but two (g and g' or e and θ_W). In a unifying scheme, θ_W must be determined from the symmetry properties of a single group.

In this scheme, the leptons and quarks of the family (I) must be combined. It is to be expected that since there must be only one common coupling constant for all the interactions in the unified scheme and the group U(1) is not factored off, "quantization" of the electric charge must be achieved in it, i.e., a connection established between the charges of the quarks and leptons. The circumstance that in the required scheme there must be just one coupling constant (or, rather, one common vertex function) is of a general nature: In the general case, the Lie algebra of the gauge group is the direct sum of a semisimple Lie algebra and an Abelian Lie algebra. If the latter is absent (and it is such a group that we seek), the gauge group is compact and, taking, naturally, the charge operator as a generator of it (corresponding to the unique massless gauge field), we arrive at the conclusion that all the constants are multiples of the electric charge. Experimentally, the coupling constants of the weak, electromagnetic, and strong interactions differ considerably. However, in the $SU(2)_L \otimes U(1)$ scheme the weak and electromagnetic interactions become comparable for $q^2 \gg M_{W,Z}$, the effective coupling constant of asymptotically free QCD decreases logarithmically, and at an energy of order $\approx 10^{15}$ GeV is comparable with the effective $SU(2)_L \otimes U(1)$ constants. It is at these energies that the required symmetry becomes exact.

It is to be hoped that in this scheme, which combines leptons and quarks, a connection can be established between the masses of the leptons and the quarks. It would also be good to have the possibility of obtaining the Cabibbo angles (and the phase δ?) from the theory.

Finally, it must be borne in mind that the Weinberg–Salam theory is not asymptotically free for $q^2 \gg M_W^2$. If we wish to preserve this property of the theory at high energies, it must be modified by new (very weak for $q^2 \lesssim M_W^2$) interactions that ensure its asymptotic freedom for $q^2 \gg M_W^2$. These interactions will be associated with a new symmetry that combines the strong, electromagnetic, and weak (and gravitational?) interactions.

Well known is Landau's idea[32] of a possible modification of quantum electrodynamics with its ghost (Landau) role by gravitation, which at energies Λ corresponding to the Landau role, $\alpha \approx (3\pi/\nu)\ln(\Lambda^2/m_e^2)$ (ν is the number of point fermions, $\alpha = 1/137$), becomes of the order of electromagnetism, which means that electrodynamics is not closed. Since the effective charge in this critical region is of order unity, Λ is of order 10^{19} GeV (the Planck mass $G_{gr}^{-1/2}$, $G_{gr} = 6\times10^{-39}\ 1/m_p^2$). From this there follows a first estimate of the number of point fermions $\nu = 12$. Many modern estimates of the number ν (see Sec. 3g below) use similar arguments to a greater or lesser extent. The question of the maximal number of elementary fermions or, in a different language, the admissible number of families of the type I (or II) is extremely topical.

A scheme in which the weak, electromagnetic, and strong interactions are manifestations of a unified fundamental interaction with one coupling constant is highly aesthetic (although the aesthetic aspect is relative and subject to strong variation in time). However, it must be borne in mind that such an approach does not take into account gravitational interactions, and this may be an important aspect that is not used by such unification schemes.

3. SU(5) MULTIPLETS. CONSEQUENCES

a) Choice of the group

The family (I) at which we have arrived has the following multiplet content of the group $SU(3)_c \otimes SU(2)$, which must serve [together with the factor U(1)] as the basis for the construction of the unified group (sometimes, such a group is denoted by G_{QHD}, where QHD stands for quantum holodynamics, from the Greek word holos, meaning complete or whole):

$$\begin{array}{l}(3,2)+2(\bar{3},1)+(1,2)+(1,1),\\ 6q^i_L,\ \ 6q^i_R=6\bar{q}^i_L,\ \ 2l_L,\ \ l_R=\bar{l}_L\end{array} \tag{3.1}$$

[under the $SU(3)_c \otimes SU(2)$ multiplets we have written the number of the corresponding particles, the superscript i labeling the quark color, $i = \mathrm{r}, \mathrm{y}, \mathrm{b}$ (red, yellow, blue)]. The rank r of G_{QHD} must be greater than or equal to four:

$$r(G_{QHD}) \geqslant r_1(G_{QCD}) + r_2(G_{QAD}) = 4$$

[for the family (II), for which $r_2 = 3, r \geqslant 5$]. We shall adhere to the principle of "minimality" and restrict ourselves to considering all simple Lie groups of rank 4.

According to Cartan's classification, there are four infinite families of Lie groups (n labels the rank of the group):

$A_n = SU(n+1)$ (linear transformations of complex matrices),

$C_n = Sp(2n)$ (linear transformations of quaternions, symplectic groups),

$B_n = SO(2n+1)$ (linear transformations of real matrices),

$D_n = SP(2n)$

and five exceptional Lie algebras, which act on octonions (a generalization of quaternions by the introduction of a third imaginary unit):

G_2, F_4, E_6, E_7, E_8 (the subscripts correspond to the rank of the group).

Thus, for $r = 4$ we must select the required group from the following candidates among the simple groups:

$$SU(5),\ Sp(8),\ SO(8),\ SO(9),\ F_4.$$

In addition, we could in principle make a choice among the nonsimple groups with $r = 4$:

$$[SU(2)]^4,\ [SO(5)]^2,\ [SU(3)]^2,\ [G_2]^2;$$

the groups $[SU(2)]^4$ and $[SO(5)]^2$ are ruled out, since they do not contain the SU(3) subgroups needed for the color group in (3.1); $[G_2]^2$, F_4, SO(9), SO(8), and Sp(8) do not contain essentially complex representations, which are needed to obtain the correct $SU(3)_c \otimes SU(2)$ content (3.1).

The group $[SU(3)]^2$ (like any group of the form $SU(3) \otimes w$, where w is any group with $r_2 = 2$) is unacceptable, since leptons and quarks having different color properties must be placed in different SU(3) representation, and then the charge generator with zero trace cannot ensure fractional charges of the quarks and nonvanishing of their sum.

This argument also shows that the leptons and quarks must necessarily be combined in representations of the required group, SU(5), remaining, as we see, as the only candidate for this group. This also means that in the corresponding QHD schemes, in which the vector fields corresponding to its generators (or Higgs bosons) carry simultaneously quark and lepton properties (quantum numbers)—fractional electric charge, color, baryon (and lepton) charge—baryon (and lepton) charge conservation will be violated to some extent.

b) Fitting the lepton and quark families in the representations of SU(5)

This can be done simply by considering the $SU(3)_c \otimes SU(2)$ content of some SU(5) multiplets.

The spinor (basic) representation **5** of SU(5) is decomposed with respect to $SU(3)_c \otimes SU(2)$ representations as follows:

$$\mathbf{5} = (3,1) + (1,2).$$

The antisymmetric SU(5) representation $\mathbf{10}$($\mathbf{5}\times\mathbf{5} = \mathbf{15}c + \mathbf{10}a$) has the following $SU(3)_c \otimes SU(2)$ content:

$$\mathbf{10}a = (3,1) + (3,2) + (1,1),$$

so that the reducible representation $\bar{\mathbf{5}} + \mathbf{10}$ has the content

$$\bar{\mathbf{5}} + \mathbf{10}a = (3,2) + 2(\bar{3},1) + (1,2) + (1,1),$$

i.e., it reproduces the family (I) of 15 quarks and leptons.

The quarks and leptons are placed in the antiquintet $\bar{Q}^a_L$ and the decuplet $(D_L)_{ab}(a, b = 1, 2, \ldots, 5)$ as follows:

$$\bar{Q}^a_L = (\underbrace{\bar{q}^r_C,\ \bar{q}^y_C,\ \bar{q}^b_C}_{(\bar{3},1)};\ \underbrace{l_C,\ l_A}_{(1,2)})_L, \tag{3.2}$$

$$(D_L)_{ab} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & \bar{q}^b_A & -\bar{q}^y_A & -q^r_A & -q^r_C \\ -\bar{q}^b_A & 0 & \bar{q}^r_A & -q^y_A & -q^y_C \\ \bar{q}^y_A & -\bar{q}^r_A & 0 & -q^b_A & -q^b_C \\ q^r_A & q^y_A & q^b_A & 0 & -l^+_C \\ q^r_C & q^y_C & q^b_C & l^+_C & 0 \end{pmatrix}_L. \tag{3.3}$$

The symmetry we have found obviously satisfies "natural" conservation of quark flavor (and, as we shall see below, satisfies accordingly the principle of "minimality" in the Higgs sector; among all the known grand synthesis schemes it has in this sector the simplest and most economic structure).

Correctly reproducing the 15 members of the family (I), it automatically satisfies the principle of anomaly cancellation (the contributions of the representations $\bar{\mathbf{5}}$ and **10** cancel). Among groups of *arbitrary rank* with the correct $SU(3)_c \otimes SU(2)_L \otimes U(1)$ content, SU(5) is the only group that naturally places the 15 members of the family (I) without anomalies.

One can say that if nature chooses the family of 15 quarks and leptons (I) as the "matrix" for successive "reproduction", then SU(5) is unique. Note that if right-handed neutrinos exist, the above assertions also apply to the group SO(10) with $r=5$:

$$\mathrm{SO}\,(10)^{30}\supset \mathrm{SU}\,(3)_c\otimes \mathrm{SU}\,(2)_L\otimes [\mathrm{U}\,(1)]^2 \sim \mathrm{SU}\,(3)_c\otimes \mathrm{SU}\,(2)_L\otimes \mathrm{SU}\,(2)_R\otimes \mathrm{U}\,(1)$$

with the family (II) (see Sec. 3d below).

We here make a comment. L. B. Okun' has noted that every particle of the decuplet (3.3) can be constructed from two *right-handed* quintet particles and the gravitino—the massless neutral particle of spin 3/2 that plays an important part in the theory of supergravity.[32] This is readily seen on the basis of the fact that the charge—color content of the particles in (3.3) can be exactly reproduced by multiplication of the right-handed quintets Q_{Ra}:

$$(D_R)_{ab}=\frac{1}{\sqrt{2}}(Q_{Ra}Q_{Rb}-Q_{Rb}Q_{Ra}).$$

By means of the gravitino, the (left-handed) $(D_R)_{ab}$ can be carried into $(D_L)_{ab}$ in (3.3).

Is this remark of the nature of a mnemonic or do we have here something more serious relating to the intrusion of gravity or, more precisely, supergravity and the geometry of superspace, that can provide the basis for the theory of a superfield which includes all interactions and all particles of nature?

c) "Quantization" of charge. On nonconservation of the baryon number

Since the charge operator Q is a generator of the group SU(5), its trace is zero for an arbitrary representation of the group: $(\mathrm{Tr}\,Q)=0$ (it is this condition that leads to the canceling of the Adler anomalies).

It follows immediately that

$$(\mathrm{Tr}\,Q)_5=-3Q_{q_C}+Q_{l_C}=0,$$
$$(\mathrm{Tr}\,Q)_{10}=-3Q_{q_A}+2Q_{\bar{l}_C}=0,$$
$$Q_{q_C}=\frac{1}{3}Q_{l_C},\quad Q_{q_A}=\frac{2}{3}Q_{\bar{l}_C},$$

i.e., we obtain a "quantization" of the charge that correctly establishes the connection between the quark and lepton charges.

Note that if we were to proceed from the color group $\mathrm{SU}(n)_c$, the obtained connection between Q_q and Q_l would be

$$Q_{q_C}=\frac{1}{n}Q_{l_C},\quad Q_{q_A}=\left(1-\frac{1}{n}\right)Q_{\bar{l}_C},$$

i.e., we obtain a connection between the charges of the quarks and color.

Bearing in mind what we said concerning the nonconservation of the baryon, B, and lepton, L, numbers when quarks and leptons are combined in representation of G_{QHD} at the end of Sec. 3a, we can speak of a synthesis of the following very important aspects: quantization of the electric charge, renormalizability (no anomalies), fractional charge of the quarks, and nonconservation of B and L.

d) Vector bosons of SU(5). Consequences of exact symmetry. Symmetry breaking

At the energies accessible today, SU(5) symmetry is strongly broken. As we have already noted, the coupling constants g_1, g_2, and g_3 of the groups U(1), SU(2), and $\mathrm{SU}(3)_c$ have a chance of becoming comparable only at very high energies $\approx 10^{14}-10^{15}$ GeV (Fig. 2), and only then can one speak of exact SU(5) symmetry.

At present-day energies ($q^2\lesssim 100\ \mathrm{GeV}^2$) $\alpha_3\equiv\alpha_s(q^2)\approx 1/5$, $\alpha_2=\alpha/\sin^2\theta_W\approx 1/26$, $\alpha_1=\alpha/C^2\cos^2\theta_W\approx 1/67$ [$C^2=3/5$ (see below), $\alpha_i\equiv g_i^2/4\pi$].

In the limit of exact symmetry,

$$\bar{\alpha}(M)=\alpha_1=\alpha_2=\alpha_3=\alpha\cdot\frac{8}{3}=0.019.$$

It can be seen from this that $\alpha_3=\alpha_s(\mu^2)$ can become equal to $8/3\alpha$ only at energies of order 10^{15} GeV if the strong interactions do not change their group structure right up to these energies. The interactions corresponding to the groups U(1), $\mathrm{SU}(2)_L$, and $\mathrm{SU}(3)_c$ have the form

$$\left.\begin{array}{l} g_1B_\mu\sum\limits_f \bar{f}C\frac{Y_f}{2}\gamma_\mu f,\\ g_2W_\mu\left[\sum\limits_q \bar{q}_L\gamma_\mu\frac{\tau}{2}q_L+\sum\limits_l \bar{l}_L\gamma_\mu\frac{\tau}{2}l_L\right],\\ g_3G_\mu^a\sum\limits_q \bar{q}\gamma_\mu\frac{\lambda^a}{2}q;\end{array}\right\}\qquad(3.4)$$

where f symbolizes q_L, q_R, l_L, l_R; B_μ is the $U(1)$ gauge field that generates the weak hypercharge, which is twice the mean charge field that generates the weak hypercharge, which is twice the mean charge of the multiplet ($B_\mu=\cos\theta_W Z_\mu^0+\sin\theta_W A_\mu$, where A_μ is the photon field); $W_\mu^\pm$ are the charged vector bosons [$W_\mu^\pm=(1/\sqrt{2})\,(W_1\pm iW_2)_\mu$]; and G_μ^a are the gluons ($a=1,\ldots,8$), which with allowance for the particles of the decuplet interact, as is readily seen, with both q_C and q_A quarks and not only with the left-handed but also with the right-handed quarks. The parameter C is a constant that occurs in the relation between the charge Q and the normalized generators T and T_0 of the groups SU(2) and U(1), respectively: $Q=T_3+CT_0$ ($\mathrm{Tr}\,T_\alpha T_\beta=N\delta_{\alpha\beta}$). The constant C is introduced in the U(1) interaction to make the normalizations of all three vertices the same [$\lambda^a/2$ and $\tau/2$ are generators of SU(5), $Y/2$ is only proportional to a generator]. Then, noting that for each generator T_a of the group the trace over the representation $\mathrm{Tr}(T_a)^2$ (no summation over a) does not depend on the number a of the generator,

$$\mathrm{Tr}\left(\frac{CY}{2}\right)^2=\mathrm{Tr}\left(\frac{\tau_i}{2}\right)^2=\mathrm{Tr}\left(\frac{\lambda^a}{2}\right)^2=\frac{1}{2},$$

and calculating $\mathrm{Tr}(Y/2)^2$ on the left-handed quintet $\underline{5}$, $\mathrm{Tr}(Y/2)^2=(1/3)^2 3+(1/4)\times 2=5/6$, we obtain $C^2=3/5$. Since $\tan\theta_W=g'/g=g_1C/g_2$ and in the limit of SU(5) symmetry $g_1=g_2=g_3$, it follows that $\sin^2\theta_W=3/8$ in exact symmetry. We see that we require a strong renormalization of θ_W from the energies at which SU(5) is an ex-

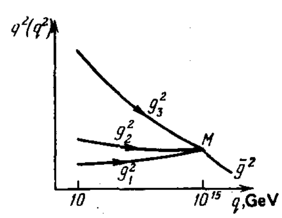

FIG. 2.

act symmetry to present-day energies, where, as we know, $\sin^2\theta_W \approx 0.2$. It is obvious that larger values of $(\theta_W)_{sym}$ [for example, in the unification on the basis of E_7 symmetry, $\sin\theta_W = 3/4$ (see Sec. 3d)] require even higher energies for the unifying symmetry to become exact. $W^{\pm}_{\mu}, Z^0_{\mu}, A_{\mu}, G^{a}_{\mu}$ are the 12 bosons which we know from the standard model and QCD, and which, of course, are massless in exact SU(5). However, there are a further 12 gauge bosons in the SU(5) symmetry which occur with the previous 12 bosons in the SU(5) vector representation **24**; these are triplets with respect to $SU(3)_c$ and doublets with respect to SU(2) [(3, 2)] and have charges $Q = \pm 4/3$ and $Q = \pm 1/3$. They are denoted $X^i_{\pm 4/3}$ and $Y^i_{\pm 1/3}$ and have nonzero baryon and lepton numbers ("leptoquarks").

They lead to effective nonconservation of the baryon and lepton numbers, causing the transitions

$$\bar{d}^i \rightleftarrows e^- (X^i_{\pm 4/3}), \quad \bar{d}^i \rightleftarrows \nu (Y^i_{\pm 1/3}).$$

In complete analogy with the breaking of SU(3) [SU(3) $\otimes$SU(3)] to SU(2) [SU(2)$\otimes$SU(2)], which transforms in accordance with the adjoint representation of the group, the breaking of SU(5) to $SU(3)_c \otimes SU(2)_L \otimes U(1)$ transforms in accordance with the **24**-plet of SU(5).

The 12 bosons $X^i_{\pm 4/3}$ and $Y^i_{\pm 1/3}$ acquire a mass (which, as we shall see, must be huge to ensure the experimental stability of the nucleon) by means of Higgs bosons [which must then form a **24**-plet of SU(5) scalar mesons] or by some other (which?) mechanism. The next stage of the breaking, $SU(3)_c \otimes SU(2)_L \otimes U(1) - SU(3)_c \otimes U(1)$, is standard and, as in the standard model, by means of the SU(5) representation $\mathbf{5} \supset (\mathbf{3}, \mathbf{1}) + (\mathbf{1}, \mathbf{2})$.

The choice of the spinor representation **5** for the Higgs mesons corresponds, as we have seen in Sec. 1a, to "natural" conservation of flavor by the neutral currents. We therefore assume that the Higgs mesons of the representation **5** also give the fermions masses.

Of course, the quintet and decuplet fermions need not acquire mass from the SU(5) representation **5**; they could also acquire mass by means of Higgs bosons from the representation $\overline{\mathbf{45}}$:

$$\bar{5} \times 10 = 45 + 5, \quad 10 \times 10 = \overline{45} + 5 + 50.$$

However, it is natural to take the representation **5** for the scalar mesons that generate the masses of the leptons and quarks in complete correspondence with the "minimality" of the Higgs sector and the natural conservation of flavor and by analogy with the manner in which in the standard $SU(2)_L \otimes U(1)$ model the breaking of the symmetry down to U(1) and the generation of the masses of the leptons proceeds through the spinor representation of the group.

As a result of this choice, the q^i_C quarks and l_C leptons obtain equal masses [mass term $(\bar{\mathbf{5}})_f \times (\mathbf{10})_f (\bar{\mathbf{5}})_H$]:

$$m_{q_C} = m_{l_C}, \quad \text{i.e.,} \quad m_d = m_l, \; m_s = m_{\mu}, \; m_b = m_{\tau}. \tag{3.5}$$

Note that if we were to take the representation **45** instead of **5** we would obtain the relation $m_{q_C} = m_{l_C}/3$ in place of $m_{q_C} = m_{l_C}$ in the limit of exact SU(5) symmetry.

A mass of the quarks q^i_A also arises, but it is not connected to the lepton mass [mass term of the structure $(\mathbf{10})_f \times (\mathbf{10})_f \times (\bar{\mathbf{5}})_H$]. Thus, despite the unification of $SU(3)_C$, SU(2), and U(1) in a single group, the masses of the anoquarks (like the Cabibbo angles θ_i) remain arbitrary parameters in SU(5) symmetry, i.e., they are external parameters for the theory.

The origin of this shortcoming of the scheme with regard to the Cabibbo angles is clear and due to the fact that the families (I) are "decoupled", and therefore only pair flavors of quarks with different charges within a family are coupled by transformations of the unified group. The generators of the group do not couple families of different generations (e^- family, μ^- family, τ^- family).

The situation is here not simple, though at the first glance it appears that there are not so many possibilities to choose among. If, for example, one introduces fermions of different generations in one multiplet of some unified group [as, for example, is done in the group E_7 (see Sec. 4)], the natural conservation of flavor by the neutral currents is lost, and the experimentally observed suppression of the flavor-nondiagonal neutral currents must be enforced by, for example, making the neutral vector mesons corresponding to the generators that couple the different flavors very heavy.

Could it be that natural conservation of flavor by the neutral currents, which to a large extent predetermines the choice of the family (I) as the basic unit for the successive "reproduction" of leptons and quarks, is a Procrusteam bed for the theory and one must proceed differently? It should be borne in mind that the conservation of flavor in the quark and lepton sectors has different levels of experimental justification, and natural flavor conservation restricts the structure of the quark sector especially strongly.

It is also true that all possibilities for searching for symmetry properties that couple quark flavors "horizontally" (see Sec. 5)[34-37] may not have been fully exploited.

It is important to know whether the obtained relations (3.5) between the masses of quarks and leptons within a family mean that the τ—μ—e and b—s—d splittings have the same origin.

It is very important to understand the nature of the coupling between the Higgs bosons and the fermions, since it is this that determines not only the masses but also the Cabibbo type angles, which establish bridges between the families. In particular, if the Higgs scalars satisfy the conditions of natural conservation of flavor by the neutral currents, then, irrespective of the quark masses, the Cabibbo angles are determined by symmetry properties that link the flavors horizontally.[37]

e) From SU(5) energies to "contemporary" energies

The existence of one constant $\bar{\alpha} = \alpha_1 = \alpha_2 = \alpha_3$ in exact SU(5), the equalities (3.5) of the lepton and quark masses, and the value $\sin^2\theta_W = 3/8$ of the Weinberg angle might appear devoid of physical content at first glance, since they hold at energies that, at least for the present, are only of academic interest.

However, this is not so, since we have at our disposal the renormalization group formalism, which enables us to take into account the change of the physical parameters as q^2 varies from M^2 [where M is the mass at which the SU(5) symmetry is exact] to μ^2, where μ is the scale of the energies at present of interest to us ($\mu^2 \approx 10$ GeV^2).[38]

The basis of the variation of the quantities in which we are interested is the variation of the constants $\alpha_i(q^2)$ with q^2, of which we have already spoken (see Sec. 3d, Fig. 2), and which is determined by the renormalization group equations

$$\mu \frac{\partial}{\partial\mu} g_i(\mu) = \beta_i(g_i(\mu)). \tag{3.6}$$

We take μ fairly large, i.e., $m \ll \mu < M$, where m are the masses of the particles, so that Eqs. (3.6) hold and $g_i(\mu^2) \ll 1$. The functions β are calculated on the "observable" group, i.e., $\beta_i = -b_i g_i^2$, where b_i are taken from the calculation of the corresponding loops of the fermions and the "ordinary" ($W^\pm$, Z^0) vector bosons. (The contribution of the Higgs scalars and the superheavy vector bosons is ignored.)

Such an approach is based on a theorem[39,38] which states that all matrix elements with external particles with masses and momenta much less than the masses M_i of the heavy particles can be calculated effectively as if the heavy particles were not present at all.

The fields of the heavy particles are effectively decoupled from those of the light particles, and the part they play reduces to a renormalization of the coupling constants, leading to a dependence of the latter on M_i. In other words, at momenta small compared with the masses of the heavy particles, the dynamics is effectively determined by the sector of the light particles of the renormalized theory.

It is a consequence of this theorem applied to the considered problem that the three groups SU(3)$_c$, SU(2), and U(1) are "decoupled" from one another at "laboratory" energies, and we are dealing with three corresponding β functions that satisfy their own renormalization group equations.

In asymptotically free SU(3)$_c$, $b_3 = 11 - 2f/3$ (Ref. 40), where the first term arises from the vector (in the given case gluon) loops and the second from the fermion loops; f is the number of quark (lepton) flavors [i.e., twice the number of families (I)]. In exact SU(5), of course, $b_2 = b_3$. However, at our energies it is broken down to SU(3)$_c\otimes$SU(2)$\otimes$U(1). Therefore, the intermediate bosons of the group SU(2) contribute to b_2, and the first term in the relation is not 11 but 22/3 [if the weak group were SU(n), then one would have $11n/3$]. The second term in b_2, which is associated with the fermion loops, is the same as for b_3, i.e., $b_3 - b_2 = 11/3$.

For the constant b_1 of the group U(1) we shall have a contribution from only the fermion loops, since B_μ is the hypercharge generator, and the gluons and intermediate bosons of SU(2) do not carry hypercharge. Therefore $b_1 = -2f/3$.

Of course, all this reflects the fact that the non-Abelian groups SU(3) and SU(2) are asymptotically free,[40] while U(1) is a Landau-role group[41] (see also Ref. 42).

As a result,

$$\alpha_i^{-1}(M) = \alpha_i^{-1}(\mu) + \frac{b_i}{2\pi} \ln\frac{M}{\mu},$$

where the values of the constants at $q^2 = \mu^2$ are

$$\alpha_1(\mu) = \frac{1}{67}, \quad \alpha_2(\mu) = \frac{1}{26}, \quad \alpha_3 \approx \frac{1}{5}$$

(see Sec. 3d). It follows that ($C^2 = 3/5$, and we take f equal to 6)

$$\left.\begin{aligned} &\ln\frac{M}{\mu} = \frac{\pi}{11\alpha}\left(1 - \frac{8}{3}\frac{\alpha}{\alpha_s(\mu)}\right), \\ &\sin^2\theta_W(\mu) = \frac{1}{6} + \frac{5}{9}\frac{\alpha}{\alpha_s(\mu)}, \\ &\bar\alpha \equiv \bar\alpha(M) = \alpha_1 = \alpha_2 = \alpha_3 = \frac{8}{3}\alpha = 0.019. \end{aligned}\right\} \tag{3.7}$$

Hence, choosing $\alpha_s(\mu) = 0.2$, we obtain $M/\mu = 2.2\times 10^{15}$ (irrespective of θ_W) and $\sin^2\theta_W = 0.19$ [at one time the value $\sin^2\theta_W = 0.2$ was regarded as a serious argument against the SU(5) scheme, since the experiments tended rather to the value $\sin^2\theta_W = 0.30$. The experiments now give values of $\sin^2\theta_W$ grouped around 0.2]. The renormalization of $\sin^2\theta_W$ is not very sensitive to the choice of $\alpha_s(\mu)$ (and the number f). Taking, for example, $\alpha_s(\mu^2) = 0.3$, we obtain $\sin^2\theta_W = 0.18$. The masses of the X and Y leptoquarks must be fairly close to M, since at the threshold of their production only the derivatives of the constants α_i with respect to $\ln M$ change strongly, while the α_i themselves change little. Further, $M_{X,Y} \le M$ to ensure that all the constants α_i follow the same variation with q^2 determined by the group SU(5) [$b_5 = (11/3)5 - (2/3)f > b_2, b_3$]. As will be seen below, $M_{X,Y} \gtrsim 10^{14}$ GeV on the basis of the experimental stability of the proton.

One can also follow the renormalization of the fermion masses from their values in exact SU(5) to the scale μ of our energies. It is determined by the anomalous dimensions $\gamma_i (i = 1, 2, 3)$ of the fermion mass operators of the groups SU(3)$_c$, SU(2), and U(1).

In the single-loop approximation, again ignoring the contribution $\approx m_q^2/q^2$ and that of the Higgs mesons, we obtain[38,27]

$$\left.\begin{aligned} &\frac{m_{q_A}(\mu)}{m_{q_A}(M)} = \left(\frac{\alpha_s(\mu)}{\bar\alpha}\right)^{4/b_3}\left(\frac{\alpha_2(\mu)}{\bar\alpha}\right)^{9/4b_2}\left(\frac{\alpha_1(\mu)}{\bar\alpha}\right)^{-1/5b_1}, \\ &\frac{m_{q_C}(\mu)}{m_{q_C}(M)} = \left(\frac{\alpha_s(\mu)}{\bar\alpha}\right)^{4/b_3}\left(\frac{\alpha_2(\mu)}{\bar\alpha}\right)^{9/4b_2}\left(\frac{\alpha_1(\mu)}{\bar\alpha}\right)^{1/10b_1}, \\ &\frac{m_{l_C}(\mu)}{m_{l_C}(M)} = \left(\frac{\alpha_2(\mu)}{\bar\alpha}\right)^{9/4b_2}\left(\frac{\alpha_1(\mu)}{\bar\alpha}\right)^{-9/10b_1}. \end{aligned}\right\} \tag{3.8}$$

The expressions (3.8) show that the masses of the quarks increase appreciably with decreasing μ while the masses of the leptons hardly change, so that at $\mu^2 \approx 10$ GeV^2 the masses of the leptons and quarks have diverged considerably, reaching their "observed" values.

Figure 3 shows this schematically.

We cannot use the first of the relations (3.8), since we do not know the connection between the mass of the q_A quarks and the q_C quarks or q_{l_C} leptons in exact symmetry.

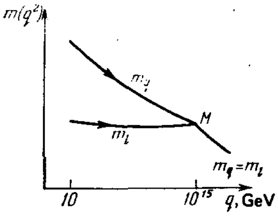

FIG. 3.

From the second and third relations, using $m_{q_C}(M) = m_{l_C}(M)$, we find

$$\frac{m_{q_C}(\mu)}{m_{l_C}(\mu)} = \left(\frac{\alpha_s(\mu)}{\bar{\alpha}}\right)^{4/b_s} \left(\frac{\alpha_1(\mu)}{\bar{\alpha}}\right)^{1/b_1} \approx 3.5\,(f=6). \tag{3.9}$$

This simple formula can be corrected by taking into account the contributions $\sim m_f^2/q^2$ associated with the finiteness of the quark masses, the contribution of the Higgs mesons,[27,43] and the contributions $\sim m_W^2/q^2$.

Let us consider the complete set of input and output data. For example, assuming

$\alpha_s(\mu^2) = \dfrac{12\pi}{25\ln(\mu^2/\Lambda^2)}$ in the interval 0.19 (charmonium)–0.32 (electroproduction) (corresponding to $\Lambda^2 = 0.005$–$0.09\ \mathrm{GeV}^2$),

$m_\mu = 0.105$ GeV, $m_\tau = 1.9$ GeV, and $f = 6$, we obtain

$\bar{\alpha} = 0.022$, $\sin^2\theta_W = 0.20$,

$M = (0.9-3.7)\times 10^{16}$ GeV, $m_s = 0.38-0.5$ GeV, $m_b = 5-5.9$ GeV. The value of $\sin^2\theta_W$ and, to a lesser extent, $\bar{\alpha}$, is very stable against variation of the input parameters (and variation of the approximations). The expression (3.5) and its improved version lead to a good prediction for m_b. The value of m_b is not very sensitive to the SU(5) details, and it is related to the variation of the constants $\alpha_1, \alpha_2, \alpha_3$ for $q^2 < M^2$ and not to the manner in which they merge into a single constant near M. To a large degree, this last is, as we have seen (see Sec. 3d), determined by the value of the parameter C^2 [=3/5 in SU(5), $(n-2)/n$ in SU(n)]; however, C enters the mass ratio m_b/m_τ through the factor in (3.5) associated with the group U(1), which itself is small {we note here that to make the coupling constants of the strong and weak interactions equal in magnitude at masses lower than $M_{\mathrm{Pl}} \approx G_{\mathrm{gr}}^{-1/2}$ it is necessary to "coordinate" the "weak" and "strong" groups appropriately. If, for example, the "weak" group is taken to be SU(3), then it is also necessary to increase the dimension of the "strong" group [to SU(4)], since the rate of approach of the constants is determined by the difference between the corresponding β functions[38,27]}.

The number of flavors f plays an important part in the derivation of the output data. For $f=8$, m_b is increased by 10%, and for $f=10$ by 30%.[27] If we also bear in mind that allowance for the two-loop corrections to the β and γ functions also increases the quark masses[27,45] (m_b by 10% and m_s by 25%), we arrive at the conclusion that f cannot be greater than 8. We shall return to this question below. At this point, we recall that if we had taken the representation 45 rather than 5 for the Higgs multiplet, then in the limit of SU(5) symmetry we would have $m_{q_C} = m_{l_C}/3$.

As a result, the relation corresponding to (3.9) would lead to masses for m_s and m_b that are three times smaller, which strongly contradicts the experiments.

We now also mention the conclusions of the SU(5) scheme concerning the quark masses, which do not correspond to generally adopted ideas and experiments. If, proceeding as above, we wish to obtain information about the d quark, it leads to a unsatisfactory conclusion: $m_d(\mu)/m_s(\mu) = m_e/m_\mu \sim 1/200$, although, of course, it would be preferable to have a value ~1/20 for this ratio of the quark masses at short distances ($\mu \gg 1$ GeV).[46] The mass of the strange quark is found to be rather overestimated [see (3.9)]. The higher corrections increase this discrepancy with the experiments even more. The unsatisfactory conclusion of the SU(5) scheme for the ratios m_s/m_μ (m_d/m_e) and m_d/m_s is also characteristic of a number of other similar unification schemes [for example, with SO(10)]. The reason for this difficulty with the extrapolation to the masses of the light fermions is not clear. It is quite possible that an important part is here played by the Higgs bosons in renormalizing the quark masses.

f) Instability of the nucleon

We have already noted that the SU(5) scheme, which combines quarks and leptons in one multiplet, can lead to nonconservation of baryon (and lepton) charge. In SU(5), the difference between the baryon and lepton numbers is conserved, i.e., the proton can go over into an antilepton, but not into a lepton. The superheavy vector bosons $X^i_{\pm 4/3}$ and $Y^i_{\pm 1/3}$ ("leptoquarks"), which generate exotic lepton—quark and diquark currents, are responsible for the nonconservation of the baryon (lepton) number.

The elementary processes that are induced by these currents and lead to different instability channels of the proton (or neutron) are given in the diagrams in Fig. 4.

At the quark level, the "elementary" Lagrangian has the following baryon-number nonconserving effective four-fermion ($M_X \approx M_Y$) form:

$$\mathscr{L}_{\mathrm{eff}} = \frac{\tilde{G}}{\sqrt{2}}\varepsilon_{ijk}\{\bar{u}_L^{k^c}\gamma_\mu u_L^j\,[\bar{e}_L\gamma_\mu d_L^i - \bar{e}_R\gamma_\mu d_R^i + \bar{\mu}_L\gamma_\mu s_L^i - \bar{\mu}_R\gamma_\mu s_R^i] + \bar{u}_L^{k^c}\gamma_\mu d_L^j\,(\bar{e}_L\gamma_\mu u_L^i) + \bar{u}_L^{k^c}\gamma_\mu d_L^j\,(\bar{\nu}_{eR}^c\gamma_\mu d_R^i + \bar{\nu}_{\mu R}^c\gamma_\mu s_R^i)\} + \mathrm{h.c.} \tag{3.10}$$

here, $\tilde{G}/\sqrt{2} \equiv \pi\bar{\alpha}/2M_X^2$, and the index c denotes charge conjugation. The diagram of Fig. 4a corresponds to the transition of a quark in a proton with nonconservation of the baryon number; Figs. 4b and 4b′ correspond to a diquark transition, and Fig. 4c to a triquark transition. Examination shows[47] that the most important contribution to the instability of the nucleon comes from the

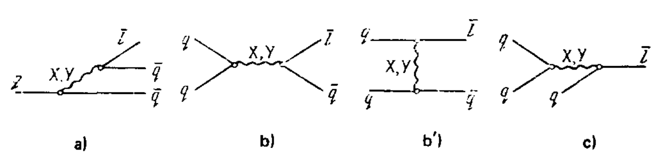

FIG. 4.

diagrams in Figs. 4b and 4b′; the diagram in Fig. 4c reduces the proton lifetime by only a few percent. The contribution of the diagram in Fig. 4a to τ_p is less than 30%, so that it is sufficient to calculate the diagrams in Figs. 4b and 4b′.

The calculation of the proton lifetime includes calculation of the transition $q+q\to\bar{l}+\bar{q}$ for the two-quark wave function $\psi(0)$ in the proton $[|\psi(0)|^2\approx(\pi R^3)^{-1}\approx 10^{-2}$ GeV^3, where R is proton radius ≈ 0.75 F] and renormalization by strong interactions from energies $\approx M_X$ to energies of order $\mu\approx$ a few GeV by a procedure analogous to that described in Sec. 3e for the constants and the fermion masses.

Assuming that the masses of all the quarks are small compared with M_X and M_W, the renormalization of the proton decay matrix element can be obtained from the anomalous dimension of the three-quark operators in the same way as is done in the calculation of the matrix elements of nonleptonic decays of hyperons.

This renormalization has approximately the form ($m_f < \mu$) (Ref. 27)

$$A\approx\left[\frac{\alpha_s(\mu^2)}{\bar{\alpha}}\right]^{2/b_0}. \tag{3.11}$$

If we wish to investigate the dependence of τ_p on f, (the number of quark flavors,) more realistically and we recall that $m_f\gtrsim\mu$ for $f>3$, we can introduce an enhancement factor corrected for this effect for the matrix element in, for example, the form[47]

$$A_s(f)=\prod_{i=4}^{f}A_s'(i)\,A_s(f)\,A_4, \tag{3.12}$$

where

$$A_s(f)=\left(\frac{\alpha_s(m_f^2)}{\bar{\alpha}}\right)^{2/b_0},$$

$$A_s'(i)=\left[\frac{\alpha_s(m_{i+1}^2)}{\alpha_s(m_i^2)}\right]^{2/[11-(2/3)i]},$$

$$A_4=\left[\frac{\alpha_s(\mu^2\approx 1\ \mathrm{GeV}^2)}{\alpha_s(m_b^2)}\right]^{6/25}$$

(for the estimate, we shall take $m_{i+1}/m_i=3$). For an estimate, we can take $R=0.75$ F, $m_q=m_p/3$, $\alpha_s(\mu^2)\approx 0.5$, $\bar{\alpha}=0.02$, $f=6$. Then

$$\tau_p\approx(10^3-10^4)\frac{m_X^4}{m_p^5}.$$

Comparing this with the experimental lower limit of τ_p (2×10^{30} years[48]), we see that $M_X\gtrsim 10^{14}$ GeV, which is by no means trivial, if we take it into account that $M\approx 10^{16}$ GeV.

It should be borne in mind that the lower limit on τ_p is associated experimentally with the muonic decay channel of the proton, which, as estimates show, makes about a 2% contribution to the proton decay probability. It should be borne in mind that our consideration of the proton instability in this section is of a fairly general nature and the use of the concrete Lagrangian (3.10) is not critical. It is reasonable to think that the general structure of (3.10) is characteristic of many grand unification schemes, differences being obtained only in the numerical factors of order unity and the actual distribution of the helicities of the quarks and leptons.

TABLE I. Lifetimes of the nucleon as a function of f (in years).

f	6	8	10	12	14
τ_f (p)	$0.8\cdot 10^{33}$	$6\cdot 10^{32}$	$4\cdot 10^{32}$	$2.2\cdot 10^{32}$	$4\cdot 10^{31}$
τ_f (n)	10^{33}	$6.7\cdot 10^{32}$	$5\cdot 10^{32}$	$2.5\cdot 10^{32}$	$4.5\cdot 10^{32}$

In conclusion, we give (in Table I) realistic values of the proton and neutron lifetimes (in years) [see (3.12)] for different values of the numbers f of quark flavors (the contribution of the muonic channel to the proton and neutron decay probability is ~2%).[47] The fastest channel is associated with the transition $\mathrm{ud}\to \mathrm{e}^+\bar{\mathrm{u}}$, which leads to the decays $\mathrm{p}\to\pi^0,\eta,\rho^0,\omega+\mathrm{e}^+$, $\mathrm{n}\to\pi^-,\rho^-+\mathrm{e}^+$ (relative probability $\approx$80%). The greatest uncertainties in the calculation attach to the parameters $\alpha_s(\mu^2)$, M_X, $|\psi(0)|^2$. The dependence of the lifetime on $|\psi(0)|^2$ and M_X is obvious $[\tau^{-1}\sim|\psi(0)|^2, \tau\sim M_X^4]$. The values given for τ_p correspond to the value $\Lambda=0.2$ GeV. An increase or decrease of Λ in the interval $0.07\leqslant\Lambda\leqslant 0.45$ increases or decreases τ_p by the factors 18 or 4, respectively.

Examination of Table I shows that it is extremely important to increase the accuracy of experiments of the Reines–Crouch[48] type. Currently under discussion and in the planning stage are experiments to look for proton decay using 2×10^5 tons of matter, which will make it possible to raise the lower limit on τ_p to 10^{35} years.[49]

g) Is there a finite number of quark flavors?

As we have seen in the previous section, τ_p depends strongly on f (see Table I). For example, if f is taken equal to 16, beyond which the asymptotic freedom of $SU(3)_c$ is lost (Ref. 40),[3)] we obtain $\tau_p\approx 10^{26}$ years, which is clearly unacceptable.

In Sec. 3e, we have also seen that an increase in f leads to unacceptable values of the masses of the s and b quarks, especially if allowance is made for various corrections to the analysis (Refs. 27, 44, 45, and 43) (allowance for the two-loop corrections to the β functions and the anomalous dimensions of the fermion mass operators, allowance for threshold effects when the external momenta of the loops are in the region of the masses of the particles in the loops).

We have already discussed this at the end of Sec. 3e. We summarize the discussion in Table II (Ref. 45), which shows the influence of the number f and allowance for the necessary corrections to the program described in Sec. 3e; here, $m^{(0)}$ and $m^{(1)}$ are the masses in the single- and two-loop approximations, respectively. Table II shows that if one believes the scheme, then $f=6$ is the most probable value and $f=8$ and especially $f=10$ are ruled out.

[3)]In Ref. 50, a study is made of the restrictions on the fermion and Higgs representations in different unification schemes under which asymptotic freedom holds for the gauge coupling constant. For the construction of a completely asymptotically free SU(5) scheme, see Ref. 51.

TABLE II. Masses of b and s quarks as functions of Λ^2.

Λ^2, GeV^2	$\alpha_s^{(0)}$, 10 GeV^2	$\alpha_s^{(1)}$, 10 GeV^2	b						s					
			$m_b^{(0)}$			$m_b^{(1)}$			$m_s^{(0)}$			$m_s^{(1)}$		
			$f=6$	$f=8$	$f=10$	$f=6$	$f=8$	$f=10$	$f=6$	$f=8$	$f=10$	$f=6$	$f=8$	$f=10$
0.09	0.32	0.23	5.3	5.8	6.9	6.5	8.0	12.5	0.47	0.56	0.63	0.81	0.95	1.7
0.03	0.26	0.19	5.0	5.5	6.5	6.0	7.8	11.2	0.43	0.50	0.57	0.63	0.63	0.78
0.005	0.19	0.16	4.5	5.0	5.9	5.2	6.3	9.4	0.36	0.43	0.47	0.48	0.57	0.90

Special attention should be paid to a circumstance related to the dependence of the physical consequences of the SU(5) scheme on the number f of quark flavors.[4)] In all the papers considered above, the contribution of the Higgs boson to the renormalization of the coupling constants was ignored. The neglect of the contribution of at least the doublet of "light" colorless Higgs bosons—the lowest components of the quintet $[\mathbf{5}\supset(\mathbf{3},\mathbf{1})+(\mathbf{1},\mathbf{2})]$—is for $q^2 \gtrsim M_W^2$ a fairly crude approximation. Allowance for them leads to modifications (in, for example, the estimates of τ_p) that are more important than the change in the value of f. Roughly speaking, each additional quintet of the Higgs field would reduce τ_p by an order of magnitude. The prediction for $\sin^2\theta_W$ would also change.

It should also be borne in mind[5)] that the mass M at which the unification occurs does not depend on f, since the slowing down in the decrease of the constant α_s with the introduction of new families is exactly compensated by the acceleration in the growth of the constant α. Only the value of $\bar{\alpha}$ depends on f.

There are a number of other estimates of the number f of quark flavors. Many of them essentially repeat Landau's arguments,[32] which we have already discussed, updating them with allowance for asymptotic freedom, color, the fractional charge of quarks, W mesons, Higgs scalars and so forth.[52]

There are arguments of a different nature based on an instability of the gluon vacuum which is manifested in the existence of an unstable vacuum state lying lower than the "perturbation-theoretic" vacuum.[53] On the basis of the results of Refs. 53, Nielsen and Olesen[54] obtained for f the estimate $f \lesssim 8$ and related it to the condition for the decay of this new level into electric vortices that ensure confinement.

There are arguments of an astrophysical nature associated with an upper limit on the number of neutrinos with zero or low mass.[55]

The argument goes as follows.

In the early stage of the expansion of the universe, massless particles (or particles with mass much less than kT) make the dominant contribution to the energy density. The addition of an extra particle with zero mass (or small mass <100 keV) living for more than a few seconds would during the first few seconds increase the energy density by $\rho'/\rho = 1 + \alpha\Delta N_\nu \equiv \xi^2$, where ΔN_ν is the number of additional particles (of the type of neutrinos with $m_\nu \approx 0$), and this would change the expansion rate $t_p' = t_p/\xi$ (t_p is the time which has elapsed since the start of expansion), since for the "free" expansion characteristic of the initial stage of the evolution of a (homogeneous and isotropic) universe $\rho = \varkappa a T^4 = 3/32\pi G_{gr} t_p^2$ (α is a constant).[6)] The expansion is accelerated ($\xi > 1$), with the consequence that the rates of primordial nucleosynthesis change and the main weak processes determing the ratio of the numbers of neutrons and protons ($e^+ + n \rightleftharpoons p + \bar{\nu}_e$, $e^- + p \rightleftharpoons n + \nu_e$, $n \rightleftharpoons p + e^- + \bar{\nu}_e$) get out of equilibrium (because of the increase in the expansion rate) earlier, at higher temperatures [the time of establishment of equilibrium is $\tau = 1/\sigma n c$ (n is the number density of neutrons and protons) $\sim T^{-5}$, since $\sigma \sim T^2$, and $n \sim R^{-3} \sim T^3$ (R is the scale factor of the universe), so that for $t > \tau$ equilibrium ceases to be established]. The neutrons are "frozen" earlier (at $T_f \sim N_\nu^{1/6}$) and the ratio of the number of neutrons to the number of protons is greater than would be the case without this extra particle with small m_ν. This leads to an increase in the fraction of d, ^{3}He, and ^{4}He and ultimately to an excess of ^{4}He as the most stable in the process of the subsequent evolution of the system. The "freezing" temperature is found to be of order 1 MeV $\approx 10^{10}$ °K, giving for the helium excess values of order 0.23–0.25. As a result, from the observed upper limit for the excess of primordial helium over hydrogen (^{4}He/H < 26%)[7)] one can obtain[4,55] (with allowance for possible right-handed neutrinos for which the cross section is $\sigma_R < 10^{-9}\sigma_L$; see Ref. 56) an upper limit $N_\nu < 3$–4 for the number of neutrinos, which corresponds to $f < 6$–8 in the considered scheme.

Thus, the chain of reasoning to the effect that "new" neutrinos (neutral leptons with low mass) → increase in the radiation density → faster rates of the early expansion → increase in the ^{4}He/H helium excess leads to a fairly stringent restriction on f.

We see that cosmological arguments become important for grand synthesis schemes. Below, we shall encounter a stronger and weightier intrusion of cosmology into unification schemes of the SU(5) symmetry type.

h) Higgs sector of the SU(5) scheme. The hierarchy problem

In SU(5) symmetry, the Higgs sector is the simplest as compared with other unifying groups.

We have already noted that the breaking of SU(5) to the level $SU(3)_c \otimes SU(2) \otimes U(1)$ occurs with the minimal Higgs system Φ of the $\underline{24}$-plet of the group: $\Phi = \sum_{a=1}^{24} \Phi_a \lambda^a/\sqrt{2}$. The Higgs field Φ gives the mass to the 12 superheavy bosons $X_{\pm 4/3}^i$ and $Y_{\pm 1/3}^i$. The components of Φ that are not used for this are physical. The interaction with the vector fields has the coupling constant

[4)]This circumstance was pointed by M. B. Voloshin.

[5)]See the previous footnote.

[6)]a is the constant that appears in the density of black body radiation, and $\varkappa = [n_b + (7/8)n_f]/2$, where n_b and n_f are the total numbers of degrees of freedom of the various bosons and fermions, respectively.

[7)]The estimates of the relative abundance of residual helium must be regarded with caution (see, for example, Ref. 57, p. 584 of the Russian translation).

$\tilde{g}$. It determines the value of these superheavy masses M_X and M_Y. In the exposition below, we shall follow Ref. 44. The most general form of the Higgs potential (with the discrete symmetry $\Phi \to -\Phi$) has the form

$$V(\Phi) = -\frac{\mu^2}{2}\,\mathrm{Tr}\,(\Phi^2) + \frac{a}{4}\,[\mathrm{Tr}\,\Phi^2]^2 + \frac{b}{2}\,\mathrm{Tr}\,\Phi^4. \tag{3.13}$$

For $b>0$ and $a> -(7/15)bV(\Phi)$, there is a nonzero vacuum expectation value and it corresponds to the symmetry breaking $SU(5) \to SU(3)_c \otimes SU(2) \otimes U(1)$:

$$\langle 0|\Phi|0\rangle = v_{24}\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -3/2 & 0 \\ 0 & 0 & 0 & 0 & -3/2 \end{pmatrix} \tag{3.14}$$

here v_{24} is determined by the relation

$$\mu^2 = \left(\frac{19}{2}a + \frac{7}{2}b\right) v_{24}^2. \tag{3.15}$$

The masses of the X and Y bosons are

$$M_X^2 = M_Y^2 = \frac{25}{8}\,\tilde{g}^2 v_{24}^2. \tag{3.16}$$

After the masses have been given to the $X^i_{\pm 4/3}$ and $Y^i_{\pm 1/3}$ bosons, there remain 12 Higgs bosons with masses of the order of M_X and $M_Y(\sim b^{1/2}v_{24})$ [a color octet with zero charge and four color singlets—two neutral and two with charge ± 1 SU(2) doublet]: $H_8, H', H'', H^{\pm}$.

The second stage of the breaking, $SU(2)\otimes U(1) \to U(1)$, is again in the minimal manner—through a quintet with $SU(3)_c \otimes SU(2)$ content $(3,1)+(1,2)$:

$$H = \begin{pmatrix} H^i_{1/3} \\ H^+ \\ H^0 \end{pmatrix}, \quad i = r,\ y,\ b.$$

Here, H contains a color triplet with charge 1/3 and an SU(2) doublet

$\begin{pmatrix} H^+ \\ H^0 \end{pmatrix}$, which breaks SU(2);

$$H^+ = (H^i_{-1/3},\ H^-,\ H^{0*});$$

Here, as in the ordinary Weinberg–Salam scheme, the neutral component $(H^0 + H^{0*})/\sqrt{2}$ gives masses to the vector bosons $W^{\pm}$ and Z^0, having a nonzero vacuum expectation value $v_5 \approx 10^2$ GeV. The H^+, H^-, and $(H^0 - H^{0*})/\sqrt{2}$ are used on the third components of the $W^{\pm}$ and Z^0 mesons.

In the language of the corresponding Higgs potential

$$V(\underline{H}) = -\frac{v^2}{2} H^+H + \frac{\lambda}{4}(H^+H)^2, \tag{3.17}$$

$$v_5^2 = \frac{2v^2}{\lambda}, \quad m_W^2 = \frac{\tilde{g}^2}{4} v_5^2. \tag{3.17^1}$$

The mass of the physical Higgs boson H^0 is

$$m^2_{H^0} = \frac{\lambda v_5^2}{2} = v^2.$$

The remaining three Higgs bosons H_3 (H_3^+) with charge $+1/3$ $(-1/3)$ can in principle have a large mass ($\lesssim M_X$) and lead, like the $X^i_{\pm 4/3}$ and $Y^i_{\pm 1/3}$ bosons, to nonconservation of baryon charge. However, the coupling of H_3 to the fermions is weak ($g_{H_3 t\bar{t}'} \approx g m_t/m_W$), as for the SU(2) Higgs mesons, and therefore their contribution to the probability of proton decay is negligibly small compared with the contribution of the X and Y bosons:

$$\frac{\Gamma(p \xrightarrow[H_3]{} \bar{l} + \text{anything})}{\Gamma(p \xrightarrow[X,Y]{} \bar{l} + \text{anything})} \approx \frac{m^4_{X,Y}}{m^4_{H_3}}\,\frac{m_t^4}{m_W^4} \ll 1.$$

Hitherto, the Higgs scalars have essentially been isolated from one another [although we have implicitly envisaged a coupling between them when we spoke of physical Higgs bosons with mass not absorbed in the vector bosons X, Y, $W^{\pm}$, Z. Without this coupling (see below), these scalar mesons, for example, H_3, would not have masses, which would lead to the undesirable possible existence of light bound states of them and quarks].

However, Φ and H are in principle coupled through exchanges of vector bosons and their loops. In principle, this coupling can have the consequence that our hierarchy of symmetry-breaking interactions or vacuum expectation values

$$\frac{v_{24}}{v_5} \sim \frac{\mu}{v} \sim \frac{M_X}{m_W} \sim 10^{12}-10^{13}$$

can disappear.

We take this coupling between Φ and H into account phenomenologically by means of the simple potential

$$U(\Phi,\ H) = \alpha H^+H\,\mathrm{Tr}\,(\Phi^2) + \beta H^+\Phi H^+. \tag{3.18}$$

In this case, $\langle 0|\Phi|0\rangle$ no longer has the form (3.14), since SU(2) is broken when allowance is made for (3.17) and (3.18), so that we can write

$$\langle 0|\Phi|0\rangle = v_{24}\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -\frac{3}{2}\left(1+\frac{\varepsilon}{2}\right) & 0 \\ 0 & 0 & 0 & 0 & -\frac{3}{2}\left(1-\frac{\varepsilon}{2}\right) \end{pmatrix}. \tag{3.19}$$

It is to be expected that the breaking of SU(2) due to the Higgs field Φ is much less than its breaking due to the "doublet" breaking H in accordance with the experimental confirmation[6] of $I=\frac{1}{2}$ for the Higgs boson in SU(2) (see Sec. 1a). It is therefore natural to seek a minimum of the total potential (3.13), (3.17), and (3.18) in which $\varepsilon \to 0$ when $\alpha, \beta \to 0$.

Then to accuracy v_5/v_{24},

$$\varepsilon = \frac{3}{20}\,\frac{\beta}{b}\left(\frac{v_5}{v_{24}}\right)^2.$$

The conditions (3.15) and (3.17^1) are then replaced by

$$\mu^2 = \left(\frac{19}{2}a + \frac{7}{2}b\right)v_{24}^2 + \left(\alpha + \frac{3\beta}{10}\right)v_5^2, \tag{3.20}$$

$$v^2 = \frac{\lambda v_5^2}{2} + \left(19\alpha + \frac{\beta}{2} - 3\varepsilon\beta\right)v_{24}^2. \tag{3.21}$$

As a result, the masses of a number of the Higgs mesons are also strongly changed, for example, $m^2_{H_3} = -(5/2)\beta v_{24}^2 + O(v_5^2)$, so that its mass in principle need not be very small compared with the mass of the X and Y bosons (α and β are small compared with a, b, λ).

We see that the coupling of the Φ and H scalar fields can, in general, lead to a significant change of the hierarchy, and the ratio $\mu/v \sim M_X/m_W$ is reduced appreciably if there are no special cancellations in (3.21).

It was at one time asserted that the hierarchy is in fact broken at both the tree and loop level of the coupling between Φ and H.[58]

However, with regard to the tree approximation, a correct study showed[59-60] that the assertion in Ref. 58 that the condition of the minimum of the potential with allowance for the coupling between the Higgs fields imposes restrictions on the relationship between the masses of the vector mesons that characterize the breaking is incorrect and associated with the unjustified restrictive assumptions made in Ref. 58.

The situation is more complicated in the case of the radiative corrections (loops)—at first glance it appears that the stronger the gauge hierarchy, the greater is the number of loop diagrams that must be included in the treatment, which may change the hierarchy that arises in the tree approximation because of the large ratio of the masses of the Higgs boson and fix it at a level determined by the coupling constant[58] ($\sim\bar{\alpha}^{-\frac{1}{2}}$).

One can speculate here about the part played by gravitation, through which we have a superstrong symmetry breaking ($\sim G_{\rm gr}^{-1/2}$) (but then this does not have a direct bearing on the considered scheme) or assume that at one of the stages the breaking is dynamical and that the theory contains one mass (superhigh) scale, and not two. Well below this scale, the theory contains only fermions and gauge fields; the second stage of the breaking is dynamical and determined by the mass at which the constant of the strong interaction is fairly high, so that perturbation theory is inapplicable.

With regard to the concrete investigation of the part played by the loop corrections, one can in principle[44,59-61] ensure a hierarchy between the masses of the vector fields corresponding to the "independent" groups at not only the tree level but also the single-loop level.

Here, it is necessary to distinguish two aspects of the question.

1) Can one ensure the required hierarchy in terms of the bare parameters of the effective potential in each order of perturbation theory? This possibility appears rather unnatural, since it implies the existence of cancellations among infinitely many terms, although the possibility cannot be ruled out that this could be achieved through the presence in the model of additional symmetries (supersymmetry?).

2) One could of course require that the renormalized vacuum expectation values (with allowance for all loops) v_{24} and v_5 satisfy the required hierarchy as renormalization conditions. Then the impossibility, in general, of ensuring the conditions of the hierarchy in each perturbation order should be regarded, not as a restriction on the hierarchy, but as an indication of the breakdown of perturbation theory.

It was on the basis of such a position that Weinberg[62] approached this question (see also Ref. 61); he showed that by imposing conditions on the renormalized parameters of the effective potential one can always ensure the necessary hierarchy. In a number of cases, these conditions do not restrict the hierarchy at all, and in other cases, which are evidently the most interesting, restrictions arise that are entirely suitable for the phenomenology of the unified theories considered here.[62] It would be extremely interesting to understand whether one could obtain these conditions on the renormalized parameters of the theory, which reduce to the masses of certain scalars being sufficiently small, in perturbation theory with symmetry.

i) The cosmological aspect of unification schemes

As we have already seen, SU(5) symmetry, which unifies the strong interactions with the electroweak interactions and "explains" a number of connections between fundamental physical parameters of the theory, leads to a rather restricted number of predictions capable of experimental verification, the main one being the instability of the proton. Of course, it is very important and appealing to observe the instability of the proton in the laboratory through a considerable increase in the amount of matter and accuracy in an experiment of the type of Reines and Crouch[48] (see Ref. 49).

However, the proton instability, which arises naturally in the scheme, has far-reaching cosmological aspects, which also provide a kind of verification of the scheme.

It must be borne in mind that the cosmological "test" of unified theories that we shall discuss below is based on the popular hot big-bang model of the expansion of the universe.[57,63] [We have already had recourse to this model in our discussion in Sec. 3g of an upper limit—important for the SU(5) scheme—on the number of neutrinos with low mass that can be deduced from the relative abundance of helium of cosmological origin, which together with the microwave background radiation gives information about the early stages in the evolution of the universe.]

Modern data[57,63,64] indicate that the number of baryons N_B in the observable part of the universe is of the order of $10^{-9}-10^{-10}$ of the number N_γ of the microwave photons (the total number of baryons is $N_B\approx 10^{78-79}$ and of photons $N_\gamma=10^{88}$), whereas the abundance of antimatter is negligible.

One could of course assume that the observed excess of baryons is due, not to nonconservation of the baryon number, but to a nonzero initial baryon charge of the universe. This can then remove the need to explain anything.

The baryon excess can be attributed to asymmetric absorption of baryons and antibaryons by black holes.[65] The asymmetry of the absorption is due here to a difference between the partial widths resulting from the violation of C and CP invariance in these processes with black holes.

But if one insists on the idea of nonconservation of the baryon number due to certain interactions, as is the case in unified theories of the type considered here (or nonconservation of the baryon number through the instanton mechanism,[66] which we shall not discuss here), then the observed baryon excess must be determined, on the one hand, by the initial conditions which obtained when the universe arose and the history of its evolution, especially during the early stages, and, on the other, by the actual processes that do not conserve the baryon number and the interactions of the particles that participate in these processes.

We adopt initial conditions in the spirit of the hot big-bang model: "random initial conditions" described by a density matrix $\hat{\rho}$ diagonal with respect to the baryon number B and symmetric under the substitution $B \to \bar{B}$, in other words, an initial equilibrium state (the result of the big bang) arose because of a B-, C-, and CP-invariant interaction.

The possibility of subsequent nonconservation in certain processes of the baryon number alone is inadequate for the occurrence of a baryon excess ΔB. These processes must be asymmetric under the substitution $B \to \bar{B}$, since otherwise the sign of the effect will be random at different stages of the evolution and cancel, so that N_B will be of order $N^{1/2} \approx 10^{44}$ and $N_B/N_\gamma \approx 10^{-44}$, and not $\approx 10^{-10}$ (Ref. 67). It is obvious that for the same reason one requires not only C but also CP violation (the possibility of generation of baryon asymmetry in the hot big-bang model through processes in which the baryon charge is not conserved and CP invariance is simultaneously violated was considered for the first time in Ref. 68 by Sakharov).

Independently of C and CP invariance, nonconservation of the baryon charge does not lead to a nonzero baryon number if the system is in statistical equilibrium.[68,69] This is a consequence of the CPT theorem[68,70,71,67] [$\hat{\rho}(t)$ is the equilibrium density matrix, $\hat{\theta} = \hat{C}\hat{P}\hat{T}$]:

$$\langle \hat{B} \rangle = \mathrm{Tr}\,(\hat{\rho}B) = \mathrm{Tr}\,(\hat{\theta}\hat{\rho}\hat{\theta}^{-1}\hat{\theta}\hat{B}\hat{\theta}^{-1}) = -\mathrm{Tr}\,(\hat{\rho}\hat{B}) = -\langle \hat{B} \rangle$$

(since the CPT-conjugate state has the same energy but opposite baryon charge). Therefore, during the earliest equilibrium stage in the evolution of the universe ($T \approx G_{gr}^{-1/2} \approx 10^{19}$ GeV, $t \approx 1.5 \times 10^{-43}$ sec) there are equal numbers of baryons and antibaryons. As the universe expands, the energy density of the primordial plasma of quarks and leptons falls [initially, in the epoch when radiation dominates over matter, in proportion to the fourth power of the temperature ($\rho \sim T^4 \sim 1/t^2$, where t is the time which has elapsed since the start of expansion)], and the rate of expansion $\dot{R}/R$ of the universe in the initial stage of expansion is proportional to T^2:

$$\dot{R}/R = 1/2t \sim T^2 \qquad (\text{in units of } \hbar = c = G_{gr} = 1). \qquad (3.22)$$

If the rate of the processes that do not conserve the baryon charge, $\sigma n v$, changes with the temperature in such a way that the rate of expansion is greater than the rate of the processes, then with the course of time the processes that do not conserve the baryon charge get out of statistical equilibrium, and the universe can in principle acquire a nonzero baryon charge.

In other words, bearing in mind what we have said above, the processes that do not conserve the baryon charge (and also violate C and CP invariance) must be "adjusted" to the Hubble expansion in such a way as to ensure the observed baryon excess N_B/N_γ. The CPT theorem imposes additional conditions[67] on the time evolution of the occurrence of the nonzero $\langle \hat{B}(t) \rangle$ [$\langle \hat{B}(t) \rangle = -\langle \hat{B}(-t) \rangle$, i.e., $\langle \hat{B}(t) \rangle$ is an odd function of t] and also on the time dependence of the rate of variation of $\langle \hat{B}(t) \rangle$ with the time: [$\langle \dot{\hat{B}}(t) \rangle_{t\to 0} \to 0$], from which it follows that the baryon excess is produced during the early stage rather slowly, $\langle \hat{B}(t) \rangle \sim t^3$ at best. In addition, it follows from these same considerations that $\langle \hat{B} \rangle$ arises only in the second order in the baryon-charge nonconserving interaction, which is farily natural. It should be borne in mind that in a closed nonexpanding universe with fixed temperature it is impossible to have a baryon excess over a long time due to processes that do not conserve the baryon charge.

As we see, the nonzero mean baryon number depends on the stage of evolution at which the B nonconserving (and C and CP violating) processes get out of equilibrium.

Now this is determined not only by the scale of the elementary event but also by fundamental properties of the theory such as renormalizability.[67]

Indeed, all renormalizable theories (i.e., theories with dimensionless coupling constant) have a reaction rate I_R proportional to the temperature; for example, for the SU(5) model for $T \gtrsim M$

$$I_R \sim \bar{\alpha}^{-2} T. \qquad (3.23)$$

An analogous dependence will hold for the rates of the processes corresponding to the strong, weak, and electromagnetic interactions at the corresponding temperatures. It follows from (3.22) and (3.23) that renormalizable interactions (QCD, weak and electromagnetic interactions, and the unified SU(5)-symmetric interaction above M) will not be in equilibrium if $T^2 > \bar{\alpha}^2 T$. This means that from $T = 10^{19}$ GeV to $T = 10^{15}$ GeV we have a nonequilibrium period (of course, gravitation will always be in a nonequilibrium state for $T < T_{Pl} = G_{gr}^{-1/2} \approx 10^{19}$ GeV since $I_G \sim G_{gr} T^3$).

With regard to nonrenormalizable interactions with dimensional coupling constant, the situation is different, because of the stronger dependence of I_{NR} on T.

For example, for the local four-fermion interaction, to which the interactions induced by the exchange of superheavy vector mesons or Higgs bosons of the SU(5) (or some other) grand unification scheme effectively reduce for $T < M \approx 10^{15}$ GeV, we have

$$I_F \sim \bar{G}^2 T^5, \qquad (3.24)$$

where $\bar{G} \approx \bar{\alpha}/M_X^2$ for the exchange of X or Y mesons and $\bar{G} \approx \bar{\alpha}/m_H^2 (m_t/m_W)^2$ for the exchange of Higgs scalar mesons. It follows from (3.22) and (3.24) that for $T < 10^{15}$ GeV the baryon-charge nonconserving interactions induced by the eschange of X and Y mesons get out of equilibrium, whereas the analogous interactions due to the exchange of Higgs scalars are in equilibrium to the temperature

$$T_c^3 = m_H^4 \left(\frac{m_W}{m_t}\right)^4 \bar{\alpha}^{-2} \qquad (\text{in units of the Planck mass}),$$

i.e., (choosing $m_H \lesssim M_X/20 \approx 10^{14}$ GeV, $m_t/m_W \approx 10^{-1}$) to temperatures of order 10^{10} GeV.

Approximate estimates of the ratio N_B/N_γ based on the above ideas are given in Refs. 67, 71, and 72.

However, in the estimate of the temperature (the time) at which the generation of the baryon excess commences it is necessary, as A. D. Dolgov has noted,[73] to take into account an important circumstance which has the consequence that the generation of baryon charge must commence comparatively late, when the density of the primordial plasma is already low, which significantly reduces the ratio N_B/N_γ compared with the estimates of Refs. 67 and 71.

As Dolgov has shown,[73] the generation of baryon charge by the reaction that does not conserve it commences, not at the time when this reaction gets out of equilibrium, but significantly later, when all the processes in which at least one of the species of the particles participating in the reaction get out of equilibrium. For example, in one of the characteristic reactions responsible for nonconservation of the baryon charge, $q+q\to\bar{q}+\bar{l}$, generation of baryon charge commences when the temperature of the quarks begins to differ from the temperature of the leptons. However, at this epoch the residual ("frozen") concentration n of leptons[74] and (or) quarks[75] in the universe is already too low to ensure the observed N_B/N_γ ratio.

Dolgov gives the following estimate for this ratio:

$$r_H\frac{\Delta\sigma}{\sigma_{\min}},$$

where $\Delta\sigma$ is the $C(CP)$-odd part of the cross section of the B-nonconserving reaction, $\sigma_{\min}$ is the smallest of the total interaction cross sections of the particles participating in this reaction, and r_H is the concentration of the heavy particles (i.e., such that $m_H>kT$) participating in the given process. All the quantities are taken at the time determined by the condition of the processes with the cross section $\sigma_{\min}$ getting out of equilibrium ($\sigma_{\min}vnt\approx 1$). For the considered reaction $q+q\to\bar{l}+\bar{q}$, $\sigma_{\min}$ is the ordinary weak cross section, so that the baryon charge begins to be generated later, when $T\lesssim m_l/2\lesssim 1$ GeV (Ref. 74), and not when $T\lesssim 10^{15}$ GeV. A similar situation obtains for many unification schemes.

As can be seen from these arguments, generation of baryon charge can commence sufficiently early to ensure a noninfinitesimal concentration of the particles participating in the baryon-charge nonconserving reactions if there exist processes in which there participates at least one "sterile" particle, i.e., a particle that has neither strong nor the ordinary electroweak interactions, so that all processes with its participation get out of equilibrium early. (Could this particle be the right-handed neutrino with $\sigma_R\lesssim 10^{-9}\sigma_L$ mentioned in Sec. 3g and which has already been considered for restricting the number of neutrinos?)

Whatever the true situation, we see how closely the ideas from the theory of elementary particles are here interwoven with cosmology.

In this connection, it is also worth pointing out that not every CP violation, which, as we have seen, is needed to generate a baryon excess $\langle\Delta B\rangle$, is sufficient to ensure a nonzero $\langle\Delta B\rangle$.

In particular, spontaneous violation of CP, which is lifted at high temperatures, cannot, generally speaking, combine with nonconservation of the baryon number sufficiently "early" to ensure an appreciable $\langle\Delta B\rangle$.

Rather, it must be a "firmly" introduced violation of the Kobayashi–Maskawa type,[16] which remains in the Lagrangian describing the scale of masses associated with the unified symmetry. Of course, this does not rule out the possibility that the observed CP violation in K^0 decays has the nature of a spontaneous symmetry breaking.

In conclusion, we emphasize that, in the light of the development of unified theories, data on the ratio N_B/N_γ provide, together with the data on the microwave background radiation and the abundance of primordial helium, very valuable information about the early history of the universe. It is remarkable that on the basis of grand unification schemes this ratio can be expressed in terms of observable parameters of elementary-particle physics.

4. ON SOME OTHER GRAND UNIFICATION SCHEMES

Let us now consider briefly other grand unification schemes.[8)] Most closely related to SU(5) symmetry is the SO(10) symmetry scheme[30] (see also Ref. 77), which corresponds to rank five. It is based on the family of 16 particles (II) in Sec. 1a, including a right-handed neutrino ν_R; the neutrinos may be massive. There are various chains of breaking of this symmetry group:

a) $SO(10)\supset SU(5)\supset SU(3)_c\otimes SU(2)_L\otimes U(1)$, which demonstrates the natural connection between the SO(10) and SU(5) schemes. The family (II) is placed in the *irreducible* 16-plet representation of SO(10): $\underline{16}\supset\underline{10}+\underline{5}+\underline{1}$.

b) $SO(10)\supset SU(3)_c\otimes SU(2)_L\otimes SU(2)_R\otimes U(1)$ [$SU(2)_L\otimes SU(2)_R\otimes SU(3)_c$ is the content of the multiplet $\underline{16}\supset(\underline{2},\underline{1},\underline{3})+(\underline{2},\underline{1},\underline{1})+(\underline{1},\underline{2},\underline{\bar{3}})+(\underline{1},\underline{2},\underline{1})$, which contains the electroweak group $SU(2)_L\otimes SU(2)_R\otimes U(1)$, which, as we have already said, was a serious alternative to the Weinberg–Salam model before confirmation of parity violation in neutral currents. In the ordinary "symmetric" form, it contradicts the experiments, but if it is assumed that the vector mesons $W_R^\pm$ corresponding to the group $SU(2)_R$ are very heavy compared with the left-handed $W_L^\pm$, the contradiction can be avoided at present-day energies].

Parity violation is a result of spontaneous breaking, and parity is asymptotically conserved in weak processes.

c) If the chain of symmetry breaking is

$$SO(10)\supset SU(2)_L\otimes SU(2)_R\otimes SU(4)_c,$$

then accordingly $16\supset(\underline{2},\underline{1},\underline{4})+(\underline{1},\underline{1},\underline{\bar{4}})$; here, $SU(4)_c$ is a color group that includes the lepton number as a fourth color. It was considered comparatively long ago by Pati and Salam.[29] In SO(10) symmetry, as in the SU(5) scheme, the value of $\sin^2\theta_W$ in unbroken symmetry is 3/8.

Its renormalization to our energies, analogous to the renormalization in the SU(5) scheme [cf (3.7)], is somewhat smaller than in SU(5) (Ref. 27):

[8)]A complete classification of the theories, including color singlets and triplets of fermions in the framework of simple unifying gauge groups, is given in Ref. 76.

$$\sin^2\theta_W(\mu)=\frac{1}{4}+\frac{\alpha}{3\alpha_s(\mu)}\approx 0.26-0.27^{9)} \qquad (\alpha_s=0.2;\ 0.1).$$

For M, we obtain[27]

$$\ln\frac{M}{\mu}=\frac{3\pi}{22\alpha}\left[1-\frac{8}{3}\frac{\alpha}{\alpha_s(\mu)}\right].$$

Here, the Higgs sector is more complicated than in SU(5). The breaking of SO(10) to $SU(3)_c\otimes SU(2)\otimes U(1)$ can be ensured[77] by a definite combination of the SO(10) representations $\underline{45}$ and $\underline{16}$ as a multiplet of Higgs bosons. The Higgs multiplets that give masses to the fermions of the 16-plet of SO(10) are in the complex representations $\underline{10}$ and $\underline{126}$ (Ref. 77).

As was noted in Sec. 3a, the introduction of several Higgs multiplets leads, in general, to flavor-changing neutral currents. In this scheme their contribution is small but observable, especially in the muon sector.[77]

At the same time, in contrast to SU(5), there is in principle a connection between the masses of the q_A quarks and l_A leptons (because of the irreducibility of the fermion representation), but this connection is unsatisfactory, giving, for example, $m_{\nu_\tau}\approx m_t$. The connection between the masses of the q_C quarks and l_C leptons can here be made the same as in exact SU(5).

The exceptional groups have also been used for unifying leptons and quarks.

The group E_7 (Ref. 31) contains the maximal subgroup $SU(6)\otimes SU(3)_c$. The leptons and quarks are put in one basic multiplet $\underline{56}$, this being a left-handed two-component spinor with respect to the Lorentz group. Here, there are no families I (II), and the fermions of all flavors are put in one multiplet. The flavor group is SU(6). The $SU(6)\otimes SU(3)_c$ content of this multiplet is

$$\underline{56}=(6.3)+(\bar{6}.\bar{3})+(20.1)$$

—a sextet (six flavors) of color quarks, a sextet of color antiquarks, and 10 two-component leptons and antileptons (color singlets).

The electric charge is a generator of SU(6) (the sum of the charges of the quarks is zero). The group SU(6) includes the subgroup SU(3) and, thus, the charges of the sextet of quarks are 2/3, −1/3, −1/3 and 2/3, −1/3, −1/3, i.e., the sixth quark is not the t quark but a quark with $Q=-1/3$ (b' or h quark). As we have already said, the leptons and antileptons are put in an SU(6) 20-plet.

There are four negatively charged leptons [$e^-, \mu^-, \tau^-, \tau^{-\prime}(?)$] and four corresponding antileptons and 12 neutral leptons (and antileptons) $\nu_e, \nu_\mu, \nu_\tau, \nu_1, \nu_2, \nu_3$. They form four doublets, two triplets, and six singlets of SU(2). One each of the negatively charged leptons are put in two doublets and two triplets; the positively charged leptons are treated similarly. The lepton number, like the baryon number, is not conserved here.

The vector representation of E_7 contains 133 gauge vector bosons with the following $SU(6)\otimes SU(3)_c$ content:

$$\underline{133}=\underbrace{(\underline{1},\underline{8})}_{\text{octet of color gluons}}+\underbrace{(\underline{35},\underline{1})(W^\pm,Z,\gamma)}_{\text{and a further 31 unknown vector particles}}+\underbrace{(\underline{15},\underline{3})+(\overline{\underline{15}},\underline{3})}_{\text{90 color particles of unknown use.}}$$

The Weinberg angle is here determined at the level of the SU(6) symmetry: $\sin^2\theta_W=3/4$, and since there are no grounds for believing that the corresponding gauge bosons are heavier than M_X and M_Y in the SU(5) scheme, it follows in accordance with what we have said above concerning the renormalization of θ_W that it must be small, so that it is impossible to ensure agreement with the experiments for $\sin^2\theta_W$, in contrast to the SU(5) scheme [and even the SO(10) scheme]. For the same reason, it is also difficult to ensure the "correct" renormalization of the constants α_3 and α_2 from their values in exact symmetry to their values at $q^2\approx\mu^2$ (see Sec. 3e).

In contrast to SU(5) [and even SO(10)], the Higgs sector is here fairly complicated—to obtain sensible breaking it is necessary to introduce at least the multiplets **133**, **912**, and **1463** of Higgs scalars. An advantage of the scheme is that it gives quite definite predictions, in contrast to SU(5) and SO(10), about the total number of quarks and leptons, and predicts an unequal number of quarks with $Q=2/3$ and $Q=-1/3$, which could be verified experimentally.

At the same time, it must be borne in mind that any model that, like E_7, includes fermions from different families in one multiplet of the unifying group must, in general, lead to violation of the natural flavor conservation by the neutral currents. In such schemes, the experimentally observed absence of neutral currents with $|\Delta S|=1$ and $|\Delta C|=1$ is explained by the assumption of very high masses of the corresponding neutral vector bosons.

5. CONCLUSIONS

Let us summarize. We have described in detail the SU(5) scheme as a possible basis for unifying the weak, electromagnetic, and strong interactions not for the reason that it appears to us "correct", containing naturally the "correct" electroweak group $SU(2)_L\otimes U(1)$, but for the reason that it is the simplest and most natural among similar schemes that do not include gravitation.

In this, in particular, one can see a shortcoming of the scheme (besides a number of other obvious shortcomings: the disappearance from view of the q_A quarks, the impossibility of predicting the Cabibbo type angles, the poor predictions for the masses of the light quarks, and so forth).

It could be that something is indicated in this connection by the fact that in the SU(5) [or SO(10)] scheme the mass M at which the interactions are unified is of order $\bar{\alpha}^2 M_{Pl}\approx\bar{\alpha}^2 G_{gr}^{-1/2}$. One could imagine that it is gravitational attraction that determines the spontaneous breaking of the SU(5) [or SO(10)] symmetry in the first stage, reducing it to $SU(3)_c\otimes SU(2)\otimes U(1)$ symmetry.

9) This value for $\sin^2\theta_W$ at one time provided an argument in favor of the SO(10) scheme as opposed to the SU(5) scheme, in which $\sin^2\theta_W=0.19-0.20$.

Whatever the truth, this is an indication that in such a program gravitation cannot be ignored. It is entirely possible that only the introduction of gravity (supergravity, a superfield combining gravitons, gravitinos, vector bosons, fermions, and Higgs scalars) in a unified scheme can put many things in their place.

With regard to one of the main shortcomings of the SU(5) scheme [or SO(10)]—the absence of an explanation in it for the number f of flavors—we have already noted in Sec. 3d that the symmetry properties relating the fermion families (I) and (II) "horizontally" have not yet been fully exploited. For example, one can introduce a so-called "horizontal" group, with respect to which the families of quarks and leptons transform.

What symmetry corresponds to this group?

Must it be discrete,[34] or a global or local gauge symmetry?[35-37]

In principle, one could also introduce a fourth undiscovered family of quarks and leptons of the type I (II) and transform the four families into each other in accordance with a new gauge group $SU(2)_H$ ("horizontal" group), this being done in such a way that there are no transitions between neighboring families of quarks (leptons) of the kind $s \leftrightarrow d, c \leftrightarrow u, t \leftrightarrow Q_A, b \leftrightarrow Q_C$ but there are transitions which "skip one family", i.e., $u \leftrightarrow t, d \leftrightarrow b, c \leftrightarrow Q_A$, $s \leftrightarrow Q_C$ (where Q_A and Q_C are the quarks of the fourth family with charge 2/3 and −1/3, respectively).[36]

The transitions between the families could be "organized," not at the doublet, but at the triplet SU(2) level, a restriction being made to the three known families of quarks (and leptons).[35]

Of course, this leads to new interactions and to new and heavier vector bosons already at the $SU(2)_L \otimes U(1) \otimes SU(2)_H$ level, and for the unification with chromodynamics one requires at least a group of sixth rank (E_6?).

It is however possible that the introduction of the families (I) and (II) is premature until we have understood the flavor group (or the electroweak group).

It is entirely reasonable to believe that $SU(2)_L \otimes U(1)$ is only a subgroup of a larger flavor group $G^f \supset SU(2)_L \otimes U(1)$, which has new gauge bosons with mass greater than m_W. It should be borne in mind that the parity violation in $SU(2)_L \otimes U(1)$ is introduced *a priori*, whereas it would be more natural to have spontaneous breaking of parity by breaking of the group $G^f \to SU(2)_L \otimes U(1)$. The smallness of the *CP* noninvariance, which in introduced in $SU(2)_L \otimes U(1)$ purely phenomenologically, could then be attributed to the very large masses of these new gauge bosons.

In the framework of this larger group G^f, one could calculate theoretically the masses and Cabibbo angles as, for example, the result of radiative corrections.

As examples, we give some conceivable possibilities:

$$G^f = SU(2)_L \otimes SU(2)_R \otimes U(1) \xrightarrow{M_R \gtrsim 100\,\mathrm{GeV}} SU(2)_L \otimes U(1) \xrightarrow{M_{W,Z}} U(1),$$
$$G^f = SU(3) \otimes U(1) \xrightarrow{M \gtrsim 100\,\mathrm{GeV}} SU(2)_L \otimes U(1) \xrightarrow{M_{W,Z}} U(1).$$

Note that the first possibility would, when G^f is combined with $SU(3)_c$, correspond to synthesis on the basis of the group SO(10). We should also not forget the possibility that the group G^s of strong interactions could be larger than the $SU(3)_c$ of QCD, so that G^f and G^s could be unified in the framework of some group G whose consequences would be applicable directly to present-day energies.

Finally, it must be emphasized that the unification schemes considered in the present review are based on the very strong assumption that in the interval from 10^2 to 10^{15} GeV there is no "new physics" differing from that with which we have been concerned at energies below 10^2 GeV. It is possible that over this wide span new phenomena arise in connection with a new mass scale in this interval of energies, and these could radically modify the entire logic of the unification and its physical consequences.

The lessons that we have extracted from the SU(5) scheme could be helpful in searches for ways of combining in a single correct scheme all the interactions of the elementary particles.

I should like to take this pleasant opportunity of expressing my thanks to M. B. Voloshin and L. B. Okun' for much valuable advice, discussions, and helpful criticism during the writing of this review.

[1]S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967); A. Salam, in: Proc. of Eighth Nobel Symposium, Stockholm (1968).
[2]L. M. Barkov and M. S. Zolotarev, Pis'ma Zh. Eksp. Teor. Fiz. 27, 379 (1978) [JETP Lett. 27, 357 (1978)].
[3]C. Prescott *et al.*, Phys. Lett. B77, 347 (1978).
[4]S. Weinberg, Rapporteur Talk at Tokyo Conference (1978).
[5]D. A. Ross and M. Veltman, Nucl. Phys. B95, 135 (1975).
[6]L. M. Sehgal, Aachen Preprint PITHA-102 (1978); L. F. Abbott and R. M. Barnett, Phys. Rev. Lett. **40**, 1303 (1978); K. Winter, in: Proc. of Intern. Lepton-Photon Symposium, Batavia (1979); see also P. R. Ermolov and A. I. Mukhin, Usp. Fiz. Nauk **124**, 385 (1978) [Sov. Phys. Usp. **21**, 185 (1978)], V. M. Shekhter, Usp. Fiz. Nauk **119**, 593 (1976) [Sov. Phys. Usp. **19**, 645 (1976)].
[7]P. O. Hung and J. J. Sakurai, Nucl. Phys. **B143**, 81 (1978); H. Fritzsch, Invited Talk at the Tokyo Conference, CERN Preprint TH-2579; H. Harari, Summary Talk at the Kyoto Summer Institute, 1978; SLAC-PUB-2221.
[8]J. L. Susskind, SLAC-PUB-2142 (1978).
[9]S. I. Glashow, J. Iliopoulos, and L. Maiani, Phys. Rev. **D2**, 1285 (1970); see also: V. I. Zakharov, V. L. Ioffe, and L. B. Okun', Usp. Fiz. Nauk 117, 227 (1975) [Sov. Phys. Usp. **18**, 757 (1975)].
[10]S. L. Adler, Phys. Rev. 177, 2426 (1969); J. S. Bell and R. Jackiw, Nuovo Cimento 51, 47 (1969).
[11]J. C. Flugge, DESY Preprint 78/42 (1978); Bacino *et al.*, Phys. Rev. Lett. **42**, 749 (1979); Ya. P. Azimov, L. L. Frankfurt, and V. A. Khoze, Usp. Fiz. Nauk **124**, 459 (1978) [Sov. Phys. Usp. 21, 225 (1978)].
[12]S. Herb *et al.*, Phys. Rev. Lett. **39**, 252 (1977); W. R. Innes *et al.*, Phys. Rev. Lett. **39**, 1240 (1977); Ch. Berger *et al.*, Phys. Lett. **B76**, 243 (1978); C. W. Darden *et al.*, Phys. Lett. **B76**, 247 (1978).
[13]E. F. Tret'yakov *et al.*, Preprint No. 15 [in Russian], Institute of Theoretical and Experimental Physics, Moscow

(1976); Particle Data Group, Rev. Mod. Phys. 48, 51 (1976).
[14]S. S. Gershteĭn and Ya. B. Zel'dovich, Pis'ma Zh. Eksp. Teor. Fiz. 4, 174 (1966) [JETP Lett. 4, 120 (1966)]; R. Cowsik and J. McClelland, Phys. Rev. Lett. 29, 669 (1972); A. S. Szaloy and G. Marx, Astron. Astrophys. 49, 437 (1976); M. I. Vysotskiĭ, A. D. Dolgov, and Ya. B. Zel'dovich, Pis'ma Zh. Eksp. Teor. Fiz. 26, 200 (1977) [JETP Lett. 26, 188 (1977)].
[15]G. I. Feldman, SLAC-PUB-2138 (1978).
[16]M. Kobayashi and K. Maskawa, Prog. Theor. Phys. 49, 652 (1973).
[17]S. Weinberg, Phys. Rev. Lett. 37, 657 (1976); T. D. Lee, Phys. Rep. 9, 143 (1973).
[18]R. N. Mohapatra, Phys. Rev. D6, 2023 (1972); H. Fritzsch and P. Minkowski, Phys. Lett. B68, 421 (1976).
[19]J. Ellis, M. K. Gaillard, and D. V. Nanopoulos, Nucl. Phys. B109, 213 (1976).
[20]G. Goldhaber, SLAC-LBL Collaboration; see: H. Harari, Phys. Rep. C42, No. 4 (1978).
[21]C. Baltay, Report to the Recontre de Moriond Flaine, March 1977; in: Proc. of Tokyo Conference (1978).
[22]S. L. Glashow and S. Weinberg, Phys. Rev. D15, 1958 (1977).
[23]E. A. Paschos, Phys. Rev. D15, 1966 (1977).
[24]H. Georgi and D. V. Nanopoulos, HUTP-78/A055.
[25]S. Weinberg, Phys. Rev. Lett. 37, 657 (1976); A. A. Ansel'm and D. I. D'yakonov, Zh. Eksp. Teor. Fiz. 68, 1614 (1975) [Sov. Phys. JETP 41, 809 (1975)].
[26]J. E. Bjorken and S. Weinberg, Phys. Rev. Lett. 38, 622 (1977).
[27]M. S. Chanowitz, J. Ellis, and M. K. Gaillard, Nucl. Phys. B128, 506 (1977).
[28]H. Georgi and S. L. Glashow, Phys. Rev. Lett. 32, 438 (1974).
[29]J. C. Pati and A. Salam, Phys. Rev. D8, 1240 (1973); 10, 275 (1974).
[30]M. Fritzsch and P. Minkowski, Ann. Phys. 93, 193 (1975).
[31]F. Gursey and P. Sikive, Phys. Rev. Lett. 36, 775 (1975); P. Ramond, Nucl. Phys. B110, 214 (1976).
[32]L. D. Landau, in: Niels Bohr and the Development of Physics, Pergamon Press, London (1955) [Russian translation published by IL, Moscow (1958)].
[33]See, for example, S. Deser and B. Zumino, Phys. Lett. B62, 335 (1976); V. I. Ogievetskiĭ and L. Mezincescu, Usp. Fiz. Nauk 117, 637 (1975) [Sov. Phys. Usp. 18, 960 (1975)].
[34]F. Wilczek and A. Zee, Phys. Lett. B70, 418 (1977); S. Weinberg, Trans. N. Y. Acad. Sci. Ser. II, 38, 185 (1977); A. De Rujula, H. Georgi, and S. Glashow, Ann. Phys. 109, 258 (1977); S. Pakvasa and H. Sugawara, Phys. Lett. B73, 61 (1978); H. Fritzsch, Phys. Lett. B70, 436 (1977); 73, 317 (1978); R. Gatto, G. Morchio, and F. Strocchi, Phys. Lett. B80, 265 (1979).
[35]F. Wilczek and A. Zee, Phys. Rev. Lett. 42, 421 (1979).
[36]C. L. Ong, Univ. of Toronto Preprint (1978).
[37]R. Gatto *et al.*, Last reference in Ref. 34.
[38]H. Georgi, H. R. Quinn, and S. Weinberg, Phys. Rev. Lett. 33, 451 (1974).
[39]T. Appelguist and T. Carrazone, Phys. Rev. D11, 2856 (1975).
[40]D. J. Gross and F. Wilczek, Phys. Rev. Lett. 30, 1343 (1973); H. D. Politzer, Phys. Rev. Lett. 30, 1346 (1973).
[41]L. D. Landau and I. Ya. Pomeranchuk, Dokl. Akad. Nauk SSSR 102, 489 (1955); E. S. Fradkin, Zh. Eksp. Teor. Fiz. 28, 750 (1955) [Sov. Phys. JETP 1, 604 (1955)].
[42]V. B. Berestetskiĭ, Usp. Fiz. Nauk 120, 439 (1976) [Sov. Phys. Usp. 19, 934 (1976)].
[43]D. A. Ross, CERN Preprint Th. 2469-CERN (1978).
[44]A. Buras *et al.*, Nucl. Phys. B135, 66 (1978).
[45]D. V. Nanopoulos and D. A. Ross, CERN Preprint Th 2536 (1978).
[46]S. Weinberg, Second reference in Ref. 34.
[47]C. Jarlskog and F. J. Yndurain, Nucl. Phys. B149, 29 (1979).
[48]F. Reines and M. F. Crouch, Phys. Rev. Lett. 32, 493 (1974).
[49]Proc. of Seminar on Proton Stability, Univ. of Winconsin, Madison, Wisconsin, December 8, 1978.
[50]M. T. Vaughn, DESY Preprint 78/78.
[51]E. S. Fradkin and O. K. Kalashnikov, Phys. Lett. B59, 159 (1975).
[52]G. Parisi, Phys. Rev. D11, 909 (1975); L. Maiani, G. Parisi, and R. Petronzio, Nucl. Phys. B136, 115 (1978); A. Zee, COO-2220-131 (1978).
[53]S. G. Matinyan and G. K. Savvidy, Nucl. Phys. B134, 539 (1977); G. K. Savvidy, Phys. Lett. B71, 133 (1977).
[54]N. K. Nielsen and P. Olesen, Phys. Lett. B79, 304 (1978).
[55]V. F. Shvartsman, Pis'ma Zh. Eksp. Teor. Fiz. 9, 315 (1962) [JETP Lett. 9, 814 (1969)]; G. Steigman, S. N. Schramm, and T. E. Gunn, Phys. Lett. B66, 202 (1977).
[56]M. Beg, B. Marciano, and M. Ruderman, See Ref. 4.
[57]S. Weinberg, Gravitation and Cosmology, J. Wiley, New York (1972) [Russian translation published by Mir, Moscow (1975), Ch. 15].
[58]E. Gildener, Phys. Rev. D14, 1667 (1976).
[59]O. K. Kalashnilov and V. V. Klimov, Phys. Lett. B80, 75 (1978).
[60]R. N. Mohapatra and G. Sejanovic, CCNY-HEP-78/6 (revised).
[61]I. Bars and M. Serdaroglu, Yale Report COO-3075-188 (1978) (revised).
[62]S. Weinberg, Harvard Preprint HUTP-78/A060.
[63]Ya. B. Zel'dovich, Usp. Fiz. Nauk 89, 647 (1966) [Sov. Phys. Usp. 9, 602 (1967)]; Ya. B. Zel'dovich and I. D. Novikov, Stroenie i évolyutsiya vselennoĭ (Structure and Evolution of the Universe), Nauka, Moscow (1975).
[64]G. Steigman, Ann. Rev. Astron. Astrophys. 14, 339 (1976).
[65]A. Hawking, Commun. Math. Phys. 43, 199 (1975); Ya. B. Zel'dovich, Pis'ma Zh. Eksp. Teor. Fiz. 24, 29 (1975) [JETP Lett. 24, 25 (1976)]; R. M. Wald, Commun. Math. Phys. 45, 9 (1975).
[66]A. A. Belavin, A. M. Polyakov, A. S. Schwartz, and Yu. S. Tyupkin, Phys. Lett. B59, 85 (1975); G. t'Hooft, Phys. Rev. Lett. 37, 8 (1976); Phys. Rev. D14, 3432 (1976).
[67]L. Susskind and S. Dimopoulos, Phys. Rev. D18, 4500 (1979).
[68]A. D. Sakharov, Pis'ma Zh. Eksp. Teor. Fiz. 5, 32 (1967) [JETP Lett. 5, 24 (1967)].
[69]L. B. Okun and Ya. B. Zeldovich, Commun. Nucl. Part. Phys. 6, 69 (1976).
[70]A. Yu. Ignatiev, N. Y. Krasnikov, V. A. Kuzmin, and A. N. Tavkhelidze, Phys. Lett. B76, 436 (1978).
[71]N. Yoshimura, Phys. Rev. Lett. 41, 281 (1978).
[72]A. D. Sakharov, Zh. Eksp. Teor. 76, 1172 (1979) [Sov. Phys. JETP 49, 594 (1979)].
[73]A. D. Dolgov, Pis'ma Zh. Eksp. Teor. Fiz. 29, 254 (1979) [JETP Lett. 29, 228 (1979)].
[74]M. I. Vysotskii, A. D. Dolgov, and Ya. B. Zel'dovich, Pis'ma Zh. Eksp. Teor. Fiz. 26, 200 (1977) [JETP Lett. 26, 188 (1977)]; P. Hut, Phys. Lett. B69, 85 (1977); B. W. Lee and S. Weinberg, Phys. Rev. Lett. 39, 165 (1977).
[75]Ya. B. Zel'dovich, L. B. Okun', and S. B. Pikel'ner, Usp. Fiz. Nauk 87, 113 (1965) [Sov. Phys. Usp. 8, 702 (1965)].
[76]M. Gell-Mann, P. Ramond, and R. Slansky, Rev. Mod. Phys. 50, 721 (1978).
[77]H. Georgi and D. V. Nanopoulos, Harvard Preprint HUTP-79/A001.

Translated by Julian B. Barbour