

SLAC – PUB – 3664

May 1985

T/E

THE ELECTROWEAK LOOP EFFECTS:
WHAT ACCURACY OF EXPERIMENT
IS NEEDED TO TEST THEM ?*

KEN-ICHI AOKI[†] AND HIDEAKI AOYAMA

*Stanford Linear Accelerator Center
Stanford University, Stanford, California, 94305*

ABSTRACT

The necessary precision that future high-energy experiments have to realize in order to establish loop effects in the electroweak theory is discussed. The recent claim that the mass difference $M_Z - M_W$ is effective for analysis is criticized. The case with three observables, M_Z , M_W and the Weinberg angle, is analyzed by the use of a generalized condition for the allowed errors. It is found that the information on the Weinberg angle with even 2.5%(5%) error improves the requirement for the errors of M_Z and M_W significantly in the $2\sigma(1\sigma)$ -confidence test of the loop effects.

Submitted to Nuclear Physics B

* Work supported by the Department of Energy, contract DE – AC03 – 76SF00515.

† On leave from Research Institute for Fundamental Physics, Kyoto University, Kyoto 606, Japan

1. Introduction

The Glashow-Weinberg-Salam theory^[1] has been accepted as the electroweak theory, especially after the remarkable discovery of the weak bosons.^[2] Many calculations of higher order effects in the electroweak theory have been made. One of the most interesting subjects in this field is the experimental confirmation of loop effects, since this is indispensable for establishing the electroweak theory as a quantum field theory. However, almost all the data in the low energy regime (compared to the weak scale) can be explained by the tree level electroweak theory. In fact, loop corrections to various cross sections and other observables at low energies are negligible compared to experimental errors.^[3]

In contrast, direct observations of the weak gauge bosons provide a good chance for confirming the loop effects. For this reason, many studies on the gauge boson masses, M_Z and M_W , have been carried out.^[4-6] As a result, the conditions on the experimental precisions of M_Z , M_W measurements necessary for detection of the loop effects have been given, and it has been shown that the confirmation is possible in a realistic future.^[7,8] Recently Grzadkowski et. al.^[9] have argued that, in analyzing the mass difference ($M_Z - M_W$), less stringent conditions apply. Their argument, however, can be shown not to be applicable to usual experimental situation. In this paper we clarify this point and further determine the necessary experimental precision for testing loop effects for the three observables, M_Z , M_W and $\sin^2 \theta_W$.

It is important to specify the property of the conditions we can obtain. We have two predictions; the zeroth order prediction and the prediction including the loop effects. We want to establish criteria by which one can say that the higher order calculation is a significantly better fit to the data than the leading order. This is achieved when, at a given confidence level, the experimental data is in agreement with at most one of the predictions. Thus we are interested in the level of accuracy required in order that this must necessarily so. As long as this condition is satisfied, the experiments provide useful information: It is of course

still interesting even if both predictions are rejected, since that would indicate the need for new ingredients in the theory (e.g. superpartners^[10,11]).

The electroweak theory has three basic parameters, which are to be determined by experimental data. (We neglect parameters in fermion and Higgs sectors, since they typically have relatively small effects.^[12] One exception is heavy-fermions' contribution, which may become large due to the corresponding large Yukawa coupling constants.^[13] This is rather a separate issue and will not be discussed in this paper.) We have two data with quite high accuracy, the fine structure constant α and the muon life time τ_μ . We adopt these two quantities as input and hereafter regard them as definite parameters with negligible errors. Accordingly one parameter remains to be determined. Therefore at least two measurements are required to further test the theory.

The observables we consider first are the weak-boson masses, M_Z and M_W . Due to the remaining free parameter, the theory can predict only the inter-relation between M_Z and M_W . At the tree level it is given as

$$M_W^{(0)} = M_Z \sqrt{\frac{1}{2} \left(1 + \sqrt{1 - \frac{2\sqrt{2}\pi\alpha}{M_Z^2 G_F}} \right)} \quad (1.1)$$

where G_F is the Fermi coupling constant determined by

$$\tau_\mu^{-1} = \frac{G_F^2 m_\mu^5}{192\pi^3} \left(1 - \frac{8m_e^2}{m_\mu^2} \right). \quad (1.2)$$

The higher order corrections are dominated by the leading logarithms. Summing the leading logarithm leads to the following relation,^[6]

$$M_W = M_W^{(0)} \left(1 + \frac{M_Z^2 - M_W^{(0)2}}{2(2M_W^{(0)2} - M_Z^2)} \left(1 - \frac{\alpha(M_W^{(0)})}{\alpha} \right) \right), \quad (1.3)$$

where $\alpha(\mu)$ is the running QED coupling constant defined by

$$\alpha(\mu) = \alpha \left\{ 1 + \frac{2\alpha}{3\pi} \sum_f Q_f^2 \ln \left(\frac{m_f}{\mu} \right) \right\}^{-1}. \quad (1.4)$$

For the full result of the one-loop plus leading-log, see e.g. Ref. 3.

Illustrated in Fig. 1 are the tree relation (1.1) and the leading-log (1.3) plus one-loop relation. Recent experimental results^[2,8] on M_Z and M_W (in GeV) are

$$\begin{aligned} \text{UA1} : M_Z &= 95.6 \pm 1.5 \pm 2.9, & M_W &= 80.9 \pm 1.5 \pm 2.4, \\ \text{UA2} : M_Z &= 92.7 \pm 1.7 \pm 1.4, & M_W &= 83.1 \pm 1.9 \pm 1.3. \end{aligned} \tag{1.5}$$

The current errors are rather large for the purpose of confirmation of higher order effects. For illustration of this situation, a circle is drawn in Fig. 1 for errors of $\Delta M_Z = \Delta M_W = 2 \text{ GeV}$. It is evident that the experimental precision has to be greatly improved. By adding other observables in the analysis, the precision required for M_Z and M_W can be reduced. Low energy data give some information on the Weinberg angle,^[5] which limits the allowed region to be intervals in both lines as seen in Fig. 1.

This paper is organized as follows. In section 2, we investigate the case where M_Z and M_W are the only data to be observed. We follow and criticize the argument of Ref. [9] and find a correct condition. We further show that, by analyzing the data in (M_Z, M_W) -plane, two kinds of conditions are obtained. The one obtained before is the weakest of them. Section 3 is devoted to numerical analyses with three-observables (M_Z , M_W and the Weinberg angle) by the use of general formulas. We examine how much improvement (over the condition given in section 2) can be obtained by using this additional input. The appendix A gives a brief description of the multi-variate confidence region. In appendix B we derive N-kinds of conditions suitable for the case with N-observables, which is used in section 3. The contents similar to the appendices may be found in appropriate mathematical literatures, but is presented here for completeness.

2. Conditions on M_Z , M_W Measurements

The electroweak theory (either at tree level or including loop corrections) gives a relation between the W-boson mass M_W and the Z-boson mass M_Z . Consequently, there are two ways to analyze the conditions on allowed errors. One is to project all data on one of the masses (or a particular linear combination of them), the other is an analysis on the (M_Z, M_W) -plane. In this section we will discuss both methods.

2.1 ERROR ANALYSIS

Consider predicting M_W by the use of an experimental value of M_Z (using a formula at the tree level or including loop effects). The error in this “predicted” value of M_W due to an error ΔM_Z ($\equiv M_Z^{(\text{true})} - M_Z^{(\text{obs.})}$) in M_Z measurement is

$$\frac{\partial M_W}{\partial M_Z} \Delta M_Z (\equiv a \Delta M_Z), \quad (2.1)$$

within the first order approximation. The slope a may be considered to be common for two predictions, since the difference is effectively of higher orders in the approximation we adopted in Eq.(2.1). The measurement of M_W also suffers from an error ΔM_W . The error of the observed M_W relative to the predicted value M_W ($M_Z^{(\text{obs.})}$) is then

$$\Delta = a \Delta M_Z - \Delta M_W. \quad (2.2)$$

We assume that the errors ΔM_Z and ΔM_W are regarded as statistical variables, for which the normal distributions have variances σ_Z^2 and σ_W^2 respectively. In cases when ΔM_Z and ΔM_W are independent, the variance of the statistical variable Δ is

$$\sigma_{ZW}^2 = a^2 \sigma_Z^2 + \sigma_W^2, \quad (2.3)$$

Denoting by d_W the difference between the tree-level and loop-corrected predic-

tions of M_W , a criterion for allowed experimental errors is given by the following,

$$d_W > 4\sigma_{ZW}. \quad (2.4)$$

This enables the experiment to reject (at least) one of the predictions at the 95%(2σ)-confidence level, since this assures that there is no region of M_W simultaneously belonging to the 2σ -confidence territories of both predictions.*

Grzadkowski et. al. claim that by using an another variable ($M_Z - M_W$) a constraint *weaker* than (2.4) is obtained. However, this statement is rather against intuition, and in fact can be proven to be wrong under the assumption stated above that the measurements of the M_Z and M_W are uncorrelated. Consider using a variable M_c instead of M_W ,

$$M_c \equiv M_W - cM_Z, \quad (2.5)$$

where c is a constant parameter. In the same manner as before, the relative error between the observed M_c and predicted M_c is

$$\Delta_c = (a - c)\Delta M_Z - \Delta M_c. \quad (2.6)$$

Note that ΔM_Z and ΔM_c are *not* statistically independent under the current assumption. In any case, since $\Delta_c = \Delta$, the variance of Δ_c is exactly equal to σ_{ZW}^2 given in (2.3). Since the difference d_c between two predictions of M_c is equal to d_W , one obtains the exactly same condition (2.4). The constraint is independent of the parameter c .

* In [9] the coefficient of the right hand side of Eq. (2.4) is given as 3 instead of 4. This is because it is assumed that the data point is (with 100% probability) located in 1σ region of the loop prediction. However, such an assumption is statistically invalid and the correct 2σ criterion should be Eq.(2.4).

Next let us assume a (quite unlikely) case where M_c is measured independently (i.e. directly). In such a case, the corresponding condition is

$$d_W > 4\sqrt{(a-c)^2\sigma_Z^2 + \sigma_c^2}. \quad (2.7)$$

The $c = 1$ case of this condition is given in Ref. [9]. In such a case, $(a-c)^2$ becomes very small compared to a^2 , since the actual numerical value of a is about 1.19. Accordingly as is demonstrated in Ref. [9], the allowed area in (σ_Z, σ_c) -plane is much wider than that of (σ_Z, σ_W) -plane determined by (2.4). One should note, however, that the above improvement is very superficial and does not immediately mean the superiority of the variable $(M_Z - M_W)$ to M_W : The conditions (2.4) and (2.7) apply to totally different experiments and therefore does not allow direct comparison with each other without dealing with the details of “actual” experimental situations.

A case where the same systematic error contributes to measurements of both M_Z and M_W can also be discussed within the present framework. The errors ΔM_Z and ΔM_W consist of two parts respectively:

$$\Delta M_Z = \Delta M_Z^{\text{st}} + \Delta M^{\text{sys}}, \quad \Delta M_W = \Delta M_W^{\text{st}} + \Delta M^{\text{sys}}. \quad (2.8)$$

Also in this case the relative error in M_c is independent of the parameter c ,

$$\Delta_c = a\Delta M_Z^{\text{st}} - \Delta M_W^{\text{st}} + (a-1)\Delta M^{\text{sys}}, \quad (2.9)$$

and the condition is written as

$$d_W > 4\sqrt{a^2\sigma_Z^2 + \sigma_W^2 + (a-1)^2\sigma_{\text{sys}}^2} \quad (2.10)$$

where σ_{sys}^2 is the variance of ΔM^{sys} .

2.2 ANALYSIS IN THE (M_Z, M_W) -PLANE

It is convenient to introduce a vector notation in the (M_Z, M_W) -plane and denote the errors as $\vec{x} = (\Delta M_Z, \Delta M_W)$. We assume a general covariance matrix $V_{ij} \equiv \langle x_i x_j \rangle$ for these variables. The probability distribution function is then given by the following,

$$P(x)d^2x = \frac{\sqrt{\det K}}{2\pi} \exp\left(-\frac{1}{2}\vec{x}K\vec{x}\right)d^2x, \quad (2.11)$$

where K is the inverse matrix of V and $\vec{x}K\vec{x} \equiv \sum_{i,j} x_i K^{ij} x_j$. In the region we are concerned with, the predictions of the tree level and loop-corrected electroweak theory give parallel straight lines (ℓ_0 and ℓ_1 in Fig. 2).

A condition can be obtained by neglecting the variable along these parallel lines. Namely, we decompose the vector \vec{x} into the direction parallel to the lines ℓ_0 and ℓ_1 (given by a unit vector \vec{u}) and the orthogonal direction (given by a unit vector \vec{v}),

$$\vec{x} = s\vec{u} + t\vec{v}. \quad (2.12)$$

Integrating $P(x)$ over the variable s , we can obtain the ‘‘projected’’ probability distribution function \tilde{P} of t ,

$$\tilde{P}(t)dt = \frac{\sqrt{\det' L}}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\vec{v}L\vec{v}\right)dt, \quad (2.13)$$

where

$$L \equiv K - \frac{K\vec{u} \cdot \vec{u}K}{\vec{u}K\vec{u}}, \quad \det' L \equiv \frac{\det K}{\vec{u}K\vec{u}}. \quad (2.14)$$

The matrix L has a zero eigenvalue since

$$L\vec{u} = \vec{u}L = 0. \quad (2.15)$$

The prime on the determinant means that this zero eigenvalue is excluded from the calculation of the determinant. Since the two predictions are projected to two

points in the \vec{v} -subspace, the condition at $n\sigma$ -confidence level is that the distance between these two points is greater than $2n \times$ [standard deviation calculated from L]. Because of (2.15), we can write this condition using *any* vector \vec{d} (not necessarily parallel to \vec{v}) that runs between the lines ℓ_0 and ℓ_1 as follows,

$$\vec{d}L\vec{d} > 4n^2. \quad (2.16)$$

It should be stressed that this condition is invariant under general linear transformation of the coordinates (i.e. data) $x \rightarrow Ux$, since by definition K is simultaneously transformed as

$$K \rightarrow U^{-1t}KU^{-1}. \quad (2.17)$$

In particular, the previous analysis on the variable M_c corresponds to the analysis in the coordinate system obtained by the transformation

$$U = \begin{pmatrix} 1 & 0 \\ -c & 1 \end{pmatrix}. \quad (2.18)$$

Therefore it is obvious that the parameter c is irrelevant.

Using the gradient a defined in (2.1), the unit vector \vec{u} is given by

$$\vec{u} = \frac{1}{\sqrt{1+a^2}} \begin{pmatrix} 1 \\ a \end{pmatrix}. \quad (2.19)$$

In this case, we find that

$$\begin{aligned} \vec{x}L\vec{x} (= t^2\vec{v}L\vec{v}) &= \det 'L \left(\frac{x_2 - ax_1}{\sqrt{1+a^2}} \right)^2 \\ &= \frac{1}{V_{22} - 2aV_{12} + a^2V_{11}} (x_2 - ax_1)^2. \end{aligned} \quad (2.20)$$

For the errors given by (2.8), the covariance is

$$V = \begin{pmatrix} \sigma_Z^2 + \sigma_{\text{sys}}^2 & \sigma_{\text{sys}}^2 \\ \sigma_{\text{sys}}^2 & \sigma_W^2 + \sigma_{\text{sys}}^2 \end{pmatrix}. \quad (2.21)$$

A distance vector \vec{d} is given by

$$\vec{d} = \begin{pmatrix} 0 \\ d_W \end{pmatrix}. \quad (2.22)$$

Using Eqs. (2.20), (2.21) and (2.22), we find that in this case the condition (2.16) for $n = 2$ reproduces the condition (2.10) exactly.

The variable “ s ” along the direction parallel to the lines ℓ_0 and ℓ_1 contains meaningful information. Actually, it corresponds to the third parameter, for example, the Weinberg angle. The projection discussed above disregards this information. If one is interested in finding the value of s , one needs to use the 2-dimensional $n\sigma$ -confidence region in (M_Z, M_W) -plane, which is an ellipsoidal region defined by

$$\vec{x}K\vec{x} \leq R_{2,n}^2. \quad (2.23)$$

The definition and some values of the constants $R_{2,n}$ are given in the appendix A. If one of the lines ℓ_0 or ℓ_1 passes this region, the (tree or loop-corrected) prediction with the values of s belonging to the region (2.23) is accepted in the test at $n\sigma$ -confidence level. Thus, in order to have only one of the predictions to be valid, the confidence region (2.23) has to be such that the lines ℓ_0 and ℓ_1 cannot pass this region simultaneously. From elementary analysis, we find the width h of this ellipsoidal region (2.23) in the \vec{v} -direction to be

$$h = 2R_{2,n} \sqrt{\frac{\vec{u}K\vec{u}}{\det K}} = 2R_{2,n} \sqrt{\frac{1}{\vec{v}L\vec{v}}}. \quad (2.24)$$

Thus by using (2.15) again we can express the appropriate condition in terms of an arbitrary difference vector \vec{d} as

$$\vec{d}L\vec{d} > 4R_{2,n}^2. \quad (2.25)$$

Note that the left-hand side of the above coincides with that of (2.16). (In the appendix B we prove this to be true in general.) Since $R_{2,n} > n$, (2.25) is a more

stringent condition than (2.16). An experiment has to satisfy this condition (2.25) ultimately to provide a definitive test of the loop effects.

There is an alternative way to derive the condition (2.25). That is to regard the matrix K as a “metric” in the (M_Z, M_W) -plane and obtain the “proper distance” between lines ℓ_0 and ℓ_1 under this metric. The condition is then that the proper distance should be bigger than the “proper diameter” $R_{2,n}$ of the region (2.23). This method gives a simpler derivation of (2.25) than the above elementary method. It is also easily generalizable and is therefore used in the appendix B.

3. Analysis with Additional Observables

The analysis in the last section can be generalized to the situation of a theory with M free parameters when $N (\geq M + 1)$ quantities are to be experimentally measured. As before, the predictions can be judged in a projected subspace. The direction of the projection has to be chosen so that the resulting condition is most effective in discriminating the two predictions. Since the projected subspaces can be of 1 to N dimensional, N kinds of conditions result (at a given confidence level). In appendix B, we examine the above procedure and derive the following condition for the $n\sigma$ -confidence test in k -dimensional optimized subspace,

$$D_{(N)} > 2R_{k,n} \quad (k = 1 \sim N). \quad (B13)$$

The general definition of $D_{(N)}$ (the statistical distance between predictions) is given in Eq. (B3). For $M=1$, (B13) is a simple generalization of (2.16) and (2.25), since $D_{(N)}^2$ is simply equal to (N -dimensional) $\vec{d}L\vec{d}$. The constant $R_{k,n}$ is the radius of the $n\sigma$ -confidence region of the k -dimensional multi-variate normal distribution (see appendix A and Table 1). The test in a smaller subspace requires a weaker condition for errors and thus is more effective than a test in

a bigger subspace (note Eq.(B14)), since it uses less number of independent observables for testing. The weakest condition is given by the most effective test for discriminating two predictions, the $k = 1$ test ($R_{k,n} \geq R_{1,n} = n$).

Now let us return to the electroweak theory and apply the general formula (B13) to evaluate how the conditions on σ_Z and σ_W are loosened by introducing additional variables to the analysis. For practical reason we assume that the errors are independent,

$$V_{ij} \equiv \delta_{ij}\sigma_i^2. \quad (3.1)$$

We denote a vector that gives the direction of the predicted parallel straight lines (ℓ_0 and ℓ_1) by $(a_i)^*$, and a difference vector by (d_i) . Then the condition (B13) is reduced to

$$D_{(N)}^2 = \frac{\sum_{i>j} \frac{\gamma_{ij}^2}{\sigma_i^2 \sigma_j^2}}{\sum_i \frac{a_i^2}{\sigma_i^2}} > 4R_{k,n}^2, \quad (k = 1 \sim N), \quad (3.2)$$

where $\gamma_{ij} \equiv a_i d_j - a_j d_i$. In this expression, it is evident that the component of d parallel to a is irrelevant. The normalization of the vector (a_i) is also seen to be irrelevant. We observe that by dividing (3.2) by $R_{k,n}^2$ the left hand side becomes a function of $R_{k,n}\sigma_i$'s (for fixed a and d). Thus in the following we plot the numerical results for the variables $R_{k,n}\sigma_i$'s. As the third variable we adopt the Weinberg angle defined by^[15]

$$\sin^2 \theta_W = 1 - \frac{M_W^2}{M_Z^2}, \quad (3.3)$$

and analyze the case with three variables (M_Z , M_W and $\sin^2 \theta_W$). Needless to say that the Weinberg angle defined above is to be measured independently, not being calculated from M_Z and M_W , so as to satisfy the assumption (3.1).

* This corresponds to the vector p in the appendix B, but with wrong normalization.

For later convenience, we rescale all the variables x_i 's so that they represent relative errors in the percentage unit at a reference point. The reference point to normalize the variables is chosen as

$$M_Z = 94.00 \text{ Gev}, \quad (3.4)$$

which determines other reference values through tree relations,[†]

$$M_W = 84.28 \text{ Gev}, \quad (3.5)$$

$$\sin^2 \theta_W = 0.196. \quad (3.6)$$

To get a parallel vector, we differentiate the variables (normalized by the above reference values) with respect to M_Z (/1GeV). The results read

$$a = \begin{pmatrix} a_Z \\ a_W \\ a_\theta \end{pmatrix} = \begin{pmatrix} 1.06 \\ 1.41 \\ -2.81 \end{pmatrix}, \quad (3.7)$$

where we have used Eqs.(1.1) and (3.3).

For a difference vector d , we need a result of the loop corrections. Since any difference vector suffices, the corrections evaluated in any renormalization condition can be adopted as a difference vector. We take the renormalization condition in which M_Z is fixed. As for the loop-corrections we include all the leading logarithmic terms^[16] in addition to the pure one-loop terms.^[3] These terms dominate the corrections to the gauge boson masses and also the Weinberg angle defined by (3.3).^[17] We take 83.36 Gev^[6] as the corrected M_W in this

[†] We use the pure tree relations which does not include even the so-called QED corrections to the muon decay width.

case. The resulting difference vector is

$$d = \begin{pmatrix} d_Z \\ d_W \\ d_\theta \end{pmatrix} = \begin{pmatrix} 0.00 \\ -1.09 \\ 8.90 \end{pmatrix}. \quad (3.8)$$

The numerical results are presented in Figs.3~5 by the use of the scaled variables $R_{k,n}\sigma_i$'s. (Subscripts n and k are omitted in the figures.) In the following we mainly refer to the most effective $k = 1$ tests for the confidence levels of $n=2$ (95% confidence) and $n=1$ (68% confidence). It should be noted that in a very lucky situation in which the mean of the experimental data coincides with the one prediction, our conditions (for $n=2$ and $n=1$) effectively correspond to those of two times higher confidence levels ($n=4$ and $n=2$ respectively) to reject the other prediction.

In Fig.3, allowed regions for errors of M_Z , M_W and $\sin^2 \theta_W$ are illustrated. Horizontal sections of Fig.3 are expressed in Fig.4. It is readily seen that with an input of the information of the Weinberg angle, the requirement for errors are loosened. For example, one sees that at a precision of $R_{k,n}\sigma_\theta = 5$ in the Weinberg angle, the allowed errors for M_Z and M_W are drastically increased. This corresponds to 2.5% error of the Weinberg angle measurement for 2σ -test and 5.0% error for 1σ -test. Since the relevant error range of M_Z and M_W are about 0.2% ($n=2$), it is at first surprising that an additional input with an order of magnitude larger error (2.5%, $n=2$) improves the situation. This, however, is explained by the fact that in $(M_Z, M_W \text{ and } \sin^2 \theta_W)$ -space the plane defined by the two parallel lines, ℓ_0 and ℓ_1 , is almost parallel to the $\sin^2 \theta_W$ -axis, whereas the lines ℓ_0 and ℓ_1 themselves are not parallel to any of the axis. In fact, we have

$$\vec{a} \times \vec{d} = \begin{pmatrix} 9.49 \\ -9.43 \\ -1.15 \end{pmatrix}, \quad (3.9)$$

and Eq. (3.7). The vector (3.9) gives the coefficients γ_{ij} in (3.2). Since γ_{12} is smaller than the other components by an order of magnitude and all a_i 's are of

the same order, a σ_θ can contribute to $D_{(N)}$ with an amount comparable to that of ten-times smaller σ_Z and σ_W .

In order to illustrate the magnitudes of improvement more clearly, we take the following two special cases: 1. $\sigma_Z = \sigma_W$, which roughly corresponds to the present experimental situation, 2. $R_{k,n}\sigma_Z = 0$, which gives almost the same results with the $R_{k,n}\sigma_Z \sim 0.2$ ($\sigma_Z \sim 0.1\%$ for $n=2$) case expected in future experiments at Z-factories.^[7,8,11] In both cases we define the improvement factor as the following

$$I(R_{k,n}\sigma_\theta) = \frac{\sigma_W(R_{k,n}\sigma_\theta)}{\sigma_W(\infty)}. \quad (3.10)$$

where $\sigma_W(R_{k,n}\sigma_\theta)$ represents the maximum allowed error for M_W as a function of the error of the Weinberg angle. (Consequently, $\sigma_W(\infty)$ is the maximum allowed error without additional information on the Weinberg angle.) The function $I(R_{k,n}\sigma_\theta)$ is plotted in Fig.5. In case 1 the improvement becomes remarkable if $\sigma_\theta < 2.0\%$ ($\sigma_\theta < 4.0\%$) for $n=2$ ($n=1$) test, where the improvement factor is larger than 2.3. Similar results are obtained in case 2. (The result for $R_{k,n}\sigma_Z \sim 0.2$ is the same with this case within the thickness of the line in Fig. 5.) In this case the error of M_W becomes irrelevant if $\sigma_\theta < 2.2\%$ ($\sigma_\theta < 4.4\%$) in $n=2$ ($n=1$) test. This indicates that in such region the two predictions can be completely discriminated only by the data of M_Z and the Weinberg angle.

In conclusion we have evaluated the requirements for errors of M_Z and M_W which can establish the loop effects in the electroweak theory. By addition of the data on the Weinberg angle whose error is less than 2.5% in $n=2$ test, or equivalently 5.0% in $n=1$ test, the allowed error for M_W is loosened by factor 1.5 (in $\sigma_Z = \sigma_W$ case) to 2.3 (in $\sigma_Z \sim 0\%$ case).

ACKNOWLEDGEMENTS

We are very grateful to Zenrō Hioki for motivating us to start this work and for useful correspondence. We would like to thank Helen Quinn for careful reading of the manuscript and Makoto Sakuda for comments. Thanks are also due to Michael Peskin for discussions on the polarization experiments at *Z*-factories.^[11] One of the authors (K.-I. A.) is supported by the Japan Society for the Promotion of Science.

APPENDIX A

In this appendix, we give definitions of constants $R_{N,n}$ for completeness. The $n\sigma$ confidence region of an N -dimensional distribution $P(x)$ is defined by

$$x_i K^{ij} x_j \equiv R^2(x) \leq R_{N,n}^2, \quad (\text{A1})$$

where the constants $R_{N,n}$ is the solution of the following equation,

$$\int_{R(x) \leq R_{N,n}} P(x) d^N x = q_n \quad . \quad (\text{A2})$$

The constant q_n is the $n\sigma$ -confidence level defined through the one-dimensional standard normal distribution:

$$q_n \equiv \frac{1}{\sqrt{2\pi}} \int_{-n}^n e^{-\frac{1}{2}x^2} dx \quad . \quad (\text{A3})$$

After appropriate diagonalization and scaling, we find that Eq. (A2) is reduced to

$$q_n = \frac{1}{2^{\frac{N-2}{2}} \Gamma\left(\frac{N}{2}\right)} \int_0^{R_{N,n}} e^{-\frac{1}{2}r^2} r^{N-1} dr \quad . \quad (\text{A4})$$

Some of the numerical results are given in Table 1.

APPENDIX B

In this appendix, we derive the conditions for the generalized case involving N -observables and M free parameters (in theory). The covariance and the probability distribution functions for errors are as in section 2. We define a (positive definite) metric in N -space by the inverse-covariance matrix K :

$$xy \equiv x_i y^i \equiv x_i K^{ij} y_j, \quad x^2 \equiv xx. \quad (\text{B1})$$

(Summation over repeated indices is always assumed.) This metric is appropriate in the sense that the resulting proper distance is an effective difference (statistical distance) determined by the errors of the corresponding direction.* Hereafter geometrical properties, e.g., orthonormality, are to be defined through this metric.

We have two predictions, each of which has M free parameters. These predictions define two M -dimensional subspaces, which are assumed to be parallel (under the linear approximation as is explained in section 2) and characterized by an orthonormal basis $\{p_\alpha\}$ ($p_\alpha p_\beta = \delta_{\alpha\beta}$) and a displacement vector, (d_i) .

In order to test a theory, one should determine in what space or by which variables a test will be performed. Tests are classified by the dimension k of the relevant space. A test in k -dimensional space (k -test) means that $(N-k)$ -dimensional set of variables are integrated out and remaining k -variables are to be used for the test. In other words the original N -space is projected onto k -space. Since we are working in the linear approximation, it is sufficient to consider only linear projection. Hence we have N -kinds of tests performed in $1 \sim N$ -dimensional spaces.

In a selected k -space, the necessary precision is obtained by requiring the statistical distance $D_{(k)}$ between the projected subspaces is larger than $n\sigma$ -radius,

* This metric is invariant under general linear-transformation of coordinates due to (2.17). By choosing an appropriate U , the metric K^{ij} is transformed to δ^{ij} . The following formulas may be understood intuitively in such a coordinate system.

$R_{k,n}$ (its definition and numerical values are given in Table 1). The power to discriminate theories can be maximized by choosing the direction of the projection so that $D_{(k)}$ is maximized.

First let us consider N-test or a test in N-space without projection. The distance $D_{(N)}$ between the two M-spaces is obtained by minimizing the norm of a vector $d + s_\alpha p_\alpha$ with respect to the parameters $\{s_\alpha\}$. One finds that the resulting distance vector is obtained by taking the component δ of d normal to vectors p_α ,

$$\delta = d - (p_\alpha d) p_\alpha, \quad (\text{B2})$$

and the distance is

$$D_{(N)}^2 = \delta^2 = d^2 - (p_\alpha d)(p_\alpha d). \quad (\text{B3})$$

Thus the N-test condition for avoiding the situation that neither of two predictions can be abandoned, is given by

$$D_{(N)} > 2R_{N,n}. \quad (\text{B4})$$

Next we proceed to a k-test. The induced probability distribution function is obtained by integrating out the variables spanned by u_l ($l = 1 \sim N - k$, $u_l u_m = \delta_{lm}$) and is a function of the remaining variables v in the k -space;

$$\tilde{P}(v) d^k v = \frac{\sqrt{\det' L}}{(2\pi)^{k/2}} \exp\left(-\frac{1}{2} v L v\right) d^k v, \quad (\text{B5})$$

where the reduced metric L^{ij} is defined by

$$L^{ij} = K^{ij} - u_i^i u_i^j, \quad (\text{B6})$$

and $\det' L$ means the determinant taken in k -space. This reduced metric has

(N-k)-zero-eigenvectors:

$$Lu_l = 0. \quad (\text{B7})$$

The “induced” distance vector $\tilde{\delta}$ in the k-space is obtained as a component of δ normal to the vectors p_β in the sense of L-norm:

$$\tilde{\delta} = \delta - (\delta L p_\alpha)(p_\alpha L p_\beta)^{-1} p_\beta, \quad (\text{B8})$$

where $(p_\alpha L p_\beta)^{-1}$ means the matrix inverse with respect to suffices α and β . Zero eigenmodes of the matrix $(p_\alpha L p_\beta)$ are understood to be excluded in the sum of α and β . The distance $D_{(k)}$ between the two subspaces projected onto k-space is the L-norm of $\tilde{\delta}$,

$$D_{(k)}^2 = (\tilde{\delta} L \tilde{\delta}) = \delta^2 - (\delta u_m)(\delta u_l) M_{ml}, \quad (\text{B9})$$

where the matrix,

$$M_{ml} \equiv \delta_{ml} + (u_m p_\beta)(u_\beta L p_\alpha)^{-1} (p_\alpha u_l), \quad (\text{B10})$$

is positive definite, since $(p_\beta L p_\alpha)$ is positive definite after excluding zero modes. Therefore we obtain the inequality,

$$D_{(k)}^2 \leq \delta^2 = D_{(N)}^2. \quad (\text{B11})$$

The equality holds only if the direction of projection satisfies

$$\delta u_l = 0 \quad (\text{for all } l = 1 \sim (N - k)). \quad (\text{B12})$$

For a given k , one can always choose u_l 's to satisfy the condition (B12) and obtain the maximized distance in k-space. (In $N = 2$, $k = 1$ case discussed in section 2, the unique solution of (B12) is $u = p$.)

In conclusion the optimized k-test imposes the condition,

$$D_{(N)} > 2R_{k,n}. \quad (\text{B13})$$

A trivial inequality

$$R_{k,n} < R_{k',n} \text{ for } k < k', \quad (\text{B14})$$

suggests that a k-test with a smaller k gives a weaker condition on the allowed errors. Therefore the weakest condition for discriminating two predictions is obtained by the case $k = 1$:

$$D_{(N)} > 2R_{1,n} = 2n. \quad (\text{B15})$$

The $k = 1$ test obtained above is related to that referred to in the mathematical literature as the most powerful test.^{[14]*}

It should be noted that k-test with N-observables is not equivalent to k-test with N+N'-observables. In fact, any addition of new data increases the distance ($D_{(N+N')} \geq D_{(N)}$), due to the inequality (B11). For example, when an (N+1)th independent datum with the distance δ_{N+1} and the variance σ_{N+1}^2 is added, one finds

$$D_{(N+1)}^2 = D_{(N)}^2 + \frac{\delta_{N+1}^2}{\sigma_{N+1}^2 - p_{N+1}^2}, \quad (\text{B16})$$

where δ_{N+1} and p_{N+1} are the N+1-th components of δ and p respectively which are properly determined in N+1-space. Thus one can always weaken the condition for the necessary precision in a fixed-k-test by adding new data. If the new data have large errors, then the improvement on the condition remains small as is seen in Eq. (B16).

* Strictly speaking, the most powerful test is the k=1 test but with the confidence region that is infinite on one side, while our condition (B15) is given by the usual symmetric confidence region. Accordingly, our condition is a little stronger than that for the most powerful test.

$N \backslash n$	1	2	3	4	5
1	1	2	3	4	5
2	1.515	2.486	3.439	4.397	5.361
3	1.878	2.833	3.763	4.697	5.640
4	2.172	3.117	4.031	4.950	5.890

Table 1. Numerical values of $R_{N,n}$.

FIGURE CAPTIONS

1. The tree and loop-corrected $M_Z - M_W$ relations are drawn. The solid section shows the 1σ -range of masses determined by low energy $\sin^2 \theta_W$ data. The circle roughly represents the present measurements of masses: The center is at the average of the UA1 and UA2 results and the radius (3 GeV) corresponds to 1σ region resulting from $\Delta M_Z = \Delta M_W = 2\text{Gev}$ (note that in Table 1, $R_{2,1} \cong 1.5$).
2. Illustration of the quantities defined in the section 2. R is an example of a confidence region (2.23).
3. The region defined by the condition (3.2) for the three-variable case is illustrated. Every coordinate represents the scaled variable $R_{k,n}\sigma_i$, and its maximum value is 8. The thin lines show the sections of the surface at intervals of 2 for $R_{k,n}\sigma_Z$ and $R_{k,n}\sigma_W$, and at intervals of 1 for $R_{k,n}\sigma_\theta$.
4. A projection of Fig. 3 onto the $(R_{k,n}\sigma_Z, R_{k,n}\sigma_W)$ plane.
5. Improvement factor $I(R_{k,n}\sigma_\theta)$ defined by (3.10) for the two cases. For $\sigma_Z = 0$, I diverges for $R_{k,n}\sigma_\theta \leq 4.4$.

REFERENCES

1. S. L. Glashow, Nucl. Phys. 22 (1961) 579. S. Weinberg, Phys. Rev. Letters 19 (1967) 1264. A. Salam, in *Elementary Particle Theory*, ed. N. Svartholm (Almqvist and Wiksell, Stockholm, 1968), p. 367. S. L. Glashow, J. Iliopoulos and L. Maiani, Phys. Rev. D2 (1970) 1285. M. Kobayashi and T. Maskawa, Prog. Theor. Phys. 49 (1973) 652.
2. UA1 Collaboration: G. Arnison *et al.*, Phys. Letters 122B (1983) 103; 126B (1983) 398; 129B (1983) 273; 134B (1984) 469. UA2 Collaboration: M. Banner *et al.*, *ibid.* 122B (1983) 476; P. Bagnaia *et al.*, *ibid.* 129B (1983) 130; Z. Phys. C24 (1984) 1.
3. K.-I. Aoki, Z. Hioki, R. Kawabe, M. Konuma and T. Muta, Prog. Theor. Phys. Suppl. No. 73 (1982), and references cited therein. M.A.B. Bég and A. Sirlin, Phys. Rep. 88 (1982) 1, and references cited therein. B. W. Lynn and J. F. Wheeler, ed: *Workshop on Radiative Corrections in $SU(2)_L \times U(1)$* , Trieste, 1983 (World Scientific, Singapore, 1984).
4. Z. Hioki, Prog. Theor. Phys. 68 (1982) 2134. L. Maiani, in: Proceedings of the International Conference on Unified Theories and Experimental Tests, Venice, 1982 (Istituto Naz. Fis. Nucl. Italy, 1983) p. 55.
5. A. Sirlin, Phys. Rev. D22 (1980) 971. W. Wetzel, Z. Phys. C11 (1981) 117. W. J. Marciano and A. Sirlin, Phys. Rev. D29 (1984) 945; D31 (1985) 213 Erratum.
6. Z. Hioki, Prog. Theor. Phys. 71 (1984) 663.
7. M. Davier, Proceedings of the LEP Summer Study (CERN 79-01), Vol. 1, p. 61. W. J. Marciano, in: Proceedings of the 1983 International Symposium on Lepton and Photon Interactions at High Energies, Cornell, 1983, eds. D. G. Cassel and D. L. Kreinick (Newman Laboratory of Nuclear Studies, Cornell University, 1983), p. 80. Z. Hioki, University of

- Tokushima preprint TOKUSHIMA 85-02 (Talk at the International Symposium on Physics of Proton-Antiproton Collision, March 13-15, 1985, Ibaraki, Japan).
8. P. Langacker, in: Proceedings of the XXII International Conference on High Energy Physics, Leipzig, 1984, eds. A. Meyer and E. Wieczorek (Akademie der Wissenschaften der DDR, Institut für Hochenergiephysik, 1984), p. 215 and references therein.
 9. B. Grzadkowski, P. Krawczyk, J. Pawelczyk and S. Pokorski, CERN-TH-4041/84 (1984).
 10. L. Alvarez-Gaumé, J. Polchinski and M. B. Wise, Nucl. Phys. B221 (1983) 495. R. Barbieri and L. Maiani, *ibid.* B224 (1983) 32. E. Eliasson, Northeastern University preprint NUB No. 2621; Phys. Letters 147B (1984) 65. C. S. Lim, T. Inami and N. Sakai, Phys. Rev. D29 (1984) 1488. J. A. Grifols and J. Solá, Phys. Letters 137B (1984) 257. K.H.G. Schwarzer, Oxford preprint Ref: 40/84. Z. Hioki, University of Tokushima preprint TOKUSHIMA 85-01 (to appear in Prog. Theor. Phys. 73 (1985) No. 5).
 11. B. W. Lynn and M. E. Peskin, SLAC preprint in preparation.
 12. A. C. Longhitano, Phys. Rev. D22 (1980) 1166. J. J. Van der Bij and M. Veltman, Nucl. Phys. B231 (1984) 205. J. J. Van der Bij, *ibid.* B248 (1984) 141.
 13. M. Veltman, Nucl. Phys. B123 (1977) 89. M. B. Einhorn, D.R.T. Jones and M. Veltman, *ibid.* B191 (1981) 146. M. Consoli, S. Lo Presti and L. Maiani, Nucl. Phys. B223 (1983) 474. F. Halzen, Z. Hioki and M. Konuma, Phys. Lett. 126B (1983) 129. Z. Hioki, Nucl. Phys. B229 (1983) 284. S. Bertolini and A. Sirlin, *ibid.* B248 (1984) 589.
 14. W. T. Eadie, D. Drijard, F. E. James, M. Roos and B. Sadoulet, *Statistical Methods in Experimental Physics* (North-Holland, Amsterdam, 1971).

15. A. Sirlin, Phys. Rev. D22 (1980) 971, W. J. Marciano and A. Sirlin, Phys. Rev. D22 (1980) 2695.
16. W. J. Marciano, Phys. Rev. D20 (1979) 274. S. Dawson, J. S. Hagelin and L. Hall, *ibid.* D23 (1981) 2666. F. Antonelli and L. Maiani, Nucl. Phys. B186 (1981) 269. S. Bellucci, M. Lusignoli and L. Maiani, *ibid.* B189 (1981) 329.
17. A. Sirlin, Phys. Rev. D29 (1984) 89.

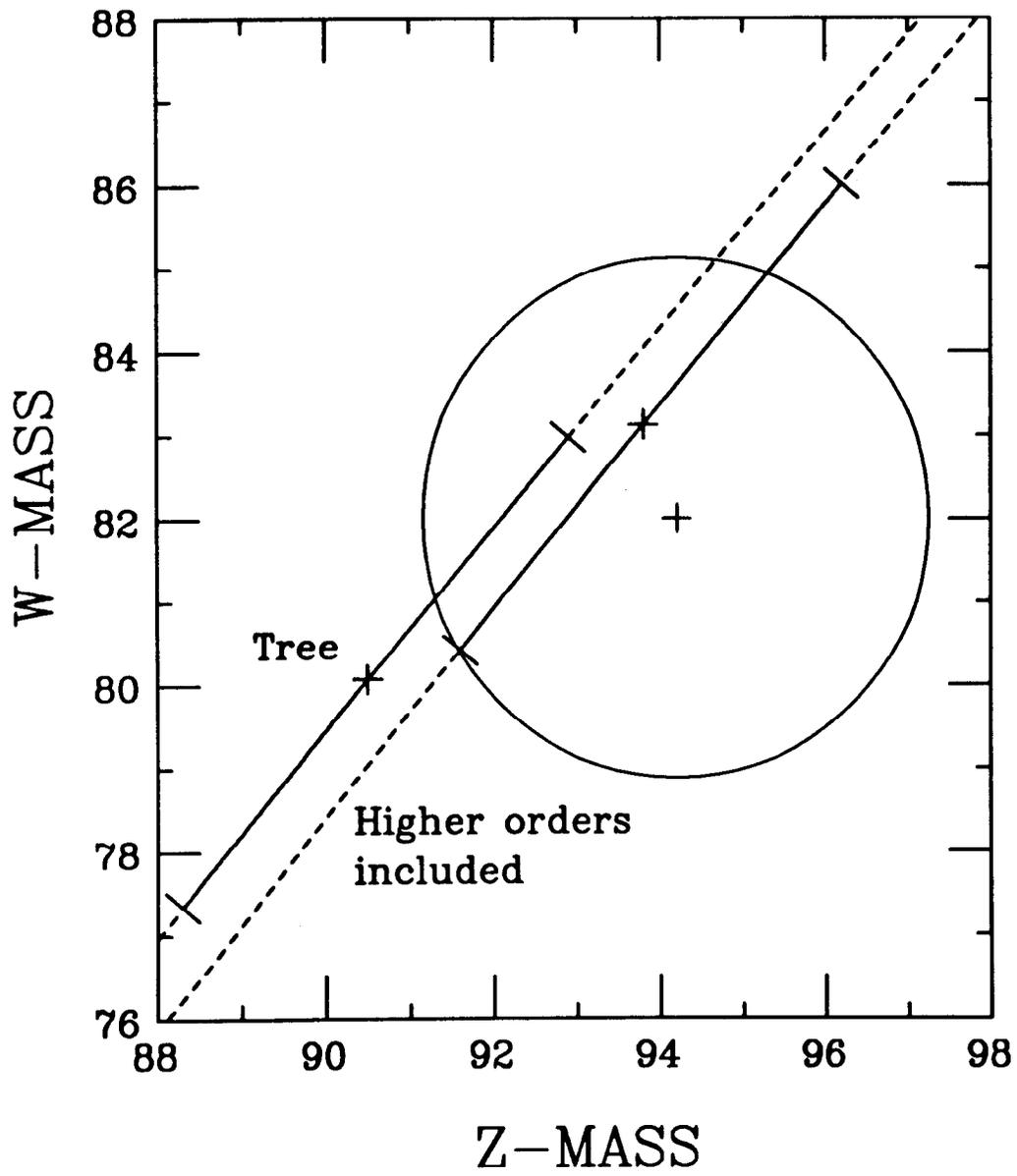


Fig.1

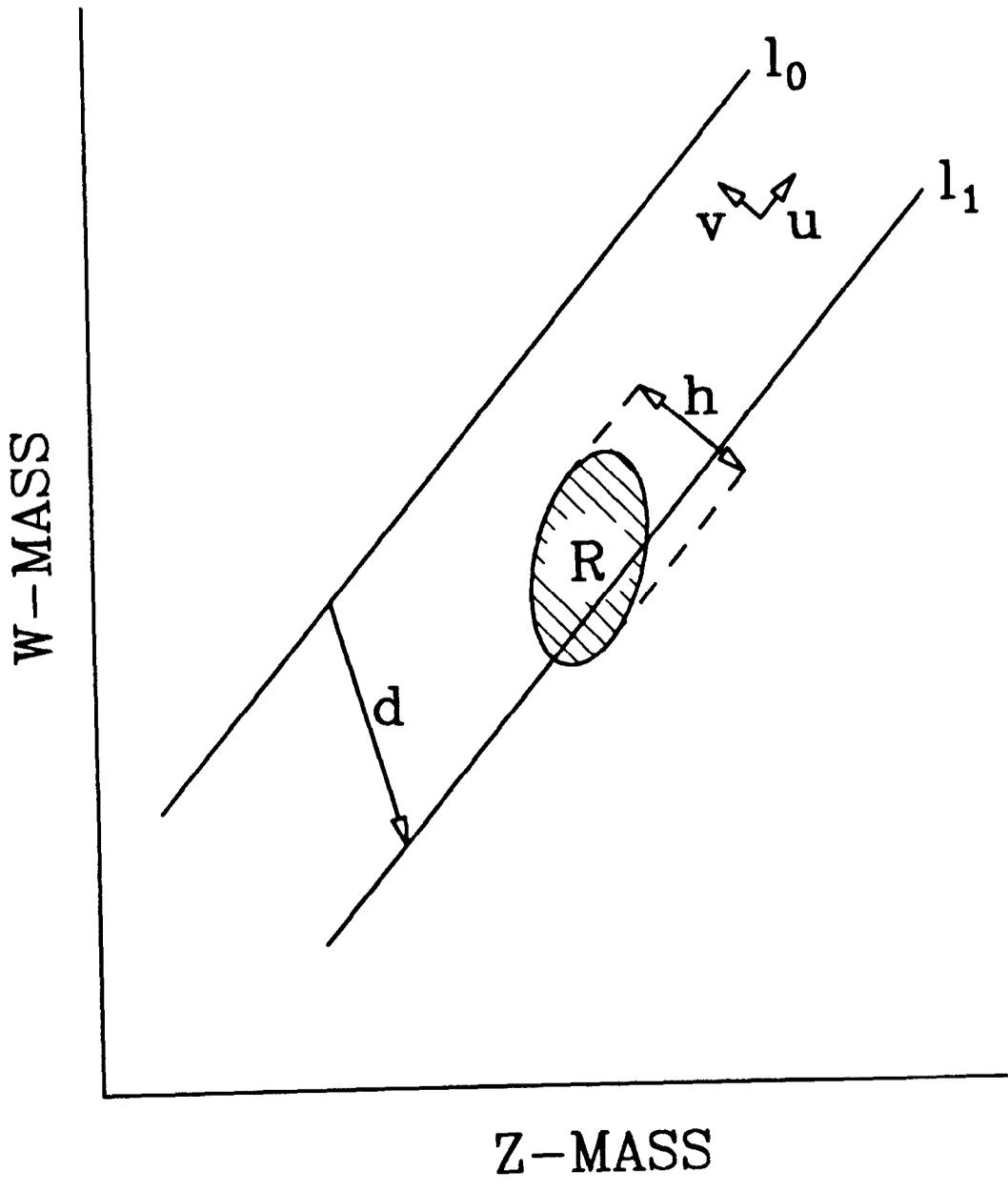


Fig.2

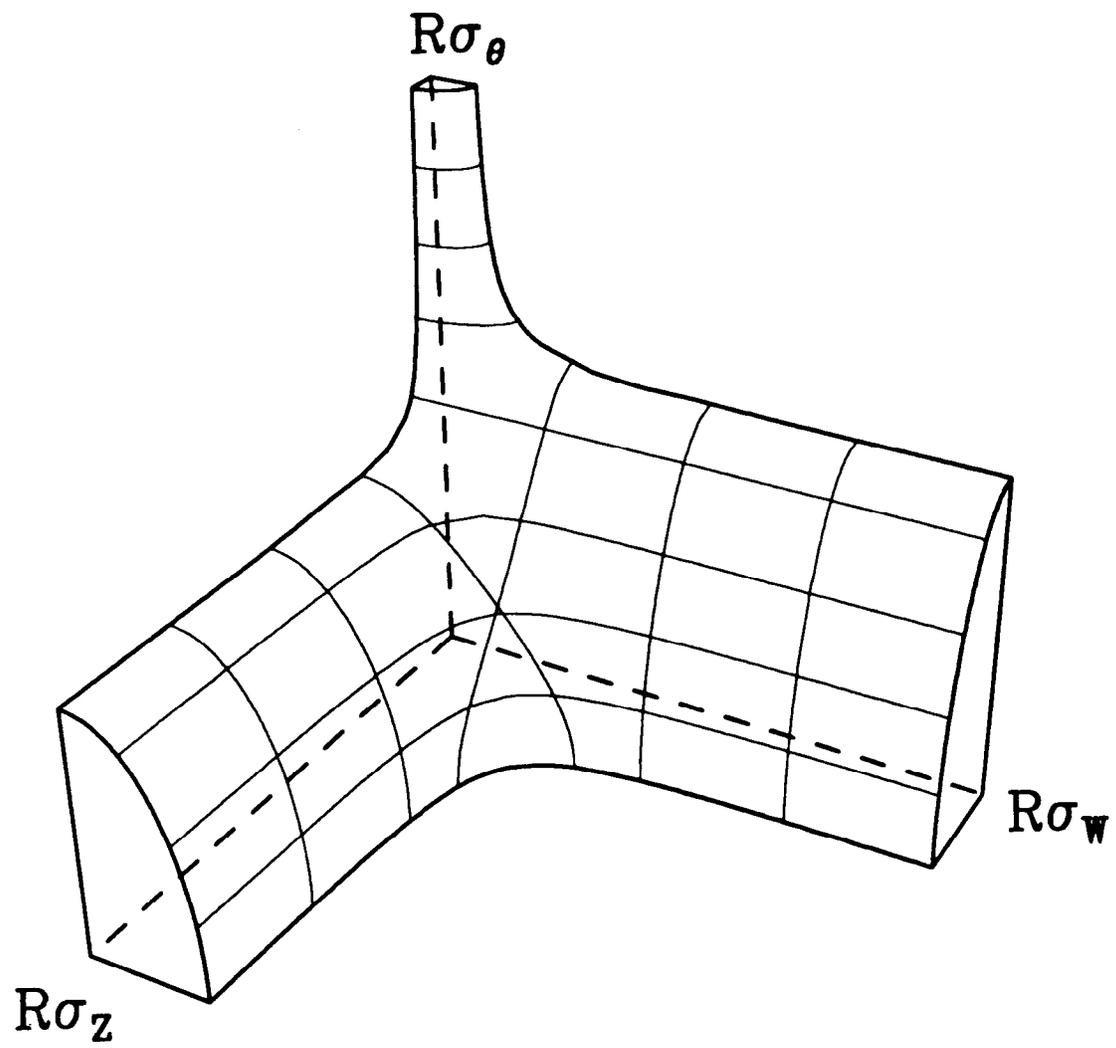


Fig.3

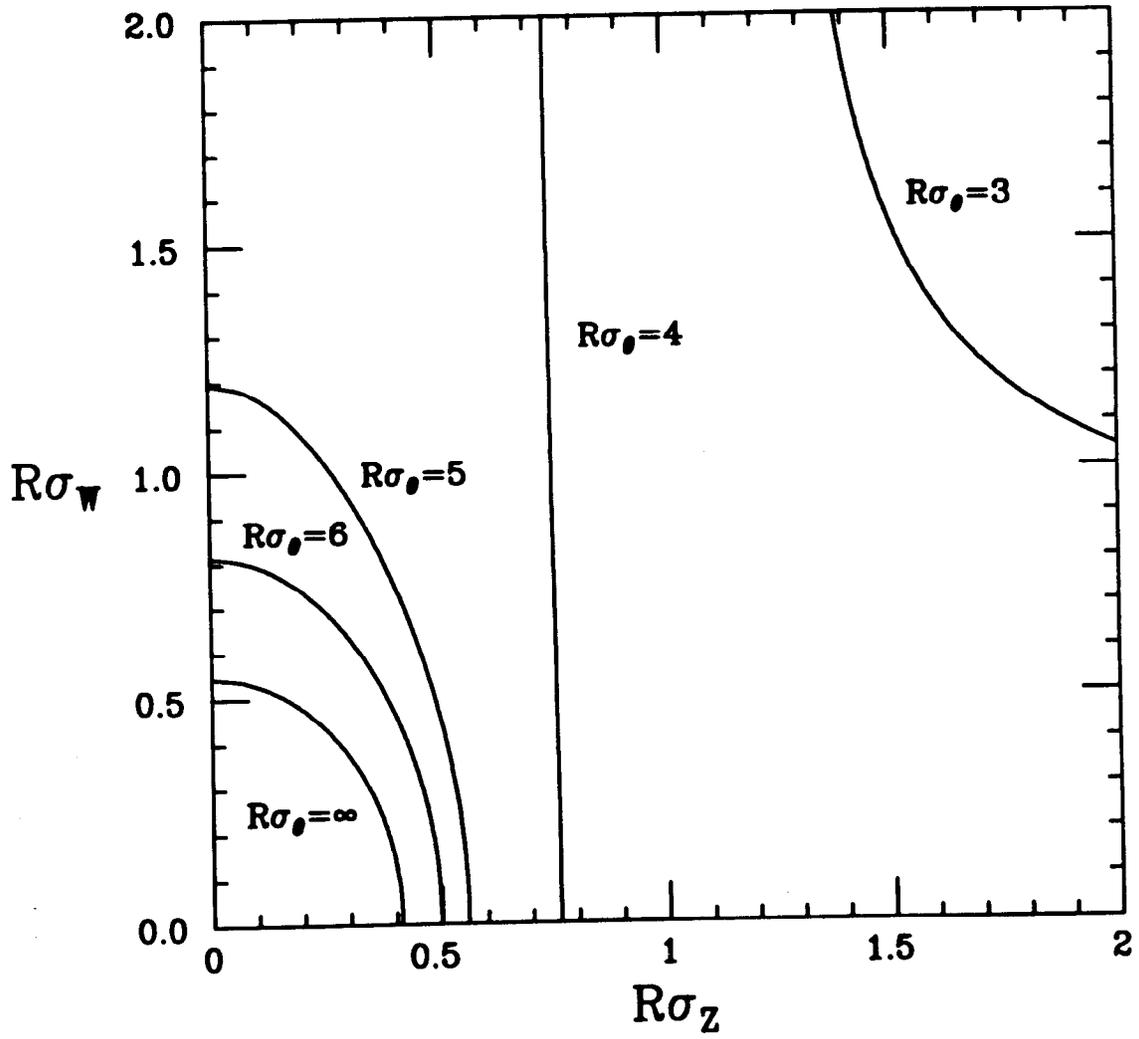


Fig.4

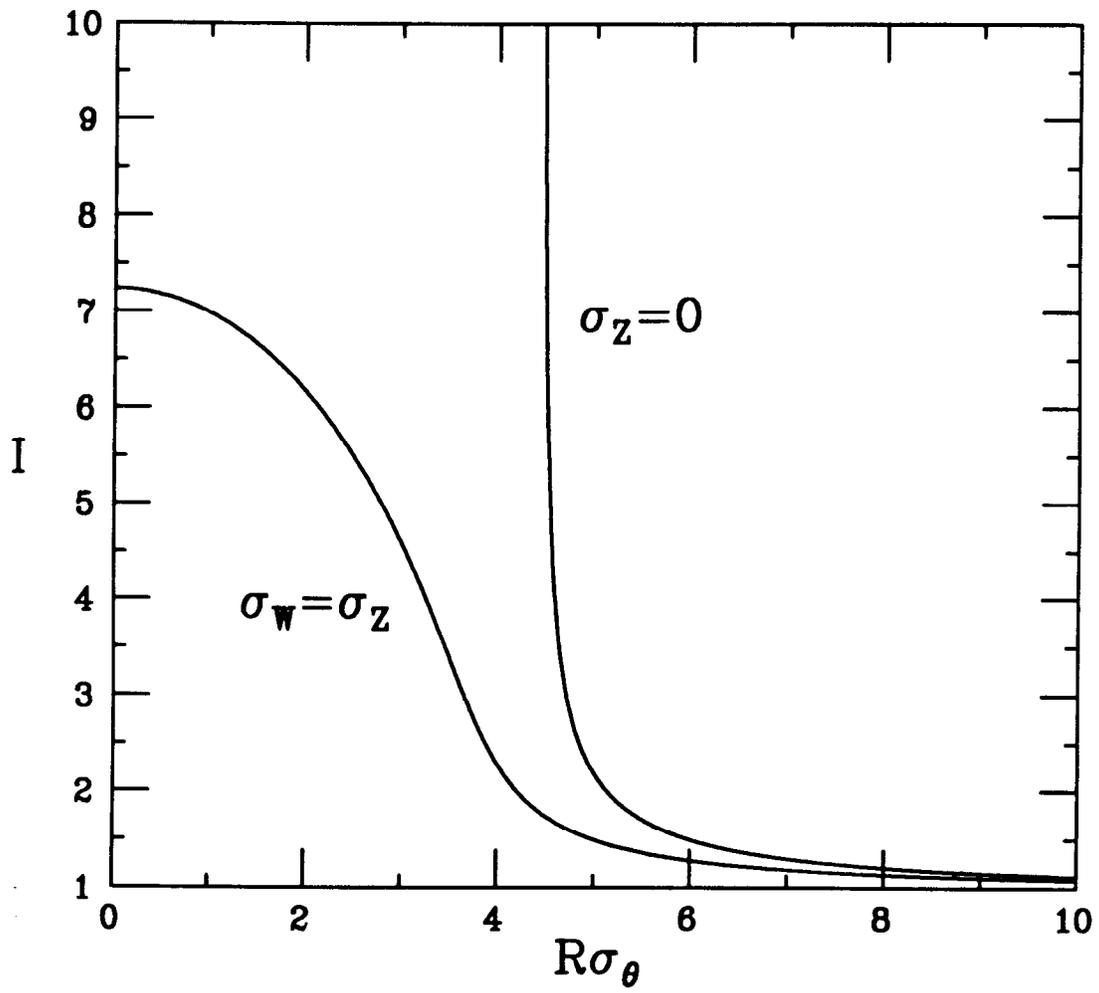


Fig.5