Vol. B15 (1984) ACTA PHYSICA POLONICA No 3

THE GENERATION PROBLEM*, **

By G. Ecker

Institut für Theoretische Physik, Universität Wien***

(Received August 9, 1983)

Evidence for the generation structure of quarks and leptons is reviewed. The two main aspects of the generation problem are emphasized. The concept and possible problems of horizontal symmetries are discussed. Two different mechanisms for horizontal symmetries are considered leading to a generalized permutation symmetry in $SU(2)_L \times U(1)$ in one case. The second mechanism uses the discrete unbroken subgroup of an axial U(1) with hypercolour anomalies in composite models. A concrete realization in the rishon model is investigated. The two different approaches produce almost identical quark mass matrices for three generations. In addition to a correct prediction for the Cabibbo angle the models yield a very small Kobayashi-Maskawa mixing angle θ_3 and thus provide for a natural explanation of the smallness of CP violation.

PACS numbers: 11.30.Iw, 11.30.Ly

1. Introduction

The fermions which are at present considered as fundamental seem to fall into certain groups called families or generations. Let us recall the theoretical and experimental evidence for such a generation structure.

A) The neutral current interactions conserve flavour to a very high degree (d $\leftrightarrow$ s and u $\leftrightarrow$ c transitions are strongly suppressed). The only natural way to understand this phenomenon is to invoke a generalized GIM mechanism [1]. For the gauge group SU(2) $\times$ U(1) this implies [2] that all fermions of the same charge and helicity have the same eigenvalues of T_3 and $\vec{T}^2$ where $\vec{T}$ is the weak isospin.

B) For all popular gauge models including GUTs the necessary cancellation of gauge anomalies holds for each generation separately (for SU(2) $\times$ U(1) the condition is $\mathrm{Tr}\ QT_3^2 = 0$ where Q is the fermion charge matrix).

---

* Dedicated to the memory of Walter Konetschny whose sudden tragic death terminated a long friendship and collaboration.

** Presented at the XXIII Cracow School of Theoretical Physics, Zakopane, May 29-June 12, 1983.

*** Address: Institut für Theoretische Physik, Universität Wien, A-1090 Wien, Boltzmanngasse 5, Austria.

C) The fermion masses clearly distinguish between the generations. This is especially evident for the charged leptons. Of course, the neutrinos play a special rôle in this respect.

D) In contrast with the neutral currents, the charged currents mediate between generations, but these generation changing transitions are comparatively weak. This is expressed by the observed smallness of weak mixing angles which parametrize the charged current interactions.

From the first two arguments it would seem that one generation is completely sufficient as far as the vertical gauge structure is concerned. The first part of the generation problem can therefore be formulated in the following way: why does Nature choose such a repetition of nearly identical structures and how many generations are there altogether? As with most philosophical questions we should probably not expect to find a definite answer, but we may hope to gain new insights while pondering over the question.

The last two arguments for a family structure involve the second, more practical aspect of the generation problem: can we understand the structure of fermion mass matrices? Both the masses and the weak mixing angles are derived from these mass matrices which according to our present understanding emerge together with the spontaneous breaking of the electroweak gauge symmetry. We may therefore conjecture that at high energies the distinction between generations disappears and a so-called horizontal symmetry is restored.

To investigate the horizontal structure, the most straightforward approach employs the framework of low energy gauge theories based on [3] $SU(2)_L \times U(1)$ or its left-right symmetric extension [4] $SU(2)_L \times SU(2)_R \times U(1)_{B-L}$. Many proponents of GUTs have suggested to postpone the issue to energies $\gtrsim 10^{15}$ GeV. However, I find it difficult to imagine that the pattern of fermion masses in the GeV range or less is determined by the structure of the theory at such enormous energies. Of course, this prejudice is just another facet of the well-known hierarchy problem.

In order to give a precise meaning to the concept of horizontal symmetry let us consider the standard gauge group $SU(2)_L \times U(1)$. The interactions of only gauge fields and quarks[1] are given by

$$\mathscr{L} = \mathscr{L}_{\text{gauge}} + \bar{L}_\alpha i\gamma^\mu \left(\partial_\mu + \frac{i}{2} g\vec{\tau}\vec{W}_\mu + \frac{i}{6} g' B_\mu\right) L_\alpha$$
$$+ \overline{p_{\alpha R}} i\gamma^\mu \left(\partial_\mu + \frac{2i}{3} g' B_\mu\right) p_{\alpha R} + \overline{n_{\alpha R}} i\gamma^\mu \left(\partial_\mu - \frac{i}{3} g' B_\mu\right) n_{\alpha R},$$
$$\alpha = 1, \ldots, n_G, \tag{1.1}$$

where $L_\alpha = \begin{pmatrix} p_\alpha \\ n_\alpha \end{pmatrix}_L$, $p_{\alpha R}$, $n_{\alpha R}$ are the left- and right-handed quark fields (weak eigenfields), $\vec{W}_\mu$, B_μ are the $SU(2) \times U(1)$ gauge fields and n_G denotes the number of generations. Although (1.1) cannot be the whole story most particle physicists will agree upon this Lagrangian as the practically established part of electroweak interactions in the hadron

[1] Leptons will not be considered here.

sector (for $n_G = 3$). In addition to the gauge symmetry the Lagrangian (1.1) exhibits the symmetry

$$H_0 = U(n_G)_L \times U(n_G)_{p_R} \times U(n_G)_{n_R} \tag{1.2}$$

which is the maximal horizontal (= flavour) symmetry of the model. The three factors $U(n_G)$ refer to the possibility of independent unitary transformations on L, p_R, n_R and the colour anomaly has been neglected for the moment.

Of course, the huge flavour group H_0 must be broken somehow. The standard approach is to break H_0 explicitly and completely by adding Yukawa interactions of fermions and scalar fields to (1.1). This leads to the unsatisfactory situation of having as many independent parameters in $\mathscr{L}_{\text{Yukawa}}$ as there are masses and mixing angles. Keeping in mind that the scalar Higgs fields may well be only effective fields reflecting some unknown underlying dynamics, we should not overlook the possibility that H_0 is initially broken to a subgroup H only. At $E \simeq G_F^{-1/2}$ the group $SU(2)_L \times U(1) \times H$ is then broken in the usual way to $U(1)_{EM}$ and we may look for traces of H in the fermion mass matrices.

At this point, one faces the problem of possible Goldstone bosons associated with the breaking of H. Three cases must be distinguished:

(*i*) H is gauged [5] and must therefore be free of anomalies. The horizontal gauge bosons induce flavour changing neutral interactions [6]. Thus, H cannot persist down to $E \simeq G_F^{-1/2}$. In my opinion, no convincing model exists.

(*ii*) H is a global continuous symmetry [7]. In order to suppress the flavour changing neutral interactions induced by the associated Goldstone bosons, H must be broken [7] already at $E \gtrsim 10^{10}$ GeV. Therefore, H can influence the generation of fermion masses only rather indirectly and again no appealing model is known.

(*iii*) H is a discrete symmetry and, consequently, there are no Goldstone bosons at all.

Although we would like to stay in the framework of low energy gauge theories it is obvious that an understanding of H requires a theory at a more fundamental level. Since experiment has not given us any hint so far what this new theory should be we have to introduce some speculations in order to proceed. This is especially the case for the scalar Higgs sector which is probably not described by fundamental scalar fields. When talking about Yukawa interactions the possibility of an effective interaction with composite scalar fields should always be in the back of our minds. But in that case there is really no reliable way of calculating scalar vacuum expectation values (VEVs). In particular, I do not consider it meaningful to make predictions based on the special properties of a carefully designed classical ϕ^4 potential. Consequently, no assumptions about Higgs potentials will be made.

The most important aspect of H as far as mass matrices are concerned is related to the calculability of any possible relation between fermion masses and mixing angles. In other words, only if a relation obtained at the tree level is due to a symmetry of the Lagrangian can we be sure that the relation persists in higher orders of perturbation theory up to finite, calculable radiative corrections. Moreover, for such natural relations to exist the Lagrangian must contain all renormalizable interaction terms permitted by the symmetry.

Even if we restrict ourselves to discrete H infinitely many possibilities remain. Together with Walter Konetschny and Walter Grimus we have therefore attempted to classify [8, 9] quark mass matrices for all possible H. The unexpected and not generally appreciated result was that for all the cases we investigated only surprisingly few inequivalent and phenomenologically acceptable quark mass matrices exist.

Instead of reviewing the general classification of horizontal symmetries I intend to motivate in two quite different ways two specific choices for H which yield interesting mass matrices. In Sect. 2 I shall introduce a generalized permutation symmetry for H and show that it leads to the unique model [8] for $SU(2)_L \times U(1)$ and $n_G = 2$ with a definite prediction for the Cabibbo angle in agreement with experiment. For three generations, no general analysis exists within $SU(2)_L \times U(1)$ but the generalization of the successful four-flavour case is straightforward. In Sect. 3 two special examples will be discussed how horizontal symmetries arise within composite models of quarks and leptons. The realization of one of those mechanisms in the framework of the rishon model [10] will be considered in some detail.

For lack of both time and competence, no general review of suggested solutions of the generation problem will be attempted. Among the many topics omitted the generation of fermion masses via radiative corrections [11] and the possible rôle of supersymmetry [12] deserve special mentioning.

2. *Generalized permutation symmetry in $SU(2)_L \times U(1)$*

Faced with the problem of choosing a particular H among infinitely many possibilities one may appeal to intuition for guidance. An especially suggestive choice [13, 14] is the symmetric group S_n (the group of permutations of n elements) where one would like to equate n to the number n_G of generations. The drawback is that S_{n_G} does in general not possess an n_G-dimensional irreducible representation (irrep). In particular, for $n_G = 2$ S_2 is abelian and we know that the quark mass matrices are not restricted in this case [15]. A possible way out is to set [13] $n = n_G + 1$. However, it is considerably more difficult to argue that the permutation group of $n_G + 1$ elements is a natural choice for n_G generations. Moreover, S_3 gives definitely a wrong prediction [13] for θ_C ($n_G = 2$).

Another possibility would be to consider U(1) subgroups as horizontal symmetries or, in other words, phase transformations on the various fields. Again, the abelian nature of H precludes the existence of natural constraints [15] within $SU(2)_L \times U(1)$.

Instead, I propose to combine S_{n_G} and discrete U(1) subgroups in a specific way. Rather than postulating a certain group, I shall try to motivate in an admittedly speculative way the possible appearance of such a horizontal symmetry.

How could the maximal flavour group H_0 in (1.2) be broken without explicit scalar fields which is, of course, always a trivial possibility? Imagine this breaking occurs through the appearance of an operator in terms of only quark fields where the breaking can be either explicit (in an effective Lagrangian) or dynamical (through a non-vanishing VEV). Since we want a Lorentz scalar which conserves $SU(2)_L \times U(1)$ and baryon number B we need at least four quark fields. Among several possibilities consider the operator

$$A = \sum_{\alpha=1}^{n_G} \varepsilon_{ij} \bar{L}_{\alpha i} p_{\alpha R} \bar{L}_{\alpha j} n_{\alpha R} + \text{h.c.}, \tag{2.1}$$

where i, j are SU(2) indices. Assuming

$$\langle 0|A|0\rangle \neq 0 \tag{2.2}$$

the maximal flavour symmetry $\mathrm{H}_0 = \mathrm{SU}(n_\mathrm{G})^3 \times \mathrm{U}(1)^3$ gets broken. The abelian part

$$\mathrm{U}(1)^3 = \mathrm{U}(1)_\mathrm{B} \times \mathrm{U}(1)_Y \times \mathrm{U}(1)_\mathrm{PQ}, \tag{2.3}$$

where Y is the weak hypercharge, contains a so-called Peccei-Quinn symmetry [16]

$$\mathrm{U}(1)_\mathrm{PQ}: L \to e^{-i\phi/2}L,\ p_\mathrm{R} \to e^{i\phi}p_\mathrm{R},\ n_\mathrm{R} \to e^{i\phi}n_\mathrm{R} \tag{2.4}$$

which has a colour anomaly and may serve to resolve the strong CP problem. By construction, (2.2) conserves B and Y, but it breaks $\mathrm{U}(1)_\mathrm{PQ}$ implying the existence of an axion [17]. This axion can be made invisible [18] if the breaking scale associated with (2.2) is large enough (astrophysical arguments suggest [19] $\gtrsim 10^8$ GeV).

The non-abelian part $\mathrm{SU}(n_\mathrm{G})^3$ is broken by (2.2) to a subgroup which can be characterized as follows: the unitary matrices V_L, $V_{p\mathrm{R}}$, $V_{n\mathrm{R}} \in \mathrm{SU}(n_\mathrm{G})$ acting on the quarks must all be monomial matrices with exactly the same structure, i.e.

$$|V_{L,\alpha\beta}| = |V_{p\mathrm{R},\alpha\beta}| = |V_{n\mathrm{R},\alpha\beta}| = \delta_{\mathrm{P}(\alpha),\beta}, \tag{2.5}$$

where $\mathrm{P}(\alpha)$ denotes a permutation of the generation labels $\alpha = 1, \ldots, n_\mathrm{G}$; furthermore,

$$V_{p\mathrm{R},\alpha\beta}V_{n\mathrm{R},\alpha\beta} = (V_{L,\alpha\beta})^2. \tag{2.6}$$

Thus, $\mathrm{SU}(n_\mathrm{G})^3$ is broken to a subgroup of permutations and phase transformations, but this subgroup is still a continuous group. In order to avoid possible problems with Goldstone bosons (see Introduction) I assume that some further mechanism breaks this group to the discrete group

$$\mathrm{H} = (\mathrm{Z}_N \times \ldots \times \mathrm{Z}_N) \otimes \mathrm{S}_{n_\mathrm{G}}, \tag{2.7}$$

where $\otimes$ denotes a semi-direct product and there are $n_\mathrm{G}-1$ factors of cyclic groups Z_N in (2.7). This group which I call generalized permutation group $P^N_{n_\mathrm{G}}$ with $n_\mathrm{G}!N^{n_\mathrm{G}-1}$ elements is defined to be the group of $\mathrm{SU}(n_\mathrm{G})$ matrices

$$V_i(p_1, \ldots, p_{n_\mathrm{G}-1}) = \begin{pmatrix} \cdots\cdots \exp 2\pi i p_1/N \cdots\cdots\cdots \\ \cdots \exp 2\pi i p_2/N \cdots\cdots\cdots\cdots \\ \vdots \\ \cdots\cdots\cdots\cdots \pm\exp 2\pi i(-p_1 - \ldots - p_{n_\mathrm{G}-1})/N \end{pmatrix} \tag{2.8}$$

$$i = 1, \ldots, n_\mathrm{G}!, \quad p_j = 0, 1, \ldots, N-1, \quad j = 1, \ldots, n_\mathrm{G}-1,$$

where N is an arbitrary even number.

We now construct a $P^N_{n_G}$ invariant Yukawa interaction

$$-\mathscr{L}_Y = \bar{L}_\alpha \Gamma^i_{p,\alpha\beta} p_{\beta R} \phi^i_p + \bar{L}_\alpha \Gamma^j_{n,\alpha\beta} n_{\beta R} \phi^j_n + \text{h.c.} \tag{2.9}$$

with (possibly effective) Higgs fields ϕ^i_p, ϕ^j_n. $P^N_{n_G}$ invariance restricts the Yukawa coupling matrices Γ^i_p, Γ^j_n and, consequently, the quark mass matrices

$$\begin{aligned} M_p &= \Gamma^i_p v^i_p & M_n &= \Gamma^j_n v^j_n \\ \begin{pmatrix} v^i_p \\ 0 \end{pmatrix} &= \langle 0|\phi^i_p|0\rangle & \begin{pmatrix} 0 \\ v^j_n \end{pmatrix} &= \langle 0|\phi^j_n|0\rangle . \end{aligned} \tag{2.10}$$

For $n_G = 2, P^N_2$ is the dihedral double point group ${}^{(d)}D_{N/2}$ with $2N$ elements [20]. This group has $N/2-1$ inequivalent 2-dimensional irreps and is therefore much better suited as a horizontal symmetry than S_2. Without going into details [21] of irreps and Clebsch-Gordan coefficients we can assign all quark fields and two Higgs fields to doublets[2] of P^N_2 such that

$$-\mathscr{L}_Y = g_n \sum_{\alpha=1}^{2} \bar{L}_\alpha n_{\alpha R} \phi_\alpha + g_p \sum_{\alpha=1}^{2} \bar{L}_\alpha p_{\alpha R} \tilde{\phi}_\alpha + \text{h.c.}, \tag{2.11}$$

with coupling constants g_n, g_p and $\tilde{\phi} = i\sigma_2 \phi^*$ is the most general P^N_2 invariant Yukawa interaction involving these fields. Upon spontaneous symmetry breaking (2.11) yields the mass matrices

$$M_p = g_p \begin{pmatrix} v_1^* & 0 \\ 0 & v_2^* \end{pmatrix}, \quad M_n = g_n \begin{pmatrix} v_1 & 0 \\ 0 & v_2 \end{pmatrix} \tag{2.12}$$

and therefore

$$m_u/m_c = m_d/m_s, \quad \theta_C = 0 \tag{2.13}$$

in striking disagreement with experiment. The minimal and only successful remedy is to introduce one additional Higgs field ϕ_3 which is a P^N_2 singlet and which for group theoretical reasons [21] can only couple to one quark charge sector, e.g. to the n-quarks. $\mathscr{L}_Y$ now contains an additional piece

$$h(\bar{L}_1 n_{2R} + \varepsilon \bar{L}_2 n_{1R})\phi_3 + \text{h.c.} \quad (\varepsilon = \pm 1). \tag{2.14}$$

Altogether, we obtain the mass matrices

$$M_p = g_p \begin{pmatrix} v_1^* & 0 \\ 0 & v_2^* \end{pmatrix}, \quad M_n = \begin{pmatrix} g_n v_1 & h v_3 \\ \varepsilon h v_3 & g_n v_2 \end{pmatrix}. \tag{2.15}$$

The main point of considering $n_G = 2$ in such detail is that the mass matrices (2.15) define the *unique* four-flavour model [8] based on $SU(2)_L \times U(1)$ and any horizontal

---

[2] All Higgs fields are, of course, doublets under $SU(2)_L$.

symmetry with a natural prediction for θ_C in agreement with experiment. As a matter of fact, from (2.15) one obtains

$$|m_d/m_s - m_u/m_c|^{1/2} \lesssim \theta_C \lesssim (m_d/m_s + m_u/m_c)^{1/2} \tag{2.16}$$

to first order in quark mass ratios. Using only the mean values of these ratios [22] one finds $0.217 \lesssim \theta_C \lesssim 0.234$ in excellent agreement with [23] $\theta_C^{exp} = 0.230 \pm 0.011$.

The quark mass hierarchy in this model is not due to the coupling constants g_p, g_n, h, but rather to a hierarchy of VEVs $|v_1| \ll |v_2|$. Furthermore, although the mass matrices (2.15) are unique up to basis transformations they can be implemented with any P_2^N for $N \geq 8$. This is a specific example for the more general result [8, 9] that, given the gauge group, n_G and the maximal number of scalar fields, there are only a few inequivalent, phenomenologically acceptable cases in spite of infinitely many possible choices for H.

For $n_G = 3$, the generalized permutation group P_3^N is sometimes called the "dihedral-like" subgroup [24] $\Delta(6N^2)$ of SU(3). It has irreps of dimensions 1, 2, 3 and 6 of which the $2(N-1)$ inequivalent 3-dimensional irreps are of special interest for us. Extending the $n_G = 2$ case in a straightforward way [21], we put all quarks and altogether six scalar fields into P_3^N triplets to obtain the invariant Yukawa Lagrangian

$$\begin{aligned} -\mathscr{L}_Y = {} & g_p \sum_{\alpha=1}^{3} \bar{L}_\alpha p_{\alpha R} \tilde{\phi}_\alpha + g_n \sum_{\alpha=1}^{3} \bar{L}_\alpha n_{\alpha R} \phi_\alpha \\ & + h\{(\bar{L}_1 n_{2R} + \varepsilon \bar{L}_2 n_{1R})\chi_1 + (\bar{L}_2 n_{3R} + \varepsilon \bar{L}_3 n_{2R})\chi_2 \\ & + (\bar{L}_3 n_{1R} + \varepsilon \bar{L}_1 n_{3R})\chi_3\} + \text{h.c.} \quad (\varepsilon = \pm 1). \end{aligned} \tag{2.17}$$

With an obvious notation for the scalar VEVs, the quark mass matrices are

$$M_p = g_p \begin{pmatrix} v_1^* & 0 & 0 \\ 0 & v_2^* & 0 \\ 0 & 0 & v_3^* \end{pmatrix}, \quad M_n = \begin{pmatrix} g_n v_1 & hw_1 & \varepsilon hw_3 \\ \varepsilon hw_1 & g_n v_2 & hw_2 \\ hw_3 & \varepsilon hw_2 & g_n v_3 \end{pmatrix}. \tag{2.18}$$

Note that for $w_2 = w_3 = 0$ the third generation decouples and we get the $n_G = 2$ model (2.15) with the additional approximate mass relation $m_t \sim m_b m_c/m_s$. I shall not discuss (2.18) in full generality here because the mass matrices contain 7 independent parameters and thus the mixing angles can not all be expressed in terms of only mass ratios. From M_p in (2.18) we conclude that up to permutations $|v_1| \ll |v_2| \ll |v_3|$. It is therefore rather plausible that either $|w_3| \ll |w_2|$ or vice versa. Moreover,

$$|g_n v_1| = \left|\frac{g_n}{g_p}\right| m_u \simeq \frac{m_b}{m_t} m_u \ll m_d \tag{2.19}$$

and so I shall set $w_3 = g_n v_1 = 0$ for the following discussion.

In this case, the diagonalization of M_n is easily performed and one finds approximately

$$s_1 \simeq \sqrt{m_d/m_s} \qquad s_2 \simeq \sqrt{|\eta|}$$

$$s_3 \simeq s_2 m_s/m_b \qquad \sin\delta \simeq \frac{m_b m_c}{m_t m_s}\frac{\operatorname{Im}\eta}{|\eta|} \tag{2.20}$$

for the Kobayashi-Maskawa mixing angles [25], where

$$|m_s/m_b - m_c/m_t| \lesssim |\eta| \lesssim m_s/m_b + m_c/m_t, \qquad \eta = (hw_2/m_b)^2. \tag{2.21}$$

From (2.20) and (2.21) one gets

$$s_2 s_3 |\sin\delta| \lesssim \frac{m_s m_c}{m_b m_t}\left(1 + \frac{m_b m_c}{m_t m_s}\right) \lesssim 3 \cdot 10^{-3} \tag{2.22}$$

and thus the smallness of CP violation is naturally explained in the model by the small value for s_3. In addition to reproducing the successful prediction for $\theta_1 = \theta_C$ one finds with the standard quark masses [22]

$$0 \leqslant s_2 \lesssim 0.18 f(m_t)$$

$$0 \leqslant s_3 \lesssim 0.006 f(m_t), \qquad f(m_t) = (1 + 43\ \mathrm{GeV}/m_t)^{1/2}. \tag{2.23}$$

These predictions agree with the results of the most recent analysis [23] which gives $s_2 = 0.12$, $s_3 = 0$ as best fit values and $0.05 \lesssim s_2 \lesssim 0.3$, $0 \leqslant s_3 \lesssim 0.15$ as approximate 1 st. dev. error domains.

As for $n_G = 2$, the quark mass hierarchy reflects a hierarchy of scalar VEVs and there are infinitely many groups P_3^N consistent with (2.18). In addition, the model for $n_G = 3$ contains the attractive feature that all fields are grouped in 3-dimensional irreps of the horizontal group.

3. Composite quarks and leptons

The multitude of quarks and leptons and the associated generation problem were among the main motivations for speculating about the possible existence of more fundamental constituents. It is therefore an appropriate question what the recent investigations of composite models have taught us with respect to the generation problem. Only two of the proposed solutions will be discussed here which are in certain sense two extreme approaches to the problem.

Almost all composite models postulate a new strong (hypercolour) gauge interaction which confines the fundamental constituents (preons) into quarks and leptons below a confinement scale Λ_{HC}. In complete analogy to (1.1), the original Lagrangian of only preons and hypercolour gauge bosons possesses a large chiral (= flavour) symmetry if the preons are assumed to be massless. This chiral symmetry is called upon to solve two problems at the same time: it should explain why quarks and leptons are so much lighter than the natural scale Λ_{HC} ($\gtrsim$ 500 GeV) and, secondly, it should provide the key to the generation problem.

As for the maximal flavour symmetry H_0 of the standard model, the crucial question is not the existence of a chiral symmetry, but rather how much of this symmetry survives down to the level of the observed fermions. The general situation is as follows [26]: part of the chiral symmetry will be broken for $E < \Lambda_{HC}$ implying a number of (possibly pseudo-) Goldstone bosons which either turn into the longitudinal components of massive gauge bosons or, if they remain massless, must be made invisible [18, 27] by choosing Λ_{HC} large enough. For the remaining chiral symmetry that survives confinement the 't Hooft consistency condition [26] requires that the anomalies for the unbroken chiral generators match at the constituent and the composite level. Note that this condition only applies to continuous, but not to discrete chiral symmetries.

How do composite models account for generations? At least three possibilities may be envisaged.

A) The higher generations are radial and/or orbital excitations

This scenario is generally considered unacceptable because one expects that an interaction confined to a volume of radius $R \sim \Lambda_{HC}^{-1}$ will lead to excitation energies of the order Λ_{HC}. Such states may exist but they have nothing to do with the observed fermions.

B) Generations are distinguished at the preon level

The most systematic study of solutions of the 't Hooft consistency conditions has been performed by Bars [28]. His approach is extreme in the following sense: in addition to many other requirements such as the persistent mass condition [29], the full chiral symmetry is assumed to survive preon confinement. Among the models satisfying all requirements consider the following typical example [28] which is based on the hypercolour gauge group $G_{HC} = SU(4) \times SU(4)$ with right-handed preons

$$\begin{aligned} 1 &\sim R_1 = (4,4), \\ 4 &\sim R_2 = (\bar{4},1), \\ 10 &\sim R_3 = (1,\bar{4}), \\ 6 &\sim \bar{R}_3 = (1,4), \end{aligned} \tag{3.1}$$

where the multiplicities of the representations R_i of G_{HC} are indicated by the first number in each row. The set (3.1) is free of G_{HC} anomalies and implies the chiral flavour symmetry

$$G_F = SU(4) \times SU(10) \times SU(6) \times U(1)^2 \times Z^2, \tag{3.2}$$

where the two cyclic groups are the remains of two axial U(1) groups with hypercolour anomalies (see below under *C*). The anomaly conditions for (3.1) and (3.2) enforce the existence of exactly four generations of quarks and leptons. The number four is due to the multiplicity of preons of type R_2. In other words, the four generations differ by the kind of R_2 preon they contain.

The model of Bars provides a possible answer to the first part of the generation problem: there are exactly four generations because otherwise some basic requirements of quantum field theory [26] would have to be violated. On the other hand, it is probably premature to worry about the structure of quark mass matrices in the model. After all, counting the number of fermionic degrees of freedom (Weyl fields) one finds 96 fundamental constituents, but only 64 composite quarks and leptons. Without any hint from experiment, it is a matter of taste whether to consider models of this kind as serious candidates for a fundamental theory of matter.

C) Generations differ by preon pair excitation

The original chiral symmetry of a composite model always contains at least one axial U(1) with a hypercolour anomaly (there are even two in (3.2)). If one accepts the common belief that the physically relevant instanton solutions have integer topological charge, the anomalous U(1) is not broken completely but a certain cyclic subgroup Z_{2K} remains unbroken where K is related to the number of fundamental constituents. This discrete axial group was suggested by Harari and Seiberg [30] to distinguish between the generations. At the same time, such discrete groups may be the only chiral symmetries [31] left after preon confinement. Contrary to scenario B) there are no anomaly conditions to be fulfilled in this case.

At the level of composite fermions we thus have an axial horizontal symmetry Z_{2K}. Experience with horizontal symmetries suggests that with an abelian H we need a left--right symmetric gauge theory at the composite level to obtain any natural constraints for fermion mass matrices. As a prominent example consider the rishon model [10] with gauge group $SU(3)_{HC}\times SU(3)_C\times U(1)_{EM}$ and with preons (= rishons) $T_{L,R}(3,3)_{Q=1/3}$, $V_{L,R}(3,\bar{3})_{Q=0}$ transforming as indicated under the gauge group. The naive chiral symmetry of the model is $U(1)^4$, but the symmetry $U(1)_X$ associated to the current

$$X_\mu = \bar{T}\gamma_\mu\gamma_5 T + \bar{V}\gamma_\mu\gamma_5 V \tag{3.3}$$

is anomalous. Due to instantons, the chiral charge X changes by [32]

$$\Delta X = 12(\nu_{HC}+\nu_C) \tag{3.4}$$

with integer topological charges ν_{HC}, ν_C. The number 12 is due to the existence of two colour triplets of both helicities for the hypercolour anomaly and vice versa for the colour anomaly. Therefore, $U(1)_X$ is broken to the discrete subgroup Z_{12} or, in other words, X is only conserved mod 12.

In the rishon model one quark generation is given by

$$\begin{aligned} p_L &= T_L T_L V_L & p_R &= T_R T_R V_R \\ n_L &= V_L^C V_L^C T_L^C & n_R &= V_R^C V_R^C T_R^C \end{aligned} \tag{3.5}$$

with $X = \mp 3$ for the left (right)-handed fields. According to Harari and Seiberg [10, 30], further generations may be obtained from (3.5) by appending composite rishon operators which have vacuum quantum numbers except for Z_{12}. The simplest such operators are $T_L T^C{}_L V_L V^C{}_L$ and $T_R T^C{}_R V_R V^C{}_R$ with $X = \mp 4$.

The structure of the theory at low energies ($E < \Lambda_{HC}$) is assumed to be that of a left-right symmetric $SU(2)_L \times SU(2)_R \times U(1)_{B-L}$ gauge theory. In particular, there will be composite scalar fields ϕ (and their charge conjugates $\tilde{\phi} = \sigma_2 \phi^* \sigma_2$) transforming as quark bilinears $\overline{q_R} q_L$ which give rise to an effective Yukawa interaction

$$-\mathscr{L}_Y = \overline{q_L} \Gamma_i \phi_i q_R + \overline{q_L} \Delta_i \tilde{\phi}_i q_R + \text{h.c.} \tag{3.6}$$

with n_G-dimensional coupling matrices Γ_i, Δ_i (hermitian because of left-right symmetry). Upon spontaneous symmetry breaking the quark mass matrices

$$\begin{aligned} M_p &= v_i \Gamma_i + w_i^* \Delta_i \\ M_n &= w_i \Gamma_i + v_i^* \Delta_i \end{aligned}, \qquad \langle 0|\phi_i|0\rangle = \begin{pmatrix} v_i & 0 \\ 0 & w_i \end{pmatrix} \tag{3.7}$$

will be constrained if we impose the residual chiral symmetry Z_{12} on (3.6).

With the composite rishon operators mentioned above there are different ways to assign the discrete quantum number X to the different generations. For $n_G = 3$ there are only three non-trivial, inequivalent possibilities [33] for $X_{q_L} (X_{q_R} = -X_{q_L})$:

$$\begin{aligned} &(i) && X_{q_L} = -3, 5, 1 \\ &(ii) && X_{q_L} = -3, -3, 5 \qquad \pmod{12} \\ &(iii) && X_{q_L} = -3, 5, 5. \end{aligned} \tag{3.8}$$

Since the effective scalar fields transform as $\overline{q_R} q_L$ we conclude from (3.8) that there are only three different types of fields ϕ_1, ϕ_2, ϕ_3 with $X = -2, 2, 6 \pmod{12}$, respectively. Thus, all scalars have $X \neq 0 \pmod{12}$ and, consequently, Z_{12} keeps all quarks massless as long as it is a good symmetry.

The various cases in (3.8) can now be investigated along standard lines. The only interesting model appears [33] in case *iii*) with two fields ϕ_1, ϕ_3 together with their charge conjugates. In a convenient choice of basis, Z_{12} restricts the Yukawa coupling matrices to be of the form

$$\Gamma_1 = \begin{pmatrix} g_1 & 0 & g_{13} \\ 0 & 0 & 0 \\ g_{13}^* & 0 & g_3 \end{pmatrix}, \quad \Delta_1 = \begin{pmatrix} 0 & h_1 & 0 \\ h_1^* & 0 & h_2 \\ 0 & h_2^* & 0 \end{pmatrix} \tag{3.9}$$

$$\Gamma_3 \sim \Delta_3 \sim \begin{pmatrix} 0 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

A similar model was found earlier in the context of a general analysis of quark mass matrices [9].

In the limit of very large m_t we obtain from (3.9) the quark mass matrices

$$M_p \simeq \begin{pmatrix} g_1 v_1 & 0 & 0 \\ 0 & k_p & 0 \\ 0 & 0 & g_3 v_1 \end{pmatrix}, \qquad M_n \simeq \begin{pmatrix} 0 & h_1 v_1^* & 0 \\ h_1^* v_1^* & k_n & h_2 v_1^* \\ 0 & h_2^* v_1^* & g_3 w_1 \end{pmatrix}. \tag{3.10}$$

Surprisingly, these mass matrices have the same structure as those of (2.18) (for $g_n v_1 = w_3 = 0$). There are two differences, however: the mass hierarchy is due to a hierarchy of Yukawa couplings rather than of scalar VEVs and k_p, k_n are arbitrary so that s_2 cannot be predicted in the present case. However, the main features

$$s_1 \simeq \sqrt{m_d/m_s}, \qquad s_3 \simeq s_2 m_s/m_b \tag{3.11}$$

and, consequently, naturally small CP violation hold for the mass matrices (3.10) as well.

For finite m_t, the mass matrices of this model [33] differ from those in (2.18). The corrections to (3.11) are inversely proportional to m_t and amount to at most 9% for s_1. Although the corrections for s_3/s_2 can be as large as 90% the crucial prediction of a very small s_3 ($s_3 \lesssim 0.06 s_2$) is preserved.

4. *Conclusion*

With two completely different approaches we have obtained almost the same quark mass matrices for three generations. In addition to a successful relation for the Cabibbo angle the models predict a very small Kobayashi-Maskawa mixing angle θ_3 and therefore provide for a natural explanation of the size of CP violation.

In a more general spirit, the main conclusion can be formulated as a recommendation: do not neglect discrete symmetries.

I wish to thank A. Białas and K. Zalewski for the cordial invitation to Zakopane. I am also grateful to many participants of the School for very enlightening discussions concerning not only physics but also matters of a more general nature.

REFERENCES

[1] S. L. Glashow, J. Iliopoulos, L. Maiani, *Phys. Rev.* **D2**, 1285 (1970).
[2] S. L. Glashow, S. Weinberg, *Phys. Rev.* **D15**, 1958 (1977); E. A. Paschos, *Phys. Rev.* **D15**, 1966 (1977).
[3] S. L. Glashow, *Nucl. Phys.* **22**, 579 (1961); S. Weinberg, *Phys. Rev. Lett.* **19**, 1264 (1967); A. Salam, Proc. 8th Nobel Symposium, N. Svartholm, ed., Almqvist and Wiksell, Stockholm 1968.
[4] J. C. Pati, A. Salam, *Phys. Rev.* **D10**, 275 (1975); R. N. Mohapatra, J. C. Pati, *Phys. Rev.* **D11**, 566 and 2558 (1975); G. Senjanović, R. N. Mohapatra, *Phys. Rev.* **D12**, 1502 (1975).
[5] T. Maehara, T. Yanagida, *Progr. Theor. Phys.* **60**, 822 (1978) and **61**, 1434 (1979); S. Barr, A. Zee, *Phys. Rev.* **D17**, 1854 (1978); F. Wilczek, A. Zee, *Phys. Rev. Lett.* **42**, 421 (1979); C. L. Ong, *Phys. Rev.* **D19**, 2738 (1979); A. Davidson, M. Koca, K. C. Wali, *Phys. Rev.* **D20**, 1195 (1979); J. Chakrabarti, *Phys. Rev.* **D20**, 2411 (1979).

[6] R. N. Cahn, H. Harari, *Nucl. Phys.* **B176**, 135 (1980); G. L. Kane, R. Thun, *Phys. Lett.* **94B**, 513 (1980).

[7] D. B. Reiss, *Phys. Lett.* **115B**, 217 (1982); F. Wilczek, *Phys. Rev. Lett.* **49**, 1549 (1982); G. B. Gelmini, S. Nussinov, T. Yanagida, ICTP Trieste preprint IC/82/157.

[8] G. Ecker, W. Konetschny, *Z. Phys.* **C3**, 155 (1979) and **C4**, 353 (1980); *Phys. Lett.* **91B**, 225 (1980).

[9] G. Ecker, W. Grimus, W. Konetschny, *Phys. Lett.* **94B**, 381 (1980); *Nucl. Phys.* **B177**, 489 (1981) and **B191**, 465 (1981).

[10] H. Harari, N. Seiberg, *Phys. Lett.* **98B**, 269 (1981); *Nucl. Phys.* **B204**, 141 (1982).

[11] E.g., L. E. Ibánez, Proc. of Int. Conf. on High Energy Physics, Lisbon 1981, eds. J. Dias de Deus and J. Soffer, p. 1003; A. Masiero, D. V. Nanopoulos, K. Tamvakis, *Phys. Lett.* **126B**, 337 (1983).

[12] W. Buchmüller, R. D. Peccei, T. Yanagida, *Phys. Lett.* **124B**, 67 (1983); *Nucl. Phys.* **B227**, 503 (1983).

[13] S. Pakvasa, H. Sugawara, *Phys. Lett.* **73B**, 61 (1978) and **82B**, 105 (1979); Y. Yamanaka, H. Sugawara, S. Pakvasa, *Phys. Rev.* **D25**, 1895 (1982).

[14] E. Derman, *Phys. Rev.* **D19**, 317 (1979); E. Derman, H. S. Tsao, *Phys. Rev.* **D20**, 1207 (1979).

[15] R. Barbieri, R. Gatto, F. Strocchi, *Phys. Lett.* **74B**, 344 (1978); D. Wyler, *Phys. Rev.* **D19**, 330 (1979).

[16] R. D. Peccei, H. Quinn, *Phys. Rev. Lett.* **38**, 1440 (1977); *Phys. Rev.* **D16**, 1791 (1977).

[17] S. Weinberg, *Phys. Rev. Lett.* **40**, 223 (1978); F. Wilczek, *Phys. Rev. Lett.* **40**, 279 (1978).

[18] J. E. Kim, *Phys. Rev. Lett.* **43**, 103 (1979); M. Dine, W. Fischler, M. Srednicki, *Phys. Lett.* **104B**, 199 (1981).

[19] M. Fukugita, S. Watamura, M. Yoshimura, *Phys. Rev.* **D26**, 1840 (1982).

[20] J. S. Lomont, *Applications of Finite Groups*, Academic Press, New York-London 1959.

[21] G. Ecker, to be published.

[22] J. Gasser, H. Leutwyler, *Phys. Rep.* **87C**, 77 (1982).

[23] E.g., K. Kleinknecht, B. Renk, *Z. Phys.* **C20** 67 (1983).

[24] W. M. Fairbairn, T. Fulton, W. H. Klink, *J. Math. Phys.* **5**, 1038 (1964); A. Bovier, M. Lüling, D. Wyler, *J. Math. Phys.* **22**, 1543 (1981).

[25] M. Kobayashi, T. Maskawa, *Progr. Theor. Phys.* **49**, 652 (1973).

[26] G. 't Hooft, in *Recent Developments in Gauge Theories*, eds. G. 't Hooft et al., Plenum Press, N.Y. 1980; Y. Frishman et al., *Nucl. Phys.* **B177**, 157 (1981); S. Coleman, B. Grossman, *Nucl. Phys.* **B203**, 205 (1982).

[27] H. Georgi, Proc. 4th Kyoto Summer Institute on Grand Unified Theories and Related Topics, eds. M. Konuma and T. Maskawa, World Scient. Publ., Singapore 1981; D. B. Reiss, Ref. [7]; G. Ecker, *Z. Phys.* **C18**, 173 (1983).

[28] I. Bars, *Nucl. Phys.* **B208**, 77 (1982); Yale Univ. preprint YTP 82-12, to appear in Proc. Rencontres de Moriond 1982.

[29] J. Preskill, S. Weinberg, *Phys. Rev.* **D24**, 1059 (1981).

[30] H. Harari, N. Seiberg, *Phys. Lett.* **102B**, 263 (1981).

[31] S. Weinberg, *Phys. Lett.* **102B**, 401 (1981); E. Guadagnini, K. Konishi, *Nucl. Phys.* **B196**, 115 (1982); K. Yamawaki, T. Yokota, *Phys. Lett.* **113B**, 293 (1982); H. Harari, N. Seiberg, *Phys. Lett.* **115B**, 450 (1982).

[32] G. 't Hooft, *Phys. Rev. Lett.* **37**, 8 (1976); *Phys. Rev.* **D14**, 3432 (1976); R. Jackiw, C. Rebbi, *Phys. Rev. Lett.* **37**, 172 (1976); C. G. Callan, R. F. Dashen, D. J. Gross, *Phys. Lett.* **63B**, 334 (1976).

[33] G. Ecker, *Phys. Lett.* **121B**, 326 (1983).