Variations on a theme: The ansatz of Stech for the quark mass matrices

L. Lavoura*
Instituto Nacional de Investigação Científica, Centro de Física da Matéria Condensada, Avenida do professor Gama Pinto, 2, P-1699 Lisboa Codex, Portugal
(Received 19 September 1989)

I analyze the *Ansatz* of Stech for the quark mass matrices, showing that it is equivalent to a system of three linear equations among the squared moduli of the Cabibbo-Kobayashi-Maskawa matrix elements. I show that in the *Ansatz* of Stech the roles of M_U and M_D may be exchanged, and suggest a generalization of the *Ansatz* of Stech in which the top-quark mass may be higher than 100 GeV.

INTRODUCTION

In 1983 Stech suggested[1] an *Ansatz* for the quark mass matrices which was inspired by the symmetry properties of the Yukawa coupling matrices in some grand unified theories. The *Ansatz* of Stech (AS) became rather popular. Unified with the other major *Ansatz* for the mass matrices, the one of Fritzsch,[2] it gave rise to the Gronau-Johnson-Schechter scheme.[3] However, by now the finding[4] that the top quark is certainly heavier than 50 GeV has eliminated the AS from consideration.

The AS is interesting because its main idea is different from the one behind the *Ansatz* of Fritzsch. Fritzsch viewed[5] the quark mass matrices as arising from a step-by-step chiral-symmetry breaking; each new step of the breaking gives mass to the quarks of one further generation, and makes those quarks participate in the weak mixing, which is expressed by the Cabibbo-Kobayashi-Maskawa[6] (CKM) matrix V. From this point of view, the smallness of that mixing (the smallness of the off-diagonal matrix elements of V) is related to the strong hierarchy of the masses. This is also the idea behind the "democratic family mixing"[7] scheme. The idea which inspires the AS is quite different. In that *Ansatz* the smallness of the mixing is due to an approximate proportionality between the up-type (charge $\frac{2}{3}$) and down-type (charge $-\frac{1}{3}$) quark mass matrices, M_U and M_D. Though this approximate proportionality will in general lead to a low-mass top quark, following the approximate equation $m_t \approx m_b m_c / m_s$, this is indeed an interesting idea, which deserves further investigation.

In this paper I first analyze the AS. I show that, for three generations, it is equivalent to a set of three equations among the physical quantities—quark masses and CKM matrix parameters—alone, equations which are linear in the squared moduli $U_{ij} \equiv |V_{ij}|^2$. As a consequence of this fact, in the AS we can easily calculate all the U_{ij} from one of them, which is taken as an input quantity, and from the quark masses. In that calculation we do not have to deal with any quantity which is not directly measurable. This is an advantage of the AS over the *Ansatz* of Fritzsch, where in the calculation of the CKM matrix one uses, in addition to the quark masses, two arbitrary phases, which are parameters not directly measurable and without a clear physical meaning. I also emphasize that the *Ansatz* of Stech contains discrete ambiguities in the signs of the quark masses. These ambiguities were usually overlooked in the existing literature, but are quite important, for they destroy most of the predictive power of the AS.

Second I note that in the *Ansatz* of Stech the roles played by M_U and M_D may be exchanged between these two matrices. The resulting *Ansatz* does not fit the data as easily as the AS, but still works, and it may be considered to be as natural, from the technical point of view, as the AS. Unfortunately this variation of the AS shares with it the need for a light top quark, and is thus also excluded by our present experimental knowledge.[4]

Third, I suggest and discuss a generalization of the AS which bridges the gap between it and its counterpart with $M_U \leftrightarrow M_D$. This generalization, though technically quite amusing, is at first sight uninteresting for, just as its two limiting cases, it yields a light top quark. However, contrary to the two limiting cases, that extension may be successfully implemented as an *Ansatz* for the Hermitian mass matrices $H_{U,D} \equiv M_{U,D} M_{U,D}^\dagger$ instead of for the mass matrices $M_{U,D}$. This is very interesting, for in the context of the standard model[8] the matrices $H_{U,D}$ do not contain some of the spurious information contained in the matrices $M_{U,D}$ and in particular they do not have the quark-mass-sign ambiguity. Furthermore, as an *Ansatz* for the H matrices this generalization of the AS easily fits the data on the $|V_{ij}|$ with a high top-quark mass, and indeed we find that in this scheme the top-quark mass may be as high as 160 GeV, and thus much higher than in the existing *Ansätze* for the mass matrices.

THE *Ansatz* OF STECH

Stech postulated that there exists a weak basis in which M_U is a real symmetric matrix and

$$M_D = pM_U + A \ , \tag{1}$$

where p is a real number and A is a Hermitian and antisymmetric matrix. Because of the presence of the matrix A in M_D, the up-type and down-type quark masses are not proportional, and there is weak mixing.

In this paper I will use the method of the mass-matrix invariants,[9-11] which when applied to the AS affords interesting insights and is very straightforward. First, from the antisymmetry of A and the symmetry of M_U it follows that

$$\mathrm{tr}(M_U M_D)=\frac{\mathrm{tr}M_D}{\mathrm{tr}M_U}\mathrm{tr}M_U^2 , \tag{2a}$$

$$\mathrm{tr}(M_U^2 M_D)=\frac{\mathrm{tr}M_D}{\mathrm{tr}M_U}\mathrm{tr}M_U^3 . \tag{2b}$$

We might proceed to higher powers of M_U, and consider $\mathrm{tr}(M_U^3 M_D)$, etc; however, the resulting equations would not be independent from Eqs. (2), because of the identity, obeyed by any 3×3 matrix C,

$$C^3-C^2\mathrm{tr}C+C\,\mathrm{csi}C-\det C=0 . \tag{3}$$

Here, "csi C" is my notation for the second-degree invariant of C, i.e., the coefficient of $(-\lambda)$ in the characteristic equation $\det(C-\lambda)=0$.

The matrix A besides being antisymmetric is also pure imaginary. Therefore,[10]

$$\det A=\det\left[M_D-\frac{\mathrm{tr}M_D}{\mathrm{tr}M_U}M_U\right]=0 . \tag{4}$$

Contrary to Eqs. (2), which explicitly display the difference of the roles of M_U and M_D in the *Ansatz* of Stech, Eq. (4) remains unchanged under $M_U\leftrightarrow M_D$. Using Eq. (3) and the tracelessness of A, Eq. (4) may be rewritten as

$$\{(\mathrm{tr}^2M_D)(\mathrm{tr}M_U)\mathrm{tr}(M_U^2M_D)-(\mathrm{tr}^3M_D)[\det M_U -(\mathrm{tr}M_U)(\mathrm{csi}M_U)]\}-(M_U\leftrightarrow M_D)=0 . \tag{5}$$

It was first noticed in Ref. 10 that Eq. (4) is a common feature of the Fritzsch and the Stech *Ansätze*. The *Ansätze* that I shall suggest in this paper also share this feature. In a recent paper[12] I suggested an *Ansatz* which has $\det A\neq 0$.

It is easy to show that Eqs. (2) and (4) are not only necessary, but also sufficient, for the *Ansatz* of Stech to hold.

Equations (2) and (5) are important because they are equations among the physical quantities alone.[10] As M_U and M_D are both Hermitian,[9]

$$\mathrm{tr}(M_U^a M_D^b)=\sum_{i,j=1}^{3} u_i^a d_j^b U_{ij} , \tag{6}$$

where the u_i and d_j denote the (real) eigenvalues of M_U and M_D, respectively, i.e., the quark masses, taken however with *a priori* arbitrary signs. The exponents a and b are arbitrary. Therefore, Eqs. (2) and (5) are constraints that the *Ansatz* of Stech enforces on the quark masses and mixing parameters alone, and no quantities which are not directly measurable are present in those equations, except for the arbitrary, and physically meaningless, signs of the eigenvalues of M_U and M_D. Those signs constitute an ambiguity that we may, or may not, be able to eliminate in the confrontation with experiment.

As a consequence of the normalization of the rows and columns of V, only four out of the nine U_{ij} are independent. Moreover,[9] we may take U_{12}, U_{13}, U_{21}, and U_{23} as the parameters of V. Then, Eqs. (2) yield

$$(u_1+u_2+u_3)[(d_2-d_1)U_{12}+(d_3-d_1)U_{13}] =u_1d_2+u_1d_3-u_2d_1-u_3d_1 , \tag{7a}$$

$$(u_1+u_2+u_3)[(d_1-d_2)U_{21}+(d_3-d_2)U_{23}] =u_2d_1+u_2d_3-u_1d_2-u_3d_2 . \tag{7b}$$

Equation (5) too is a linear equation on these four U_{ij}, and its coefficients are functions of the u_i and d_j. We may solve the resulting system of three linear equations to determine three of the U_{ij} as functions of the fourth one and of the quark masses. For instance, taking $U_{12}=|V_{us}|^2$ as an input quantity, together with the quark masses and their signs, we obtain for the other U_{ij} the following exact equations:

$$U_{11}=-\frac{d_3-d_2}{d_3-d_1}U_{12}+\frac{u_1d_2+u_1d_1-u_3d_3-u_2d_3}{(d_1-d_3)\mathrm{tr}M_U} , \tag{8a}$$

$$U_{13}=-\frac{d_2-d_1}{d_3-d_1}U_{12}+\frac{u_1d_3+u_1d_2-u_3d_1-u_2d_1}{(d_3-d_1)\mathrm{tr}M_U} , \tag{8b}$$

$$\begin{aligned}U_{21}=\frac{1}{(u_3-u_2)(d_3-d_1)}\Bigg\{&(u_3-u_1)(d_3-d_2)U_{12}-\frac{\mathrm{tr}M_U\det M_D}{(d_2-d_1)\mathrm{tr}M_D}\\ &+\frac{1}{(d_2-d_1)\mathrm{tr}^2M_U}[-u_1^3d_2^2-u_2^3d_1^2-u_3^3d_2d_1+u_3^2u_2d_1(d_1+2d_3)\\ &+u_3^2u_1d_2(d_2+2d_3)+u_3u_2^2d_1(d_2+2d_3)\\ &+u_3u_1^2d_2(d_1+2d_3)+u_2^2u_1(-d_1^2+d_2d_1+\mathrm{csi}M_D)\\ &+u_2u_1^2(-d_2^2+d_2d_1+\mathrm{csi}M_D)\\ &-(d_1^2+d_2^2+2d_3^2+d_3d_2+d_3d_1)\det M_U]\Bigg\} ,\end{aligned} \tag{9a}$$

$$U_{22}=-\frac{u_3-u_1}{u_3-u_2}U_{12}+\frac{1}{(u_3-u_2)(d_3-d_2)(d_2-d_1)}\left\{\frac{\mathrm{tr}M_U}{\mathrm{tr}M_D}\det M_D - \frac{1}{\mathrm{tr}^2M_U}[-u_1^3d_2^2-u_2^3d_2^2 -u_3^3(d_3d_2+d_2d_1-d_3d_1) +u_3^2(u_2+u_1)(d_2^2+2d_3d_1) +u_2u_1(u_2+u_1)d_2(-d_2+2d_3+2d_1) +u_3(u_2^2+u_1^2)\mathrm{csi}M_D -2(d_3^2+d_1^2+d_3d_1)\det M_U]\right\}, \tag{9b}$$

$$U_{23}=\frac{1}{(u_3-u_2)(d_3-d_1)}\left\{(u_3-u_1)(d_2-d_1)U_{12}-\frac{\mathrm{tr}M_U\det M_D}{(d_3-d_2)\mathrm{tr}M_D} +\frac{1}{(d_3-d_2)\mathrm{tr}^2M_U}[-u_1^3d_2^2-u_2^3d_3^2-u_3^3d_3d_2+u_3^2u_2d_3(d_3+2d_1) +u_3^2u_1d_2(d_2+2d_1)+u_3u_2^2d_3(d_2+2d_1) +u_3u_1^2d_2(d_3+2d_1)+u_2^2u_1(-d_3^2+d_3d_2+\mathrm{csi}M_D) +u_2u_1^2(-d_2^2+d_3d_2+\mathrm{csi}M_D) -(d_3^2+d_2^2+2d_1^2+d_3d_1+d_2d_1)\det M_U]\right\}, \tag{9c}$$

and finally

$$U_{3i}=U_{2i}(u_2\leftrightarrow u_3) \quad \text{for } i=1,2,3 . \tag{10}$$

Thus, the confrontation of the *Ansatz* of Stech with experiment is quite straightforward. We use as input the values of the quark masses[13] and of $|V_{us}|$ (Ref. 14), together with the assumed signs for the u_i and d_j, and then use Eqs. (8)–(10) to derive all the other moduli from that input. We try to fit those moduli into their experimental ranges.[14] We should moreover beware that the output moduli be consistent with the unitarity of V. If they are not, that simply means that the *Ansatz* of Stech is unable to sustain the particular input values taken for the quark masses, and their signs, and $|V_{12}|$. Thus, we have first of all to check that all the output U_{ij} are non-negative and moreover that[9]

$$4U_{12}U_{13}U_{22}U_{23}\geq(1-U_{12}-U_{13}-U_{22}-U_{23}+U_{12}U_{23}+U_{13}U_{22})^2 . \tag{11}$$

Let us now consider Eqs. (7) more carefully. Because of the strong mass hierarchy in both the up and down sectors, we may write

$$d_2U_{12}+d_3U_{13}\approx -d_1 , \tag{12a}$$

$$u_3(-d_2U_{21}+d_3U_{23})\approx u_2d_3-u_3d_2 . \tag{12b}$$

Take Eq. (12a). If U_{13} is small enough to agree with experiment, then $d_2U_{12}\gg d_3U_{13}$, and therefore $U_{12}\approx -d_1/d_2$ will hold. Thus we learn that, first, the ratio d_1/d_2 must be negative if the *Ansatz* of Stech is to work, and second, the *Ansatz* of Stech yields $|V_{us}|\approx\sqrt{m_d/m_s}$ provided Cabibbo mixing holds approximately.[15] We might try to reverse this argument and argue that, given the known values of $|V_{us}|$ and of m_d/m_s, the *Ansatz* of Stech predicts $|V_{ub}|$ to be very small. This is true, but we easily find that the *Ansatz* of Stech cannot predict that U_{13} takes such a small value as indeed it does. The bound of the AS on $|V_{ub}|$ is weaker than the experimental bound $U_{13}\leq 0.0001$. This is to be contrasted with the *Ansatz* of Fritzsch, which gives a bound on U_{13} stronger than the experimental one.

Now for Eq. (12b). Experiment tells us that both terms in its left-hand side (LHS) are much smaller than the second term of the RHS. Therefore, the two terms in the RHS must approximately cancel. Two conclusions may be drawn from that fact. First, if u_3 and d_3 are, without loss of generality, taken to be both positive, then the signs of u_2 and of d_2 must be equal; second, the top-quark mass should satisfy $m_t\approx m_cm_b/m_s$.

Thus, if we take $u_3=m_t>0$ and $d_3=m_b>0$, the AS can only work if we make one of the following four choices for the signs of the other quark masses:

$$u_2>0,\quad d_2>0,\quad u_1>0,\quad d_1<0 , \tag{13a}$$

$$u_2>0,\quad d_2>0,\quad u_1<0,\quad d_1<0 , \tag{13b}$$

$$u_2<0,\quad d_2<0,\quad u_1>0,\quad d_1>0 , \tag{13c}$$

$$u_2<0,\quad d_2<0,\quad u_1<0,\quad d_1>0 . \tag{13d}$$

The possibility (13c) coincides with the quark-mass signs in the *Ansatz* of Fritzsch and is, for some mysterious reason, the only one that is considered in most of the literature (the notable exception being the original paper of Stech). Indeed, it is easily checked that the *Ansatz* of Stech works perfectly well with any other of the choices in (13).

What are the predictions of the AS? I use the current quark mass values[13]

$$d_3=5.3\pm0.1 \text{ GeV}, \quad |u_2|=1.35\pm0.05 \text{ GeV} ,$$
$$\left|\frac{d_2}{d_1}\right|=19.6\pm1.6, \quad \left|\frac{d_1}{u_1}\right|=1.76\pm0.13 , \tag{14}$$
$$|d_1|=0.0089+0.0026 \text{ GeV} ,$$

together with the information on U_{ij} (Ref. 14):

$$U_{12}=0.0484\pm0.0009, \quad U_{23}=0.0024\pm0.0005 ,$$
$$U_{13}\leq 0.0001 . \tag{15}$$

Using the exact equations (8)–(10), I find that the *Ansatz* of Stech requires m_t to be less than 85 GeV. This value for the current top-quark mass at $\mu=1$ GeV corresponds to a physical top-quark mass of about 50 GeV. Such a low value for m_t is by now experimentally excluded.[4] Moreover, in the AS a high m_t is strongly correlated with a low m_s. As I emphasized before, $|V_{ub}|$ is not predicted. There are, indeed, no further predictions of the AS: though its output depends strongly on the input ratio $|d_2/d_1|$, this dependence works differently for different signs of the quark masses. It is, therefore, fair to say that the only prediction of the AS is the low mass of the top quark. Because that very clear-cut prediction failed, the AS is now eliminated.

The relation $m_t \approx m_b m_c/m_s$ also holds in the *Ansatz* of Fritzsch,[2] though in that case the freedom in the choice of m_t is much larger (m_t may be as high as 115 GeV in that *Ansatz)*. The final reason for that relation is that in both *Ansätze* the natural prediction for $|V_{cb}|$ is $|V_{cb}|\approx\sqrt{m_s/m_b}$ (Ref. 16). The fact that $|V_{cb}|$ is much smaller than $\sqrt{m_s/m_b}$ requires that in both *Ansätze* there be an approximate cancellation $u_3 d_2 \approx u_2 d_3$. If m_t turns out to be larger than 115 GeV, we should also discard the *Ansatz* of Fritzsch and look for *Ansätze* where $|V_{cb}|$ is naturally very small.[16]

I want to emphasize the important role played by Eq. (4), which is another common feature of the *Ansätze* of Fritzsch and Stech.[10] This relation was first studied in Ref. 17, where it was shown that, for the usual signs of the quark masses (13c), it leads to "maximal *CP* violation." Here, maximal *CP* violation has the following well-defined sense:[18] for fixed values of $|V_{us}|$, $|V_{cb}|$, and $|V_{ub}|$, the value of $|V_{cs}|$ is such that the rephasing-invariant source of *CP* violation,[19] $|\delta_{\mathrm{KM}}|$, is maximized. Thus, it is because Eq. (4) holds in both the *Ansätze* of Fritzsch and Stech that both of them present an approximate maximal *CP* violation. However, as was also noticed in Ref. 17, Eq. (4) may also apply for other choices of the signs of the quark masses, and then it will not lead to maximal *CP* violation. This will be the case, in particular, if all the quark masses are taken with the same sign (see below). Thus, Eq. (4) has no interesting predictive power unless it is supplemented with some particular assignment for the signs of the eigenvalues of M_D and M_U.

ONE VARIATION OF THE *Ansatz* OF STECH

As I emphasized before, the main ingredient of the *Ansatz* of Stech is the approximate proportionality of M_U and M_D. Another important ingredient is the interplay between a symmetric and an antisymmetric matrix in M_U and M_D; that interplay is the reason why the AS has less parameters than the number of physical quantities which it tries to fit, and therefore the reason why it has some predictive power. Now, we may try to use these two ingredients in a different way from the one in which Stech did it. That is what I will do next, in one straightforward variation of the AS.

In that variation I exchange the roles of M_U and M_D in the AS. I postulate that in some weak basis M_D is real and symmetric, while $M_U=pM_D+A$, p being a real number and A a Hermitian and antisymmetric matrix. Let us investigate whether this *Ansatz* has any chances of being able to fit the data. To obtain the relevant equations we take the equations for the *Ansatz* of Stech, and make the changes $u_i \leftrightarrow d_i$ and $U_{ij}\rightarrow U_{ji}$. We find that the analogues of Eqs. (12) are

$$d_3(u_2 U_{21}+u_3 U_{31})\approx u_3 d_1 , \tag{16a}$$
$$d_3(-u_2 U_{12}+u_3 U_{32})\approx u_3 d_2-u_2 d_3 . \tag{16b}$$

In Eq. (16b), the smallness of U_{12} and the fact that $U_{32} \ll |d_2/d_3|$ are such that the two terms in the RHS should cancel almost completely. Then, inserting $u_3 d_2 \approx u_2 d_3$ into Eq. (16a), we find $U_{21}\approx d_1/d_2$. We thus conclude that if, without lack of generality, we set u_3 and d_3 positive, this *Ansatz* may work if the signs of u_2, d_2, and d_1 are all the same. Moreover, the top-quark mass will, just as in the AS, obey $m_t\approx m_b m_c/m_s$.

To go beyond these guidelines, we must take the exact equations, which are found from Eqs. (8)–(10) by applying the transformations $u_i \leftrightarrow d_i$ and $U_{ij}\rightarrow U_{ji}$. We then readily find out that this *Ansatz* usually yields a value for $|V_{ub}|$ too high. However, this problem can be solved in the case in which all the quark masses have the same sign, say, they are all positive, by choosing a small enough $|V_{cb}|$. Thus, the "inverted AS" works perfectly well when all the quark masses are taken positive, and it predicts a value for $|V_{cb}|$ near the experimental lower bound 0.0019, a high value for $|V_{ub}|/|V_{cb}|$, and also a very low value for $|V_{td}|$, which in this *Ansatz* is lower than $|V_{ub}|$. However, this *Ansatz,* just as the AS, is eliminated because of its prediction $m_t\approx m_b m_c/m_s$. Also notice that this *Ansatz* does not exhibit "maximal *CP* violation," though Eq. (4) still applies.

A GENERALIZATION OF THE AS

We saw that the *inverted Ansatz of Stech* works almost as well as the AS itself. On the other hand, both the AS and its inverted counterpart have a disagreeable asymmetry in their treatment of M_U and M_D. It is therefore natural to try to write down an *Ansatz* which, on the one hand, bridges the gap between the AS and its counterpart with M_U and M_D interchanged, and, on the other hand, is symmetric under $M_U \leftrightarrow M_D$. That is what I will do next. Of course, we should not expect the new *Ansatz* to give much better results than the AS itself in what concerns the crucial matter of the low mass of the top quark; this is because the other limiting case of the new *Ansatz*, the inverted AS, also has problems with that point. However, it turns out that the *Ansatz* that I will now suggest may be applied in a slightly different context, then yielding quite interesting results.

I postulate that in some weak basis[20]

$$M_U = S + qA \ , \tag{17a}$$

$$M_D = pS + A \ , \tag{17b}$$

where S is a real symmetric matrix, A is a Hermitian and antisymmetric matrix, and the numbers p and q are real. It is clear that when $q=0$ this *Ansatz* is identical to the AS, and when $q \to \infty$ this *Ansatz* becomes identical to the inverted AS. The roles of M_U and M_D in Eqs. (17) are clearly symmetric; indeed,

$$M_U \leftrightarrow M_D \Longrightarrow p \to \frac{1}{p}, \quad q \to \frac{1}{q} \ . \tag{18}$$

The price to be paid for this symmetry between M_U and M_D is the extra parameter q, and the ensuing smaller predictive power of the *Ansatz*.

I use the method of the mass-matrix invariants to work out the predictions of this *Ansatz*. I first calculate the traces of M_U, M_U^2, M_U^3, M_D, M_D^2, M_D^3, and $(M_U M_D)$. The resulting equations are easily inverted to give

$$p = \frac{\mathrm{tr} M_D}{\mathrm{tr} M_U} \ , \tag{19a}$$

$$q = -\frac{\mathrm{tr} M_U \mathrm{tr}(M_U M_D) - \mathrm{tr} M_U^2 \mathrm{tr} M_D}{\mathrm{tr} M_D \mathrm{tr}(M_U M_D) - \mathrm{tr} M_D^2 \mathrm{tr} M_U} \ , \tag{19b}$$

together with other equations which give the traces of S, S^2, S^3, A^2, and (SA^2) as functions of the physical quantities. Notice that Eqs. (19) satisfy the condition (18). If we now calculate the trace of $(M_U^2 M_D^2)$, then we will also be able to write down the trace of A^4 as a function of the physical quantities. Up to this point, we have found no predictions.

However, in this *Ansatz*, the traces of $(M_U^2 M_D)$ and of $(M_U M_D^2)$ are not independent from the other ones. Indeed, we easily find that

$$\mathrm{tr}(M_U^2 M_D) = \left[\mathrm{tr} M_U + \frac{F}{\mathrm{tr} M_D} \right] \mathrm{tr}(M_U M_D) + L_U \ , \tag{20a}$$

$$\mathrm{tr}(M_U M_D^2) = \left[\mathrm{tr} M_D + \frac{F}{\mathrm{tr} M_U} \right] \mathrm{tr}(M_U M_D) + L_D \ , \tag{20b}$$

where F, L_U, and L_D are functions of the quark masses and their signs only:

$$F \equiv \frac{\mathrm{tr}^3 M_U \det M_D - \mathrm{tr}^3 M_D \det M_U}{\mathrm{tr} M_D^2 \mathrm{csi} M_U - \mathrm{tr} M_U^2 \mathrm{csi} M_D} \ , \tag{21}$$

$$L_U \equiv \mathrm{tr} M_D \left[-\mathrm{csi} M_U + 3 \frac{\det M_U}{\mathrm{tr} M_U} \right] - F \frac{\mathrm{tr} M_U^2}{\mathrm{tr} M_U} \ , \tag{22a}$$

$$L_D = L_U (M_U \leftrightarrow M_D) \ . \tag{22b}$$

Notice that Eqs. (20), which are the constraints that this *Ansatz* enforces on the physical quantities, are symmetric under $M_U \leftrightarrow M_D$, as they should. Moreover, we can easily verify that they are also satisfied in the AS. Finally, as a consequence of Eqs. (20), Eq. (5) also holds in this *Ansatz*.

The crucial point is the fact that Eqs. (20) are linear equations among the "mixed traces" $\mathrm{tr}(M_U^a M_D^b)$ with a and b different from zero; this implies, via Eq. (6), that this *Ansatz* enforces equations among the physical quantities which are linear in the U_{ij}, and therefore can easily be solved with respect to those quantities. Thus, just as in the AS all the U_{ij} can be written as linear functions of one of them, the coefficients of those functions being functions of the quark masses; similarly in this *Ansatz* we may write down all the U_{ij} as linear functions of, e.g., U_{13} and U_{23}, the coefficients being functions of the quark masses (and their assumed signs).

If we feed those exact equations to a computer and use it to find out the predictions of this *Ansatz*, we get what we expected: this *Ansatz* yields a low value for the top-quark mass. It does not fare better than the AS in that respect. Moreover, the quark-mass sign ambiguity is worse in this *Ansatz* than in the AS: when $d_2/d_1 < 0$ this *Ansatz* approaches the AS, and works well for low values of q; when $d_2/d_1 > 0$ the parameter q will be large and we will have similar results to the ones of the inverted AS. From these points of view, the new *Ansatz* is worth being discarded. However, we will see nest that this *Ansatz*, when applied in a different way, takes us far beyond the AS.

AN *Ansatz* FOR THE MATRICES H

I have emphasized the existence of ambiguities in the AS related to the signs of the eigenvalues of M_U and M_D. It is well known that in the standard model not all the information contained in M_U and M_D is physically meaningful. Indeed, at the classical level of the theory (and thus forgetting about quantum effects such as strong *CP* violation[21]), all the physically relevant information in M_U and M_D is also present in the "Hermitian mass matrices" $H_{U,D} \equiv M_{U,D} M_{U,D}^\dagger$. Contrary to the matrices $M_{U,D}$ which in general are not Hermitian, the matrices $H_{U,D}$ are, by definition, Hermitian. Moreover, in a weak basis in which M_U and M_D are Hermitian, the signs of their eigenvalues are in general arbitrary; on the other hand, the eigenvalues of H_U and H_D, which are the squared quark masses, do not contain any sign ambiguity. This is only an instance of the fact that the matrices M_U and M_D contain some unphysical information which is no longer present in the matrices H_U and H_D, which, on the other hand, contain all the physical information in M_U and

M_D, about the quark masses and the CKM matrix. Though the matrices M_U and M_D are more fundamental, the matrices $H_{U,D}$ are closer to the physical information.

Surprisingly enough, it turns out that the *Ansatz* of Eqs. (17) is successful; i.e., it fits well the experimental data, if it is taken as an *Ansatz* not for the matrices M_U and M_D, but for the matrices H_U and H_D. The *Ansatz* then reads in the following way: there is a weak basis in which the symmetric and antisymmetric parts of the matrices H_U and H_D are separately proportional to each other. This double proportionality implies equations among the physical quantities which are the analogous of Eqs. (20)–(22), with the matrices M_U and M_D substituted by the matrices H_U and H_D. Those equations, contrary to the original Eqs. (20)–(22), do not contain any ambiguity in the signs of the quark masses. Indeed, to develop those equations into sum rules for the U_{ij} we should now use, instead of Eq. (6),

$$\operatorname{tr}(H_U^a H_D^b) = \sum_{i,j-1}^{3} u_i^{2a} d_j^{2b} U_{ij} , \tag{23}$$

in which there are no unphysical quantities or signs present. This is a very important advantage of this *Ansatz* over the AS.

The fact that we have an *Ansatz* for the matrices H which works well is quite remarkable. Remember that the idea of the *Ansätze* for the mass matrices followed from the observation, due to Weinberg,[22] that

$$U_{12} \equiv |V_{us}|^2 \approx \frac{m_d}{m_s} . \tag{24}$$

Indeed, the approximate equality (24) holds in both the *Ansätze* of Fritzsch and of Stech. The idea following from (24) was that maybe all the U_{ij} are some functions of the quark mass ratios alone, i.e., without the presence of any nonphysical parameters. However, if we work with the matrices H instead of with the matrices M, the quark mass ratios are substituted by the squared quark mass ratios. And, because of the strong mass hierarchy, these squared ratios are much smaller than the ratios themselves, in such a way that it becomes difficult to understand how a quantity as large as U_{12} might be related to the squared quark mass ratios. This apparent paradox disappears if we take into account that we are now working with three generations, and not two, as in Weinberg's original work;[22] and that our *Ansätze*, insofar as they parametrize the mass matrices (M or H) by more than six parameters, will in general not really determine the U_{ij} completely (i.e., without the presence of any nonphysical quantities in the corresponding equations) as functions of the quark mass ratios, rather they will only yield some sum rules relating the different U_{ij} among themselves and with the quark masses. Thus, there is no reason why a simple *Ansatz* for H_U and H_D should not work.

At present, the idea of an *Ansatz* for the matrices H_U and H_D is difficult to justify. This is because it is the matrices M which follow directly from the fundamental theory,[8] i.e., from the Yukawa coupling matrices. Moreover, an *Ansatz* which is not enforced by means of some symmetries of the underlying theory will as a consequence be unstable under the radiative corrections, and therefore meaningless. Clearly, it is difficult to find symmetries which give rise to *Ansätze* which are simpler when written in terms of H_u and H_D than when written in terms of M_U and M_D. Anyway, the fact that the weak-basis arbitrariness problem is milder in the *Ansätze* for the matrices H renders those *Ansätze* rather appealing, and this offsets, in my opinion, the fact that the naturalness problem is more difficult to solve in those *Ansätze*. An *Ansatz* for the matrices H should at present be considered to have a more phenomenological character than an *Ansatz* for the matrices M, just as the matrices H are closer to the experimentally measurable quantities, while the matrices M are closer to the underlying theory.

Because of Eq. (12a) the AS does not work if d_1, d_2, and d_3 are all positive. Thus, the AS could never be transformed into an *Ansatz* for the matrices H. The same cannot be said of the inverted AS which, as I have pointed out, works well if all the eigenvalues of M_U and M_D are taken to be positive. However, it turns out that the inverted AS too is unable to fit the data if it is taken as an *Ansatz* for the matrices $H_{U,D}$. The generalization of the AS which I worked out in the previous section has just enough flexibility to do the job. On the other hand, it is easily found that, whenever that *Ansatz* for H_U and H_D fits well the data, the q value of Eq. (19b) (with $M_{U,D} \rightarrow H_{U,D}$) is of order 10^3, i.e., quite high, showing that we are not far from the inverted AS.

The *Ansatz* for the matrices $H_{U,D}$ that I am suggesting has one further attractive feature: it easily allows the top-quark mass to be very high. This is very different from what happened in the corresponding *Ansatz* for the matrices $M_{U,D}$. Feeding the exact equations to the computer one easily finds that, if U_{13} and U_{23} are taken, together with the quark masses, as input quantities, then there are two ranges of values of u_3 for which the output U_{12} agrees with the experimentally measured value. In one of these ranges we have, just as in the AS, $m_t \approx m_b m_c / m_s$. In the other range, m_t is very high. From now on, I will only consider the *Ansatz* with m_t taken in that interesting range of high values.

We easily find that, provided we allow U_{23} to be near to its experimental lower bound 0.0019, and U_{13} to be near to its upper bound 0.0001, then the top-quark mass (renormalized at the scale $\mu = 1$ GeV) is about 175 GeV, which corresponds to a physical m_t about 105 GeV. This value is obtained for the central values[13] of the masses of all the other quarks; if we stretch the values of the quark masses to the one-standard deviation limits given in Ref. 13, then we easily obtain a top-quark mass as high as 160 GeV. Such a high value for m_t is beyond the reach of all the previous *Ansätze* for the mass matrices; for instance, the *Ansatz* of Fritzsch yields, for the same one-standard deviation quark masses, $m_t < 115$ GeV.

This high top-quark mass also suggests that this *Ansatz* may suffer important corrections if it is postulated to hold at some high-energy scale, say, the grand-unified-theory (GUT) scale, and then the mass matrices are allowed to evolve to the Fermi scale following the renormalization-group equations for the Yukawa couplings.[23]

Let me finally state briefly the predictions of this *Ansatz* for the matrices H_U and H_D. As I said before, I only take into account the branch of the *Ansatz* which fits U_{12} by means of a high top-quark mass. The predictions are then the following: (1) the top-quark mass may be as high as 160 GeV, and its "central value" is of about 110 GeV; (2) $|V_{cb}|$ is rather low: $U_{23} \equiv |V_{cb}|^2 < 0.0022$; (3) $|V_{ub}|$ is rather high: $U_{13} \equiv |V_{ub}|^2 > 0.00008$; (4) $|V_{td}|$ is extremely close to zero.

As a consequence of the last prediction, though m_t is very high, in this *Ansatz* the high B_d-B_d mixing[24] cannot be explained by the top-quark box diagrams alone; we must assume the presence of some extra virtual particles in the B_d-B_d transition diagrams. This should not, however, be considered as a great handicap for this *Ansatz,* nor for any other one. Indeed, the *Ansätze* are unstable under quantum corrections if they are not protected by some symmetry of the underlying theory (if they are, they are not *Ansätze* anymore, but models). Any such more complete theory will most likely contain extra particles and couplings, and for that reason it is not, in general, a sound procedure to use the results of the calculation of the standard-model loop diagrams alone to derive constraints on the *Ansätze.*

CONCLUSIONS

I have shown that the predictions of the *Ansatz* of Stech can be easily worked out by using the method of the mass-matrix invariants. These predictions can be written down in the form of three linear equations on the squared moduli of the CKM matrix elements, the coefficients of those equations being well-defined functions of the quark masses and their signs. I emphasized the existence of ambiguities in the AS, and that those ambiguities are responsible for the fact that that *Ansatz* is only able to predict a low mass for the top quark.

I showed that the two main features of the AS, the interplay of a symmetric and an antisymmetric matrix, and the approximate proportionality of M_U and M_D, can be implemented in other equally successful *Ansätze.* The simplest of these consists simply in interchanging the roles played by M_U and M_D in the AS.

I built and analyzed an extension of the AS which is symmetric when M_U and M_D are interchanged. That *Ansatz* does not have any advantage over the AS when it is taken to be an *Ansatz* for M_U and M_D. However, it may be successfully taken as an *Ansatz* for H_U and H_D. It is indeed the first *Ansatz* suggested directly for those matrices. That *Ansatz* for H_U and H_D allows the top-quark mass to be very high, $m_t < 160$ GeV, and predicts $|V_{td}|$ to be almost vanishing, $|V_{cb}|$ to be close to its experimental lower bound, and a rather high ratio $|V_{ub}|/|V_{cb}|$.

ACKNOWLEDGMENTS

I thank Professor Walter Grimus for reading and criticizing the manuscript. This work benefited greatly from early collaboration on related subjects, and various conservations, with Professor Gustavo C. Branco.

---

*Present address: Institut für Physik, Universität Dortmund, D-4600 Dortmund 50, Federal Republic of Germany.
[1]B. Stech, Phys. Lett. **130B**, 189 (1983).
[2]H. Fritzsch, Phys. Lett. **73B**, 317 (1978); Nucl. Phys. **B155**, 189 (1979); L. F. Li, Phys. Lett. **84B**, 461 (1979).
[3]M. Gronau, R. Johnson, and J. Schechter, Phys. Rev. Lett. **54**, 2176 (1985).
[4]J. Yoh, recent results of the Collider Detector at Fermilab (CDF) Collaboration, talk presented at the 1989 DESY Theory Workshop (unpublished). The bound is $m_t > 72$ GeV, 95% C.L.
[5]H. Fritzsch, Report No. MPI-PAE/PTh 26/89, 1989 (unpublished).
[6]N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963); S. L. Glashow, J. Iliopoulos, and L. Maiani, Phys. Rev. D **2**, 1285 (1970); M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).
[7]H. Harari, H. Haut, and J. Weyers, Phys. Lett. **78B**, 459 (1978).
[8]S. L. Glashow, Nucl. Phys. **22**, 579 (1961); S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967); A. Salam, in *Elementary Particle Theory: Relativistic Groups and Analyticity (Nobel Symposium No. 8),* edited by N. Svartholm (Almqvist and Wiksell, Stockholm, 1968).
[9]G. C. Branco and L. Lavoura, Phys. Lett. B **208**, 123 (1988).
[10]G. C. Branco and D. D. Wu, Phys. Lett. B **205**, 353 (1988).
[11]C. Jarlskog and A. Kleppe, Nucl. Phys. **B286**, 245 (1987); C. Jarlskog, Phys. Rev. D **35**, 1685 (1987).
[12]L. Lavoura, Phys. Lett. B **228**, 245 (1989).
[13]J. Gasser and H. Leutwyler, Phys. Rep. **87**, 77 (1982).
[14]K. R. Schubert, in *Proceedings of the 1988 Summer School in High Energy Physics and Cosmology,* edited by J. C. Pati and Q. Shafi (World Scientific, Singapore, 1989). In this paper I do not take into account the recent analysis of the ARGUS Collaboration which indicates that $|V_{ub}|$ is nonzero, but is much smaller than the previous upper bound. See M. Danilov, in *Proceedings of the XIV International Symposium on Lepton and Photon Interactions,* Stanford, California, edited by M. Riordan (World Scientific, Singapore, in press). See also CLEO Collaboration, R. Fulton *et al.*, Report No. CLNS-89/951 (unpublished).
[15]G. Ecker, Z. Phys. C **24**, 353 (1984).
[16]X. G. He and W. S. Hou, Phys. Rev. D **41**, 1517 (1990).
[17]M. Gronau, V. Gupta, R. Johnson, and J. Schechter, Phys. Rev. D **33**, 3368 (1986).
[18]M. Gronau and J. Schechter, Phys. Rev. Lett. **54**, 385 (1985); **54**, 1209(E) (1985).
[19]C. Jarlskog, Phys. Rev. Lett. **55**, 1039 (1985); I. Dunietz, O. W. Greenberg, and D. D. Wu, Phys. Rev. Lett. **55**, 2935 (1985).
[20]A model which may yield this *Ansatz* was proposed by G. Ecker, W. Grimus, and H. Neufeld, Nucl. Phys. **B247**, 70 (1984).

[21]For a review, see R. Peccei, in *CP Violation,* edited by C. Jarlskog (World Scientific, Singapore, 1989).
[22]S. Weinberg, Trans. N.Y. Acad. Sci. Series II, **38** (1977).
[23]K. Sasaki, Z. Phys. C **32**, 149 (1986); B. Grzadkowski and M. Linder, Phys. Lett. B **193**, 71 (1987); C. H. Albright and M. Lindner, *ibid.* **213**, 347 (1988).
[24]H. Albrecht *et al.*, Phys. Lett. B **192**, 245 (1987); M. Artuso *et al.*, Phys. Rev. Lett. **62**, 2233 (1989).