Lawrence Berkeley National Laboratory
Recent Work

Title
New perspectives in physics beyond the standard model

Permalink
https://escholarship.org/uc/item/0xk388xm

Author
Weiner, Neal J.

Publication Date
2000-09-09

eScholarship.org Powered by the California Digital Library
University of California

LBNL-45782

ERNEST ORLANDO LAWRENCE
BERKELEY NATIONAL LABORATORY

New Perspectives in Physics Beyond the Standard Model

Neal J. Weiner

Physics Division

September 2000

Ph.D. Thesis

REFERENCE COPY
Does Not
Circulate

Bldg. 50 Library - Ref.
Lawrence Berkeley National Laboratory

Copy 1

LBNL-45782

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof, or The Regents of the University of California.

Ernest Orlando Lawrence Berkeley National Laboratory
is an equal opportunity employer.

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

LBNL-45782

New Perspectives in Physics Beyond the Standard Model

Neal Jonathan Weiner
Ph.D. Thesis

Department of Physics
University of California, Berkeley

and

Physics Division
Ernest Orlando Lawrence Berkeley National Laboratory
University of California
Berkeley, CA 94720

September 2000

This work was supported by the Director, Office of Science, Office of High Energy and Nuclear Physics, Division of High Energy Physics, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098, and by National Science Foundation Grant No. PHY-95-14797.

LBNL-45782

New Perspectives in Physics Beyond the Standard Model

by

Neal Jonathan Weiner

B.A. (Carleton College) 1996

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Physics

in the

GRADUATE DIVISION
of the
UNIVERSITY of CALIFORNIA at BERKELEY

Committee in charge:

Professor Lawrence Hall, Chair
Professor Hitoshi Murayama
Professor Nicolai Reshetikhin

2000

New Perspectives in Physics Beyond the Standard Model

Copyright © 2000

by

Neal Jonathan Weiner

The U.S. Department of Energy has the right to use this document
for any purpose whatsoever including the right to reproduce
all or any part thereof.

Abstract

New Perspectives in Physics Beyond the Standard Model

by

Neal Jonathan Weiner

Doctor of Philosophy in Physics

University of California at Berkeley

Professor Lawrence Hall, Chair

In 1934 Fermi postulated a theory for weak interactions containing a dimensionful coupling with a size of roughly 250GeV. Only now are we finally exploring this energy regime. What arises is an open question: supersymmetry and large extra dimensions are two possible scenarios. Meanwhile, other experiments will begin providing definitive information into the nature of neutrino masses and CP violation. In this paper, we explore features of possible theoretical scenarios, and study the phenomenological implications of various models addressing the open questions surrounding these issues.

Professor Lawrence Hall
Dissertation Committee Chair

To all my teachers and my family,

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Particle physics theory has enjoyed a long stretch of tremendous success. Up until recently, there was no known experimental deviation from the standard model except gravity. Precision measurements have repeatedly confirmed standard model predictions, such as measurements of the anomalous magnetic moment of the electron (accurate at one part in 10^8), the muon anomalous magnetic moment (accurate at one part in 10^5), and the consistency of M_W, M_Z and $\sin^2\theta_W$, as well as many others.

Nonetheless, particle theorists generally consider the standard model merely an effective theory, valid in the energy regime below some scale Λ. It leaves open the question of the origin of CP violation, has nineteen undetermined parameters, does not include neutrino masses or an understanding of their small size, and is unstable against radiative corrections.

All of these things suggest that physics at some scale greater than presently tested regimes is responsible for these things. Extensions of the standard model introduce new

questions that need answering, and, in answering them, often give new experimental signals that will, hopefully, allow us to distinguish them in the near future. Moreover, new relationships between these issues appear as we look at particular extensions beyond the standard model.

1.1 The Standard Model

The Standard Model (SM) of particle physics is the $SU(2) \otimes U(1)$ theory of weak and electromagnetic interactions combined with the $SU(3)$ theory of strong interactions between quarks.

The action for the theory is

$$\begin{aligned} S = \int d^4x \quad & \overline{L}^i \gamma^\mu D_\mu L_i + \overline{E}^{c\,i} \gamma^\mu D_\mu E^c_i + \overline{Q}^i \gamma^\mu D_\mu Q_i \\ + \quad & \overline{U}^{c\,i} \gamma^\mu D_\mu U^c_i + \overline{D}^{c\,i} \gamma^\mu D_\mu D^c_i \\ + \quad & \lambda^u_{ij} Q^i U^{c\,j} \tilde{H} + \lambda^d_{ij} D^{c\,i} Q^j H + \lambda^e_{ij} L^i E^{c\,j} H \\ + \quad & (D^\mu \phi)^* D_\mu \phi + \mu^2 \phi^* \phi - \lambda (\phi^* \phi)^2 \\ - \quad & \frac{1}{4} (F^i_{\mu\nu})^2_{3,2,1} \end{aligned} \tag{1.1}$$

In this expression, i and j are generation indices (u,c,t for up type quarks, d,s,b for down type quarks, e, μ, τ for the charged leptons and ν_e, ν_μ and ν_τ for the neutrinos). D is the gauge covariant derivative,

$$D_\mu = \partial_\mu - ig' Y B_\mu - ig T_i^{SU(2)} A^i_\mu - ig_s T_i^{SU(3)} G^i_\mu. \tag{1.2}$$

Field	$SU(3)$	$SU(2)$	$U(1)$
Q	**3**	**2**	1/6
U	**3**	**1**	-1/3
D	**3**	**1**	2/3
L	**1**	**2**	-1/2
E	**1**	**1**	1
H	**1**	$\overline{\mathbf{2}}$	-1/2

Table 1.1: Transformation properties of SM fields under $SU(3) \otimes SU(2) \otimes U(1)$.

The gauge field strength tensor is defined as

$$[D_\mu, D_\nu] = -gF^i_{\mu\nu}T_i, \tag{1.3}$$

and the T^i are the generators of the group in question. The representations and charges of the fields are shown in table 1.1. $\tilde{H} = iT_2^{SU(2)}H^*$ carries the proper gauge numbers for Yukawas for up type fermions.

This action is invariant under local gauge transformations. However, the potential for the scalar field ϕ has a negative mass squared, making it energetically favorable for ϕ to take on a vacuum expectation value (vev), spontaneously breaking the $SU(2) \otimes U(1)_Y$ symmetry down to $U(1)_{EM}$. The vev of ϕ generates masses for three of the four electroweak gauge bosons, making their effects much weaker than those of the massless photon or gluons. The vev also generates masses for the fermions via the Yukawa terms.

This action is remarkable in that it contains all known interactions (except gravity) and masses for all known particles (except neutrinos). In fact, it is nearly the most general,

renormalizeable action that can be written with the given field content[1] This is one of the great successes of the Standard Model. Baryon number violating operators are four-fermion operators, which are dimension six and nonrenormalizable. Many flavor changing processes such as $\mu \to e\gamma$ which are not observed correspond to nonrenormalizable operators as well. Their experimental absence is a great success of the standard model, but as we shall see, it also makes the possible observation of such things a great signal for physics beyond the standard model.

1.2 Effective Field Theories

For a great while, nonrenormalizable operators were considered a serious defect in a theory. In a renormalizable theory, all divergences arising in quantum corrections can be absorbed into observed couplings or precisely cancelled by other divergences in the theory. Nonrenormalizable terms, by contrast, introduce divergences that cannot be cancelled, even by a finite number of new operators, without introducing new divergences.

Within the context of effective field theories, nonrenormalizable operators are not only not considered a defect, they are generally considered to be present unless forbidden by some symmetry. Effective theories are only considered valid below some energy scale Λ. Since nonrenormalizable operators come with dimensionful coefficients, these coefficients will have a size typically of the order of Λ to some power, corresponding to the exchange of some particles with a mass roughly equal to Λ. The divergences of the theory will be cut off at Λ when the new particles can be resolved.

[1] An additional $\tilde{G}G$ operator can be written for the $SU(3)$ fields, which can violate CP in the strong interactions.

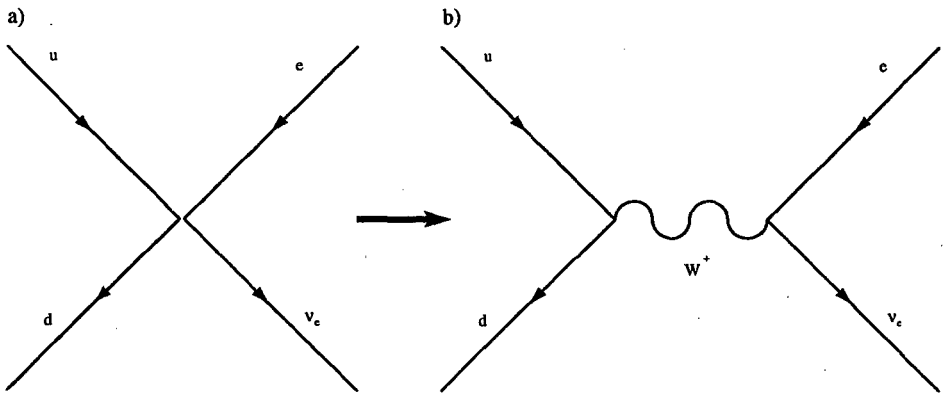

Figure 1.1: In 1.1a we see the low energy effective interaction. At higher energies, we can resolve this interaction to be the exchange of a W boson.

Consider as an example the theory of weak interactions. If we restrict our attention to processes $s < 1\text{GeV}$, our theory will consist of mesons, baryons and light leptons. In particular, we will observe the decay $\pi^+ \to e^+\nu_e$. This process can be understood in terms of the four fermion operator $\overline{u}\gamma_\mu d\overline{e}\gamma^\mu\nu_e$, which we show in figure 1.1a. Such an operator is nonrenormalizable and generates divergences, such as through the diagram shown in figure 1.2. However, we cannot understand the decay through any particle that would be produced at the energy scales in question.

Does this mean that the theory is pathological? Of course not, the given operator is just an *effective* operator, giving the low energy effects of an exchange of a W-boson. When integrating over the momenta flowing through the loop in figure 1.2, at a scale of the order M_W^2, the local four-fermion interaction is resolved to be the interaction in 1.1b, and the diagram becomes a controlled divergence in a renormalizable theory.

Thus, the presence of new physics at some energy scale motivates the inclusion of nonrenormalizable operators suppressed by the scale of the new physics. If we can motivate the need for new physics at some nearby energy scale, then we can generically expect the

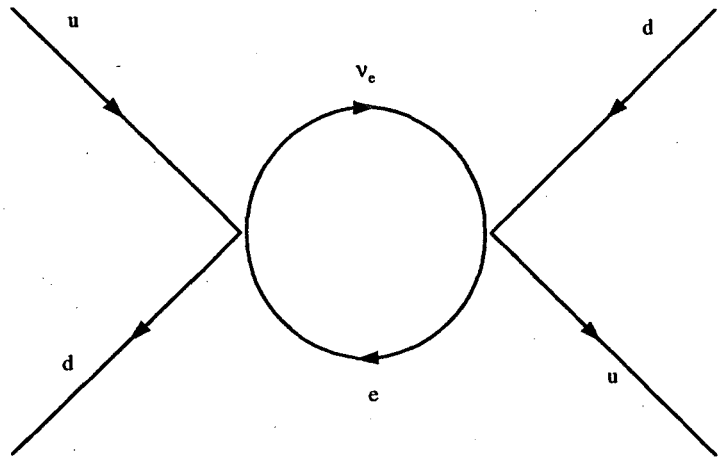

Figure 1.2: The effective operator generates loop divergences, such as the one shown here.

presence of nonrenormalizable operators, and thus signals of the new physics. Employing notions of effective theory is will be instrumental in what lies ahead. Sometimes we have a clear idea of what the more fundamental theory is, but sometimes we may not even *have* a field theory above the scale, and instead something like string theory. If this is the case, effective field theory is all we have to work with.

1.3 CP Violation

C,P and T were thought for a long time to be fundamental symmetries of nature. Quite simply, they correspond to exchange of particle and antiparticle (C), inversion of space via $x \to -x$ (P), and time reversal via $t \to -t$ (T). It a theorem of local quantum field theory that their product, CPT, is a symmetry of the theory. The maximal parity violation present in the weak interactions showed that C and P were not fundamental symmetries. However, it appeared that the product, CP, remained a good symmetry of the theory. The discovery in 1964 of a CP violating decay of K_L required the inclusion of CP violation in

any theory of elementary particles.

Within the Standard Model, it is possible to include such effects. By allowing the Yukawa matrices to be complex, there is generically a phase that cannot be removed by field redefinitions. This phase will can explain the observed CP violating processes. However, if new physics is present near the weak scale, it could easily contribute to, or be the source of the observed signals, at least in principle.

We are motivated to ask whether the standard model itself must be CP violating. Could all CP violation be generated in some new physics, in a so-called "superweak" theory? We shall see that recent improvements to the limit of ΔM_{B_s} imply that pure superweak theories, while not excluded, no longer provide a good fit to the data. We will introduce class of general superweak theories in which all flavor changing interactions are governed by an approximate flavor symmetry which gives a "3 mechanism". These theories are in good agreement with data, and predict low values for $|V_{td}|, |V_{ub}/V_{cb}|, B(K^+ \to \pi^+ \bar{\nu}\nu), \epsilon'/\epsilon$ and CP asymmetries in B decays, and high values for ΔM_{B_s} and $f_B\sqrt{B_B}$. An important example of such a theory is provided by weak scale supersymmetric theories with soft CP violation. The CP violation originates in the squark mass matrix, and, with phases of order unity, flavor symmetries can yield a correct prediction for the order of magnitude of ϵ_K.

1.4 Flavor

One of the most disappointing features of the standard model is the presence of nineteen undetermined parameters. Given the similarity in quantum numbers of u and t, for example, it seems unnatural that there should be five orders of magnitude difference

in their Yukawas, and similarly for b and d, τ and e. Since the fundamental theory seems to distinguish them, it begs us to ask the question: what relates these particles? What allows the top to be so much heavier than the up, for instance? Furthermore, the CKM angles are generally small. If the yukawas were random matrices, why does the top decay predominantly to bottom, charm to strange, etc?

One of the most appealing ideas is that of flavor symmetries [1]. Suppose that there is some global symmetry acting on the different generations. For instance, suppose there is a $U(1)$ symmetry for which the heavy generation has charge 0, the intermediate generation has charge 1 and the lightest generation has charge 2. Then further assume that there is a scalar field ϕ with charge -1. We then expect nonrenormalizable operators to be generated at some scale M_F

$$(M_F^{-1}\phi(t^c q_2 h + c^c q_3 h) + M_F^{-2}\phi^2(t^c q_1 h + u^c q_3 h + c^c q_2 h) + \dots . \tag{1.4}$$

If further the field ϕ takes on a vev ϵM_F, then we will generate a texture

$$\lambda^u \sim \begin{pmatrix} \epsilon^4 & \epsilon^3 & \epsilon^2 \\ \epsilon^3 & \epsilon^2 & \epsilon \\ \epsilon^2 & \epsilon & 1 \end{pmatrix} \tag{1.5}$$

where it is understood that there are undetermined coefficients of order unity multiplying each entry of the texture. Such a texture then gives both small mixing angles and hierachical masses.

When employing flavor symmetries within the context of an effective theory, we have a very powerful tool with which to address possible flavor changing signals. Non-

renormalizable operators which are absent in the standard model should appear, and flavor symmetries will give us an idea of what size we should expect. In extensions of the standard model, we will see observable consequences of particular symmetries, and the utility of non-Abelian flavor symmetries.

1.5 The Hierachy Problem

Quantum corrections to parameters in our action are generally expected in a quantum field theory. For instance, in QED, there is a correction to the mass of the electron given by

$$\delta m_e = \frac{3\alpha}{4\pi} m_e \log\left(\frac{\Lambda^2}{m_e^2}\right) \tag{1.6}$$

There are two interesting features of this expression: first, it is proportional to m_e. This is the result of a chiral symmetry of our theory in which we can, absent mass terms, separately transform the right- and left-chiral components of the electron. Secondly, the correction to the mass of the electron depends only logarithmically on the cutoff. The end result of this is that a small mass for the electron is technically natural: if we input a small mass into the theory at tree level, we do not expect quantum corrections to radically alter it.

In contrast, scalar fields have no such chiral symmetry protecting their masses. Within the standard model there are quantum corrections to the mass squared of the higgs, which are expected to be of the order of the cutoff of the theory (figure 1.3). Thus, if we input a weak-scale mass for the higgs, quantum corrections should give a mass of the

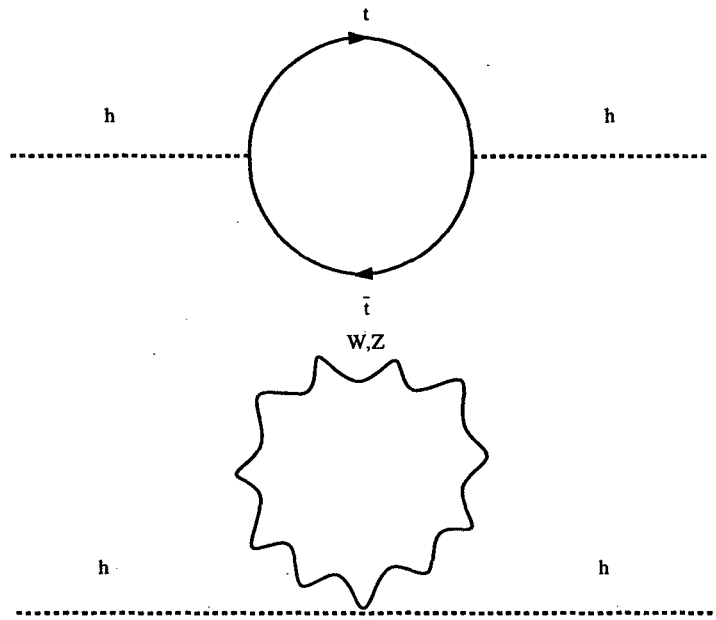

Figure 1.3: Two diagrams which contribute to the quadratically divergent corrections to the Higgs mass.

order of the cutoff of the theory. Alternatively, if the physical mass of the higgs boson is to be of the order of the weak scale, and if we expect the cutoff of the theory to be much higher than that, there must be a fine cancellation between the tree-level parameter and the quantum corrections. (For a cutoff at the Planck scale, we need a tuning of one part in 10^{26}!) Furthermore, it is not sufficient to tune the tree level quantity against the one-loop correction, as at each order in perturbation theory we will introduce new divergences which will need tuning.

The naturalness of such a procedure is known as the hierarchy problem, or what stabilizes the weak scale against radiative corrections? Generally it is assumed not that there is a fine cancellation to be explained, but rather that there is new physics near the weak scale that cuts off the divergence.

The three most popular explanations of this are supersymmetry, technicolor, and

large extra dimensions. We will not consider the case of technicolor (for a review, see [2]), but will briefly review the other two alternatives.

1.6 Supersymmetry

Probably the most popular extension of the standard model is supersymmetry (SUSY), in which there is postulated a partner fermion for every boson in the theory, and likewise a partner boson for every fermion [3]. This solves the hierarchy problem in the following manner: the SM diagrams that generate divergences in the higgs mass are cancelled by the new diagrams in which superpartners propagate in the loop. More conceptually, supersymmetry relates the mass of the higgs boson to the mass of the higgsino, its SUSY partner, which is protected against radiative corrections by an approximate chiral symmetry (broken by the Higgsino mass μ).

Of course, we do not observe partners to SM fields, so supersymmetry must be broken. The cancellation between fermion and boson loops will only occur at a scale higher than the splitting of the masses of the SM fields from their spartners. If this splitting is of the order of the weak scale, then this naturally solves the hierarchy problem. Of course, what remains is the question of what sets the scale of supersymmetry breaking. We will explore this question later.

Furthermore, the presence of new particles introduces many new parameters into the theory. Each new particle has SUSY-breaking masses as well as new mixing angles. If these quantities appear without any organizing principle, then we would expect a tremendous amount of FCNC and CP violating processes that have not been observed. There have

been many explanations given of why these processes do not occur, including non-Abelian flavor symmetries [4], gauge-mediated supersymmetry breaking [5], and others.

The most convenient formulation of supersymmetry (in particular supersymmetry with only one fermionic generator, so-called $N = 1$ supersymmetry) is in terms of superfields. (For a review of superfields, see, for instance, [6].) The action of the Minimal Supersymmetric Extension of the Standard Model (MSSM) can be written

$$\begin{aligned} S &= \int d^4x\, d^2\theta(\lambda^d_{ij}Q^jU^jH_u + \lambda^d_{ij}Q^iD^jH_d \\ &+ \lambda^e_{ij}L^iE^jH_d + \frac{1}{16\pi g^2}Tr(W_{\alpha,i}W^\alpha_i) + \mu H_uH_d) + h.c. \\ &+ \int d^4x\, d^2\theta d^2\overline{\theta}(Q^\dagger e^{V_Q}Q + U^\dagger e^{V_U}U + D^\dagger e^{V_D}D \\ &+ L^\dagger e^{V_L}L + E^\dagger e^{V_E}E + H_d^\dagger e^{V_{H_D}}H_D + H_u^\dagger e^{V_{H_u}}H_u) \\ &+ \int d^4x\, 1/2m_u^2h_uh_u^* + 1/2m_d^2h_dh_d^* \\ &+ 1/2ml_{ij}^2l^il^{j*} + 1/2me_{ij}^2e^ie^{j*} + 1/2mq_{ij}^2q^iq^{j*} + 1/2mu_{ij}^2u^iu^{j*} + 1/2md_{ij}^2d^id^{j*} \\ &+ (\mu Bh_uh_d + A^u_{ij}q^iu^jh^u + A^d_{ij}q^id^jh^u + A^e_{ij}l^ie^jh^d + h.c.) + m_g\overline{\lambda}^g_a\lambda^g_a \end{aligned} \tag{1.7}$$

where i and j are flavor indices, g is a group index, and $V_X = T^a_XV^a$ is the gives the group action on a field X.

Of course, the supersymmetry breaking operators are much smaller than the Planck scale, which prompts another question: what physics generates this small scale? The traditional answer has been dimensional transmutation. Another question is why is μ, which is a SUSY-conserving parameter, of the order of the SUSY-violating parameters? There are a variety of answers to this question. However, in the context of theories in higher

dimensions, there is a new alternative which simultaneously answers both these questions.

We will employ supersymmetric "shining" of free massive chiral superfields in extra dimensions from a distant source brane. This can trigger exponentially small supersymmetry breaking on our brane of order $e^{-2\pi R}$, where R is the radius of the extra dimensions. This supersymmetry breaking can be transmitted to the superpartners in a number of ways, for instance by gravity or via the standard model gauge interactions. The radius R can easily be stabilized at a size $O(10)$ larger that the fundamental scale. The models we will see are extremely simple, relying only on free, classical bulk dynamics to solve the hierarchy problem.

1.7 Large Extra Dimensions

At the heart of the hierarchy problem is the idea that within the standard model, there are only two dimensionful parameters: $M_W \approx 100$ GeV and $M_{pl} \approx 10^{19}$ GeV. If we take the theory to be valid up to the Planck scale, where quantum gravity effects become significant, we expect a correction to the Higgs mass of the order of the Planck scale. This presupposes that the Planck scale is a fundamental quantity in nature. As first proposed in [7], this need not necessarily be the case.

In all versions of string theory, one universal element is the presence of extra dimensions (that is, $d > 4$). Arkani-Hamed, Dimopolous and Dvali proposed that if the size of the extra dimensions (the "bulk") is much larger than the fundamental scale of the theory M_*, then M_* can be much less than gravity. If the bulk is large enough, then M_* could be of the order of the weak scale. The physics cutting off the Higgs mass divergence

could be string theory itself!

There are many constraints on these theories (which are summarized in [8]), but there is no a priori reason to reject them. However, if the cutoff of the theory is O(TeV), we need to understand if we can control the potentially harmful flavor changing operators that would otherwise be generated. In fact, generating flavor at the TeV scale while avoiding flavor-changing difficulties appears prohibitively difficult at first sight. We will see to the contrary that having such a large bulk *allows* us to lower flavor physics close to the TeV scale. Small Yukawa couplings can be generated by "shining" badly broken flavor symmetries from distant branes, and flavor and CP-violating processes are adequately suppressed by these symmetries. We will further see how the extra dimensions avoid four dimensional disasters associated with light fields charged under flavor. We construct elegant and realistic theories of flavor based on the maximal $U(3)^5$ flavor symmetry which naturally generate the simultaneous hierarchy of masses and mixing angles. All of this can be incorporated into a new framework for *predictive* theories of flavor, where our 3-brane is embedded within highly symmetrical configurations of higher-dimensional branes.

In these theories, aside from phenomenological questions, we must understand why the bulk is so much larger than the fundamental scale. We will see that in theories with (sets of) two large extra dimensions and supersymmetry in the bulk, the presence of non-supersymmetric brane defects naturally induces a logarithmic potential for the volume of the transverse dimensions. Since the logarithm of the volume rather than the volume itself is the natural variable, parameters of O(10) in the potential can generate an exponentially large size for the extra dimensions. This provides a true solution to the hierarchy problem,

on the same footing as technicolor or dynamical supersymmetry breaking. The area moduli have a Compton wavelength of about a millimeter and mediate Yukawa interactions with gravitational strength. We will see a simple explicit example of this idea which generates two exponentially large dimensions. In this model, the area modulus mass is in the millimeter range even for six dimensional Planck scales as high as 100 TeV.

1.8 Neutrino Masses

Aside from gravity, there is only one known deviation from the standard model, namely neutrino masses. Within the standard model, there is no right handed neutrino, and hence no mass term for the neutrino.

If neutrinos have a mass, and, like the quarks, their mass eigenstates are not aligned with the weak eigenstates, then production of a neutrino results in the production of a superposition of these states. Very simple quantum mechanics shows that a neutrino which begins as e can convert to μ, or μ to τ, etc. Experiments showing a deficit in the solar neutrino flux have long been interpreted as potentially such a process. Recently, more conclusive evidence from Superkamiokande has demonstrated an anomaly in atmospheric neutrinos when comparing upward-going neutrinos and downward-going neutrinos.

It seems simple to input a neutrino mass into the standard model by adding a right handed neutrino. However, the mass scale needed for solar and atmospheric neutrino experiments is $O(10^{-1} - 10^{-3} eV)$, several orders of magnitude smaller than the electron, the lightest fermion.

We can understand the smallness of the neutrino mass from the point of view of

effective field theory. Although we cannot write a renormalizable, gauge-invariant mass term for the neutrino in the standard model, we *can* write down a dimension five operator $1/M_x(i\sigma_2\nu^* H)^\dagger \nu H$, which gives a Majorana mass for the neutrino when the Higgs aquires a vev. The smallness of the observed neutrino mass is then understood by the suppression v/M_x relative to the weak scale. A significant particular example of this, explaining the scale M_x is the seesaw mechanism [9].

We will study both solar and atmospheric neutrino fluxes in the context of oscillations of the three known neutrinos. We will aim at a global view which identifies the various possibilities, rather than attempting the most accurate determination of the parameters of each scenario. For solar neutrinos we will emphasize the importance of performing a general analysis, independent of any particular solar model and we consider the possibility that any one of the techniques — chlorine, gallium or water Cerenkov — has a large unknown systematic error, so that its results should be discarded. The atmospheric neutrino anomaly is studied by paying special attention to the ratios of upward and downward going ν_e and ν_μ fluxes. Both anomalies can be described in a minimal scheme where the respective oscillation frequencies are widely separated or in non-minimal schemes with two comparable oscillation frequencies. We discuss explicit forms of neutrino mass matrices in which both atmospheric and solar neutrino fluxes are explained. In the minimal scheme we identify only two 'zeroth order' textures that can result from unbroken symmetries.

But what can we say about the structure of neutrino masses? Can we understand it together with theories of quark masses, or do we need a radical departure? If we are will to consider scenarios with a sterile neutrino, we can make connections with previous work

in flavor symmetries. A $U(2)$ flavor symmetry can successfully describe the charged fermion masses and mixings [4], and supress SUSY FCNC processes, making it a viable candidate for a theory of flavor. We will see that a direct application of this $U(2)$ flavor symmetry automatically predicts a mixing of 45^o for $\nu_\mu \Rightarrow \nu_s$, where ν_s is a light, right-handed state. The introduction of an additional flavor symmetry acting on the right-handed neutrinos makes the model phenomenologically viable, explaining the solar neutrino deficit as well as the atmospheric neutrino anomaly, while giving a potential hot dark matter candidate and retaining the theory's predictivity in the quark sector.

In quite a different direction, we will study the possibility that *no* structure is necessary for the neutrino mass matrices, so long as it is generated by a seesaw process. Such a scenario would be consistent with the observed hierarchy, even with the experimental constraints currently in place.

1.9 Summary

In this paper we shall attempt to address some of these questions before us. What do we know about neutrino masses and what can we know? Is there a relationship between the structure of the masses of the known fermions? What information do we already have about CP violation and what limits does it place on the CP violation in new physics?

More generally, are there consequences of the presence of additional dimensions, whether large or small?

While definitive answers to these questions must wait for the round of experiments in progress and in preparation, we will discuss the hints we already have, and the evidence

we expect within the context of particular scenarios.

Bibliography

[1] C.D. Froggatt and H.B. Nielsen, *Nucl. Phys.* **B147**(1979) 277.

[2] K. Lane, Lectures given at Theoretical Advanced Study Institute (TASI 93) in Elementary Particle Physics, 1993, hep-ph/9401324; C.T. Hill, In Nagoya 1996, Perspectives of strong coupling gauge theories* 54-71. e-Print Archive: hep-ph/9702320.

[3] J. Wess and B. Zumino *Phys. Lett.* **B49**:(1974) 52.

[4] R. Barbieri, G. Dvali and L.J. Hall. Phys.Lett.B377:76-82,1996, hep-ph/9512388; R. Barbieri, L.J. Hall and A. Romanino. Phys.Lett.B401:47-53,1997, hep-ph/9702315.

[5] L. Alvarez-Gaume, M. Claudson, M.B. Wise, *Nucl. Phys.* **B207** (1982) 96; M. Dine, A.E. Nelson Y. Nir and Y. Shirman, *Phys. Rev.* **D53** (1996) 2658.

[6] J. Wess and J. Bagger, Supsersymmetry and Supergravity, 2nd ed., Princeton University Press, 1992.

[7] N. Arkani-Hamed, S. Dimopoulos and G. Dvali, *Phys. Lett.* **B429**(1998) 263, hep-ph/9803315.

[8] N. Arkani-Hamed, S. Dimopoulos, G. Dvali, *Phys. Rev.* **D59** (1999) 086004, hep-ph/9807344.

[9] M. Gell-Mann, P. Ramond, and R. Slansky, in *Supergravity* (ed. P. van Nieuwenhuizen and D. Z. Freedman, North-Holland, Amsterdam, 1979), p. 315;
T. Yanagida, in *proceedings of the workshop on the unified theory and the baryon number in the universe* (ed. O. Sawada and A. Sugamoto, KEK Report No. 79-18, Tsukuba, Japan, 1979).

Chapter 2

Alternative Theories of CP Violation

2.1 CP Violation

All observed CP violation can be described by the complex parameter ϵ_K, which describes an imaginary contribution to the $\Delta S = 2$ mixing of the neutral K mesons. Such a mixing implies the existence of an effective Hamiltonian

$$\mathcal{H}_{eff}^{\Delta S=2} = \frac{1}{v^2} \sum_{ij} iC_{ij}(\bar{s}\Gamma_i d)\ (\bar{s}\Gamma_j d) \tag{2.1}$$

where $v = 247$ GeV, and i, j run over possible gamma matrix structures. The dimensionless coefficients C_{ij} are real in a basis where the standard model $\Delta S = 1$ effective Hamiltonian has a real coefficient. In the case that the dominant term is $\Gamma_i = \Gamma_j = \gamma^\mu(1 - \gamma_5)/2$,

$$C_{LL} = 4(1 \pm 0.3) \cdot 10^{-10} \frac{|\epsilon_K|}{2.3 \cdot 10^{-3}} \frac{0.75}{B_K}. \tag{2.2}$$

The two basic issues of CP violation are

- What is the underlying physics which leads to $\mathcal{H}_{eff}^{\Delta S=2}$? Is it a very small effect originating at the weak scale, as suggested by the form C/v^2, or is it a larger effect generated by physics at higher energies?

- How can the magnitude $C \approx 10^{-9}$ to 10^{-10} be understood?

2.2 The CKM Theory of CP Violation

In the standard model all information about flavor and CP violation originates from the Yukawa coupling matrices. After electroweak symmetry breaking, this is manifested in the Cabibbo-Kobayashi-Maskawa (CKM) matrix of the charged current interactions of the W boson [1]. A one loop box diagram with internal top quarks gives the dominant contribution to $\mathcal{H}_{eff}^{\Delta S=2}$ via

$$C_{LL,SM} = \frac{g^2}{32\pi^2} S_t \mathcal{I}m[(V_{td}V_{ts}^*)^2] \tag{2.3}$$

where $S_t \simeq 2.6$ is the result of the loop integration, and g is the SU(2) gauge coupling constant. For a suitable choice of the CKM matrix elements, V_{ij}, the standard model can provide a description of the observed CP violation. The fundamental reason for the size of the CP violation observed in nature remains a mystery, however, and must await a theory of flavor which can explain the values of $|V_{td}|, |V_{ts}|$ and the CKM phase. If the CKM matrix contained no small parameters one would expect $C_{LL,SM}$ to be of order 10^{-2} to 10^{-3} rather than the observed value of order 10^{-9} to 10^{-10}.

Of course, measurements of CP conserving observables have shown that $|V_{ij}|$ are small for $i \neq j$, and, given the measured values of $|V_{us}|$ and $|V_{cb}|$, it is convenient to use the

Wolfenstein parameterization[2] of the CKM matrix, in which case (2.3) becomes

$$C_{LL,SM} \simeq 20 \cdot 10^{-10}(1-\rho)\eta \tag{2.4}$$

If we *assume* that the CKM matrix does not have any other small parameters, the standard model yields a value of ϵ_K of the observed order of magnitude. While this is not a prediction, it is an important success of the standard model, and has made the CKM theory the leading candidate for CP violation. To our knowledge, there is no similar success in any published alternative to the CKM theory of CP violation, since in these theories the order of magnitude of C can only be fixed by fitting to the measured value of ϵ_K. In this letter we present such an alternative theory.

Two further measurements of $|V_{ij}|$, with $i \neq j$, would determine both ρ and η allowing a prediction of $C_{LL,SM}$ and ϵ_K. A fit to the two observables $|V_{ub}/V_{cb}|$ and ΔM_{B_d}, but *not* ϵ_K, is shown in Figure 1. For all numerical work, we use the data and parameters listed in Table 1 — for a discussion of these, and references, see [3]. Unfortunately the large uncertainties make this a very weak prediction: $\eta = 0$ is allowed even at the 68% confidence level. Hence, from this one cannot claim strong evidence for CKM CP violation.

Recent observations at LEP have improved the limit on $B_s - \bar{B}_s$ mixing, so that $\Delta M_{B_s} > 10.2\ \mathrm{ps}^{-1}$ at 95% confidence level [4]. The result of a χ^2 fit in the standard model to ρ and η using the three observables $|V_{ub}/V_{cb}|$, ΔM_{B_d} and ΔM_{B_s}, but *not* ϵ_K, is shown in figure 2. For B_s mixing the amplitude method is used [5, 3]. Comparing Figures 1 and 2, it is clear that the ΔM_{B_s} limit is now very significant. At 68% confidence level the standard model is able to predict the value of ϵ_K to within a factor of 2; however, at 90% confidence level $\eta = 0$ is allowed, so that at this level there is no prediction, only an upper bound.

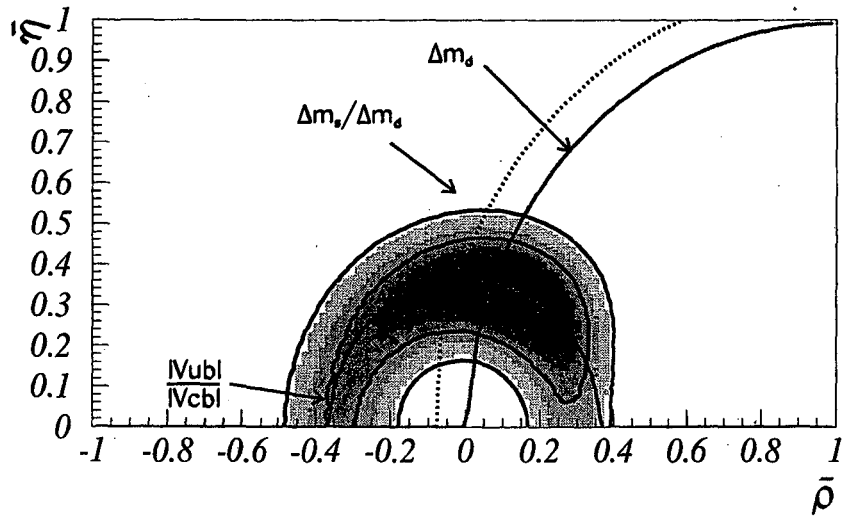

Figure 2.1: The 68% and 95% C.L. contours fits of $|V_{ub}/V_{cb}|$ and ΔM_{B_d} in the $\bar{\rho}/\bar{\eta}$ plane in the standard model. The curves correspond to constraints obtained from measurements of $|V_{ub}/V_{cb}|, \Delta M_{B_d}$ and ΔM_{B_s} (The last constraint is not included in the fit). $\bar{\rho} = \rho(1 - \lambda^2/2), \bar{\eta} = \eta(1 - \lambda^2/2)$.

While this is an important success of the CKM theory, it is still worth pursuing credible alternative theories of CP violation.

2.3 Pure superweak theories

A superweak theory [7] is one in which the CKM matrix is real, so $\eta = 0$, and $\mathcal{H}_{eff}^{\Delta S=2}$ of eq. (1) originates from physics outside the standard model. We define a *pure*

Table 2.1: Values of observables and parameters

$\|V_{ub}/V_{cb}\|$	0.080 ± 0.020
ΔM_{B_d}	0.472 ± 0.018 ps^{-1}
ΔM_{B_s}	> 10.2ps^{-1} at 95% C.L.
$f_{B_d}\sqrt{B_{B_d}}$	(200 ± 50) MeV
$f_{B_s}\sqrt{B_{B_s}}/f_{B_d}\sqrt{B_{B_d}}$	1.10 ± 0.07[6]
A	0.81 ± 0.04
$m_t(m_t)$	168 ± 6 GeV

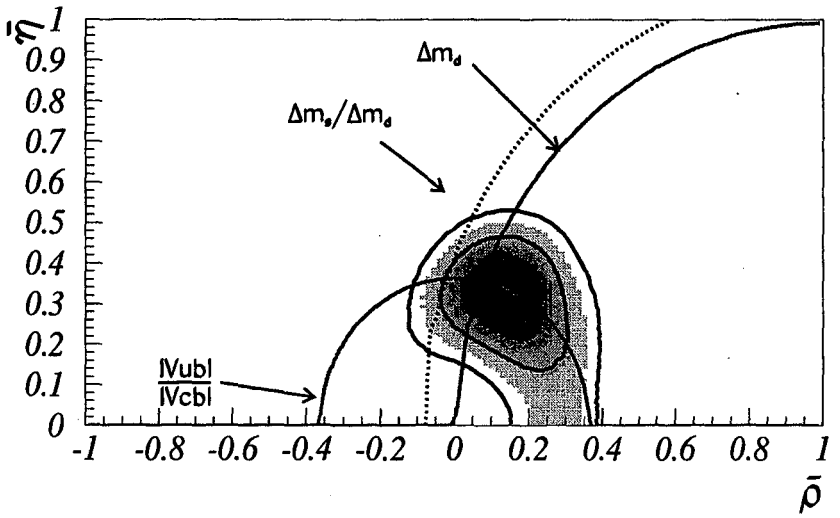

Figure 2.2: The 68% and 95% C.L. contours fits of $|V_{ub}/V_{cb}|, \Delta M_{B_d}$ and ΔM_{B_s} in the $\bar{\rho}/\bar{\eta}$ plane in the standard model. The curves correspond to constraints obtained from measurements of $|V_{ub}/V_{cb}|, \Delta M_{B_d}$ and ΔM_{B_s}.w $\bar{\rho} = \rho(1 - \lambda^2/2), \bar{\eta} = \eta(1 - \lambda^2/2)$.

superweak theory to be one where all flavor changing phenomena (other than ϵ_K) are accurately described by the real CKM matrix. Comparing Figures 1 and 2 at low η, one sees that the new limit on B_s mixing has excluded superweak theories with negative ρ. This has important phenomenological consequences for pure superweak theories.

We have computed $\chi^2(\rho)$ in pure superweak theories, using as input the three observables $|V_{ub}/V_{cb}|$, ΔM_{B_d} and ΔM_{B_s}. We find that all negative values of ρ are excluded at greater than 99% confidence level. At positive ρ only the two observables $|V_{ub}/V_{cb}|$ and ΔM_{B_d}, are relevant, and we find the most probable value of ρ to be +0.27. However, even this value of ρ corresponds to the pure superweak theory being excluded at 92% confidence level. Since the uncertainties are dominated by the theory of $f_{B_d}\sqrt{B_{B_d}}$, we take the view that this does not exclude purely superweak theories. In such theories positive values of ρ are 40 times more probable than negative values, and hence large values for $f_B\sqrt{B_B} \approx 250$

MeV and small values for $|V_{ub}/V_{cb}| \approx 0.06$ are predicted. A pure superweak description of CP violation implies

$$+0.20\ (0.13) < \rho < 0.34\ (+0.41) \qquad \text{at 68\% (95\%) confidence level} \qquad (2.5)$$

An important consequence of the new limit on B_s mixing is the strong preference for positive ρ and the resulting small values for $|V_{td}| \propto 1-\rho$. This is numerically significant: without the B_s mixing result the superweak theory can also have negative values of ρ which give $|V_{td}|$ about a factor of two larger than the positive ρ case. With the B_s result, a pure superweak theory must have $|V_{td}|$ at the lower end of the standard model range. Thus in a pure superweak theory, $\Delta M_{B_s} \propto \Delta M_{B_d}/|V_{td}|^2$ is predicted to be

$$14\ (10)\ \text{ps}^{-1} < (\Delta M_{B_s})_{PSW} < 26\ (32)\ \text{ps}^{-1} \qquad \text{at 68\% (95\%) confidence level} \qquad (2.6)$$

By comparison, in the standard model 10.5 (9.5) $\text{ps}^{-1} < \Delta M_{B_s} < 15$ (19) ps^{-1} at 68% (95%) confidence level.

In the standard model, the branching ratio $B(K^+ \to \pi^+\nu\bar{\nu})$ is given by [8]

$$B(K^+ \to \pi^+\nu\bar{\nu}) = c_1 \left((c_2 + c_3 A^2(1-\rho))^2 + (c_3 A^2 \eta)^2\right) \qquad (2.7)$$

where $c_1 = 3.9 \times 10^{-11}, c_2 = 0.4 \pm 0.06$ and $c_3 = 1.52 \pm 0.07$. In pure superweak theories, since ρ is positive and $\eta = 0$, the branching ratio is lowered to

$$B(K^+ \to \pi^+\nu\bar{\nu}) = (5.0 \pm 1.0) \cdot 10^{-11} \qquad (2.8)$$

relative to the standard model prediction of $(6.6^{+1.4}_{-1.2}) \cdot 10^{-11}$.* The recent observation of a candidate event for this decay [9] is not sufficient to exclude pure superweak theories, but further data from this experiment could provide evidence against such theories.

*This standard model result is smaller than that quoted in the literature because the improved limit on B_s mixing increases ρ even in the standard model.

2.4 General superweak theories

Pure superweak theories are artificial: they do not possess a symmetry which allows $\mathcal{H}_{eff}^{\Delta S=2}$ of eq. (1), while forbidding similar $\Delta B = 2$ operators. If ϵ_K is generated by new physics, why does this new physics not contribute to $B\bar{B}$ mixing? In general it would be expected to also contribute to $\Delta S = 1$ and $\Delta B = 1$ processes. In the absence of a fundamental theory of flavor, the relative sizes of the various flavor changing operators can be estimated only by introducing arguments based on approximate flavor symmetries.

We assume that the underlying theory of flavor possesses a flavor symmetry group, G_f, and a mass scale M_f. The breaking of G_f, whether explicit or spontaneous, is described in the low energy effective theory by a set of dimensionless parameters, $\{\epsilon\}$, each with a well defined G_f transformation. The low energy effective theory of flavor is taken to be the most general operator expansion in powers of $1/M_f$ allowed by G_f and $\{\epsilon\}$. In the case that the CKM matrix can be made real, we call these *general* superweak theories. The phenomenology of such theories depends on G_f, M_f and $\{\epsilon\}$ and will typically not coincide with the pure superweak phenomenology. The $\Delta B = 2$ operators may lead to exotic CP violation in neutral B meson decays and may contribute to ΔM_{B_d}, allowing large values of $|V_{td}|$ invalidating (2.6). Similarly the $\Delta S = 1$ operators may invalidate (2.8), and may give an observable contribution to ϵ'/ϵ.

2.5 The effective Hamiltonian for the "3 mechanism"

The dominant flavor changing neutral current (FCNC) interactions of the down sector of the standard model result from the "3 mechanism": small flavor breaking param-

eters which mix the light quarks with the heavy third generation quarks, together with a large, order unity, breaking of the flavor symmetry that distinguishes the third generation from the first two. Hence, beneath the weak scale, the standard model yields an effective Hamiltonian with dominant FCNC operators which contain a factor $V_{ti}^* V_{tj}$ for each flavor changing current $\bar{d}_i d_j$, and a factor $G_F^2 m_t^2/16\pi^2 \approx (1/16\pi^2)(1/v^2)$ from the loop integration. The relevant diagrams are all 1 loop, giving the $(1/16\pi^2)$ factor, and involve the large GIM violation of the top quark mass; since there is no small flavor violating parameter, the rest of the loop integral has an order of magnitude given by dimensional analysis as $(1/v^2)$.

Now consider physics beyond the standard model where the entire flavor structure of the theory beneath M_f is controlled by G_f and $\{\epsilon\}$ — both the Yukawa matrices of the standard model, $\lambda(\epsilon)$, and the non-standard model operators in $\mathcal{H}_{eff}(\epsilon)$. Since the dominant down sector, FCNC effects from $\lambda(\epsilon)$ are known to arise from the "3 mechanism", we assume that G_f and $\{\epsilon\}$ are chosen so that the dominant such effects from $\mathcal{H}_{eff}(\epsilon)$ are also from the "3 mechanism".

The most general parameterization of the "3 mechanism" in the down sector involves four complex parameters: $\epsilon_{L_i} = |\epsilon_{L_i}| e^{i\phi_{L_i}}$ and $\epsilon_{R_i} = |\epsilon_{R_i}| e^{i\phi_{R_i}}$, $i = 1, 2$, which describe the mixing of d_{L_i} and d_{R_i} with b_L and b_R. Assuming all phases to be of order unity, we can describe the "3 mechanism" in terms of just four real small parameters $|\epsilon_{L_i}|$ and $|\epsilon_{R_i}|$. We make the additional simplifying assumption that $|\epsilon_{L_i}| = |\epsilon_{R_i}| = \epsilon_i$, yielding the non-standard model interactions†

$$\mathcal{H}_{eff}^{(3)} = \frac{1}{M_f^2} \left[C_1 \left(\epsilon_1 \epsilon_2\right)^2 (\bar{s}d)^2 + C_2 \, \epsilon_2^2 \, (\bar{s}b)^2 + C_3 \, \epsilon_1^2 \, (\bar{b}d)^2 \right.$$

†It is straightforward to extend this Hamiltonian to the most general case of the "3 mechanism" involving four complex parameters.

$$+C_{1l}\,\epsilon_1\epsilon_2\,(\bar{s}d)(\bar{l}l) + C_{2l}\,\epsilon_2\,(\bar{s}b)(\bar{l}l) + C_{3l}\,\epsilon_1\,(\bar{b}d)(\bar{l}l) + \ldots] \qquad (2.9)$$

where C_i are complex coefficients of order unity, and l is a lepton field.[‡] A sum on possible gamma matrix structures is understood for each operator. Since the flavor changing interactions from both the standard model and the new physics are governed by the same symmetry, we can choose $\epsilon_1 = |V_{td}|$ and $\epsilon_2 = |V_{ts}|$. Such interactions can arise from many choices of G_f and $\{\epsilon\}$; the particular choice is unimportant, however, as the phenomenology rests only on three assumptions

- There is an underlying theory of flavor based on symmetry G_f and breaking parameters $\{\epsilon\}$.
- The dominant non-standard model FCNC operators of the down sector arise from the "3 mechanism".
- The symmetry breaking parameters of the down sector are left-right symmetric, and have phases of order unity.

In the standard model, the dominant FCNC of the down sector arises from the "3 mechanism", so that it is useful to describe the effective theory beneath the weak scale by eq. (2.9) with

$$\epsilon_1 = V_{td} \qquad \epsilon_2 = V_{ts}^* \qquad \frac{1}{M_f^2} = \frac{1}{16\pi^2}\frac{1}{v^2} \qquad (2.10)$$

and C_i real. This special case of the "3 mechanism" has a restricted set of gamma structures due to the left-handed nature of the weak interaction.

[‡]We do not consider lepton flavor violation in this letter.

2.6 Phenomenology of the "3 mechanism" in superweak theories

We have argued that pure superweak theories are artificial, and we now study superweak theories where FCNC interactions are generated by the "3 mechanism" and yield $\mathcal{H}_{eff}^{(3)}$ of (2.9). Why should such theories have V_{ij} real when C_i are complex? One possibility is that G_f forces the Yukawa matrices $\lambda(\epsilon)$ to have a sufficiently simple form that they can be made real by field redefinitions. Another possibility will be discussed later.

Since $\mathcal{H}_{eff}^{(3)}$ will be the origin of all CP violation, one may wonder if it could also account for all of $\Delta M_{B_{d,s}}$. This is not possible — charged current measurements, together with the unitarity of V, imply $|V_{td}|$ and $|V_{ts}|$ are sufficiently large that W exchange contributes a significant fraction of $\Delta M_{B_{d,s}}$.

Given that the FCNC of both the standard model and exotic interactions have the form of (2.9), it would appear that the exotic interactions must give a large fraction of $\Delta M_{B_{d,s}}$ since they are responsible for all of ϵ_K. This is not the case; in the standard model the $\Delta S = 2$ and $\Delta B = 2$ operators have chirality LL, whereas for a generic "3 mechanism" they will have all chiral structures. It is known that the LR, $\Delta S = 2$, operator has a matrix element which is enhanced by about an order of magnitude relative to that of the LL operator [10], and that there is no similar enhancement in the $\Delta B = 2$ case. Furthermore, the LR operator is enhanced by QCD radiative corrections in the infrared [11]; with the enhancement at 1 GeV about a factor of 3 larger than at 5 GeV. Hence we conclude *In a generic superweak theory, we expect that* $\mathcal{H}_{eff}^{(3)}$ *leads to* $\approx 3\%$ *contributions to* $\Delta M_{B_{d,s}}$. There is considerable uncertainty in this percentage because of the uncertainty in

the overall enhancement of the $\Delta S = 2$ and $\Delta B = 2$ contributions from the LR operator, and because of the unknown order unity C_i coefficients. Given this result, we must evaluate how well these generic superweak theories can account for the data, and to what extent they lead to predictions.

Let $\Delta_{d,s}$ and $\delta_{d,s}$ be the standard model and new physics contributions to

$$\Delta M_{B_{d,s}} = \Delta_{d,s} + \delta_{d,s} \tag{2.11}$$

First we consider a perturbation around the pure superweak case, where the fractional contributions from new physics $F_{d,s} = \delta_{d,s}/\Delta M_{B_{d,s}}$ are small. The central value of ρ, from ΔM_{B_d} alone, changes by $\Delta\rho = 0.5F_d$ for very small F_d ($\Delta\rho \simeq 0.3F_d$ for $F_d \simeq 0.1$). For positive F_d, this improves the fit of general superweak theories to ΔM_{B_d} and $|V_{ub}/V_{cb}|$. For example, $F_d = 0.1$ gives a central value of $\rho = 0.28$ with $\chi^2(\rho = 0.28) \simeq 2.4$, which corresponds to 68% C. L. Since the allowed range of ρ is little changed from eq. (5), the prediction of small $|V_{td}|$ persists in these general superweak theories, so that the prediction of eq. (7) for low values of $B(K^+ \to \pi^+\nu\bar{\nu})$ applies. Similarly, since ρ is little altered, the prediction for B_s mixing is $\Delta M_{B_s} = (\Delta M_{B_s})_{PSW}(1 - F_d + F_s)$, where the pure superweak prediction $(\Delta M_{B_s})_{PSW}$ is given in eq. (6). In this case the general superweak theory also predicts large values of ΔM_{B_s}, although for negative F_s, it is not quite so large as $(\Delta M_{B_s})_{PSW}$.

There is a second class of general superweak theories which is not a perturbation about the parameters of the pure superweak theories. In general superweak theories, the limit $\Delta M_{B_s} > 10.2$ ps^{-1} can be expressed as $\rho > -0.06 + 0.5(F_d - F_s)$. For negative F_d and positive F_s, the negative ρ region could become allowed. For example, $F_s = -F_d =$

0.1 (0.05) gives a theory in which ρ has a probability 25% (9%) of being negative. This class of superweak theories *requires* values of $|F_{d,s}|$ which are larger than our expectation, and appear somewhat improbable. They have $|V_{td}|$ and $B(K^+ \to \pi^+ \nu\bar{\nu})$ at the upper end of the standard model range. In these theories ΔM_{B_s} is likely to be low, although it depends on $F_{d,s}$.

2.7 Supersymmetry with a "3 mechanism"

In general, the alternative theory of CP violation of $\mathcal{H}_{eff}^{(3)}$ from the "3 mechanism" is not a strong competitor to the CKM theory of CP violation. The CKM theory, with two small measured parameters, $|V_{us}|$ and $|V_{cb}|$, yields the correct order of magnitude for ϵ_K, while superweak theories with the "3 mechanism" apparently require a new scale $M_f \approx 30v \approx 10$ TeV. However, there is the interesting possibility that the new physics generates FCNC operators only at 1 loop, as in the standard model. This would give $M_f \approx 4\pi m_f$, with the mass of the new quanta close to the weak scale at $m_f \approx 1$ TeV. We therefore take the view that the "3 mechanism" generating FCNC operators at 1 loop at the weak scale is a credible alternative to the CKM theory of CP violation. While not as minimal as the CKM theory, it correctly accounts for the order of magnitude of ϵ_K.

Let l represent d, s or b, left or right handed. New interactions of the form $\bar{l}lH$, where H is some new heavy field, will generate FCNC at tree level, whereas lHH generates them at 1 loop. Thus the exotic new heavy particles at the weak scale should possess a parity so that they appear only in pairs.

Weak scale supersymmetry allows a symmetry description of the weak scale, and leads to a successful prediction for the weak mixing angle. Furthermore, it incorporates the

economical Higgs description of flavor of the standard model. R parity ensures that superpartners appear pairwise in interactions, so that the dominant supersymmetric contributions to FCNC processes occur only at one loop. Supersymmetric theories have several new generation mixing matrices — in particular $W_{L,R}$ at the gluino interaction $(\vec{d}^{\dagger}_{L,R} W_{L,R} d_{L,R})\tilde{g}$. A flavor symmetry, G_f, can ensure that the largest contribution from superpartner exchange to FCNC occurs via the "3 mechanism"[12, 13]. If the small symmetry breaking parameters are left-right symmetric and real, this gives $\mathcal{H}^{(3)}_{eff}$ of (2.9) with

$$|W_{L,R_{31}}| \approx \epsilon_1 = |V_{td}| \qquad |W_{L,R_{32}}| \approx \epsilon_2 = |V_{ts}| \qquad \frac{1}{M_f^2} = \frac{1}{16\pi^2}\frac{1}{\tilde{m}^2} \tag{2.12}$$

where $\tilde{m}$ is the average mass of the colored superpartners in the loop. As the superpartners are at the weak scale, $\tilde{m} \approx v$, and comparing with (2.10) one finds that, with weak scale supersymmetry, it may well be that ϵ_K receives comparable standard model and supersymmetric contributions.[§]

Here we stress that weak scale supersymmetry can provide an important example of the general superweak theories discussed in this letter. The absence of CKM CP violation would be guaranteed if CP violation were *soft* — restricted to operators of dimension two and three. The Yukawa matrices would then be real, so that there would be no CP violation from diagrams with internal quarks, but the scalar mass matrices would contain phases, so that CP violation would arise from diagrams with internal squarks.[¶] Soft CP violation in

[§]Given the order of magnitude enhancement of the matrix element of the LR operator relative to the LL, and given the further order of magnitude enhancement of C_{LR} relative to C_{LL} from QCD scaling, one generically expects the supersymmetric contribution to be larger. However, these factors may be outweighed by colored superpartner masses somewhat larger than v, some degree of degeneracy between the third generation scalars and those of the lighter generations, and by W_{ij} somewhat less than V_{ij}. We note that the QCD enhancement of C_{LR} for the $\Delta S = 2$ operator[11] was not included in [12, 13, 14]

[¶]This is an alternative view to the one presented in [13], where the specific flavor symmetry forces forms for V and W matrices such that even the supersymmetric contribution to ϵ_K involves a phase originating from the Yukawa couplings.

supersymmetric theories, with FCNC operators arising from the "3 mechanism", represents a well-motivated and credible alternative to CKM CP violation, and will be explored in detail elsewhere.

2.8 Summary

Fits of the CKM matrix to $|V_{ub}/V_{cb}|$, ΔM_{B_d} and ΔM_{B_s} show that at 68% C.L. the standard model correctly predicts ϵ_K to better than a factor of two, while at 90% C.L. not even the order of magnitude can be predicted. On one hand the standard model is highly successful; on the other, there is still room for an alternative theory of CP violation.

The recent improvement on the limit on ΔM_{B_s}[4] implies that pure superweak theories with negative ρ are excluded, while at positive ρ they are somewhat disfavored. Pure superweak theories allow $0.13 < \rho < 0.41$ at 95% C.L., and predict high values for ΔM_{B_s} and $f_B\sqrt{B_B}$ and low values for $|V_{ub}/V_{cb}|$, $B(K^+ \to \pi^+\nu\bar{\nu})$ and ϵ'/ϵ.

We have argued that *pure* superweak theories are artificial, and have introduced *general* superweak theories, in which all FCNC are governed by an approximate flavor symmetry and the "3 mechanism." In this case the new physics induces other flavor changing operators in addition to the $\Delta S = 2$ operator responsible for ϵ_K; in particular, $\mathcal{O}(3)$% contributions to $B_{d,s}$ mixing are expected. There are two important classes of general superweak theories, one with positive ρ and the other with negative ρ. The first can be viewed as a perturbation about the superweak case, with an improved fit to data, while retaining the characteristic predictions mentioned above. The negative ρ possibility appears less likely, and arises only if the new physics contributes more than 10% of $\Delta M_{B_{d,s}}$. In this case future data should show a high value for $B(K^+ \to \pi^+\nu\bar{\nu})$ and low values for ΔM_{B_s},

$f_B\sqrt{B_B}$, $|V_{ub}/V_{cb}|$, and ϵ'/ϵ. All these superweak theories predict low values for the CP asymmetries in B meson decays.

Weak scale supersymmetric theories with softly broken CP can provide an important example of general superweak theories. As in the CKM theory, assuming phases of order unity yields a correct prediction for the order of magnitude of ϵ_K. In addition they have $\bar{\theta} = 0$ at tree level, and it is interesting to seek a flavor symmetry which would sufficiently protect $\bar{\theta}$ from radiative corrections to solve the strong CP problem.

Bibliography

[1] N. Cabibbo, *Phys. Rev. Lett.* **10** 531 (1963); M. Kobayashi and K. Maskawa *Prog. Theor. Phys.* **49** 652 (1973).

[2] L Wolfenstein, *Phys. Rev. Lett.* **51** 1945 (1983).

[3] P. Paganini, F. Parodi, P. Roudeau and A. Stocchi, hep-ph/9711261.

[4] B Oscillation LEP Working Group averages for Jerusalem Conf. (Aug. 1997), http://www.cern.ch/LEPBOSC.

[5] H.G. Moser and A. Roussarie, *Nucl. Inst. and Methods* **A384** (1997) 491.

[6] C. Bernard et al. (MILC Collaboration) " B Mixing on Lattice : f_B, f_{B_s} and related quantities " hep-ph/9709328; V. Giménez and G. Martinelli, " B $\overline{B}$ Mixing in HQET ", ROME prep. 96/1153, FTUV 96/26-IFIC 96/30 , hep-lat/9610024v2; Results presented in Hamburg by C. Sachradja at the Electron-Photon 1997 Conference (Hamburg).

[7] L Wolfenstein, *Phys. Rev. Lett.* **13** 562 (1964).

[8] A. Buras, M. Lautenbacher and C. Ostermaier, *Phys. Rev.* **D50** 3433 (1994).

[9] S. Adler et al., E787 Collaboration, *Phys. Rev. Lett.* **79** 2204 (1997).

[10] R. Gupta et al., *Phys. Rev.* **D47** 5113 (1993).

[11] J. Bagger, K. Matchev and R. Zhang, JHU-TIPAC-97011, hep-ph/9707225.

[12] R. Barbieri, G. Dvali and L. Hall, *Phys. Lett.* **B377** 76 (1996).

[13] R. Barbieri, L. Hall and A. Romanino, *Phys. Lett.* **B401** 47 (1997).

[14] R. Barbieri and A. Strumia, hep-ph/9704402

Chapter 3

Atmospheric and Solar Neutrinos

3.1 Introduction

The solar and atmospheric neutrino flux anomalies have both been considerably strengthened by recent observations from Super-Kamiokande. The solar neutrino flux is measured to be [1] 0.37 ± 0.03 of that expected from the 'BP95' standard solar model [2], without including any theoretical error. This is the fifth solar neutrino experiment to report results in strong disagreement with the predictions of solar models. Furthermore, using a solar model independent analysis, the measured solar fluxes are found to be in conflict with each other. For events at SuperKamiokande with visible energies of order a GeV, the ratio of 1 ring μ-like to e-like events is 0.66 ± 0.10 that expected from calculations of the flux of neutrinos produced in the atmosphere in cosmic ray showers [3]. Furthermore, the distribution in zenith angle of these 1 ring events provides striking evidence for a depletion of ν_μ which depends on the distance travelled by the neutrinos before reaching the Super-Kamiokande detector. In particular, the observed up/down ratio of the multi-GeV, μ-like events is 0.52 ± 0.07. This significantly strengthens the evidence that ν_μ oscillate as they traverse the earth.

In this paper, we interpret the solar and atmospheric neutrino flux anomalies in terms of oscillations of the three known neutrinos $\nu_{e,\mu,\tau}$. The lightness of these three neutrinos, relative to the charged fermions, can be simply understood as resulting from large $SU(2)_L \otimes U(1)_Y$ invariant masses for the right-handed neutrinos, via the see-saw mechanism. We do not consider the possibility of a fourth light neutrino, as it would have to be singlet under $SU(2)_L \otimes U(1)_Y$, and would either require a new mass scale far below the weak scale, running counter to the idea of the see-saw mechanism, or a more complicated see-saw.

Theoretical ideas about generation mixing are guided by the quark sector, where the mixing angles are all small, indicating a hierarchical breaking of horizontal symmetries in nature. A similar hierarchy of horizontal symmetry breaking in the lepton sector is also likely to yield small angles, suggesting small probabilities for a neutrino to oscillate from one flavour to another. However, the solar and atmospheric neutrino flux measurements both require neutrino survival probabilities, P_{ee} and $P_{\mu\mu}$, far from unity. Over a decade ago [4], it was realised that large angles were not necessary to account for the large suppression of solar neutrino fluxes — while ν_e have charged current interactions in the solar medium, $\nu_{\mu,\tau}$ do not, allowing a level crossing phenomena where a ν_e state produced in the solar interior evolves to a $\nu_{\mu,\tau}$ state as it traverses the sun. This simple picture can reconcile the three types of solar neutrino flux measurements with the standard solar model, for a mixing angle as small as 0.03 — a significant achievement. Could such resonant oscillations occur for atmospheric neutrinos in the earth, again allowing a small vacuum mixing angle? In this case, since the earth does not have a continuously varying density, the matter mixing

angle in the earth is much larger than the vacuum mixing angle only in a small range of energies. Hence, an oscillation interpretation of the atmospheric neutrino fluxes requires a large mixing angle, and calls into question the frequently stated theoretical prejudice in favour of small mixing angles.

In this paper, we attempt to understand both solar and atmospheric neutrino fluxes using 3-generation neutrino oscillations, aiming at a global view which identifies the various possibilities, rather than attempting the most accurate determination of the parameters of each scenario. When data from chlorine, gallium and water Cerenkov detectors are fitted to a standard solar model, standard analyses find very small regions of neutrino mass and mixing parameters. For 2-generation mixing, these are known as the "small angle MSW", "large angle MSW" and "just so" regions. This analysis has been extended to the case of three generations [5], with a single matter resonance in the sun, as suggested by the atmospheric neutrino data. The large and small angle MSW areas are found to merge into a single MSW volume of parameter space. In subsection 3.2.1, we study how this volume is enlarged when a solar model independent analysis of the solar fluxes replaces the use of a single solar model. In subsection 3.2.2 we extend our analysis to see what areas of neutrino parameter space become allowed if one of the three observational techniques to measure the solar fluxes is seriously in error.

We combine these regions of parameters with those yielding the atmospheric fluxes, and find there is still considerable allowed ranges of masses and mixing angles. This is done in section 3.3, assuming that the smallest of the two neutrino squared mass differences is too small to affect the oscillations of atmospheric neutrinos (minimal scheme). In section 3.4,

on the contrary, we allow for the possibility that the two independent neutrino squared mass differences are both large enough to affect atmospheric neutrino oscillations (non minimal schemes). For solar neutrinos, this requires that there is a serious flaw either in at least one measurement technique or in solar model analyses.

The forms of neutrino mass matrices that can lead to a large $\nu_\mu \rightleftharpoons \nu_\tau$ mixing for atmospheric neutrinos are discussed in section 3.5. In section 3.6 only two 'zeroth order' textures for neutrinos masses are identified that can account for the atmospheric and solar neutrino data in the minimal scheme and can result from unbroken symmetries.

Our conclusions are drawn in section 3.7. Based on a simple set of alternative hypotheses, we discuss how future measurements could eventually determine the two neutrino mass differences and the three mixing angles.

3.2 Solar neutrinos: model-independent analysis

In the flavour eigenstate basis, in which the charged lepton mass matrix is diagonal, the neutrino mass matrix is in general non-diagonal. It may be diagonalized by a unitary transformation:

$$^{\circ}\nu_f = V^*_{fi}\nu_i \tag{3.2.1}$$

where ν_f and ν_i are flavour and mass eigenstate fields, respectively. The leptonic analogue of the Cabibbo-Kobayashi-Maskawa mixing matrix is V^T, since the W boson couples to the charged current $\bar{\nu}_{i_L} V^T_{if} \gamma^\mu e_{f_L}$. In addition to the three Euler angles, V contains physical phases: one if the light neutrinos are Dirac, and three if they are Majorana. These flavour

and mass eigenstate fields destroy basis states which are related by

$$|\nu_f\rangle = V_{fi}|\nu_i\rangle \tag{3.2.2}$$

If some process creates a flavour eigenstate, $|\nu_f\rangle$, at time $t = 0$, then at a later time t it will have evolved to the state $|\nu_f, t\rangle = \psi_{f'}(t)|\nu_{f'}\rangle$ via the matrix Schroedinger equation

$$i\frac{d\psi}{dt} = (V\frac{m_\nu^2}{2E}V^\dagger + A_{\rm CC} + \mathcal{E})\psi \tag{3.2.3}$$

where E is the energy of the relativistic neutrino, m_ν is the diagonal neutrino mass matrix with entries m_i, $\mathcal{E}$ is an irrelevant term proportional to the unit matrix, and $A_{\rm CC}$ represents matter effects. For neutrinos propagating in matter with electron number density N_e, $A_{\rm CC}$ is a matrix with a single non-zero entry, $A_{\rm CC}^{11} = \sqrt{2}G_F N_e$.

The mixing matrix V can be written quite generally as

$$V = R_{23}(\theta_{23}) \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{i\phi} & 0 \\ 0 & 0 & 1 \end{pmatrix} R_{13}(\theta_{13}) R_{12}(\theta_{12}) \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{i\alpha} & 0 \\ 0 & 0 & e^{i\beta} \end{pmatrix} \tag{3.2.4}$$

where $R_{ij}(\theta_{ij})$ represents a rotation by θ_{ij} in the ij plane. We have chosen a sequence of rotations which frequently arises in the diagonalization of simple hierarchical forms for the neutrino mass matrix, as illustrated in section 3.6. From equation (3.2.3) we see that the phases α and β never appear in oscillation phenomena, and hence can be dropped, giving

$$V = \begin{pmatrix} c_{12}c_{13} & c_{13}s_{12} & s_{13} \\ -c_{23}s_{12}e^{i\phi} - c_{12}s_{13}s_{23} & c_{12}c_{23}e^{i\phi} - s_{12}s_{13}s_{23} & c_{13}s_{23} \\ s_{23}s_{12}e^{i\phi} - c_{12}c_{23}s_{13} & -c_{12}s_{23}e^{i\phi} - c_{23}s_{12}s_{13} & c_{13}c_{23} \end{pmatrix}. \tag{3.2.5}$$

Each R_{ij} must diagonalize a symmetric 2×2 sub-matrix determining $\tan 2\theta_{ij}$, hence, without loss of generality, we may choose $0 \leq \theta_{ij} \leq \pi/2$, while $0 \leq \phi < 2\pi$. A more convenient

choice is to keep $\theta_{12,13}$ in the first quadrant, while $0 \leq \theta_{23}, \phi \leq \pi$. We choose to order the neutrino mass eigenstates so that $\Delta m^2_{23} > \Delta m^2_{12} > 0$, where $\Delta m^2_{ij} \equiv m^2_i - m^2_j$. Notice that with this parametrization $V_{e3} \ll 1$ means θ_{13} close to 0 or to 90°.

To study solar neutrinos, we are interested only in the electron neutrino survival probability, P_{ee}, and hence in the evolution of ψ_e. This evolution does not depend on θ_{23} or on ϕ — on substituting (3.2.4) in (3.2.3), R_{23} and ϕ can be absorbed into redefined states μ' and τ'. Hence, we have shown quite generally that P_{ee} depends only on four neutrino parameters: Δm^2_{12}, Δm^2_{23}, θ_{12} and θ_{13}.

For an oscillation explanation of the atmospheric neutrino fluxes, Δm^2_{23} is sufficiently large that it does not cause a resonance transition in the sun. In the Landau-Zehner approximation, the evolution equation (3.2.3) can be solved to give [6]

$$P_{ee} = (|V_{e1}|^2, |V_{e2}|^2, |V_{e3}|^2) \begin{pmatrix} 1-P & P & 0 \\ P & 1-P & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} |V^{\rm m}_{e1}|^2 \\ |V^{\rm m}_{e2}|^2 \\ |V^{\rm m}_{e3}|^2 \end{pmatrix} \tag{3.2.6}$$

where $V^{\rm m}_{ei}$ are the mixing matrix elements in matter, and P is the transition probability between the states at resonance:

$$P = e^{-E_{\rm NA}/E}\theta(E - E_{\rm A}), \qquad E_{\rm NA} = \frac{\pi \Delta m^2_{12} \sin^2(2\theta_{12})}{4|\frac{1}{N_e}\frac{dN_e}{dx}|_1 \cos(2\theta_{12})}, \qquad E_{\rm A} = \frac{\Delta m^2_{12} \cos 2\theta_{12}}{2\sqrt{2} G_{\rm F} |N_e|_0 \cos^2\theta_{13}} \tag{3.2.7}$$

Here E is the neutrino energy, θ is the step function, the 1 subscript indicates that N_e and its gradient dN_e/dx are evaluated at the resonance point, while the 0 subscript indicates the production point. The large mass splitting Δm^2_{23} enters P_{ee} only via the matter mixing angles, and decouples from these expressions in the limit that it is much larger than $A_{11}E$, and also in the limit that θ_{13} vanishes. For most of this section we make Δm^2_{23} sufficiently

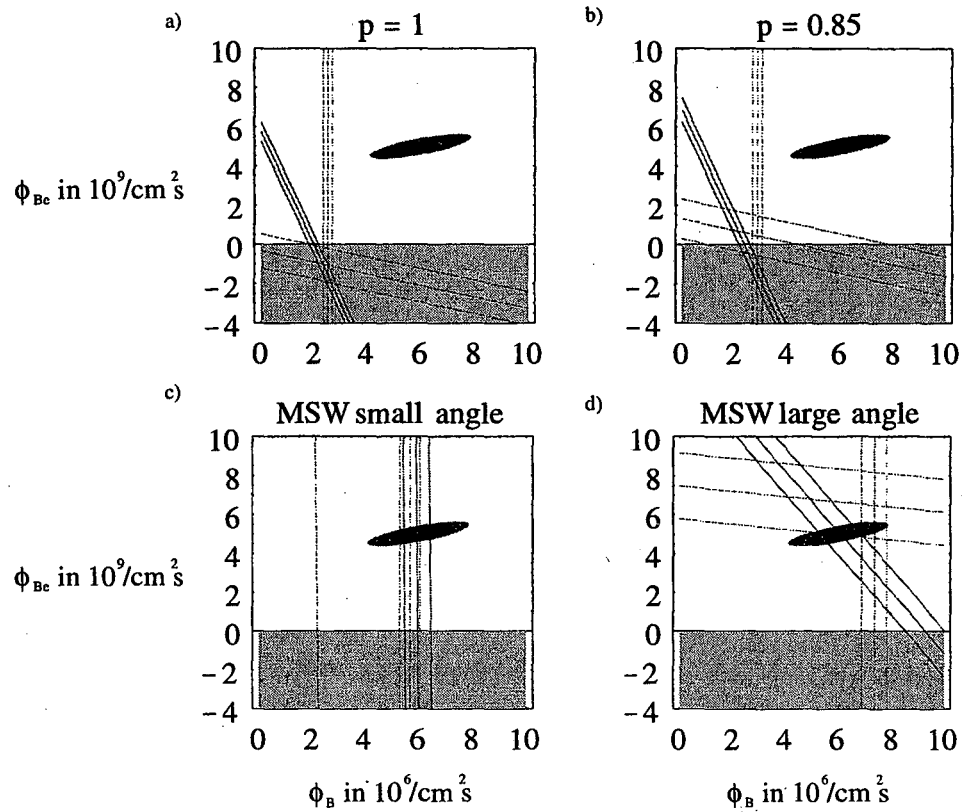

Figure 3.1: Values of $(\Phi_{^8B}, \Phi_{^7Be})$ measured by the Chlorine experiment (continuous lines), the Gallium experiment (dashed lines) and by the SuperKamiokande experiment (long dashed lines) assuming various neutrino oscillation schemes: • no oscillation in fig. 3.1a; • an energy-independent $P(\nu_e \to \nu_e) = 0.85$ in fig. 3.1b; • the best-fit point of the small-angle MSW oscillation in fig. 3.1c; • the best-fit point of the large-angle MSW oscillation in fig. 3.1d.

large that it decouples, and we comment at the end on the effect on the allowed regions of parameter space for non-zero θ_{13} and small Δm^2_{23}, where Δm^2_{23} effects may not decouple.

The signals S_i at the three types of solar neutrino experiments are

$$S_i = \int dE\, \Phi(E)[\sigma_i^e(E)P_{ee}(E) + \sigma_i^{\not e}(E)(1 - P_{ee}(E))], \quad i = \{\mathrm{SK}, \mathrm{Ga}, \mathrm{Cl}\} \tag{3.2.8}$$

where $\Phi(E)$ is the total flux of solar neutrinos with energy E, and $\sigma_i^{e,\not e}(E)$ are the interaction cross sections at experiment i for electron-type and non-electron-type neutrinos, respectively

(only the water Cerenkov detectors are sensitive to neutral currents, so $\sigma^{\ell}_{\mathrm{Ga}}(E) = \sigma^{\ell}_{\mathrm{Cl}}(E) = 0$). We will use the theoretical predictions of the various cross sections found in [7, 8]. The flux $\Phi(E)$ is broken into components in the standard way by specifying the production reaction, giving [7]

$$\Phi(E) = \sum_{\alpha} \Phi_{\alpha} f_{\alpha}(E), \qquad \text{with} \qquad \int_0^{\infty} f_{\alpha}(E)\, dE = 1 \tag{3.2.9}$$

and $\alpha = \mathrm{pp}, \mathrm{pep}, {}^{7}\mathrm{Be}, {}^{13}\mathrm{N}, {}^{15}\mathrm{O}, {}^{17}\mathrm{F}, {}^{8}\mathrm{B}, \mathrm{hep}$. At this point we follow the (nearly) model-independent treatment of the fluxes described in [9] by making the following assumptions:

1. The energy dependence $f_{\alpha}(E)$ of the single components of the neutrino fluxes predicted by solar models ([7, 2] for instance) are correct. In fact the $f_{\alpha}(E)$ do not depend on the structure of the sun, and are the same in any solar model that does not introduce non-standard electroweak effects [7].

2. The overall Φ_{α} can differ from their solar models predictions. However there are strong physical reasons to believe that the ratios $\Phi_{{}^{13}\mathrm{N}}/\Phi_{{}^{15}\mathrm{O}}$ and $\Phi_{\mathrm{pep}}/\Phi_{\mathrm{pp}}$ can be set to their solar SM values [2]. Furthermore we neglect entirely hep and ${}^{17}\mathrm{F}$ neutrinos, which we expect to be extremely rare.

3. The present total luminosity of the sun, $K_{\odot}$, determines its present total neutrino luminosity as

$$K_{\odot} = \sum_{\alpha} \left(\frac{Q}{2} - \langle E_{\nu_{\alpha}} \rangle \right) \Phi_{\alpha} \approx \frac{Q}{2} \sum_{\alpha} \Phi_{\alpha}, \tag{3.2.10}$$

where $Q = 26.73$ MeV is the energy released in the reaction $4\mathrm{p} + 2e \rightarrow {}^{4}\mathrm{He} + 2\nu_e$, and $K_{\odot} = 8.53 \cdot 10^{11}$ MeV cm^{-2} s^{-1} is the solar radiative flux at the earth. Using (3.2.10) amounts to assuming that the solar energy comes from nuclear reactions that

reach completion, and that the sun is essentially static over the 10^4 years employed by photons to random-walk out of the solar interior.

After the first assumption we have one free parameter Φ_α for each α; the second then reduces the number of free parameters to four, which we can take to be

$$\Phi_{\rm p} \equiv \Phi_{\rm pp} + \Phi_{\rm pep}, \qquad \Phi_{\rm CNO} \equiv \Phi_{^{13}{\rm N}} + \Phi_{^{15}{\rm O}}, \qquad \Phi_{^7{\rm Be}} \qquad \text{and} \qquad \Phi_{^8{\rm B}}. \tag{3.2.11}$$

The luminosity constraint allows us to eliminate $\Phi_{\rm p}$, giving

$$S_i = S_i(\Delta m^2_{12},\, \theta_{12},\, \theta_{13}; \Phi_{^8{\rm B}}, \Phi_{^7{\rm Be}}, \frac{\Phi_{\rm CNO}}{\Phi_{^7{\rm Be}}}). \tag{3.2.12}$$

Since solar models give a stable prediction for $\Phi_{\rm CNO}/\Phi_{^8{\rm B}} = 0.22$ [9], we have singled out this ratio and we will use its SSM value in our analysis. Variations of even an order of magnitude in the ratio affect negligibly our final results, since the two neutrino components have similar cross sections in existing detectors.

3.2.1 Model-independent solar analysis — all experiments

The signals now depend only on $\Phi_{^8{\rm B}}$ and $\Phi_{^7{\rm Be}}$, so that, for any given oscillation pattern $P_{ee}(E)$ it is possible to plot the three experimental results* [1, 10, 11, 12]

$$S^{\rm exp}_{\rm Cl} = (2.54 \pm 0.20)\ 10^{-36}{\rm s}^{-1} \tag{3.2.13a}$$

$$S^{\rm exp}_{\rm Ga} = (75 \pm 7)\ 10^{-36}{\rm s}^{-1} \tag{3.2.13b}$$

$$S^{\rm exp}_{\rm SK} = (2.51 \pm 0.16) \cdot 10^6\,{\rm cm}^{-2}{\rm s}^{-1} \tag{3.2.13c}$$

as three bands in the ($\Phi_{^8{\rm B}}$, $\Phi_{^7{\rm Be}}$) plane. The three bands will in general not meet, giving interesting solar model independent restrictions on the oscillations parameters.

*The SuperKamiokande experimentalists give directly the value of the flux they measure. The other experiments involve more uncertain neutrino cross sections and prefer to give the frequency of events measured per target atom in their detector. For simplicity we have omitted this detail in the text, leaving a trivial inconsistency between eq. (3.2.13c) and (3.2.8).

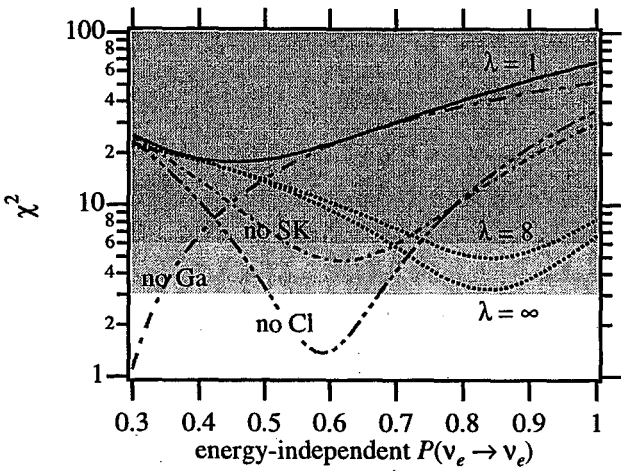

Figure 3.2: Values of the χ-square as function of an energy independent $P(\nu_e \rightarrow \nu_e)$. The parameter λ is defined in eq. (3.2.14). Also shown is the χ^2 with one experiment discarded and $\lambda = 1$.

We begin the analysis by studying the case of *no neutrino oscillations* ($P_{ee} = 1$). In this particular case the solar model independent analysis does not give a strong result. Surprisingly the three bands perfectly meet [9, 13] as shown in fig. 3.1a, but mainly in the unphysical $\Phi_{^7\mathrm{Be}} < 0$ region, with a small area in the physical region lying within 2σ of each central value. Since the physical crossing region has a negligible ^{7}Be flux, the value of $\Phi_{\mathrm{CNO}}/\Phi_{^7\mathrm{Be}}$ becomes completely irrelevant.

To discuss this case in a quantitative way and to deal with more general cases it is useful to introduce the χ-square function

$$\chi^2_\lambda \left(P_{ee}(\Delta m^2_{12}, \theta_{12}, \theta_{13}), \Phi_{^8\mathrm{B}}, \Phi_{^7\mathrm{Be}}\right) \equiv \sum_i \left(\frac{S_i - S_i^{\mathrm{exp}}}{\Delta S_i^{\mathrm{exp}}}\right)^2 + \sum_{jk} \frac{(\Phi_j - \Phi_j^{\mathrm{SSM}})(\Phi_k - \Phi_k^{\mathrm{SSM}})}{\lambda^2 \Delta\Phi^{2\,\mathrm{SSM}}_{jk}} \tag{3.2.14}$$

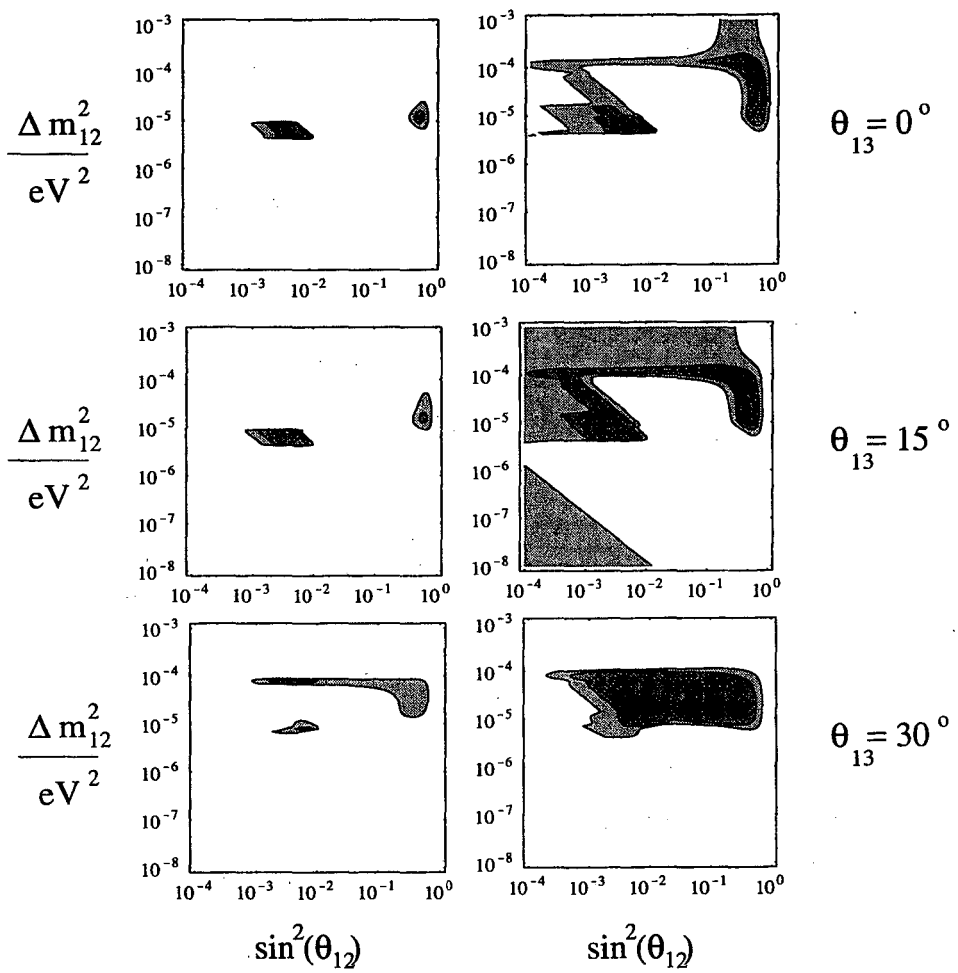

Figure 3.3: Allowed regions in the plane $(\sin^2 2\theta_{12}, \Delta m^2_{12})$ for $\theta_{13} = 0, 15°$ and $30°$. The plots on the left assume that the BP solar model is correct. The plots on the right are the result of the solar model independent analysis described in the text.

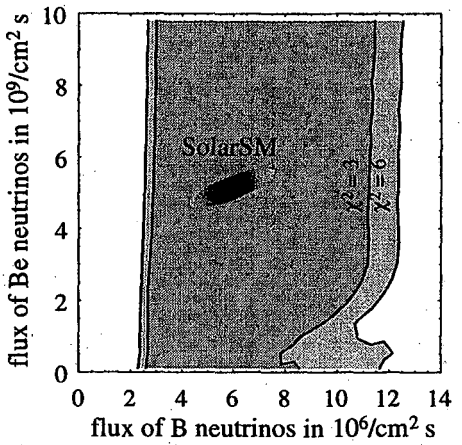

Figure 3.4: Isoplot of χ^2_8, minimized in the mixing parameters.

where $\Delta S_i^{\rm exp}$ is the 1σ uncertainty for experiment i, given in (3.2.13), $\Phi^{\rm SSM}$ is the flux prediction of the solar model [2] and $\Delta\Phi^{\rm SSM}$ is the corresponding error matrix, taken with some generosity. The 1σ ranges of $\Phi_{^8{\rm B}}$ and $\Phi_{^7{\rm Be}}$ are represented by the ellipse in fig. 3.1. We perform our analysis with two choices for $\Delta\Phi = \lambda \cdot \Delta\Phi^{\rm SSM}$. We call the first choice, $\Delta\Phi = \Delta\Phi^{\rm SSM}$, "*solar SM inspired*"). The second choice, $\Delta\Phi = 8 \cdot \Delta\Phi^{\rm SSM}$ ("*model independent*") has the same shape as the first, but is eight times as large. The part of the analysis done using this $\Delta\Phi$ is virtually free of solar physics input. The choice $\lambda = 8$ (rather than $\lambda = \infty$) avoids unnatural values of $\Phi_{^7{\rm Be}}$. This choice essentially ignores solar physics considerations, but the virtue of having a number of independent experimental results is precisely that we need no longer rely heavily on solar modelling to gain insight into the underlying particle physics.

Minimizing the χ^2 in the positive flux region we obtain $\min\chi^2_8(P_{ee} = 1) = 8.25$. The usual criterion for goodness of fit says that a χ^2 with one degree of freedom larger than 8.25 is obtained with a very small probability, $\wp \approx 0.4\%$ (a careful Monte Carlo treatment

of the $\Phi_{^7\mathrm{Be}} > 0$ constraint gives similar results [13]). We however remark that, if the sun really emits the best-fit fluxes, $\Phi_{^7\mathrm{Be}} = 0$ and $\Phi_{^8\mathrm{B}} = 2.5\ 10^6/\mathrm{cm}^2\mathrm{s}$, there is a 10% probability that statistical fluctuations produce the present experimental data.

We can just as easily investigate the slightly more general case of an energy independent P_{ee}. The dependence on the neutrino parameters Δm^2_{12}, θ_{12}, and θ_{13} arises through P_{ee}; if the survival probability is a constant, then we can minimize χ^2_λ in the positive-flux region for any value of P_{ee} to obtain $\min \chi^2_\lambda(P_{ee})$, which is plotted in fig. 3.2 for $\lambda = 1$ (SSM analysis), $\lambda = 8$ (SSM independent analysis) and $\lambda = \infty$ (completely model independent analysis). For $P_{ee} \sim 0.85$, $\min \chi^2_8$ drops to 5, but the fluxes required to get relatively small χ^2 values are disfavoured by solar physics considerations — Φ_{CNO} and $\Phi_{^7\mathrm{Be}}$ must be nearly made to vanish, as shown in fig 3.1b. When $P_{ee} \lesssim 1/2$ the (accidental?) threefold crossing no longer occurs, so that this case can be firmly excluded in a solar-model independent way [9, 14] (see fig. 3.2). However, as we shall see in subsection 3.2.2, once we allow for the possibility that one type of experiment's results should be discarded, it is possible to obtain good fits of the data for constant $P_{ee} \sim 1/2$ without having to resort to unnatural flux values.

Of course, we are interested in any points in parameter space that fit the data well, regardless of whether they lead to constant P_{ee}. For any values of Δm^2_{12}, θ_{12}, and θ_{13} we can make plots similar to fig. 3.1a. Fig.s 3.1c and 3.1d show two examples that illustrate the familiar 2-generation small and large angle MSW solutions, which evidently fit the data well if standard solar model fluxes are used.

In fig. 3.3 we show how the allowed regions in neutrino parameter space change

if we let the fluxes vary over an expanded range of values. For each point in (Δm^2_{12}, θ_{12}, θ_{13}) space, we minimize χ^2_1 and χ^2_8 by varying the fluxes within the physical region, and then we plot contours of $\min \chi^2_\lambda$ in the ($\sin^2(2\theta_{12})$, Δm^2_{12}) plane for various values of θ_{13}. The results for the "SSM inspired" and "model independent" analyses are shown in fig.s 3.3 (upper row and lower row, respectively). The contours are for $\chi^2 = 3$ and $\chi^2 = 6$.

For small θ_{13} the "*SSM inspired*" results show the standard small and large angle MSW regions. For larger values of θ_{13}, the two MSW regions join, and, as θ_{13} approaches $\pi/4$, the solutions with large θ_{12} disappear. For $\theta_{13} = \pi/4$ the region with $\min \chi^2_1 < 3$ is in fact absent entirely.

The "*model independent*" results similarly exhibit a very strong θ_{13} dependence. We see that the "model independent" analysis continues to give strong restrictions of the oscillation parameters — in particular the Δm^2_{12} values with $\min \chi^2_8 < 3$ are always in the range $\sim 10^{-(4\div5)}\,\mathrm{eV}^2$. This will not remain true when we consider the consequences of ignoring one experiment's data in subsection 3.2.2.

If $\Phi_{\mathrm{CNO}}/\Phi_{^8\mathrm{B}}$ is ten times larger than in SSM there are new allowed regions. However these possible new regions, with $\Delta m^2 = 10^{-(5\div6)}\,\mathrm{eV}^2$ and $\sin^2 2\theta_{12} \gtrsim 10^{-2}$, are excluded in a model-independent way by the non observation of a day/night asymmetry at SuperKamiokande [1, 15]. The recent data [1] on this asymmetry in fact disfavour as well the large angle MSW solution of the SSM-inspired analysis. Moreover, we have not included in our χ^2 analysis the SuperKamiokande measurement of the distortion of the ^{8}B spectrum [1, 15], because the present positive 1σ signal could be produced by a $\Phi_{\mathrm{hep}}/\Phi_{^8\mathrm{B}}$ ratio 15 times larger than the prediction of BP95 [2]. Without a very large hep flux, the present

measurement excludes an otherwise allowed region with $\Delta m^2 \approx 10^{-4}\,\mathrm{eV}^2$ and $\sin^2 2\theta_{12}$ in the range $10^{-4} \div 10^{-1}$ [13, 16].

Our model independent analysis allows us to investigate how well present experiments are able to measure the SSM-independent neutrino fluxes $\Phi_{^8\mathrm{B}}$ and $\Phi_{^7\mathrm{Be}}$. This question is answered in fig. 3.4, where we plot the values of the fluxes that can give a good ($\chi^2_8 < 6$) or very good ($\chi^2_8 < 3$) fit for *some* value of the oscillation parameters Δm^2_{12}, θ_{12} and θ_{13}. We see that the value of $\Phi_{^8\mathrm{B}}$ is currently determined with an error larger than the solar model expectation. It will be directly measured in the new on-going SNO experiment. On the contrary the value of $\Phi_{^7\mathrm{Be}}$ is at present totally unknown: in fact in the small angle MSW solution the monochromatic ^{7}Be flux can be completely converted into $\not{e}$ neutrinos, that are not detected by existing experiments. Borexino will be able to detect neutral currents effects in this range of energies and probably allow a direct determination of $\Phi_{^7\mathrm{Be}}$ [13].

As discussed above we perform our analysis under the assumption that Δm^2_{23} is large enough that its effects decouple. For any given Δm^2_{23} it is straightforward to reproduce fig. 3.3 by using the exact expressions for θ^{m}_{12} and θ^{m}_{13} in equation (3.2.6). In this way we find that for small θ_{13} ($\lesssim 15°$), our results are insensitive to Δm^2_{23} down to $\Delta m^2_{23} = 5 \cdot 10^{-4}\,\mathrm{eV}^2$. For large θ_{13}, Δm^2_{23} effects start to become noticeable when Δm^2_{23} drops below $2 \times 10^{-3}\,\mathrm{eV}^2$; for example, for $\theta_{13} = 40°$ and $\Delta m^2_{23} = 5 \cdot 10^{-4}\,\mathrm{eV}^2$, the allowed region in the SM inspired analysis is significantly smaller than in the decoupled limit, with the $\chi^2_{\mathrm{min}} < 6$ region never reaching $\sin^2(2\theta_{12}) > 0.1$. In spite of these changes for small Δm^2_{23}, the essential features of fig. 3.3 in any case remain unchanged.

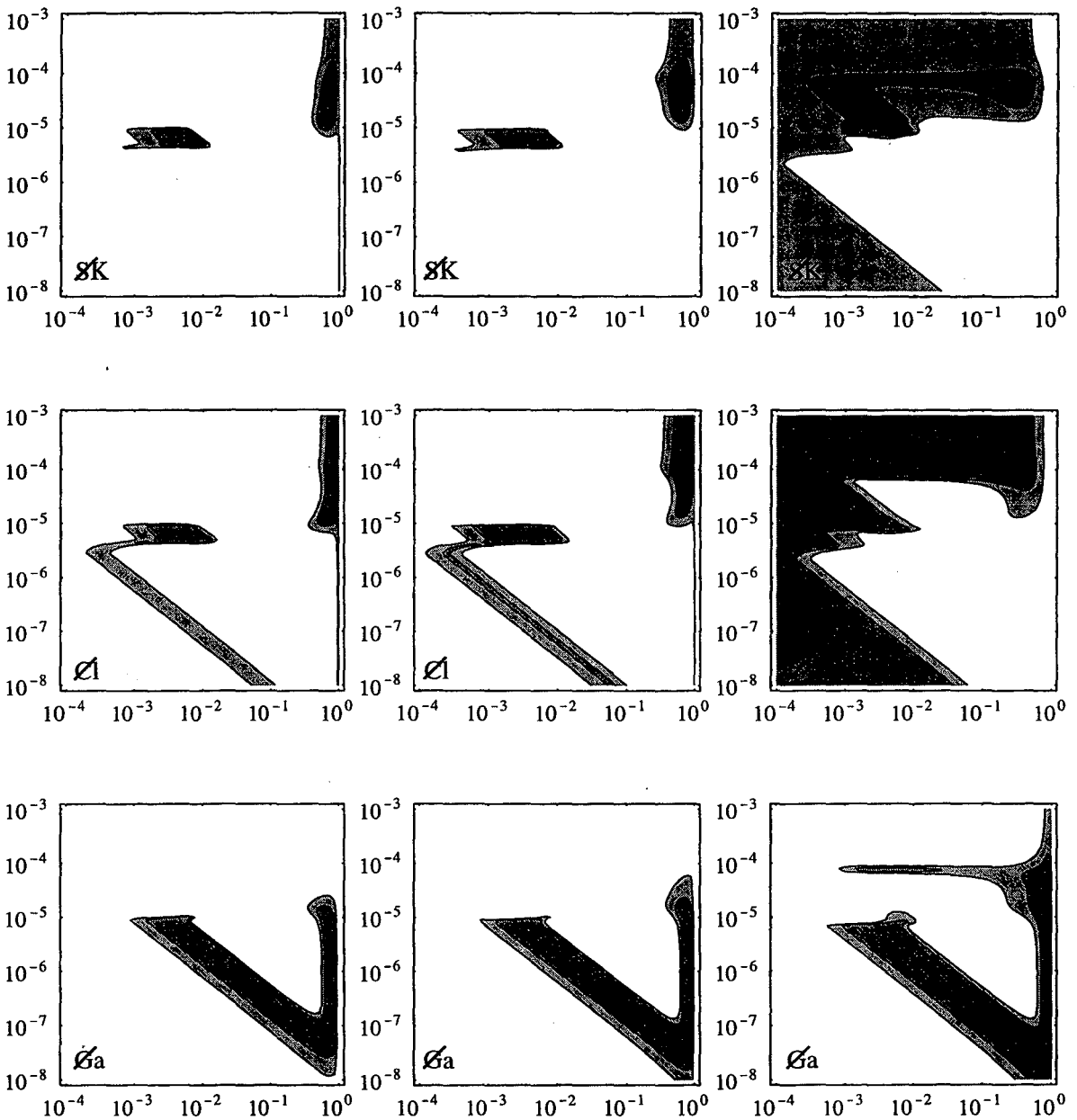

Figure 3.5: Fits of the solar data in the plane $(\sin^2 2\theta_{12}, \Delta m^2_{12}/\,\mathrm{eV}^2)$ for $\theta_{13} = 0$, 15° and 30° assuming that one of the three solar neutrino experiments has a large unknown systematic error (SuperKamiokande in the first row, Chlorine in the second and Gallium in the third) and is therefore discarded from the analysis. The contours are for $\chi^2 = 3$ and $\chi^2 = 6$.

3.2.2 Model independent solar analysis — one experiment ignored

In subsection 3.2.1, the present level of experimental evidence allowed us to omit one restriction (the solar model) and still yield interesting results. Likewise, we can choose to omit one experiment from the analysis while keeping some solar information and still yield interesting results.

The motivation for this is obvious: neutrino experiments are extremely difficult to perform and particular detection schemes may suffer from some systematic error previously not considered. We make no judgements here about the errors associated with any particular experiment. Instead we consider analyses where we do not include one class of experiment, either water-Cerenkov, gallium or chlorine, which we designate $\not{\text{SK}}$, $\not{\text{Cl}}$ and $\not{\text{Ga}}$ respectively. However, because we are losing an experiment, it is impossible to analyze the data without some level of information regarding the solar model. Consequently, we perform the analysis only within the solar SM inspired region. The results for this analysis are shown in figures 3.5 (upper row: without SuperKamiokande data, middle row: without chlorine data, and lower row: without gallium data).

The $\not{\text{SK}}$ case largely resembles the complete data set analysis, with some additional space allowed in the higher Δm^2_{12} region. In contrast, the other two cases ($\not{\text{Ga}}$ and $\not{\text{Cl}}$) show considerable differences.

For the $\not{\text{Ga}}$ case, there is a strong preference for either small Δm^2_{12} or large θ_{12} and θ_{13}. For the $\not{\text{Cl}}$ case, for both large θ_{12} and θ_{13} we have the presence of large regions with large $\Delta m^2_{12} = 10^{-4}\,\text{eV}^2$, above the level-crossing threshold, and with small Δm^2_{12}, in the non-adiabatic region. In either case, in a large portion of these regions matter

enhancements are unimportant. That is, in the absence of one of these two classes of experiment, given sufficiently large angles, the solar neutrino problem can be resolved *simply by vacuum oscillations alone!* In such a case, new experiments, such as Borexino, would see an absence of energy dependence in the electron neutrino survival probability.

3.3 Atmospheric and Solar Neutrinos: The Minimal Scheme

The simplest picture for reconciling both solar and atmospheric neutrino fluxes via oscillations of $\nu_{e,\mu,\tau}$ results when there is a hierarchy $|\Delta m^2_{23}| \gg |\Delta m^2_{12}|$, and Δm^2_{12} is too small to affect oscillations of atmospheric neutrinos. In section 3.2, we showed that in this case the solar fluxes depend only on Δm^2_{12}, θ_{12} and θ_{13}[†], and below we show that the atmospheric fluxes depend only on $\Delta m^2_{23}, \theta_{23}$ and θ_{13}. In the limit that $\theta_{13} = 0$, the two phenomena become independent, in the sense that they depend on no common parameters: solar oscillations are $\nu_e \to \nu_\mu$ at a low frequency, while atmospheric oscillations are $\nu_\mu \to \nu_\tau$ at a much higher frequency. However, solar oscillations are allowed for a wide range of parameters with large θ_{13}, and the atmospheric data does not require θ_{13} to be very small. Hence, in this section we explore this simple picture keeping θ_{13} as a free parameter. We comment on the alternative possibility — that Δm^2_{12} is large enough to contribute to atmospheric neutrino oscillations — in section 3.4.

Matter effects in the earth are important only for a relatively small fraction of the atmospheric neutrinos, those with high energy, and they are neglected here[‡]. In this case, (3.2.3) can be integrated to give oscillation probabilities $P_{ff'}(t) = |A_{ff'}(t)|^2$, where A is

[†]Although for non-zero θ_{13}, there is a dependence on Δm^2_{23} if it is small enough.
[‡]For more details see e.g. ref. [17].

given by the matrix equation

$$A(t) = Ve^{-iEt}V^{\dagger}. \tag{3.3.1}$$

Since an overall phase in A is irrelevant to P, and Δm^2_{12} effects are negligible, we may make the substitution

$$e^{-iEt} \longrightarrow \mathrm{diag}\,(1, 1, e^{-i\Delta m^2_{23}t/2E}) \tag{3.3.2}$$

Using the form (3.2.4) for V, we immediately discover that the probabilities are independent of θ_{12} and ϕ, as well as α and β. The probabilities are given by

$$P_{e\mu} = s^2_{23} \sin^2 2\theta_{13}\, S_{23} \tag{3.3.3a}$$

$$P_{e\tau} = c^2_{23} \sin^2 2\theta_{13}\, S_{23} \tag{3.3.3b}$$

$$P_{\mu\tau} = c^4_{13} \sin^2 2\theta_{23}\, S_{23} \tag{3.3.3c}$$

or equivalently, by unitarity

$$P_{ee} = 1 - \sin^2 2\theta_{13}\, S_{23} \tag{3.3.3d}$$

$$P_{\mu\mu} = 1 - 4c^2_{13}s^2_{23}(1 - c^2_{13}s^2_{23})\, S_{23} \tag{3.3.3e}$$

$$P_{\tau\tau} = 1 - 4c^2_{13}c^2_{23}(1 - c^2_{13}c^2_{23})\, S_{23} \tag{3.3.3f}$$

where $S_{23} = \sin^2(\Delta m^2_{23}t/4E)$. The parameter Δm^2_{23} can be extracted from the data by fitting to the zenith angle distribution of the events. Here we concentrate on the determination of the parameters θ_{13} and θ_{23}. These can be extracted, independent of the value of Δm^2_{23}, if we assume that the downward going neutrinos have not oscillated, while the upward going neutrinos are completely oscillated, so that S_{23} is averaged to 0.5. In view of the reported angular distribution of the multi-GeV data for 1-ring e-like, 1-ring μ-like

and partially contained (PC) events [3], this assumption appears to be valid, at least for angular cone sizes about the vertical which are not too large. For events of class i, which are induced by ν_e charged current, ν_μ charged current and neutral current interactions with relative probabilities $f^i_{eCC}, f^i_{\mu CC}$ and f^i_{NC}, the up-down ratio ρ_i is given by

$$\rho_i = \frac{N_i^\uparrow}{N_i^\downarrow} = f^i_{eCC} \cdot (P_{ee} + rP_{e\mu}) + f^i_{\mu CC} \cdot (P_{\mu\mu} + \frac{1}{r}P_{e\mu}) + f^i_{NC}. \tag{3.3.4}$$

where we have set $S_{23} = 0.5$, and $N_i^{\uparrow,\downarrow}$ are the number of upward and downward events of class i. We are interested in i being 1-ring e-like, 1-ring μ-like and PC. The overall normalization of these event numbers has considerable uncertainties due to the calculation of the neutrino fluxes produced in cosmic ray showers, hence we consider three up-down ratios

$$\rho_e = 1.23 \pm 0.29 \tag{3.3.5a}$$

$$\rho_\mu = 0.62 \pm 0.16 \tag{3.3.5b}$$

$$\rho_{PC} = 0.48 \pm 0.12 \tag{3.3.5c}$$

and two ratios of downward going fluxes

$$\frac{N_\mu^\downarrow + N_{PC}^\downarrow}{N_e^\downarrow} = \xi r = 3.0 \pm 0.6 \tag{3.3.5d}$$

$$\frac{N_{PC}^\downarrow}{N_\mu^\downarrow} = \xi' = 1.3 \pm 0.3 \tag{3.3.5e}$$

where r is the ratio of ν_μ to ν_e fluxes. The numbers give the Super-Kamiokande data, extracted from the figures of Ref. [3], with upward and downward directions defined by the azimuthal angle having $\cos\theta$ within 0.4 of the vertical direction. The parameters ξr and ξ' represent the theoretical values for the ratios of (3.3.5d) and (3.3.5e). These two downward

going ratios do not involve oscillations, and the Super-Kamiokande collaboration compute Monte Carlo values of 3.1 and 1.0, respectively, agreeing very well with the data. Since these two ratios do not probe oscillations, at least within our assumptions, we do not use them for the fits below. We do not use the sub-GeV data as the poor angular correlation between the neutrino and charged lepton directions leads to a smoothing of the up-down ratio. From the flux calculations of Honda *et al* [18], and using the measured momentum distributions for the events [3], we estimate $r = 4.0 \pm 0.5$, for this multi-GeV data near the vertical direction. A more refined analysis would use a larger value of r for PC events than for FC events.

The results of a fit of the three up/down ratios to the two free parameters θ_{23} and θ_{13} are shown in figure 3.6(a). We have obtained the fractions $f^i_{eCC,\mu CC,NC}$ from the Monte Carlo results of the Super-Kamiokande collaboration [3], and we have used the oscillation probabilities of (3.3.3c). In order to work with Gaussian distributed experimental data, we have directly fitted the six measured neutrino numbers $N_i^{\uparrow,\downarrow}$ leaving arbitrary the three overall fluxes of each type, $N_i^{\uparrow} + N_i^{\downarrow}$. The preferred region of the plot is easy to understand, since at the point $\theta_{23} = 45°$ and $\theta_{13} = 0$, the ν_e are unmixed, while there is complete $\nu_\mu \leftrightarrow \nu_\tau$ mixing, so $\rho_e \simeq 1$ and $\rho_\mu \simeq \rho_{PC} \simeq 0.5$. It is apparent from Fig. 3.6(a) that this minimal scheme is allowed for a large range of angles about this point: $\theta_{23} = 45° \pm 15°$ and $\theta_{13} = 0 \div 45°$.

If the solar neutrino fluxes, measured by all three techniques, are to agree with solar model inspired values, then the results of section 3.2 show that Δm^2_{12} is too small to affect atmospheric oscillations, it is either of order $10^{-4} - 10^{-5}\,\mathrm{eV}^2$ or of order $10^{-10}\,\mathrm{eV}^2$.

In this case, the minimal scheme for atmospheric neutrinos, described in this section, is the unique possibility using just the three known neutrinos. This observation enhances the importance of the fit of figure 3.6(a); further data will reduce the allowed region, as the three up-down ratios of (3.3.5c) have small systematic uncertainties and are statistics limited. The solar neutrino fluxes do not put extra constraints on the value of θ_{13}, although it becomes correlated with θ_{12}, as shown in figure 3.3. If the atmospheric flux measurements require $\Delta m^2_{23} > 2 \times 10^{-3}\,\mathrm{eV}^2$, then the limit on P_{ee} from the CHOOZ experiment [19] requires $\theta_{13} < 13°$.

Recent analyses [20] of SuperKamiokande data that make use of MonteCarlo predictions for the angular and energy distributions of the atmospheric neutrinos get more stringent constraints on the neutrino oscillation parameters. Our fit uses only those data — the ratio of upward and downward multi-GeV neutrinos (the ones in bins 1 and 5 of the angular distribution in [3]) — that do not depend on the spectrum of the atmospheric neutrinos nor on the precise value of Δm^2, assuming a full averaged oscillation in between. Since statistics gives presently the dominant error, we obtain weaker constraints than in [20]. If we knew that the neutrino mass difference relevant for atmospheric neutrinos were close to the center of the presently allowed region, we could add to the data to be fitted the intermediate bins 2 and 4 of [3] (the bins that contain 'oblique' neutrinos). We cannot use in any case the multi-GeV data in the intermediate bin 3, that contains 'horizontal' neutrinos. Having doubled the statistics, we would find the more stringent contours shown in fig. 3.6a′, b1′, b2′. We remind the reader that our fit does not include matter effects [4].

3.4 Atmospheric and Solar Neutrinos: Non-Minimal Schemes

In this section, we study atmospheric neutrinos when two conditions apply.

- The smallest mass splitting is large enough to affect atmospheric neutrino oscillations: $\Delta m^2_{12} > 3 \times 10^{-4}\,\mathrm{eV}^2$. For solar neutrinos, this implies that there is a serious flaw either in at least one measurement technique, or in the solar models.

- The mass splittings are hierarchical $\Delta m^2_{23} \gg \Delta m^2_{12}$. This is a simplification, which we relax at the end of the section. It includes the interesting possibility that Δm^2_{23} is large enough to induce the apparent oscillations reported by the LSND collaboration [21], while Δm^2_{12} effects are causing both solar and atmospheric oscillations.

Using (3.2.5), the ν_e survival probability is

$$P_{ee} = 1 - c^4_{13} \sin^2 2\theta_{12}\, S_{12} - s^2_{12} \sin^2 2\theta_{13}\, S_{23} - c^2_{12} \sin^2 2\theta_{13}\, S_{31} \qquad (3.4.1)$$

where $S_{ij} = \sin^2(\Delta m^2_{ij} t/4E)$. The above two conditions imply that $\Delta m^2_{23} > 2 \times 10^{-3}\,\mathrm{eV}^2$, so that, for the CHOOZ experiment, (3.4.1) should be used with $S_{23} = S_{31} = 0.5$. The CHOOZ limit, $P_{ee} > 0.9$, then gives $\theta_{13} < 0.23$. If Δm^2_{12} were also greater than $2 \times 10^{-3}\,\mathrm{eV}^2$, then for the CHOOZ experiment one also has $S_{12} = 0.5$, so that $\theta_{12} < 0.23$. However, in this case the survival probability for solar neutrinos is the same as for the anti-neutrinos at CHOOZ: $P_{ee} > 0.9$. Hence, given our two conditions, the observed solar neutrino fluxes require $\Delta m^2_{12} < 2 \times 10^{-3}\,\mathrm{eV}^2$.

It is frequently stated that the three known neutrinos cannot explain the LSND, atmospheric and solar neutrino anomalies, as this would require three Δm^2 with different orders of magnitudes. However, this argument no longer applies in the case that either a

solar neutrino measurement technique or solar models are incorrect, when a single Δm^2 could give both atmospheric and solar anomalies. Hence, we consider first the case that Δm^2_{23} is large enough to explain the observations of LSND. The oscillation probabilities induced by Δm^2_{23} are given by (3.3.3c). From the limit on P_{ee} from the Bugey reactor, one then concludes

$$\Delta m^2_{23} \;>\; 0.2\,\mathrm{eV}^2, \tag{3.4.2a}$$

and

$$\theta_{13} \;<\; 0.1 \tag{3.4.2b}$$

which is significantly stronger than the CHOOZ limit. A second possibility, θ_{13} close to 90° does not allow any significant oscillations of ν_e and is thus not acceptable to explain the solar neutrino anomaly at a relatively large frequency. For atmospheric neutrinos, both upward going and downward going, one may then use oscillation probabilities with $\theta_{13} = 0$[§], and with S_{23} and S_{13} both averaged to 0.5:

$$P_{e\mu} \;=\; c^2_{23} \sin^2 2\theta_{12}\; S_{12} \tag{3.4.3a}$$

$$P_{e\tau} \;=\; s^2_{23} \sin^2 2\theta_{12}\; S_{12} \tag{3.4.3b}$$

$$P_{\mu\tau} \;=\; -\tfrac{1}{4} \sin^2 2\theta_{23} \sin^2 2\theta_{12}\; S_{12} + \tfrac{1}{2} \sin^2 2\theta_{23} \tag{3.4.3c}$$

or equivalently, from unitarity

$$P_{ee} \;=\; 1 - \sin^2 2\theta_{12}\; S_{12} \tag{3.4.3d}$$

$$P_{\mu\mu} \;=\; 1 - \tfrac{1}{2} \sin^2 2\theta_{23} - c^4_{23} \sin^2 2\theta_{12}\; S_{12} \tag{3.4.3e}$$

[§]In which case the P_{ij} are independent of ϕ.

$$P_{\tau\tau} \;=\; 1 - \tfrac{1}{2}\sin^2 2\theta_{23} - s_{23}^4 \sin^2 2\theta_{12}\, S_{12}. \qquad (3.4.3\mathrm{f})$$

Since in these formulæ $S_{13} = S_{23}$, θ_{12} enters only via $\sin^2 2\theta_{12}$ so that, without loss of generality, we may reduce the range of θ_{12} to $0 \le \theta_{12} < \pi/4$. We again study the up-down ratios (3.3.4), as they have small systematic uncertainties. We calculate them approximately, using (3.3.4) with $f^e_{eCC} = f^\mu_{\mu CC} = f^{PC}_{\mu CC} = 1$ and all other f-factors equal to zero. A fraction, $P^{(0)}_{\mu\mu} = 1 - \sin^2 2\theta_{23}/2$, of the downward going ν_μ oscillate to ν_τ before detection, so the up-down ratios are given by

$$\rho_e \approx P_{ee} + r P_{e\mu} \qquad (3.4.4)$$

and

$$\rho_\mu \approx \frac{P_{\mu\mu} + P_{e\mu}/r}{P^{(0)}_{\mu\mu}}. \qquad (3.4.5)$$

Hence we find

$$(\rho_\mu - 1) \approx -\frac{1}{r}\frac{c_{23}^2}{1 - \frac{1}{2}\sin^2 2\theta_{23}} \cdot (\rho_e - 1). \qquad (3.4.6)$$

For the multi-GeV data, where the angular correlation is best, r is large, and (3.4.6) implies that $|\rho_\mu - 1| < (1/3)|\rho_e - 1|$, in strong disagreement with data of (3.3.5c). The same inequality holds if ρ_μ is replaced by ρ_{PC}, when the disagreement with data is even stronger.[¶] *With oscillations of the three known neutrinos, the LSND observation conflicts with the atmospheric and solar neutrino anomalies even using a model independent analysis of the solar neutrino fluxes or allowing for a systematic error in one of the solar neutrino experiments*[‖].

[¶] Even ignoring ρ_e, we find $\rho_{PC,\mu} > 0.61$.

[‖] For $\theta_{23} = 0$, this corresponds to purely $\nu_\mu \to \nu_e$ oscillations, which is therefore excluded as an explanation of the atmospheric neutrino measurements.

Does the atmospheric neutrino data allow other values of $\Delta m^2_{23} \gg \Delta m^2_{12}$? The limit from the Bugey reactor, (3.4.2b), applies for all $\Delta m^2_{23} > 0.06\,\mathrm{eV}^2$, and the up-down ratio relation, (3.4.6), applies for all $\Delta m^2_{23} > 0.1\,\mathrm{eV}^2$. Hence, $\Delta m^2_{23} > 0.1\,\mathrm{eV}^2$ is excluded. For $\Delta m^2_{23} < 0.1\,\mathrm{eV}^2$, the downward going ν_μ have not oscillated to ν_τ when they reach the Super-Kamiokande detector, so that (3.4.5) is replaced by

$$\rho_\mu \approx P_{\mu\mu} + \frac{1}{r}P_{e\mu} = 1 - \frac{1}{2}\sin^2 2\theta_{23} - \frac{c^2_{23}}{r}(\rho_e - 1). \tag{3.4.7}$$

Consistency with the data, (3.3.5c), is now possible, and requires large θ_{23}. As Δm^2_{23} drops below $0.06\,\mathrm{eV}^2$, the limit from the Bugey reactor on θ_{13} is progressively weakened, so that θ_{13} terms must be kept in P_{ij}. Furthermore, as Δm^2_{23} drops below $0.01\,\mathrm{eV}^2$, our hierarchy condition is no longer satisfied, so that P_{ij} depend also on θ_{12}. For these cases we have performed a χ squared fit of the three up-down ratios (3.3.5c) to θ_{23}, θ_{13} and θ_{12}, for various values of the mass splittings, and have found acceptable regions of parameter space. Results are shown in figure 3.6b for the case that all $S_{ij} = 0$ for downward going neutrinos, while all $S_{ij} = 0.5$ for upward going neutrinos and $\phi = 0$ (no CP violation). An equivalent fit would be obtained for $\phi = \pi$ and $\theta_{23} \to \pi - \theta_{23}$. The (relatively small) asymmetry of fig.s 3.6b under $\theta_{23} \to \pi - \theta_{23}$ shows the dependence on ϕ of the SuperKamiokande data considered here.

A comparison of figure 3.6b with figures 3.3 and 3.5 shows under what conditions this large Δm^2_{12} scheme gives consistency. If all solar measurement techniques are correct, then, from figure 3.3, θ_{13} is small and $\theta_{12} = 10^\circ \div 20^\circ$. Figure 3.6b then shows that θ_{23} is centred on $45^\circ \pm 25^\circ$, the range around $\theta_{23} = 135^\circ$ being equivalent for any ϕ since θ_{13} is small. Figure 3.5 shows that solar model inspired fits to data from two solar techniques

at large Δm^2_{12} allow larger ranges of θ_{12} and θ_{13}, and these become correlated with θ_{23} via figure 3.6b.

3.5 Large $\nu_\mu \to \nu_\tau$ Mixing For Atmospheric Neutrinos

The pattern of masses and mixings suggested by the previous considerations show peculiar features, especially if both the atmospheric and solar neutrino anomalies are accounted for in the minimal scheme of section 3.3. The mass differences are hierarchical. However a large mixing ($\theta_{23} \approx 45°$) is required between the states with the *largest* mass difference. The mixing angle θ_{12} between the states with the *smallest* mass splitting may be large or small. Finally, if $\Delta m^2_{\rm atm} \geq 2 \cdot 10^{-3}\,{\rm eV}^2$, i.e. in the CHOOZ range, the third mixing angle must be small, $\theta_{13} \leq 13°$. Therefore it looks interesting to see which mass matrix could produce this pattern and which flavour symmetries can justify it.

3.5.1 2×2 Matrix Forms

As stressed in the introduction, an important consequence of the data on atmospheric neutrino fluxes is the need for large mixing angles. Here we study four possible forms of the 2×2 Majorana mass matrix for ν_μ and ν_τ which have a large mixing angle. In subsection 3.5.2 we study whether these forms can be incorporated in 3×3 mixing schemes which also give solar neutrino oscillations, and whether 3×3 cases exist which cannot be reduced to a 2×2 form. In section 3.6 we study whether these forms may be obtained from flavour symmetries of abelian type.

In a basis with a diagonal charged lepton mass matrix, the Majorana neutrino

	small entires	small parameters	order unity parameters	$\Delta m^2/(\frac{v^2}{M})^2$
(1) Generic	none	none	A, B, C	≈ 1
(2) Determinant small	none	none	$A, B, C = B^2/A + \varepsilon$	≈ 1
(3) One diagonal small	one diagonal	$C \approx \varepsilon$	A, B	≈ 1
(4) Pseudo-Dirac	both diagonal	$A, C \approx \varepsilon$	B	$\approx \varepsilon$

Table 3.1: The four possible 2×2 matrix forms which give a large mixing angle.

mass matrix is

$$m = \frac{v^2}{M} \begin{pmatrix} C & B \\ B & A \end{pmatrix}. \tag{3.5.1}$$

This is brought into real, diagonal form by the unitary matrix

$$V = R(\theta) \begin{pmatrix} 1 & 0 \\ 0 & e^{i\alpha} \end{pmatrix} \tag{3.5.2}$$

where $\tan 2\theta = 2B/(A - C)$, and the phase α does not affect oscillations. The mass difference relevant for oscillations is $\Delta m^2 = (A + C)\sqrt{(A - C)^2 + 4B^2}$. The coefficient v^2/M is motivated by the see-saw mechanism, with v the electroweak vacuum expectation value and M the mass of a heavy right-handed neutrino.

There are four possible forms of this matrix which give $\theta \approx 1$, and these are shown in Table 1. In cases (1) and (2) the entries are all of order unity; in the generic case they are unrelated, while in case (2) they are related in such a way that the determinant is suppressed. We discuss how such a suppression can occur naturally via the seesaw mechanism in the next section. Case (3) has one of the diagonal entries suppressed, which, however, does not follow from a simple symmetry argument. For cases $(1 \div 3)$, taking $\Delta m^2 = 10^{-3}\,\mathrm{eV}^2$, one

finds

$$M = (1 \div 3) \times 10^{15}\text{GeV}, \tag{3.5.3}$$

close to the scale of gauge coupling unification in supersymmetric theories.

Finally, case (4) has both diagonal entries small, making ν_μ and ν_τ components of a pseudo-Dirac neutrino. This follows from an approximate $L_\mu - L_\tau$ symmetry, and implies that $\theta \simeq 45°$. This agrees well with data: combing ρ_μ and ρ_{PC} of (3.3.5c) gives $\theta = 45° \pm 15°$. Of the four possible cases with large mixing angle, it is only the pseudo-Dirac neutrino which allows $\nu_{\mu,\tau}$ to be the astrophysical hot dark matter, in which case one predicts $\theta = 45°$ to high accuracy.

From the viewpoint of atmospheric neutrino oscillations alone, the distinction between cases (1) and (2) is unimportant. Since case (3) does not follow from simple symmetry arguments, one is left with two main 2×2 mixing schemes: the generic and pseudo-Dirac cases.

3.5.2 3×3 Matrix Forms

There are many possibilities for 3×3 neutrino mixing giving $P_{\mu\mu} \approx 0.5$, with oscillation primarily to ν_τ. In general two independent frequencies and three Euler angles are involved.

For the case that the oscillation is dominated by a single frequency, the possibilities may be divided into two classes: "2×2-like" and "inherently 3×3." The 2×2-like cases are just the four discussed in subsection 3.5.1, with $\theta_{12,13}$ small. Even though Δm^2_{23} may not be the largest Δm^2, it is the only one which causes substantial depletion of ν_μ. More interesting are the inherently 3×3 cases, for which there is no 2×2 reduction.

Consider the case

$$m = \frac{v^2}{M}\begin{pmatrix} 0 & B & A \\ B & 0 & 0 \\ A & 0 & 0 \end{pmatrix} \tag{3.5.4}$$

with $A, B \approx 1$. This is diagonalized by $V = R_{23}(\theta_{23})R_{12}(\theta_{12} = 45°)$ giving a Dirac state of ν_e married to $c_{23}\nu_\mu + s_{23}\nu_\tau$. The mass eigenvalues are $(M, M, 0)$, which, from the viewpoint of oscillations are equivalent to $(0, 0, M)$. Hence, one immediately sees that the oscillation probabilities are given by (3.3.3c) with $\theta_{13} = 0$: $P_{\mu\tau} = \sin^2 2\theta_{23} S_{23}$ has the form of a 2×2 oscillation, even though the mass matrix has an inherently 3×3 form. This arises because (3.5.4) is governed by the symmetry $L_e - L_\mu - L_\tau$, which allows $\nu_\mu \leftrightarrow \nu_\tau$, but prevents ν_e from oscillating.

We claim that (3.5.4) is the only inherently 3×3 form for $\nu_\mu \to \nu_\tau$ *at a single frequency*, as we now show. An inherently 3×3 form must have large entries outside the 2×2 block in 23 subspace. The three possibilities are 11, 12 and 13 (and their symmetric). None of these entries work alone, even coupled to any structure in the 23 block: either one gets two comparable frequencies or one does not get $\nu_\mu \to \nu_\tau$. The same is true for $11 + 12$ or $11 + 13$, again possibly together with any 23-block. Since $11 + 12 + 13$ leads to two comparable frequencies, the only case remaining is $12 + 13$, with a relatively negligible 23 block, i.e. the 3×3 form in (3.5.4). Basic to this conclusion is the assumption of no special relations among the different neutrino matrix elements other than the symmetry of the matrix itself (for alternatives see [22]).

3.6 Models for both Solar and Atmospheric Neutrinos

In this section we construct models for the minimal scheme for atmospheric and solar neutrino oscillations, discussed in section 3.3. The mass pattern suggested by this scheme has the hierarchy $\Delta m^2_{\odot} \equiv \Delta m^2_{12} \ll \Delta m^2_{\rm atm} \equiv \Delta m^2_{23}$. We take the form of the lepton mass matrices to be determined by flavour symmetries (FS) and assume that all small entries in these matrices are governed by small flavour symmetry breaking (FSB) parameters.

The low energy effective mass matrix for the three light left-handed neutrinos can be written as the sum of two matrices: $m_{LL} = m_{\rm atm} + m_{\odot}$, where all non-zero entries of $m_{\rm atm}$ are larger than all entries of $m_{\odot}$. The form of $m_{\rm atm}$ is such that there is a large mass splitting: $\Delta m^2_{\rm atm} \approx 10^{-(2\div 3)}\,{\rm eV}^2$, and a vanishing Δm^2. Furthermore, this matrix must give a large depletion of ν_μ, and, as discussed in the last section, this could occur if it has certain 2×2-like or inherently 3×3 forms. Of the two 2×2-like forms shown in Table 1, only case (2) is acceptable: in cases (1) and (3) the two independent Δm^2 are comparable, while in case (4) the second independent Δm^2 is larger than $\Delta m^2_{\rm atm}$. Hence, we arrive at the possibility**:

$$m^{2\times 2}_{\rm atm} = \frac{v^2}{M}\begin{pmatrix} 0 & 0 & 0 \\ 0 & C & B \\ 0 & B & A \end{pmatrix} \qquad (3.6.1)$$

with $A, B \approx 1$ and $C = B^2/A$. A reason for the vanishing sub-determinant will be given shortly.

In the previous section we have proved that there is a unique form for $m_{\rm atm}$ which

**Ansätze of this type for the neutrino mass matrix, up to small corrections, to describe atmospheric and solar neutrinos are contained in ref.s [23].

is inherently 3×3:

$$m_{\rm atm}^{3\times 3} = \frac{v^2}{M}\begin{pmatrix} 0 & B & A \\ B & 0 & 0 \\ A & 0 & 0 \end{pmatrix} \tag{3.6.2}$$

with $A, B \approx 1$.

The oscillation angles in the leptonic mixing matrix, V, have contributions from diagonalization of both the neutrino mass matrix, θ_{ij}^{ν}, and the charged lepton mass matrix, θ_{ij}^{e}: $V(\theta_{ij}) = V^{e\dagger}(\theta_{ij}^{e})V^{\nu}(\theta_{ij}^{\nu})$. This requires discussing also the charged lepton mass matrix. It is not easy to construct an exhaustive list of the possible symmetries and their breaking parameters. This is partly because there are both discrete and continuous symmetries with many choices for breaking parameters; but is mainly because of a subtlety of the seesaw mechanism. Let m_{RR} and m_{LR} be the most general Majorana and Dirac mass matrices of the seesaw mechanism allowed by some approximate symmetry. On forming the mass matrix for the light states, $m_{LL} = m_{LR}m_{RR}^{-1}m_{LR}^{T}$, one discovers that m_{LL} need not be the most general matrix allowed by the approximate symmetry. This means that one cannot construct an exhaustive list by only studying the symmetry properties of m_{LL} — it is necessary to study the full theory containing the right-handed states.

A casual glance at (3.6.1) and (3.6.2) shows that the flavor symmetry we seek, from the viewpoint of $\Delta L = 2$ operators, does not distinguish l_μ from l_τ, but does distinguish these from l_e. There are many combinations of the three lepton numbers L_a, and their subgroups, acting on l_a, which have this property. As representative of this group, we choose the combination $L_e - L_\mu - L_\tau$. We find it remarkable that this symmetry group can yield both (3.6.1) and (3.6.2), depending on how it is realized.

3.6.1 $L_e - L_\mu - L_\tau$ realized in the Low Energy Effective Theory

In the effective theory at the weak scale, we impose an approximate $L_e - L_\mu - L_\tau$ symmetry, which acts on the weak doublets, $l_{e,\mu,\tau}$, and is broken by small FSB parameters, ε and ε' of charge +2 and -2, respectively, giving a neutrino mass matrix:

$$m_{LL} = \frac{v^2}{M} \begin{pmatrix} \varepsilon' & 1 & 1 \\ 1 & \varepsilon & \varepsilon \\ 1 & \varepsilon & \varepsilon \end{pmatrix} \tag{3.6.3}$$

Hereafter, the various entries of the matrices only indicate the corresponding order of magnitude, allowing for an independent parameter for each entry. This texture gives

$$m_1 \approx m_2 \approx \frac{v^2}{M} \qquad \Delta m^2_{12} \approx \frac{v^4}{M^2}(\varepsilon + \varepsilon') \tag{3.6.4a}$$

$$m_3 \approx \frac{v^2}{M}\varepsilon \qquad \Delta m^2_{23} \approx \frac{v^4}{M^2} \tag{3.6.4b}$$

and

$$\theta^\nu_{23} \approx 1 \qquad \theta^\nu_{13} \approx \varepsilon \qquad \theta^\nu_{12} = 45°. \tag{3.6.4c}$$

While the texture gives only the order of magnitude of θ^ν_{23}, it precisely predicts θ^ν_{12} to be close to 45°. If the FSB parameters ε and ε' are taken to be extremely small, this becomes an excellent candidate for the case of "just so" solar neutrino oscillations, with the prediction that $\theta_{12} = 45°$. However, from figure 3.3 it follows that this model cannot give matter neutrino oscillations in the sun, which requires $\sin 2\theta_{12} \leq 0.9$. There are several contributions to the deviation of $\sin 2\theta_{12}$ from unity, but they are all too small to reconcile the discrepancy. A hierarchy in Δm^2 requires $\varepsilon, \varepsilon' < 0.1$, and since $\sin^2 2\theta^\nu_{12} \simeq 1 - (\varepsilon - \varepsilon')^2/8$, the deviation of $\sin 2\theta^\nu_{12}$ from 1 is negligible. After performing the θ^ν_{12} rotation, there are

small $\mathcal{O}(\varepsilon)$ rotations in the 13 and 23 planes necessary to fully diagonalize m_{LL}; these are too small to affect our conclusions. The last hope is that there could be a significant contribution to θ_{12} from diagonalization of the charged lepton mass matrix. As mentioned above, the diagonalization of the charged lepton mass matrix has to be discussed anyhow.

Consistently with the symmetry structure of (3.6.3), the most general form for the charged lepton mass matrix, with a structure governed by abelian symmetries is

$$m_E = \lambda v \begin{pmatrix} \xi' & \xi\varepsilon' & \varepsilon' \\ \xi'\varepsilon & \xi & 1 \\ \xi'\varepsilon & \xi & 1 \end{pmatrix} \tag{3.6.5}$$

when left (right) handed leptons are contracted to the left (right), $\bar{e}_L m_E e_R$. $(1, \xi, \xi')$ are the relative FSB parameters of (τ_R, μ_R, e_R) with respect to some other approximate FS, needed to describe the charged lepton mass hierarchies, and λ is the absolute FSB parameter of $\bar{\tau}_R \tau_L$. Here we ignore the fact that non-abelian symmetries could modify this form, for example by requiring some entries to vanish.

Diagonalization of (3.6.5) leads to

$$\theta^e_{23} \approx 1, \qquad \theta^e_{13} \approx \varepsilon', \qquad \theta^e_{12} \approx \varepsilon'$$

Therefore, altogether

$$\theta_{23} \approx 1, \qquad \theta_{12} \approx \varepsilon + \varepsilon', \qquad \theta_{12} \approx 45^\circ. \tag{3.6.6}$$

Since $\sin^2 2\theta_{12}$ remains corrected only by quadratic terms in ε and/or ε', we conclude that $L_e - L_\mu - L_\tau$, realized as an approximate symmetry of the low energy effective theory, can explain both atmospheric and solar neutrino fluxes with a hierarchy of Δm^2, most likely *only* for the case of "just so" vacuum solar oscillations, in which case the scale of new

physics, M, is close to the gauge unification scale, and the FSB parameters are extremely small: $\varepsilon, \varepsilon' \approx 10^{-7}$. This result also applies when *any* approximate FS of the low energy effective theory yields (3.6.3). In view of (3.6.4), with $\Delta m^2_{23} \approx \Delta m^2_{\rm atm} \approx 10^{-(2\div 3)}\,{\rm eV}^2$, notice that all three neutrinos are cosmologically irrelevant. Furthermore, the smallness of the 11 entry of (3.6.3) makes the search for neutrino-less 2β-decay uninteresting.

Comparing the θ_{13} plots of figures 3.3 and 3.5, one finds that, with one experiment excluded, the case of $\theta_{12} = 45^\circ$ becomes allowed for a large range of Δm^2_{12}, giving another application for this inherently 3×3 form of the mass matrix.

3.6.2 $L_e - L_\mu - L_\tau$ realized via the Seesaw Mechanism

The seesaw mechanism [24] allows a simple origin for the vanishing of the 2×2 sub-determinant of (3.6.1). Consider a single right-handed neutrino, N, with Majorana mass M and Dirac mass term $vN(\cos\theta\nu_\tau + \sin\theta\nu_\mu)$, where $\theta \approx 1$. Integrating out this single heavy state produces a single non-zero eigenvalue in m_{LL} — giving (3.6.1) with $A = \cos^2\theta, B = \cos\theta\sin\theta$ and $C = \sin^2\theta$, so that $AC = B^2$.

How could this carry over to a theory with three right-handed neutrinos, N_a? As long as one of them, N with the above mass terms, is much lighter than the others, then it will give the dominant contribution to m_{LL}, which will have (3.6.1) as its leading term. Clearly the key is that there be one right-handed neutrino which is lighter than the others, and couples comparably to ν_μ and ν_τ.

This can be realized using $L_e - L_\mu - L_\tau$, with two small FSB parameters ε (+2)

and ε' (-2). The right-handed neutrino mass matrix is

$$m_{RR} = M \begin{pmatrix} \varepsilon' & 1 & 1 \\ 1 & \varepsilon & \varepsilon \\ 1 & \varepsilon & \varepsilon \end{pmatrix} \tag{3.6.7}$$

and the Dirac mass matrices of neutrinos and charged leptons are

$$m_{LR} = \lambda' v \begin{pmatrix} \eta' & \varepsilon\eta' & \varepsilon\eta' \\ \varepsilon'\eta & \eta & \eta \\ \varepsilon' & 1 & 1 \end{pmatrix} \quad \text{and} \quad m_E = \lambda v \begin{pmatrix} \xi'\eta' & \varepsilon\xi\eta' & \varepsilon\eta' \\ \xi'\varepsilon'\eta & \xi\eta & \eta \\ \xi'\varepsilon' & \xi & 1 \end{pmatrix} \tag{3.6.8}$$

where, in analogy with (3.6.5), we have introduced FSB parameters consistent with (3.6.7).

For ease of exposition, let us first consider the case where all the η and ξ factors are set equal to unity. The crucial point is that there is a massless right-handed neutrino in the limit $\varepsilon \to 0$. Hence, taking ε small, and doing a rotation in the 23 plane we have 2×2 sub-matrices

$$m_{RR}^{-1} = \frac{1}{M} \begin{pmatrix} 0 & 0 \\ 0 & 1/\varepsilon \end{pmatrix}, \qquad m_{LR} = \lambda' v \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \tag{3.6.9}$$

giving

$$m_{LL} = \frac{(\lambda' v)^2}{M} \begin{pmatrix} 1/\varepsilon & 1/\varepsilon \\ 1/\varepsilon & 1/\varepsilon \end{pmatrix} \tag{3.6.10}$$

where $\det m_{LL} = 0$ at this order. In a theory with right-handed neutrinos, $L_e - L_\mu - L_\tau$ leads to (3.6.1).

Extending the analysis to 3×3 matrices is straightforward. The inverse of m_{RR}

$$m_{RR}^{-1} = \frac{1}{M} \begin{pmatrix} \varepsilon & 1 & 1 \\ 1 & \varepsilon' & \varepsilon' \\ 1 & \varepsilon' & \frac{1}{\varepsilon} \end{pmatrix} \tag{3.6.11}$$

shows a pseudo-Dirac structure in the 12 subspace, which is preserved in the light neutrino mass matrix:

$$m_{LL} = \frac{(\lambda' v)^2}{M} \begin{pmatrix} \varepsilon & 1 & 1 \\ 1 & \varepsilon' & 0 \\ 1 & 0 & \frac{1}{\varepsilon} \end{pmatrix} \tag{3.6.12}$$

where we have gone to a basis which diagonalizes the 23 subspace. The parameters relevant for neutrino oscillation are

$$\theta_{23}^{e,\nu} \approx 1, \qquad \theta_{13}^{e,\nu} \approx \varepsilon, \qquad \theta_{12}^{\nu} = 45^\circ, \qquad \theta_{12}^{e} \approx \varepsilon \tag{3.6.13a}$$

and

$$\Delta m_{23}^2 \approx \frac{1}{\varepsilon^2} \frac{(\lambda' v)^4}{M^2}, \qquad \Delta m_{12}^2 \approx (\varepsilon + \varepsilon') \frac{(\lambda' v)^4}{M^2}. \tag{3.6.13b}$$

It is remarkable that $L_e - L_\mu - L_\tau$ has forced a pseudo-Dirac structure in the 12 subspace as in its previous realization, again giving θ_{12} near 45°. The crucial difference is that the pseudo-Dirac mass splitting is now a higher power in FSB than before

$$\frac{\Delta m_{12}^2}{\Delta m_{23}^2} \approx \varepsilon^2(\varepsilon + \varepsilon') \tag{3.6.13c}$$

rather than $\varepsilon + \varepsilon'$. This allows ε and ε' to be considerably larger than before, so $\sin 2\theta_{12} < 0.8$ is now possible, allowing large angle MSW solar neutrino oscillations. In this case the FSB parameters are not very small $\varepsilon, \varepsilon' \approx 0.3 \div 0.5$, so that the mass of the right-handed neutrinos is still quite close to the gauge coupling unification scale. Notice again the cosmological irrelevance of the neutrino masses. For neutrino-less 2β decay searches $(m_{LL})_{11} \approx \epsilon^3 (\Delta m_{23}^2)^{1/2} \leq 10^{-2}\,\mathrm{eV}^2$. Finally, $\varepsilon' \approx 0.1$ and $\lambda' \approx 1$ can make M exactly coincident with the unification scale.

So far we have only produced models with large θ_{12}. However $L_e - L_\mu - L_\tau$ realized

with the seesaw mechanism may also lead to small θ_{12}, using the FSB suppression factors in (3.6.8). Taking $\eta' \ll \varepsilon'$ and $\eta \approx 1$, in an appropriate 23 basis gives

$$m_{LL} = \frac{(\lambda' v)^2}{M} \begin{pmatrix} \eta'^2\varepsilon & \eta' & \eta' \\ \eta' & \varepsilon' & 0 \\ \eta' & 0 & \frac{1}{\varepsilon} \end{pmatrix} \tag{3.6.14}$$

so that eq.s (3.6.13b) and (3.6.13c) remain valid but

$$\theta_{23}^{e,\nu} \approx 1, \qquad \theta_{13}^{e,\nu} \approx \eta'\varepsilon, \qquad \theta_{12}^{e} \approx \varepsilon' \tag{3.6.15}$$

and, most importantly

$$\theta_{12}^{\nu} \approx \eta'/\varepsilon' \tag{3.6.16}$$

which can make θ_{12} small.

3.7 Conclusions

The solar and atmospheric neutrino anomalies, strengthened by the recent SuperKamiokande observations, can be interpreted as due to oscillations of the three known neutrinos. However there is still considerable allowed ranges of masses and mixing angles that can account for all these anomalies, especially if a cautious attitude is taken with regard to the theoretical analysis and/or the (difficult) experiments relevant to solar neutrinos. A further major element of uncertainty is related to the relatively large range of values for the mass splitting that can account for the atmospheric neutrino anomaly. We summarize our conclusions by considering a set of alternative hypotheses, related to these dominant uncertainties, with an eye to the experimental program that may lead to their resolution and eventually to the determination of the full set of neutrino oscillation parameters.

A critical value for Δm^2_{23} is around $2 \cdot 10^{-3}\,\mathrm{eV}^2$ mainly because for larger values CHOOZ sets a considerable constraint on the mixing pattern, but also because $(1 \div 2) \cdot 10^{-3}\,\mathrm{eV}^2$ is frequently discussed as a typical sensitivity limit for various Long-Base-Line (LBL) neutrino experiments, like the one from KEK to SK, or the ν_τ appearance experiments with a high energy beam from CERN to Gran Sasso or from Fermilab to Soudan. On the other end, a value of $\Delta m^2_{12} < 2 \cdot 10^{-4}\,\mathrm{eV}^2$, as certainly required by a standard Solar Neutrino Analysis (SNA), would make the corresponding oscillation frequency irrelevant to the SK experiment on atmospheric neutrinos. On this basis we consider the following four possibilities, none of which, we believe, can be firmly excluded at present. They are graphically represented in fig. 3.7.

1. $\Delta m^2_{23} > 2{\cdot}10^{-3}\,\mathrm{eV}^2$ and $\Delta m^2_{12} < 2{\cdot}10^{-4}\,\mathrm{eV}^2$. Here a minimal scheme to describe both solar and atmospheric neutrinos is required, as discussed in section 3.3, with $\Delta m^2_{23} \gg \Delta m^2_{12}$. Since Δm^2_{12} is too small to affect atmospheric and/or LBL experiments, in both cases eqs. 3.3.3 apply. The fit relevant to SK is given in fig. 3.6a, with the further constraint, from CHOOZ, that θ_{13} is small, $\theta_{13} \leq 13°$, and therefore $\theta_{23} = 45° \pm 15°$. In turn θ_{12}, together with Δm^2_{12}, will have to be determined by solar neutrino experiments. In this alternative, the neatest confirmation of the SK result would come from a ν_τ appearance LBL experiment. At the same time, a dominant $\nu_\mu \to \nu_\tau$ oscillation should also lead to a signal in the KEK to SK ν_μ disappearance experiment, with no appreciable ν_e appearance signal.

2. $\Delta m^2_{23} < 2 \cdot 10^{-3}\,\mathrm{eV}^2$ and $\Delta m^2_{12} < 2 \cdot 10^{-4}\,\mathrm{eV}^2$. The main difference with respect to the previous case is that now θ_{13} is not constrained by CHOOZ, and therefore, from

fig. 3.6a, it can be as large as 45°. This implies, from eqs. 3.3.3, that the results of both atmospheric and LBL experiments, with low enough ν_μ energies to permit exploration of Δm^2 lower than $2 \cdot 10^{-3}\,\mathrm{eV}^2$, may be affected by a significant $P_{\mu e} \neq 0$. By the same token, an experiment with low energy $\bar{\nu}_e$ extending the sensitivity of CHOOZ (e.g. Kam-LAND) may show a large signal if θ_{13} is indeed large. In any event $P_{\mu\tau}$ will be significant. Finally, as in case 1., decoupling of solar and atmospheric neutrino oscillations implies that θ_{12} can only be determined by solar neutrino experiments, with an analysis complicated by θ_{13} being potentially unconstrained (see fig.s 3.3, upper row)

3. $\Delta m^2_{23} > 2 \cdot 10^{-3}\,\mathrm{eV}^2$ and $\Delta m^2_{12} > 2 \cdot 10^{-4}\,\mathrm{eV}^2$. This case is possible only if SSM constraints are relaxed (fig. 3.3, lower row) and/or if one of the experimental techniques for solar neutrinos is problematic (fig. 3.5). However, as discussed in section 3.4, Δm^2_{12} must be lower than $2 \cdot 10^{-3}\,\mathrm{eV}^2$, below the CHOOZ range. Since, on the other hand, $\Delta m^2_{\mathrm{atm}} = \Delta m^2_{23}$ is in the CHOOZ range, θ_{13} is small and eq.s 3.4.3 are relevant for atmospheric and LBL experiments. The fit of the present SK results gives $\theta_{23} = 45° \pm 25°$ (the range at $\theta_{23} \approx 135°$ being equivalent since θ_{13} is small). Therefore the main difference with respect to case 1. is the possibility of a S_{12} contribution in eq. (3.4.3). While ν_τ appearance in LBL experiments must still give a positive signal, $P_{\mu e}$ could significantly deviate from zero at low enough oscillation frequencies (relevant to lower energy ν_μ LBL experiments or to reactor experiments such as Kam-LAND). The finding of such an effect, together with a positive ν_τ appearance signal, would prove, in the three neutrino oscillation picture, the inadequacy of the NSA as

it is done now.

4. $\Delta m^2_{23} < 2 \cdot 10^{-3}\,\mathrm{eV}^2$ and $\Delta m^2_{12} > 2 \cdot 10^{-4}\,\mathrm{eV}^2$. This is the relatively less constrained case (and also the relatively less likely). Here both neutrino squared mass differences are outside of the CHOOZ range, so that θ_{13} is unconstrained. Appropriate values of the mixing angles can fit the SuperKamiokande up/down ratios of atmospheric neutrinos, as shown in fig. 3.6b. In this case, the two comparable Δm^2 might lead to sizeable CP-violating effects if all the three mixing angles are large.

Measurements by SNO and Borexino will increase the number of independent observational signals of the solar fluxes, S_i, from 3 to 5; so that, from (3.2.12) with $\Phi_{\mathrm{CNO}}/\Phi_{^7\mathrm{Be}} = 0.22$, $\Delta m^2_{12}, \theta_{12}, \theta_{13}, \Phi_{^7\mathrm{Be}}$ and $\Phi_{^8\mathrm{B}}$ can all be determined. This will provide a crucial consistency check between the experimental techniques and the solar models. If θ_{13} is found to be large, $\Delta m^2_{23} < 2 \times 10^{-3}\,\mathrm{eV}^2$, giving a signal at Kam-LAND, but making it harder for LBL experiments.

In the minimal scheme, with a hierarchy amongst the Δm^2, several years of data from Super-Kamiokande will allow a fit to $\Delta m^2_{23}, \theta_{23}$ and θ_{13}. Combining with fits to the solar flux measurements, and to LBL and Kam-LAND experiments, could allow the emergence of a consistent picture for the two oscillation frequencies and the three leptonic mixing angles.

The variety of possibilities discussed above makes it uncertain which is the relevant neutrino mass matrix and, a fortiori, which are the flavour symmetries that might be responsible for it. Nevertheless, focusing on the minimal scheme for both solar and atmospheric neutrinos, the peculiar pattern of masses and mixings renders meaningful the

search for an appropriate mass matrix. As discussed in section 3.5 on general grounds, two forms of mass matrices emerge as being able to describe the data, eq.s (3.6.1) and (3.6.2). Since in the minimal scheme $\Delta m^2_{12} \ll 2 \cdot 10^{-3}\,\mathrm{eV}^2$, these forms imply that neutrino masses will not give rise to an observable neutrinoless double beta decay signal. The combination $L_e - L_\mu - L_\tau$ of the individual lepton numbers may play a role in yielding both these forms. A common feature of the resulting solutions is that the heaviest neutrino mass is determined by the oscillation length of the atmospheric neutrinos, $(\Delta m^2_{\mathrm{atm}})^{1/2}$. As such, the neutrino masses are irrelevant for present cosmology. Again quite in general, an increasing separation between the two Δm^2 requires the angle θ_{13} to become increasingly small.

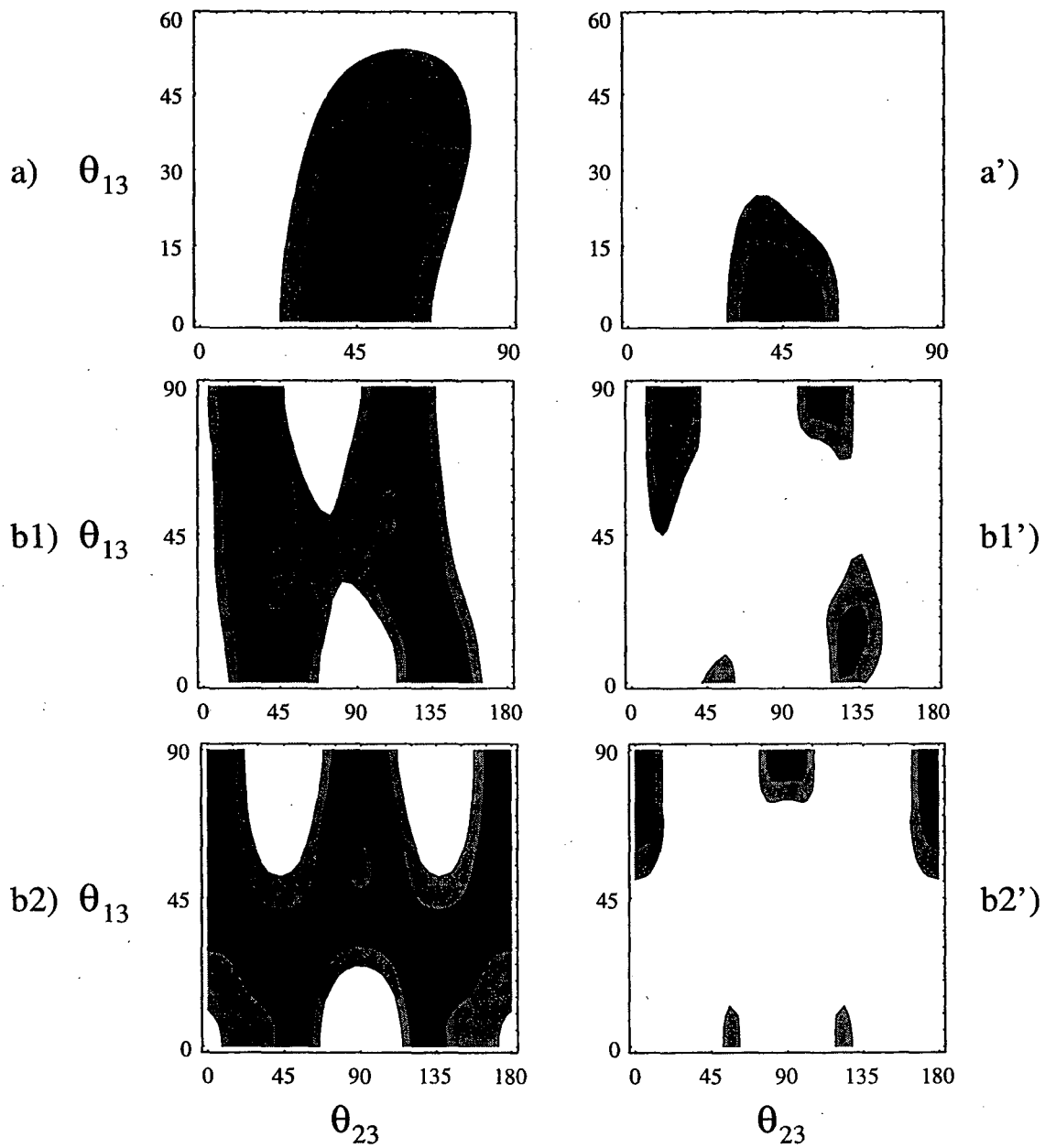

Figure 3.6: Mixing angles θ_{ij} that fit the up/down ratios (3.3.5a,b,c) of atmospheric neutrinos, assuming that (a) $\Delta m^2_{12} \ll 10^{-3}\,\mathrm{eV}^2$ and any θ_{12}, (b) $\Delta m^2_{12} \approx \Delta m^2_{23} \approx 10^{-3}\,\mathrm{eV}^2$, $\phi = 0$ and (b1) $\theta_{12} = 20°$, (b2) $\theta_{12} = 45°$. Primed figures are as above, but including in the asymmetry also the intermediate bins in the angular distribution of [3] (see text). The contours are for $\chi^2 = 3$ and $\chi^2 = 6$.

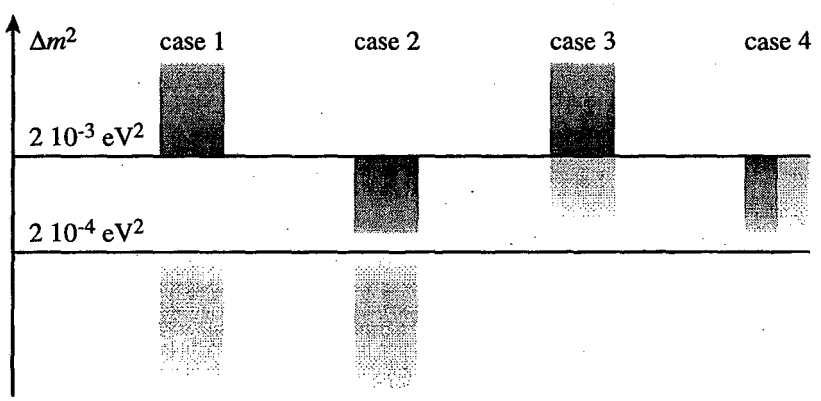

Figure 3.7: Different combinations of ranges for Δm^2_{23} (dark gray) and Δm^2_{12} (light gray) discussed in the text.

Bibliography

[1] Y. Fukuda et al. (Super-Kamiokande collaboration), *hep-ex/9805021*; Y. Suzuki (Super-Kamiokande collaboration), in Neutrino 98, proceedings of the 18th international conference on neutrino physics and astrophysics, Takayama, Japan, edited by Y. Suzuki and Y. Totsuka, to appear on Nucl. Phys. B (Proc. Suppl.).

[2] J.N. Bahcall and M.H. Pinsonneault, *Rev. Mod. Phys* **67** (1995) 781 ("BP95"). The recent revision of these fluxes, J.N. Bahcall, S. Basu and M.H. Pinsonneault, *astro-ph/9805135* ("BP98"), would not change any of our conclusions.

[3] Y. Fukuda et al. (Super-Kamiokande collaboration), *hep-ex/9805006.*

[4] L. Wolfenstein, *Phys. Rev.* **D17** (1978) 2369;
S.P. Mikheyev and A. Yu Smirnov, *Sov. J. Nucl. Phys.* **42** (1986) 913.

[5] S.P. Mikheyev and A. Yu Smirnov, *Phys. Lett.* **B200** (1988) 560;
G. Fogli, E. Lisi and D. Montanino, *Phys. Rev.* **D54** (1996) 2048 (*hep-ph/9605273*);
S. Bilenky, C. Giunti and C. Kim, *Astrop. Phys.* **4** (1996) 241 (*hep-ph/9505301*);
T. Teshima, T. Sakai and O. Inagaki, *hep-ph/9801276.*

[6] S. Parke, *Phys. Rev. Lett.* **57** (1986) 1275;
T.K. Kuo and J. Pantaleone, *Phys. Rev.* **D35** (1987) 3432.

[7] J.N. Bahcall, "*Neutrino Astrophysics*", Cambridge University Press, 1989.

[8] Solar neutrino fluxes and cross sections can be obtained from the J.N. Bahcall internet page, at www.sns.ias.edu/~jnb.

[9] For a review and references see V. Castellani et al., *Phys. Rep.* **281** (1997) 309 (*astro-ph/9606180*).

[10] B.T. Cleveland, *Nucl. Phys. (Proc. Suppl.)* **B38** (1995) 47; K. Lande et al., in Neutrino '96, proceedings of the 17th international con- ference on neutrino physics and astrophysics, Helsinki, 1996, edited by K. Huitu, K. Enqvist and J. Maalampi (World Scientific, Singapore, 1997).

[11] W. Hampel et al. (Gallex collaboration), *Phys. Lett.* **B388** (1996) 384;
Gallex Collaboration, TAUP97, Laboratori Nazionali del Gran Sasso September 1997, to appear on Nucl. Phys. B (Proc. Suppl.); V. Gavrin et al. (SAGE collaboration), *Phys. Lett.* **B328** (1994) 234; SAGE Collaboration, Neutrino 96, Helsinki June 1996 to appear on Nucl. Phys. B (Proc. Suppl.).

[12] Y. Fukuda et al., *Phys. Rev. Lett.* **77** (1996) 1683.

[13] N. Hata and P. Langacker, *Phys. Rev.* **D52** (1995) 420 (*hep-ph/9409372*);
N. Hata and P. Langacker, *Phys. Rev.* **D56** (1997) 6107 (*hep-ph/9705339*).

[14] P.I. Krastev and S.T. Petcov, *Phys. Lett.* **B395** (1997) 69 (*hep-ph/9612243*).

[15] G.L. Fogli, E. Lisi and D. Montanino, *hep-ph/9709473.*

[16] Q.Y. Liu, M. Maris and S. Petcov, *Phys. Rev.* **D56** (1997) 5991 (*hep-ph/9702361*); M. Maris and S. Petcov, *Phys. Rev.* **D56** (1997) 7444 (*hep-ph/9705392*).

[17] P. Lipari and M. Lusignuoli, *hep-ph/9803440.*

[18] M. Honda, T. Kajita, K. Kasahara, and S. Midorikawa, *Phys. Rev.* **D52** (1995) 4985.

[19] M. Apollonio et. al (CHOOZ collaboration), *hep-ex/9711002.*

[20] V. Barger, T. J. Weiler, K. Whisnant, *hep-ph/9807319*; G.L. Fogli, E. Lisi, A. Marrone, G. Scioscia, *hep-ph/9808205.*

[21] C. Athanassopoulos et al. (LSND collaboration) *Phys. Rev. Lett.* **75** (1995) 365; ibid. **77** (1996) 3082, *nucl-ex/9706006.*

[22] P. Binetruy et al., *Nucl. Phys.* **B496** (1997) 3 (*hep-ph/9610481*); M. Fukugita, M. Tanimoto and T. Yanagida, *Phys. Rev.* **D57** (1998) 4429 (*hep-ph/9709388*); Zhi-zhong Xing, *hep-ph/9804433.*

[23] M. Drees et al., *Phys. Rev.* **D57** (1998) 5335 (*hep-ph/9712392*); M. Bando, T. Kugo, K. Yoshioka, *Phys. Rev. Lett.* **80** (1998) 3004 (*hep-ph/9710417*); V. Barger, S. Pakvasa, T. J. Weiler, K. Whisnant, *hep-ph/9806387*; S. King, *hep-ph/9806440.*

[24] M. Gell-Mann, P. Ramond, and R. Slansky, in *Supergravity* (ed. P. van Nieuwenhuizen and D. Z. Freedman, North-Holland, Amsterdam, 1979), p. 315; T. Yanagida, in *proceedings of the workshop on the unified theory and the baryon*

number in the universe (ed. O. Sawada and A. Sugamoto, KEK Report No. 79-18, Tsukuba, Japan, 1979).

Chapter 4

U(2) and Neutrino Physics

4.1 Introduction

The pattern and origin of the quark and lepton masses and mixings remains a challenging question for particle physics. Although a detailed description of this pattern requires a theory of flavor with a certain level of complexity, the gross features may be described simply in terms of a flavor symmetry and its sequential breaking.

One simple flavor structure is motivated by four facts about flavor:

- The quarks and leptons fall into three generations, $\psi_{1,2,3}$, each of which may eventually have a unified description.
- The top quark is sufficiently heavy, that any flavor symmetry which acts on it non-trivially must be strongly broken.
- The masses of the two light generations imply a phenomenological description in terms of small dimensionless parameters, $\{\epsilon\}$.
- In supersymmetric theories, flavor-changing and CP violating phenomena suggest that the squarks and sleptons of the first two generations are highly degnerate.

It is attractive to infer that, at least at a phenomenological level, there is a non-Abelian flavor symmetry which divides the three generations according to

$$2 \oplus 1 : \qquad \psi_a \oplus \psi_3, \ a = 1, 2. \tag{4.1.1}$$

The four facts listed above follow immediately from such a structure, with $\{\epsilon\}$ identified as the small symmetry breaking parameters of the non-Abelian group. These control both the small values for quark masses and mixing angles, and also the small fractional non-degeneracies of the scalars of the first two generations.

The Super-Kamiokande collaboration has provided strong evidence for an anomaly in the flux of atmospheric neutrinos, which may be interpreted as large angle oscillations of ν_μ predominantly either to ν_τ or to ν_s, a singlet neutrino [1]. This observation provides a challenge to the non-Abelian $2 \oplus 1$ structure:

- ν_τ is expected to have a very different mass from that of $\nu_{e,\mu}$, and to only weakly mix with them.
- If the atmospheric oscillation is $\nu_\mu \to \nu_s$, what is the identity of this new singlet state, why is it light, and how could it fit into the $2 \oplus 1$ structure?

There are a variety of possible reactions to this challenge. One possibility is to drop the $2 \oplus 1$ idea; perhaps the CP and flavor violating problems of supersymmetry are solved by other means, or perhaps supersymmetry is not relevant to the weak scale. Another option is to retain the $2 \oplus 1$ structure for quarks, but not for leptons, where the flavor changing constraints are much weaker.

In this paper we study theories based on the flavor group $U(2)$, which immediately yields the structure (4.1.1), giving the $2 \oplus 1$ structure to both quarks and leptons [2]. The

masses and mixings of the charged fermions and scalars resulting fron $U(2)$ have been studied in detail, and significant successes have been identified [3]. We add a right-handed neutrino to each generation, and find that the symmetry structure of the neutrino mass matrix automatically chooses ν_μ to be a pseudo-Dirac state coupled to one of the right-handed neutrinos, resulting in $\nu_\mu \to \nu_s$ with a mixing angle close to 45°.

4.2 U(2) Theories of Quark and Charged Lepton Masses.

The most general $U(2)$ effective Lagrangian for charged fermion masses, at leading order in the $U(2)$ breaking fields, is

$$\mathcal{L} = \psi_3\psi_3 h + \frac{1}{M}\Big(\psi_3\phi^a\psi_a h + \psi_a(S^{ab} + A^{ab})\psi_b h\Big) \tag{4.2.2}$$

where ϕ^a is a doublet, S^{ab} a symmetric triplet, A^{ab} an antisymmetric singlet of $U(2)$, and h are Higgs doublets. Coupling constants have been omitted, and M is a flavor physics mass scale. An entire generation is represented by ψ, so that each operator contains terms in up, down and charged lepton sectors, but unification is not assumed. For example, this theory follows from a renormalizable Froggatt-Nielsen model on integrating out a single heavy vector $U(2)$ doublet of mass M (see the second of [3]).

The hierarchical pattern of masses and mixings for charged fermions is generated by breaking $U(2)$ first to $U(1)$ with vevs $\phi^2, S^{22} \approx \epsilon M$, and then breaking $U(1)$ via the vev $A^{12} \approx \epsilon' M$. The symmetry breaking

$$U(2) \overset{\epsilon}{\Rightarrow} U(1) \overset{\epsilon'}{\Rightarrow} 1 \tag{4.2.3}$$

produces the Yukawa coupling textures

$$M_{LR} = v \begin{pmatrix} 0 & \epsilon' & 0 \\ -\epsilon' & \epsilon & \epsilon \\ 0 & \epsilon & 1 \end{pmatrix}. \tag{4.2.4}$$

4.3 General Effective Theory of Neutrino Masses.

Without right-handed neutrinos, the most general $U(2)$ effective Lagrangian for neutrino masses, linear in $U(2)$ breaking fields, is

$$\mathcal{L}^{\nu}_{eff} = \frac{1}{M} l_3 l_3 hh + \frac{1}{M^2}\left(l_3 \phi^a l_a hh + l_a S^{ab} l_b hh\right). \tag{4.3.5}$$

where l_a, l_3 are lepton doublets. The term $l_a A^{ab} l_b hh$ vanishes by symmetry; hence the above vevs give the neutrino mass texture

$$M_{LL} = \frac{v^2}{M} \begin{pmatrix} 0 & 0 & 0 \\ 0 & \epsilon & \epsilon \\ 0 & \epsilon & 1 \end{pmatrix}. \tag{4.3.6}$$

so that the lightest neutrino is massless.* The mixing angle for $\nu_\mu \to \nu_\tau$ oscillations, $\theta_{\mu\tau}$, is of order ϵ — the same order as mixing of the quarks of the two heavier generations, V_{cb} — and is much too small to explain the atmospheric neutrino fluxes. However, in theories with flavor symmetries, the seesaw mechanism typically does not yield the most general neutrino mass matrix in the low energy effective theory. This apparent problem requires that we look more closely at the full theory, including the right-handed neutrinos.

*Including operators higher order in the $U(2)$ breaking fields, the lightest neutrino remains massless in a supersymmetric theory, but not in the non-supersymmetric case , where operators such as $l_a A^{ab} \phi^\dagger_b l_3 hh$ occur.

4.4 The Seesaw Mechanism: A Single Light ν_R

Adding three right-handed neutrinos to the theory, $N_a + N_3$, the texture for the Majorana mass matrix is:

$$M_{RR} = M \begin{pmatrix} 0 & 0 & 0 \\ 0 & \epsilon & \epsilon \\ 0 & \epsilon & 1 \end{pmatrix}. \tag{4.4.7}$$

with the 12 and 21 entries again vanishing by symmetry. In supersymmetric theories the zero eigenvalue is not lifted at higher order in the flavor symmetry breaking. This presents a problem for the 3×3 seesaw mechanism in $U(2)$ theories, since $M_{LL} = M_{LR} M_{RR}^{-1} M_{LR}^T$ and M_{RR} cannot be inverted.

One approach [4] is to allow further flavor symmetry breaking vevs, for example $\phi^1 \neq 0$, so that M_{RR} has no zero eigenvalues. Remarkably, taking $\phi^1/M \approx \epsilon'$, the seesaw gives $\theta_{\mu\tau} \approx 1$, as needed for the atmospheric neutrino anomaly. On the other hand, this pattern of neutrino masses cannot explain the solar neutrino fluxes, and the additional flavor breaking vevs remove two of the highly successful mass relation predictions of the quark sector.

In this paper we keep the minimal $U(2)$ symmetry breaking vevs and pursue the consequences of the light N_e state which results from (4.4.7). The singular nature of M_{RR}^{-1} is not a problem; it is an indication that N_e cannot be integrated out of the theory. However, N_τ and N_μ do acquire large masses, and when they are integrated out of the theory the low

energy 4×4 neutrino mass matrix is:

$$M^{(4)} = \begin{pmatrix} & & & 0 \\ & M_{LL} & & \epsilon' v \\ & & & 0 \\ 0 & \epsilon' v & 0 & 0 \end{pmatrix} \tag{4.4.8}$$

where M_{LL} is a 3×3 matrix in the (ν_a, ν_3) space, determined from seesawing out the two heavy right-handed states, and has one zero eigenvalue.

Because the $N_e - \nu_\mu$ mixing is weak scale, while all other couplings to ν_μ are suppressed, N_e and ν_μ are maximally mixed. Thus, we note that *a direct application of the* $U(2)$ *theory to the neutrino sector predicts a* 45^o *mixing between* ν_μ *and* ν_s*!*

There is a significant phenomenological difficulty with this model. The mass of the $N_e - \nu_\mu$ pseudo-Dirac state is of order $\epsilon' v$. Using a value for ϵ' extracted from an analysis of the charged lepton sector, this is of order 1 GeV, well in excess of the 170 keV limit obtained from direct searches. One simple solution is to restrict the couplings of the right-handed neutrinos by an additional $U(1)_N$ approximate flavor symmetry. Each N field carries N charge +1, while the symmetry is broken by a field with charge -1, leading to a small dimensionless breaking parameter ϵ_N. The entries in the neutrino mass matrices receive further suppressions

$$M_{LR} \Rightarrow \epsilon_N M_{LR} \quad M_{RR} \Rightarrow \epsilon_N^2 M_{RR} \tag{4.4.9}$$

which, for the 4×4 light neutrino matrix, simply leads to the replacement $\epsilon' v \Rightarrow \epsilon_N \epsilon' v$ in

the $N_e - \nu_\mu$ entry, giving

$$M^{(4)} = \begin{pmatrix} \frac{\epsilon'^2}{\epsilon}\frac{v^2}{M} & \epsilon'\frac{v^2}{M} & \epsilon'\frac{v^2}{M} & 0 \\ \epsilon'\frac{v^2}{M} & \epsilon\frac{v^2}{M} & \epsilon\frac{v^2}{M} & \epsilon_N\epsilon' v \\ \epsilon'\frac{v^2}{M} & \epsilon\frac{v^2}{M} & \frac{v^2}{M} & 0 \\ 0 & \epsilon_N\epsilon' v & 0 & 0 \end{pmatrix} \tag{4.4.10}$$

It is understood that all entries have unknown $O(1)$ coefficients.

Note that M_{LL} is unchanged. There is a simple reason for this. If we modify our right-handed couplings by the replacements $M_{LR} \to M_{LR}T$, $M_{RR} \to T^T M_{RR} T$, where T is any diagonal matrix, then

$$M_{LL} \Rightarrow M_{LR}T(T^T M_{RR}T)^{-1}(M_{LR}T)^T = M_{LL}. \tag{4.4.11}$$

It is interesting that the observed value of $\delta m^2_\odot$ can give the appearance that right-handed neutrinos receive GUT-scale masses, while their masses are in fact much lower.

If the $N_e - \nu_\mu$ entry dominates the mass of ν_μ, i.e. if $\epsilon_N \gg \frac{v}{M}$, this 4×4 matrix splits approximately into two 2×2 matrices, and maximal mixing is preserved. One 2×2 matrix describes the pseudo-Dirac state

$$\begin{pmatrix} \epsilon\frac{v^2}{M} & \epsilon_N\epsilon' v \\ \epsilon_N\epsilon' v & 0 \end{pmatrix} \tag{4.4.12}$$

while $\nu_e \Rightarrow \nu_\tau$ mixing is described by

$$\frac{v^2}{M}\begin{pmatrix} \frac{\epsilon'^2}{\epsilon} & \epsilon' \\ \epsilon' & 1 \end{pmatrix} \tag{4.4.13}$$

The resulting masses and mixings are given in Table 4.1.

Since ϵ and ϵ' are determined by the charged fermion masses, in the neutrino sector there are two free parameters, ϵ_N and M, which describe five important observables: $\theta_\odot$,

	m_{light}	m_{heavy}	δm^2	θ_{mix}
(1) Heavy states	$v\epsilon_N\epsilon' - \epsilon\frac{v^2}{2M}$	$v\epsilon_N\epsilon' + \epsilon\frac{v^2}{2M}$	$\frac{v^3}{M}\epsilon_N\epsilon\epsilon'$	45^o
(2) Light states	$\frac{\epsilon'^2}{\epsilon}\frac{v^2}{M}$	$\frac{v^2}{M}$	$(\frac{v^2}{M})^2$	ϵ'

Table 4.1: General Theory: the masses, mixings, and splittings of the two sets of neutrinos.

θ_{atm}, $\delta m^2_{\odot}$, δm^2_{atm} and m_ν, the mass of the pseudo-Dirac muon neutrino. However, the various predictions of the theory have varying levels of certainty. Because there are a large number of order one constants in the original formulation of the theory, we can end up with a prediction which has a coefficient of a product of some number of these quantities. To assess the level of certainty, we will include a quantity i, which we term the "stability index" of the prediction, which is simply the power of unknown order one coefficients appearing in the prediction.

Two of the three resulting predictions are the mixing angles

$$\sin\theta_{\odot} \approx \epsilon' \qquad [i=4]\,, \quad \theta_{atm} = 45^o \qquad [i=0]\,. \tag{4.4.14}$$

The postdiction of a maximal mixing angle for atmospheric oscillations is an important consequence of the $U(2)$ theory. The value of ϵ' extracted from the charged fermion sector is 0.004, within an order of magnitude of the central value $\theta_{\odot} = 0.037$ of the recent BP98 fit to the solar data, and within a factor of 4 of the minimal acceptable value of 0.016 [5]. Such a discrepancy is not a great concern, as we gain a comparable contribution from the charged lepton matrix. Furthermore, the prediction of $\theta_{\odot}$ involves the fourth power of unknown order one coefficients, thus $i = 4$, and is somewhat uncertain.

The relevant mass splitting for the $\nu_e \to \nu_\tau$ oscillations occuring in the sun is

$$\delta m^2_\odot \approx \left(\frac{v^2}{M}\right)^2. \tag{4.4.15}$$

While this is not a prediction of the theory, it is intriguing, as has been noticed elsewhere in other contexts, that if M is taken close to the scale of coupling constant unification, $\delta m^2_\odot \approx 10^{-5}$ eV^2, in the right range for either small or large angle MSW oscillations.

The final free parameter ϵ_N is fixed by the observed mass splitting for atmospheric oscillations

$$\delta m^2_{atm} \approx \epsilon\epsilon'\epsilon_N \frac{v^3}{M} \approx \epsilon\epsilon'\epsilon_N v\sqrt{\delta m^2_\odot} \tag{4.4.16}$$

giving $\epsilon_N \approx 10^{-8}$ — the $U(1)_N$ symmetry is broken only very weakly.

The final prediction is for the mass of the heavy pseudo-Dirac $\nu_\mu N_e$ state:

$$m_\nu \approx \epsilon'\epsilon_N v \approx \frac{\delta m^2_{atm}}{\epsilon\sqrt{\delta m^2_\odot}} \approx 10^{0.4}\text{eV} - 10^2\text{eV}, \qquad [i = \dot{4}] \tag{4.4.17}$$

where the given spread in mass is due to uncertainty in δm^2_{atm} and $\delta m^2_\odot$. While it is tempting to interpret this as a good candidate for hot dark matter, we will see later that KARMEN places stringent limits on the acceptable values of m_ν.

4.5 A Variant Theory

A variation on this breaking structure was explored in a particular model (see the second of [6]), and it is interesting to explore whether this same approach for neutrino masses can work within that model. In this variation, there is no S^{ab} field present, and the RR and LR masses are given by

$$M_{LR} = \begin{pmatrix} 0 & \epsilon' & 0 \\ -\epsilon' & 0 & \epsilon \\ 0 & \epsilon & 1 \end{pmatrix} \quad M_{RR} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \epsilon \\ 0 & \epsilon & 1 \end{pmatrix} \tag{4.5.18}$$

generating a light 4×4 mass matrix

$$M^{(4)} = \begin{pmatrix} \frac{\epsilon'^2}{\epsilon^2}\frac{v^2}{M} & \epsilon'\frac{v^2}{M} & \frac{\epsilon'}{\epsilon}\frac{v^2}{M} & 0 \\ \epsilon'\frac{v^2}{M} & 0 & \epsilon\frac{v^2}{M} & \epsilon_N\epsilon' v \\ \frac{\epsilon'}{\epsilon}\frac{v^2}{M} & \epsilon\frac{v^2}{M} & \frac{v^2}{M} & 0 \\ 0 & \epsilon_N\epsilon' v & 0 & 0 \end{pmatrix} \tag{4.5.19}$$

This matrix is problematic, because the 2×2 submatrix for the atmospheric neutrinos does not contain a splitting term. Of course, a splitting would be generated through interactions with the other left-handed states, we estimate

$$M^{(4)}_{\mu\mu} \approx \epsilon^2 \frac{1}{m_\nu}(\frac{v^2}{M})^2. \tag{4.5.20}$$

Consequently our atmospheric splitting is

$$\delta m^2_{atm} \approx \epsilon^2 (\frac{v^2}{M})^2. \tag{4.5.21}$$

Since we have $(\frac{v^2}{M})^2 = \delta m^2_\odot$, this would predict $\delta m^2_{atm} < \delta m^2_\odot$, which is unacceptable. One simple solution is to allow the appearance of the operators

$$(\frac{1}{M})^2 \phi^a \phi^b N_a N_b M_{GUT} \tag{4.5.22}$$

$$(\frac{1}{M})^2 \phi^a \phi^b N_a \nu_b H. \tag{4.5.23}$$

The inclusion of one or both of these operators in our Lagrangian has the same effect on our final mass matrix, inducing $M^{(4)}_{\mu\mu} \approx \epsilon^2 \frac{v^2}{M}$ and yielding the 2×2 submatrix

	m_{light}	m_{heavy}	δm^2	θ_{mix}
(1) Heavy states	$v\epsilon_N\epsilon' - \epsilon^2\frac{v^2}{2M}$	$v\epsilon_N\epsilon' + \epsilon^2\frac{v^2}{2M}$	$\frac{v^3}{M}\epsilon_N\epsilon^2\epsilon'$	45^o
(2) Light states	$(\frac{\epsilon'}{\epsilon})^2\frac{v^2}{M}$	$\frac{v^2}{M}$	$(\frac{v^2}{M})^2$	$\frac{\epsilon'}{\epsilon}$

Table 4.2: Without S field: The masses, mixings, and splittings of the two sets of neutrinos.

$$\begin{pmatrix} \epsilon^2\frac{v^2}{M} & \epsilon_N\epsilon' v \\ \epsilon_N\epsilon' v & 0 \end{pmatrix} \tag{4.5.24}$$

describing the pseudo-Dirac state, while $\nu_e \Rightarrow \nu_\tau$ mixing is now described by

$$\frac{v^2}{M}\begin{pmatrix} (\frac{\epsilon'}{\epsilon})^2 & \frac{\epsilon'}{\epsilon} \\ \frac{\epsilon'}{\epsilon} & 1 \end{pmatrix} \tag{4.5.25}$$

The resulting masses and mixings are given in table 4.2.

The mixing angles in this variation are predicted to be

$$\sin(\theta_\odot) \approx \frac{\epsilon'}{\epsilon} \qquad [i=5]\,, \quad \theta_{atm} = 45^o \qquad [i=0] \tag{4.5.26}$$

As the pseudo-Dirac muon neutrino is still present, the atmospheric angle is unchanged. However, the solar angle is changed somewhat. We should note that values for ϵ and ϵ' extracted for a fit of this model are different than for those of the previous model. Using values from fits in the charged fermion sector, we have $\epsilon \approx 0.03$ and $\epsilon' \approx 5 \times 10^{-4}$ or $\epsilon' \approx 2.4 \times 10^{-4}$ (depending on certain signs), yielding $\theta_\odot \approx O(1.5 \times 10^{-2})$. Given the number of $O(1)$ parameters involved, this is again quite consistent with the BP98 small-angle MSW solution.

The solar splitting scale is unchanged, while the atmospheric splitting is further surpressed by a factor of ϵ.

$$\delta m^2_{atm} \approx \epsilon' \epsilon_N v \epsilon^2 \sqrt{\delta m^2_\odot} \tag{4.5.27}$$

We fit this splitting again with the free parameter $\epsilon_N \approx 10^{-6} - 10^{-7}$. The resulting muon neutrino mass is then

$$m_\nu \approx \frac{\delta m^2_{atm}}{\epsilon^2 \sqrt{\delta m^2_\odot}} \approx 10^{1.7}\text{eV} - 10^{3.5}\text{eV} \qquad [i = 5] \tag{4.5.28}$$

Thus, while the explanations of the solar and atmospheric neutrinos remain, the neutrino becomes potentially dangerous in its cosmological implications. However, given the large stability index of this prediction, there are large uncertainties in the prediction for its mass.

4.6 KARMEN and LSND

The presence of an additional sterile state makes it possible that a signal would be seen in short baseline $\nu_\mu \to \nu_e$ oscillations, such as has been reported at LSND [7]. An estimate of the LSND mixing angle from the neutrino sector gives $\epsilon\epsilon' \delta m^2_\odot / \delta m^2_{atm}$, a very small result. Hence, this mixing originates from the charged lepton sector

$$\theta_{LSND} = \sqrt{\frac{m_e}{m_\mu}} \qquad [i = 0]\,. \tag{4.6.29}$$

The precise predictions for $1-2$ mixing angles in the charged sector is an essential feature of the $U(2)$ flavor symmetry. In the quark sector it is highly successful. In the lepton sector,

$\theta_{LSND} = \sqrt{\frac{m_e}{m_\mu}}$ is only useful if the neutrino mixing is either predicted or small, as in this theory. Recently, the KARMEN experiment has placed limits on the allowed region for such oscillations, giving a limit $m_\nu \leq 0.6$ eV [8]. While the prediction for m_ν has a large stability index in both the general theory as well as the variant theory, because the initial range for m_ν is so high in the variant theory, it is disfavored by this bound.

The general theory is much safer, however. As we discuss in the appendix, the uncertainty due to order one coefficients would allow it to satisfy the KARMEN bound. Such a result would likely coincide with higher values of $\delta m^2_\odot$ and lower values of δm^2_{atm}.

4.7 Astrophysical and Cosmological Implications

There are three important cosmological implications of our theory.

1. We predict a small, but potentially significant amount of neutrino hot dark matter. The KARMEN bound limits us to a 0.6 eV neutrino, but because there are two massive states, it is still within the interesting region for HDM.

2. We predict abundances for light nuclei resulting from four light neutrino species. While newer data suggest D/H ratios lie in the low end of the range previously thought, and thus $N_\nu < 4$, this is still an open question.

3. There may be two further singlet neutrino states, dominantly N_μ and N_τ, at or below the weak scale. Successful nucleosynthesis requires that they decay before the era of nucleosynthesis. Because the mass eigenstates are slightly left-handed, the primary decay mode will be through the process shown in figure 4.1. This is similar to muon decay, which we use as a benchmark. For the lighter of the two states, we estimate its lifetime to be

$$\tau_{N_\mu} \approx \epsilon'^{-2} \left(\frac{\delta m^2_{atm}}{\delta m^2_\odot}\right)^2 \left(\frac{m_\mu}{m_{N_\mu}}\right)^5 \tau_\mu \tag{4.7.30}$$

The mass of this particle is

$$m_{N_\mu} \approx \frac{1}{\epsilon} \frac{(\delta m^2_{atm})^2}{(\delta m^2_\odot)^{3/2}} \qquad [i = 12] \tag{4.7.31}$$

for the general theory and

$$m_{N_\mu} \approx \frac{1}{\epsilon^2} \frac{(\delta m^2_{atm})^2}{(\delta m^2_\odot)^{3/2}} \qquad [i \approx 11 - 16] \tag{4.7.32}$$

in the variant theory. The stability index is approximate because it involves sums of order one coefficients of different powers. Furthermore, i will change depending on which of (4.5.23) are included.

The more dangerous case, the general theory, then has a mass $O(100MeV)$ and thus a lifetime $\tau_{N_\mu} \approx 10^3 s$, which is far too long to be acceptable. However, because the lifetime has a fifth power dependence on the mass, and because the prediction for the mass has index 12, deviations in the order one quantities could very easily push the lifetime down to an acceptable level. As we explore in the appendix, even conservatively we can only reasonably estimate the mass of this particle to be in the range (17MeV, 40GeV), which means that the lifetime could easily be $10^{-9} s$, without even beginning to push the limits of the order one quantities. The details are presented in the appendix.

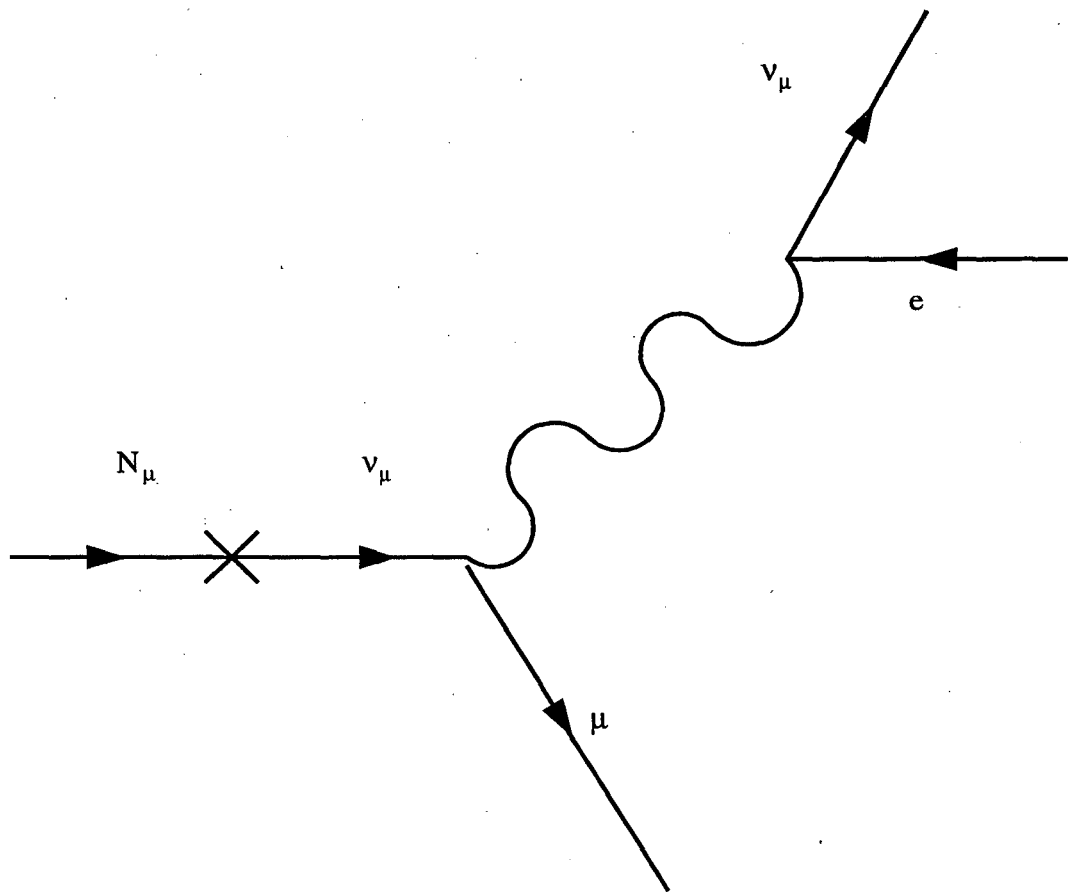

Figure 4.1: Principal decay mode for N_μ.

4.8 Models

The theory described in this paper has a low energy effective Lagrangian of (4.2.2) for charged fermion masses, while the neutrino masses arise from the $U(2) \times U(1)_N$ effective Lagrangian

$$\mathcal{W} = \frac{\phi_N}{M} N_3 l_3 h + \frac{\phi_N}{M^2} \left(N_3 \phi^a l_a h + l_3 \phi^a N_a h + N_a (S^{ab} + A^{ab}) l_b h \right)$$

$$+ \frac{\phi_N^2}{M^3} \left(N_3 N_3 M + N_3 N_a \phi^a M + N_a N_b S^{ab} M \right) \tag{4.8.33}$$

where N_3 and N_a have $U(1)_N$ charges $+1$, while ϕ_N has $U(1)_N$ charged -1. The field ϕ_N gets a vev, breaking $U(1)_N$ and establishing an overall scale for these coefficients: $\frac{<\phi_N>}{M} = \epsilon_N$. This effective theory can result from a renormalizable model by integrating out heavy states, both singlet and doublet under $U(2)$, in the Froggatt-Nielsen mechanism.

This symmetry structure on the right-handed singlet sector is far from unique. Another possibility is for N_a to carry $U(1)_N$ charge, while N_3 is neutral under $U(1)_N$. This has no effect on any of our predictions, since the form of (4.4.10) for the light neutrino mass matrix is unchanged. The only change is that N_3 has a mass of the order of the unification scale M rather than of order $\epsilon_N^2 M$.

Another possible symmetry structure for the theory is $U(2)_\psi \times U(2)_N$, where $U(2)_\psi$ acts as usual on all the matter with non-trivial $SU(3) \times SU(2) \times U(1)$ quantum numbers, while $U(2)_N$ acts only on the three right-handed neutrinos, with N_3 a singlet and N_a a doublet. The matrix M_{RR} now has the form of (4.4.7), and arises from the renormalizable interactions

$$W_{RR} = MN_3N_3 + N_3\phi^A N_A + N_A S^{AB} N_B \tag{4.8.34}$$

with vevs for S^{22} and ϕ^2 being of order ϵM and breaking $U(2)_N \to U(1)_N$. The interactions for M_{LR} are

$$W_{LR} = l_3 N_3 h + \frac{1}{M}\left(l_a \phi^a N_3 h + l_3 \phi^A N_A h + l_a R^{aA} N_A h\right) \tag{4.8.35}$$

where R^{aA} transforms as a (2,2). The vev for R^{22} is also of order ϵM, since this is the scale of breaking of $U(2)_\psi \times U(2)_N \to U(1)_\psi \times U(1)_N$. The breaking scale for $U(1)_\psi$ is $\epsilon' M$, so the vev of R^{12} takes this value. On the other hand, $U(1)_N$ is broken by R^{21}. We choose this scale to be smaller by a factor of ϵ_N, $< R^{21} > \approx \epsilon_N \epsilon' M$, giving

$$M_{LR} = v \begin{pmatrix} 0 & \epsilon' & 0 \\ \epsilon'\epsilon_N & \epsilon & \epsilon \\ 0 & \epsilon & 1 \end{pmatrix} \tag{4.8.36}$$

Integrating out the heavy states N_2 and N_3, which now have masses of order the unification scale, this theory now reproduces (4.4.10) for the mass matrix of the four light neutrinos.

The common features of these models, which are inherent to our scheme, are:

- There is a $U(2)$ symmetry, which acts on the known matter as $\psi_3 \oplus \psi_a$, and is broken sequentially at scale ϵM and $\epsilon' M$.

- A $U(2)$ symmetry also acts on the three right-handed neutrinos with N_3 a singlet and $N_{1,2}$ a doublet. This $U(2)$, together with the symmetry of the Majorana mass, implies that N_1 does not have a Majorana mass and becomes a fourth light neutrino.

- There is an addition to the flavor group, beyond the $U(2)$ which acts on ψ. At least part of this additional flavor symmetry is broken at a scale very much less than M, leading to a small Dirac mass coupling of $\nu_\mu N_e$. Such a small symmetry breaking scale could be generated by the logarithmic evolution of a scalar m^2 term.

4.9 Conclusions

There are several theories with sterile neutrinos [9, 10, 11] some of which have 4×4 textures that split into two 2×2 matrices. Such theories provide a simple picture for atmospheric oscillations via $\nu_\mu \to \nu_s$, and solar oscillations via $\nu_e \to \nu_\tau$, with $\delta m^2_\odot \approx \frac{v^2}{M} \approx 10^{-5}\text{eV}^2$ for $M \approx M_{unif}$. However, theories of this kind typically do not provide an understanding for several key points:

- Why is the Majorana mass of the singlet state ν_s small, allowing ν_s in the low energy theory?
- Why does ν_s mix with ν_μ rather than with ν_e or ν_τ?
- Why is the $\nu_s - \nu_\mu$ state pseudo-Dirac, leading to 45^o mixing?
- How can this extended neutrino sector be combined with the pattern of charged quark and lepton masses in a complete theory of flavor?
- What determines the large number of free parameters in the neutrino sector?

In the theory presented here, all these questions are answered: the key tool is the $U(2)$ flavor symmetry, motivated several years ago by the charged fermion masses and the supersymmetric flavor problem. The simplest pattern of $U(2)$ symmetry breaking consistent

with the charged fermion masses does not allow a Majorana mass for one of the three right-handed neutrinos. Furthermore, it is precisely this right-handed state that has a Dirac coupling to ν_μ but not to ν_e or ν_τ, guaranteeing that ν_μ is pseudo-Dirac with a 45^o mixing angle.

Our theory provides a unified description of both charged fermion and neutrino masses, in terms of just three small symmetry breaking parameters and a set of order unity coefficients. Some predictions, such as $|V_{ub}/V_{cb}| = \sqrt{m_u/m_c}$ and $\theta_{atm} = 45^o$ are independent of the order unity coefficients and are precise. Other predictions, such as $|V_{cb}| \approx m_s/m_b$ and $\theta_\odot \approx \sqrt{m_e m_\mu/m_\tau^2}$ involve the order unity coefficients and are approximate. In the appendix we have introduced the "stability index" which attempts to quantify the uncertainty in such predictions according to the power of the unknown order unity coefficients appearing in the prediction.

There is one further free parameter of the theory—the overall mass scale M setting the normalization of the right handed Majorana mass matrix. If M is taken to be the scale of coupling constant unification $\delta m^2_\odot \approx 10^{-5} eV^2$.

The value of δm^2_{atm} is not predicted— this is the largest deficiency of the theory. It can be described by a very small flavor symmetry breaking parameter. Once this parameter is set by the observed value of δm^2_{atm}, it can be used to predict the approximate mass range of the pseudo-Dirac ν_μ to be in the range $10^{0.4} - 10^2$eV, with significant additional uncertainty due to order one coefficients. This, even with the KARMEN bound, allows for a neutrino of cosmological interest with $\sum_i m_{\nu_i} \approx 1$ eV. Such a neutrino could be seen at short baseline experiments, and may have already been seen by LSND. Searching

Experiment	Mode	Signal
Present solar ν exp.	$\nu_e \to \nu_\tau$	All data consistent with 2-flavor MSW
SNO	$\nu_e \to \nu_\tau$	Confirm SK measurement of B8. Measure $\frac{\phi_{NC}}{\phi_{CC}} \neq 1$
Borexino	$\nu_e \to \nu_\tau$	Consistent with small-angle 2-flavor MSW
KAMLAND	$\nu_e \to \nu_\tau$	No signal
LSND,KARMEN	$\nu_\mu \to \nu_e$	$\sin^2(2\theta) = 2 \times 10^{-2}$
K2K	$\nu_\mu \to N_e$	ν_μ disappearance. No e appearance
MINOS, ICARUS	$\nu_\mu \to N_e$	ν_μ disappearance. No τ appearance.
Atmospheric ν exp.	$\nu_\mu \to N_e$	Confirm 2 flavor $\nu_\mu \to \nu_s$ with 45^o mixing.

Table 4.3: Experimental signals.

for $\nu_\mu \to \nu_e$, with $\sin^2(2\theta) = 2 \times 10^{-2}$, below the current limit of δm^2 is an important experiment for the $U(2)$ theory, since it is this prediction which differentiates $U(2)$ from several other theories with a light singlet neutrino.

Predictions of the theory for experiments sensitive to neutrino oscillations are listed in table 4.3. We expect a small angle MSW solution to the solar neutrino anomaly, through a $\nu_e \Rightarrow \nu_\tau$ oscillation. The atmospheric neutrino anomaly is from $\nu_\mu \Rightarrow \nu_s$. This will be distinguishable from $\nu_\mu \Rightarrow \nu_\tau$ through a number of means: LBL experiments will see ν_μ disappearance, but no ν_e or ν_τ appearance. Improved statistics from Super-Kamiokande will be useful in distinguishing $\nu_\mu \Rightarrow \nu_\tau$ and $\nu_\mu \Rightarrow \nu_s$, for example via inclusive studies of multi-ring events [12].

Bibliography

[1] *Study of the atmospheric neutrino flux in the multi-GeV energy range,* The Super-Kamiokande Collaboration, hep-ex/9805006, PLB436:33-41,1998; *Evidence for oscillation of atmospheric neutrinos ,* The Super-Kamiokande Collaboration, hep-ex/9807003, PRL81:1562-1567,1998.

[2] A. Pomarol, D. Tommasin, hep-ph/9507462, Nucl.Phys.B466:3-24,1996.

[3] R. Barbieri, G. Dvali and L.J. Hall. Phys.Lett.B377:76-82,1996, hep-ph/9512388; R. Barbieri, L.J. Hall and A. Romanino. Phys.Lett.B401:47-53,1997, hep-ph/9702315.

[4] C.D. Carone and L.J. Hall. Phys.Rev.D56:4198-4206,1997, hep-ph/9702430

[5] J.N. Bahcall, P.I. Krastev, A.Yu. Smirnov, Jul 1998. hep-ph/9807216.

[6] R. Barbieri, L.J. Hall, S. Raby and A. Romanino. Nucl.Phys.B493:3-26,1997, hep-ph/9610449; R. Barbieri and L.J. Hall. Nuovo Cim.110A:1-30,1997, hep-ph/9605224

[7] By LSND Collaboration (C. Athanassopoulos et al.) Phys.Rev.Lett.81:1774-1777,1998. nucl-ex/9709006.

[8] The Search for Neutrino Oscillations $\overline{\nu}_\mu \rightarrow \overline{\nu}_e$ with KARMEN, KARMEN Collabo-

ration, Proceedings Contribution to Neutrino98 in Takayama, Japan, June 4-9, 1998. hep-ex/9809007

[9] A.S. Joshipura, A.Yu. Smirnov. PRL-TH-98-005. hep-ph/9806376

[10] D.O. Caldwell and R. Mohapatra, Phys.Rev.D50:3477, 1994; E.J. Chun, A.S. Joshipura and A. Yu. Smirnov, Phys.Lett.B357:608, 1995. R. Foot and R Volkas, Phys. Rev. D52:6595, 1995; Z. Berezhiani and R.N. Mohapatra, Phys.Rev.D52:6607, 1995; E.Ma and P.Roy, Phys.Rev.D52:R4780, 1995; J.T. Peltoniemi and J.W.F. Valle, Nucl.Phys.B406:409, 1993.

[11] E.J. Chun, A.S. Joshipura and A.Yu.Smirnov, Phys.Rev.D54:4654, 1996. hep-ph/9507371; K Benakli and A.Yu. Smirnov, Phys.Rev.Lett. 79:4314, 1997. hep-ph/9703465; E.J.Chun, C.W. Kim and U.W. Lee, hep-ph/9802209.

[12] L.J. Hall, H. Murayama. hep-ph/9806218.

Chapter 5

Neutrino Mass Anarchy

5.1 Introduction

Neutrinos are the most poorly understood among known elementary particles, and have important consequences in particle and nuclear physics, astrophysics and cosmology. Special interests are devoted to neutrino oscillations, which, if they exist, imply physics beyond the standard model of particle physics, in particular neutrino masses. The SuperKamiokande data on the angular dependence of the atmospheric neutrino flux provides strong evidence for neutrino oscillations, with ν_μ disappearance via large, near maximal mixing, and $\Delta m^2_{atm} \approx 10^{-3}$ eV^2[1]. Several measurements of the solar neutrino flux can also be interpreted as neutrino oscillations, via ν_e disappearance[2]. While a variety of $\Delta m^2_\odot$ and mixing angles fit the data, in most cases $\Delta m^2_\odot$ is considerably lower than Δm^2_{atm}, and even in the case of the large angle MSW solution, the data typically require $\Delta m^2_\odot \approx 0.1 \Delta m^2_{atm}$[3]. The neutrino mass matrix apparently has an ordered, hierarchical form for the eigenvalues, even though it has a structure allowing large mixing angles.

All attempts at explaining atmospheric and solar neutrino fluxes in terms of neutrino oscillations have resorted to some form of ordered, highly structured neutrino mass

matrix[4]. These structures take the form $M_0 + \epsilon M_1 + ...$, where the zeroth order mass matrix, M_0, contains the largest non-zero entries, but has many zero entries, while the first order correction terms, ϵM_1, have their own definite texture, and are regulated in size by a small parameter ϵ. Frequently the pattern of the zeroth order matrix is governed by a flavor symmetry, and the hierarchy of mass eigenvalues result from carefully-chosen, small, symmetry-breaking parameters, such as ϵ. Such schemes are able to account for both a hierarchical pattern of eigenvalues, and order unity, sometimes maximal, mixing. Mass matrices have also been proposed where precise numerical ratios of different entries lead to the desired hierarchy and mixing.

In this letter we propose an alternative view. This new view selects the large angle MSW solution of the solar neutrino problem, which is preferred by the day to night time flux ratio at the 2σ level[2]. While the masses and mixings of the charged fermions certainly imply regulated, hierarchical mass matrices, we find the necessity for an ordered structure in the neutrino sector to be less obvious. Large mixing angles would result from a random, structureless matrix, and such large angles could be responsible for solar as well as atmospheric oscillations. Furthermore, in this case the hierarchy of Δm^2 need only be an order of magnitude, much less extreme than for the charged fermions. We therefore propose that the underlying theory of nature has dynamics which produces a neutrino mass matrix which, from the viewpoint of the low energy effective theory, displays *anarchy*: all entries are comparable, no pattern or structure is easily discernable, and there are no special precise ratios between any entries. Certainly the form of this mass matrix is not governed by approximate flavor symmetries.

There are four simple arguments against such a proposal

- The neutrino sector exhibits a hierarchy with $\Delta m^2_{\odot} \approx 10^{-5} - 10^{-3}\text{eV}^2$ for the large mixing angle solution, while $\Delta m^2_{atm} \approx 10^{-3} - 10^{-2}\text{eV}^2$,
- Reactor studies of $\overline{\nu}_e$ at the CHOOZ experiment have indicated that mixing of ν_e in the 10^{-3}eV^2 channel is small [5], requiring at least one small angle,
- Even though large mixing would typically be expected from anarchy, *maximal* or near maximal mixing, as preferred by SuperKamiokande data, would be unlikely,
- ν_e, ν_μ and ν_τ fall into doublets with e_L, μ_L and τ_L, respectively, whose masses are extremely hierarchical ($m_e : m_\mu : m_\tau \approx 10^{-4} : 10^{-1} : 1$).

By studying a sample of randomly generated neutrino mass matrices, we demonstrate that each of these arguments is weak, and that, even when taken together, the possibility of neutrino mass anarchy still appears quite plausible.

5.2 Analysis

We have performed an analysis of a sample of random neutrino matrices. We investigated three types of neutrino mass matrices: Majorana, Dirac and seesaw. For the Majorana type, we considered 3×3 symmetric matrices with 6 uncorrelated parameters. For the Dirac type, we considered 3×3 matrices with 9 uncorrelated parameters. Lastly, for the seesaw type, we considered matrices of the form $M_D M_{RR}^{-1} M_D^T$[6], where M_{RR} is of the former type and M_D is of the latter. We ran one million sample matrices with independently generated elements, each with a uniform distribution in the interval $[-1, 1]$ for each matrix type: Dirac, Majorana and seesaw.

To check the robustness of the analysis, we ran smaller sets using a distribution with the logarithm base ten uniformly distributed in the interval $[-1/2, 1/2]$ and with random sign. We further checked both of these distributions but with a phase uniformly distributed in $[0, 2\pi]$. Introducing a logarithmic distribution and phases did not significantly affect our results (within a factor of two), and hence we discuss only matrices with a linear distribution and real entries.

We make no claim that our distribution is somehow physical, nor do we make strong quantitative claims about the confidence intervals of various parameters. However, if the basic prejudices against anarchy fail in these simple distributions, we see no reason to cling to them.

In each case we generated a random neutrino mass matrix, which we diagonalized with a matrix U. We then investigated the following quantities:

$$R \equiv \Delta m^2_{12}/\Delta m^2_{23}, \tag{5.2.1}$$

$$s_C \equiv 4|U_{e3}|^2(1-|U_{e3}|^2), \tag{5.2.2}$$

$$s_{atm} \equiv 4|U_{\mu 3}|^2(1-|U_{\mu 3}|^2), \tag{5.2.3}$$

$$s_\odot \equiv 4|U_{e2}|^2|U_{e1}|^2, \tag{5.2.4}$$

where Δm^2_{12} is the smallest splitting and Δm^2_{23} is the next largest splitting. What ranges of values for these parameters should we demand from our matrices? We could require they lie within the experimentally preferred region. However, as experiments improve and these regions contract, the probability that a random matrix will satisfy this goes to zero. Thus we are instead interested in mass matrices that satisfy certain *qualitative* properties. For our numerical study we select these properties by the specific cuts

Dirac	no cuts	s_{atm}	$s_\odot$	$s_{atm} + s_\odot$
no cuts	1,000,000	671,701	184,128	135,782
s_C	145,000	97,027	66,311	45,810
R	106,771	78,303	17,538	14,269
$s_C + R$	12,077	9,067	5,656	4,375
Majorana	no cuts	s_{atm}	$s_\odot$	$s_{atm} + s_\odot$
no cuts	1,000,000	709,076	200,987	164,198
s_C	121,129	91,269	70,350	56,391
R	200,452	149,140	37,238	31,708
$s_C + R$	21,414	16,507	12,133	10,027
seesaw	no cuts	s_{atm}	$s_\odot$	$s_{atm} + s_\odot$
no cuts	1,000,000	594,823	210,727	133,800
s_C	186,684	101,665	86,511	49,787
R	643,394	390,043	132,649	86,302
$s_C + R$	115,614	64,558	53,430	31,547

Table 5.1: Mass matrices satisfying various sets of cuts for the real linear Dirac, Majorana and seesaw scenarios.

- $R < 1/10$ to achieve a large hierarchy in the Δm^2.
- $s_C < 0.15$ to enforce small ν_e mixing through this Δm^2.
- $s_{atm} > 0.5$ for large atmospheric mixing.
- $s_\odot > 0.5$ for large solar mixing.

The results of subjecting our 10^6 sample matrices, of Dirac, Majorana and seesaw types, to all possible combinations of these cuts is shown in Table 5.1. First consider making a single cut. As expected, for all types of matrices, a large percentage (from 18% to 21%) of the random matrices pass the large mixing angle solar cut, and similarly for the large mixing angle atmospheric cut (from 59% to 71%). Much more surprising, and contrary to conventional wisdom, is the relatively large percentage passing the individual cuts for R (from 10% to 64%) and for s_C (from 12% to 18%). The distribution for R is shown in Figure

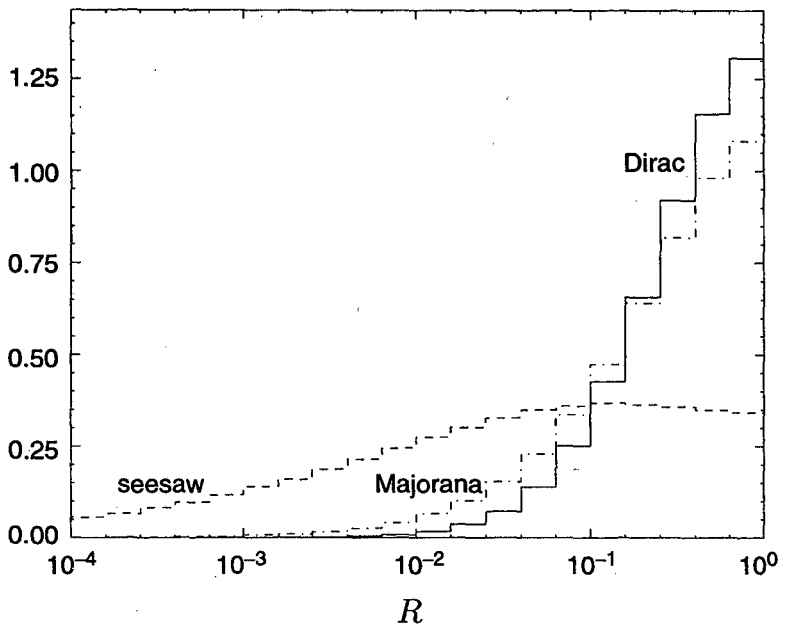

Figure 5.1: The distribution of $\Delta m^2_{\odot}/\Delta m^2_{atm}$ for Dirac (solid) Majorana (dot-dashed) and seesaw (dashed) scenarios.

5.1. Naively, one might expect that this would peak at $R = 1$, which is largely the case for Dirac matrices, although with a wide peak. In the Majorana case there is an appreciable fraction ($\sim 20\%$) that have a splitting $R \leq 1/10$, while in the seesaw scenario the *majority* of cases ($\sim 64\%$) have a splitting $R \leq 1/10$ — it is not at all unusual to generate a large hierarchy.

We can understand this simply: first a splitting of a factor of 10 in the Δm^2's corresponds to only a factor of 3 in the masses themselves if they happen to be hierarchically arranged. Secondly, in the seesaw scenario, taking the product of three matrices spreads the Δm^2 distribution over a wide range.

While one would expect random matrices to typically give large atmospheric mixing, is it plausible that they would give near-maximal mixing, as required by the SuperKamiokande data? In Figure 5.2 we show distributions of s_{atm}, which actually *peak*

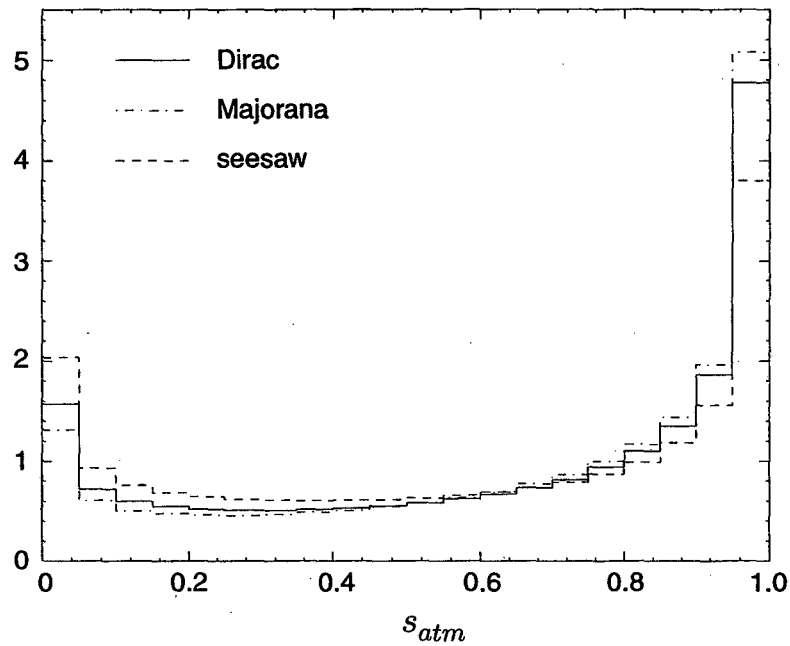

Figure 5.2: Plots of the normalized, binned distributions of s_{atm} for Dirac, Majorana and seesaw cases. Contrary to intuition, the distributions actually peak at large s_{atm}.

in the $0.95 < s_{atm} < 1.0$ bin. We conclude that it is not necessary to impose a precise order on the mass matrix to achieve this near-maximal mixing. Finally, we consider correlations between the various cuts. For example, could it be that the cuts on R and s_C selectively pass matrices which accidentally have a hierarchical structure, such that s_{atm} and $s_\odot$ are also small in these cases? From Table 5.1 we see that there is little correlation of s_{atm} with s_C or R: the fraction of matrices passing the s_{atm} cut is relatively insensitive to whether or not the s_C or R cuts have been applied. However, there is an important anticorrelation between $s_\odot$ and s_C cuts; for example, in the seesaw case roughly half of the matrices satisfying the s_C cut satisfy the $s_\odot$ cut, compared with 20% of the original set. This anticorrelation is shown in more detail in Figure 5.3, which illustrates how the s_C cut serves to produce a peak at large mixing angle in the $s_\odot$ distribution.

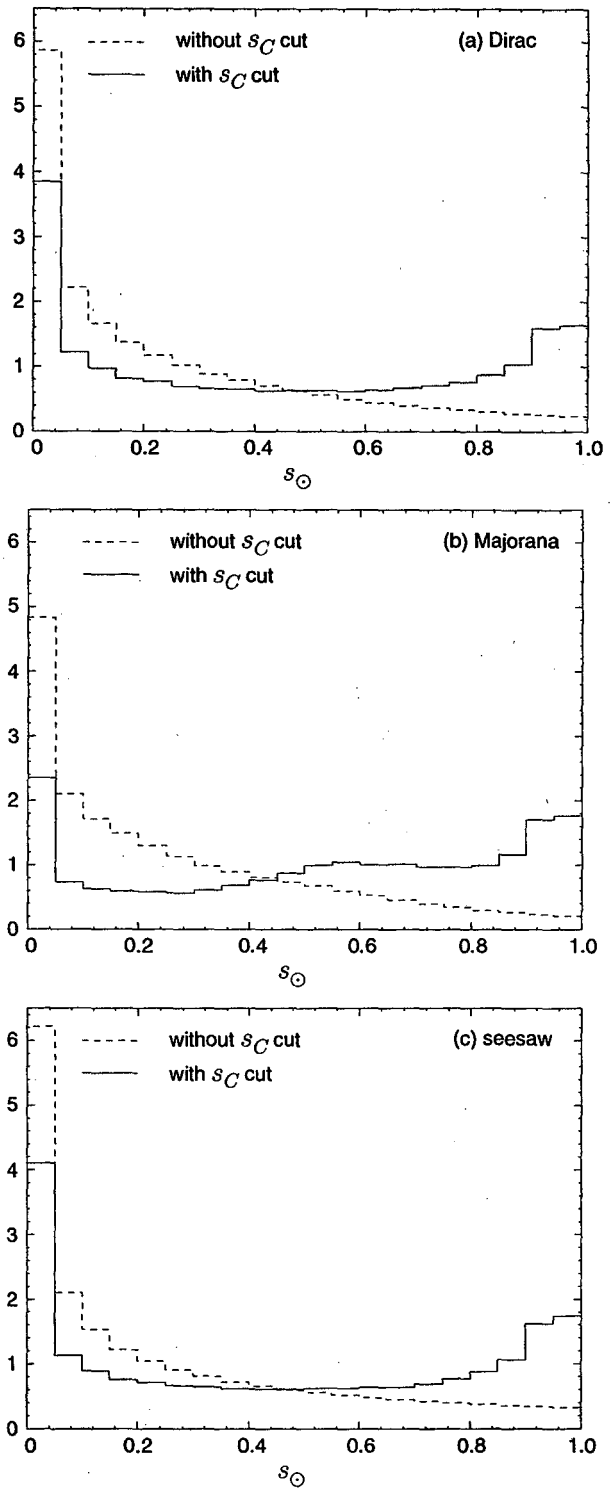

Figure 5.3: Plots of the normalized, binned distributions of $s_\odot$ for Dirac (a), Majorana (b), and seesaw (c) cases. The distribution after imposing the s_C cut (solid) shows a greater preference for large $s_\odot$ compared with the original distribution (dashed).

For random matrices we expect the quantity

$$s_C + s_\odot = 4(|U_{e1}U_{e2}|^2 + |U_{e1}U_{e3}|^2 + |U_{e2}U_{e3}|^2) \tag{5.2.5}$$

to be large, since otherwise ν_e would have to be closely aligned with one of the mass eigenstates. Hence, when we select matrices where s_C happens to be small, we are selecting ones where $s_\odot$ is expected to be large.

5.3 Right Handed Flavor Symmetries

We have argued that the neutrino mass matrix may follow from complete anarchy, however the electron, muon, tau mass hierarchies imply that the charged fermion mass matrix has considerable order and regularity. What is the origin for this difference? The only answer which we find plausible is that the lepton doublets, $(\nu_l, l)_L$, appear randomly in mass operators, while the lepton singlets, l_R, appear in an orderly way, for example, regulated by an approximate flavor symmetry. This idea is particularly attractive in SU(5) grand unified theories where only the 10-plets of matter feel the approximate flavor symmetry, explaining why the mass hierarchy in the up quark sector is roughly the square of that in the down quark and charged lepton sectors. Hence we consider a charged lepton mass matrix of the form

$$M_l = \hat{M}_l \begin{pmatrix} \lambda_e & 0 & 0 \\ 0 & \lambda_\mu & 0 \\ 0 & 0 & \lambda_\tau \end{pmatrix} \tag{5.3.6}$$

where $\lambda_{e,\mu,\tau}$ are small flavor symmetry breaking parameters of order the corresponding Yukawa couplings, while $\hat{M}_l$ is a matrix with randomly generated entries. We generated one million neutrino mass matrices and one million lepton mass matrices, and provide

cuts	none	s_{atm}	$s_{\odot}$	$s_{atm} + s_{\odot}$
none	1,000,000	537,936	221,785	126,914
s_C	222,389	102,178	99,050	50,277
R	643,127	345,427	142,789	81,511
$s_C + R$	143,713	65,875	63,988	32,435

Table 5.2: Mass matrices satisfying various sets of cuts for the real linear seesaw scenario, with additional mixing from the charged lepton sector.

results for the mixing matrix $U = U_l^{\dagger} U_{\nu}$, where U_{ν} and U_l are the unitary transformations on ν_l and l_l which diagonalize the neutrino and charged lepton mass matrices. We find that the additional mixing from the charged leptons does not substantially alter any of our conclusions – this is illustrated for the case of seesaw matrices in Table 5.2. The mixing of charged leptons obviously cannot affect R, but it is surprising that the distributions for the mixings $s_{atm,\odot,C}$ are not substantially changed.

5.4 Conclusions

All neutrino mass matrices proposed for atmospheric and solar neutrino oscillations have a highly ordered form. In contrast, we have proposed that the mass matrix appears random, with all entries comparable in size and no precise relations between entries. We have shown, especially in the case of seesaw matrices, that not only are large mixing angles for solar and atmospheric oscillations expected, but $\Delta m^2_{\odot} \approx 0.1 \Delta m^2_{atm}$, giving an excellent match to the large angle solar MSW oscillations, as preferred at the 2σ level in the day/night flux ratio. In a sample of a million random seesaw matrices, 40% have such mass ratios and a large atmospheric mixing. Of these, about 10% also have large solar mixing while having small ν_e disappearance at reactor experiments. Random neutrino mass matrices produce a narrow peak in atmospheric oscillations around the observationally preferred

case of maximal mixing. In contrast to flavor symmetry models, there is no reason to expect U_{e3} is particularly small, and long baseline experiments which probe Δm^2_{atm}, such as K2K and MINOS, will likely see large signals in $\overline{\nu}_e$ appearance. If Δm^2_{atm} is at the lower edge of the current Superkamiokande limit, this could be seen at a future extreme long baseline experiment with a muon source. Furthermore, in this scheme $\Delta m^2_{\odot}$ is large enough to be probed at KamLAND, which will measure large $\bar{\nu}_e$ disappearance.

Bibliography

[1] Super-Kamiokande Collaboration (Y. Fukuda et al.), Phys.Rev.Lett. **81** (1998) 1562, hep-ex/9807003.

[2] Y. Suzuki, talk at the "XIX International Symposium on Lepton and Photon Interactions at High Energies", Stanford University, August 9-14, 1999.

[3] For a recent fit, see M.C. Gonzalez-Garcia et al., hep-ph/9906469.

[4] For papers after the Superkamiokande result, see, for instance, C.H. Albright, S.M. Barr, Phys.Lett. **B461** (1999) 218, hep-ph/9906297; C.H. Albright, S.M. Barr, Phys.Lett. **B452** (1999) 287, hep-ph/9901318; G. Altarelli, F. Feruglio, Phys.Lett. **B439** (1998) 112, hep-ph/9807353; G. Altarelli, F. Feruglio, JHEP **9811** (1998) 021, hep-ph/9809596; G. Altarelli, F. Feruglio, Phys.Lett. **B451** (1999) 388, hep-ph/9812475; A. Aranda, C. Carone, R.F. Lebed, hep-ph/9910392; K.S. Babu, J.C. Pati, F. Wilczek, hep-ph/9812538; R. Barbieri, L.J. Hall and A. Strumia, Phys.Lett. **B445** (1999) 407, hep-ph/9808333; R. Barbieri, P. Creminelli, A. Romanino, hep-ph/9903460; T. Blazek, S. Raby, K. Tobe, Phys.Rev. **D60** (1999) 113001, hep-ph/9903340; K. Choi, E.J. Chun, K. Hwang, Phys.Rev. **D60** (1999) 031301, hep-ph/9811363; R. Dermisek, S. Raby, hep-ph/9911275; J. Ellis, G.K. Leontaris, S. Lola, D.V. Nanopou-

los, Eur.Phys.J. **C9** (1999) 389, hep-ph/9808251; J.K. Elwood, N. Irges, P. Ramond, Phys.Rev.Lett. **81** (1998) 5064, hep-ph/9807228; P.H. Frampton, A. Rasin, hep-ph/9910522; P.H. Frampton, S.L. Glashow, Phys.Lett. **B461** (1999) 95, hep-ph/9906375; H. Fritzsch, Z.Z Xing, Phys.Lett. **B440** (1998) 313, hep-ph/9808272 C.D. Froggatt, M. Gibson, H.B. Nielsen, Phys.Lett. **B446** (1999) 256, hep-ph/9811265; M. Fukugita, M. Tanimoto, T. Yanagida, Phys.Rev. **D59** (1999) 113016, hep-ph/9809554; H. Georgi, S.L. Glashow, hep-ph/9808293; M.E. Gomez, G.K. Leontaris, S. Lola and J.D. Vergados, Phys.Rev. **D59** (1999) 116009, hep-ph/9810291; L.J. Hall, N. Weiner, Phys.Rev. **D60** (1999) 033005, hep-ph/9811299; L.J. Hall, D. Smith, Phys.Rev **D59** (1999) 113013, hep-ph/9812308; N. Irges, S. Lavignac, P. Ramond, Phys.Rev. **D58** (1998) 035003, hep-ph/9802334; A.S. Joshipura, Phys.Rev. **D59** (1999) 077301, hep-ph/9808261; A.S. Joshipura, S.D. Rindani, hep-ph/9811252; S. King, Phys.Lett. **B439** (1998) 350, hep-ph/9806440; S.K. Kang, C.S. Kim, Phys.Rev. **D59** (1999) 091302, hep-ph/9811379; G.K. Leontaris, S. Lola, C. Scheich and J.D. Vergados, Phys.Rev. **D53**, 6381(1996); S. Lola, J.D. Vergados, Prog.Part.Nucl.Phys. **40** (1998) 71; S. Lola, G.G. Ross, Nucl.Phys. **B553** (1999) 81, hep-ph/9902283; R.N. Mohapatra, S. Nussinov, Phys.Rev. **D60**, (1999) 013002, hep-ph/9809415; Y. Nomura, T. Yanagida, Phys.Rev. **D59** (1999) 017303, hep-ph/9807325; T. Yanagida, J. Sato, Nucl.Phys.Proc.Suppl. **77** (1999) 293, hep-ph/9809307; M. Tanimoto, hep-ph/9807517.

[5] M. Apollonio et al., hep-ex/9907037.

[6] M. Gell-Mann, P. Ramond, R. Slansky, in Supergravity, ed. by F. van Nieuwenhuizen and D. Freedman (Amsterdam, North Holland, 1979) 315; T. Yanagida, in Proc. of the

workshop on the unified Theory and Baryon Number in the Universe, eds. O. Sawada and A. Sugamoto (KEK, Tsukuba, 1979) 95.

Chapter 6

Stabilizing Large Extra Dimensions

6.1 Introduction

It has recently been realized that the fundamental scales of gravitational and string physics can be far beneath $\sim 10^{18}$ GeV, in theories where the Standard Model fields live on a 3-brane in large-volume extra dimensions [1]. Lowering these fundamental scales close to the weak scale provides a novel approach to the hierarchy problem, and implies that the structure of quantum gravity may be experimentally accessible in the near future.

While this prospect is very exciting, two important theoretical issues need to be addressed for this scenario to be as compelling as the more "standard" picture with high fundamental scale, where the hierarchy is stabilized by SUSY dynamically broken at scales far beneath the string scale. First: what generates the large volume of the extra dimensions? And second: what about the successful picture of logarithmic gauge coupling unification in the supersymmetric standard model? The success is so striking that we do not wish to think it is an accident.

One way of generating a large volume for the extra dimensions involves considering a highly curved bulk. Indeed Randall and Sundrum have proposed a scenario where the

bulk volume can be exponentially larger than the proper size of a single extra dimension [2]. Goldberger and Wise then showed how such a dimension could be stabilized [3]. In the original proposal of [1], however, the bulk was taken to be very nearly flat. Previous attempts at stabilizing large dimensions in this framework involved the introduction of large integer numbers in the theory, such as large topological charges [4, 5] or large numbers of branes [5]. In this paper, we instead demonstrate how to stabilize *exponentially* large dimensions in the framework of [1].

The set-up needed to accomplish this meshes nicely with recent discussions of how the success of logarithmic gauge coupling unification can be maintained with large dimensions and low string scale. In [6, 7, 8, 9, 10] it was argued that logarithmic gauge coupling unification may be reproduced in theories with (sets of) two large dimensions. If various light fields propagate in effectively two transverse dimensions, then the logarithmic Green's functions for these fields can give rise to logarithmic variation of the parameters on our brane universe; in cases with sufficient supersymmetries, this logarithmic variation can exactly reproduce the logarithmic running of couplings seemingly far above the (now very low) string scale. This phenomena is another example of the bulk reproducing the physics of the desert, this time with quantitative precision. Of course, for the "infrared running" picture to work after SUSY breaking, we must assume that SUSY is not broken in the bulk but only directly on branes. This is the analogue of softly breaking SUSY at low energies in the usual desert picture.

It is interesting that these same ingredients: sets of two transverse dimensions with SUSY in the bulk, only broken on branes, can also be used to address the issue of

large radius stabilization. Indeed, in the SUSY limit, there is no bulk cosmological constant and there is no potential for the radii; they can be set at any size. The crucial point is that once SUSY is broken on branes with a characteristic scale Λ^4, locality guarantees that no bulk cosmological constant is induced, and therefore the effective potential for the radius moduli does not develop any positive power-law dependence on the volume of the transverse dimensions. For two transverse dimensions, logarithmic variation of light bulk fields can then give rise to a logarithmic potential for the size, R, of the extra dimensions:

$$V(R) \sim \Lambda^4 f(\log(RM_*)) \tag{6.1.1}$$

where M_* is the fundamental scale of the theory. This can arise, for instance, from the infrared logarithmic variation of coupling constants on branes where SUSY is broken or from inter-brane forces [7, 10, 11]. Since $\log(R)$ rather R itself is the natural variable, if the potential has parameters of $O(10)$, a minimum can result at $\log(\overline{R}) \sim 10$, thereby generating an exponentially large radius and providing a genuine solution to the hierarchy problem, on the same footing as technicolor or dynamical SUSY breaking.

This idea is appealing and general; relying only on sets of two transverse dimensions (for the logarithmic dependence) and supersymmetry in the bulk (to stably guarantee the absence of a bulk constant which would induce power-law corrections to the effective potential for the radii). It makes the existence of large extra dimensions seem plausible. However, the discussions in [7, 10, 11] have only pointed out this possibility on general grounds but have not presented concrete models realizing the idea. In this paper we remedy this situation by presenting an explicit example of a simple theory with two extra dimensions, which stabilizes exponentially large dimensions. The interaction of branes with

massless bulk scalar fields induces a logarithmic potential for the area A of the transverse dimensions of the form

$$V(A) = -f^4 + \frac{v^4}{\log(AM_*^2)} + w^4 \log(AM_*^2). \tag{6.1.2}$$

This potential is minimized for an area

$$\overline{A}M_*^2 = e^{v^2/w^2} \tag{6.1.3}$$

and so only a ratio of $v/w \sim 6$ is needed to generate an area to generate the $\sim (\text{mm})^2$ area needed to solve the hierarchy problem with $M_* \sim$ TeV. There is a single fine-tuning among the parameters v, w and f, which are all of order M_*, to set the 4D cosmological constant to zero.

6.2 The Radion Signal

Since the potential for the radii of the extra dimensions vary only logarithmically, one might worry that the mass of the radius modulus about the minimum of the potential will be too light. In fact, the mass turns out to be just in the millimeter range, and gives an observable deviation from Newton's law at sub-millimeter distances.

Consider a 6 dimensional spacetime with metric of the form

$$ds^2 = g_{\mu\nu}(x)dx^\mu dx^\nu + R^2(x)\tilde{g}(y)_{mn}dy^m dy^n, \tag{6.2.4}$$

where the geometry of $\tilde{g}$ is taken to be fixed at high energy scales; for example by brane configurations, as illustrated in the next section. The low energy 4D effective field theory involves the 4D graviton together with the radion field, $R(x)$, which feels the potential of eq. (1). After a Weyl rescaling of the metric to obtain canonical kinetic terms, the radion

is found to have a mass

$$m_R^2 \sim \frac{R^2 V''(R)}{M_{Pl}^2} \sim \frac{\Lambda^4}{M_{Pl}^2} f''(\log R) \sim \left(\frac{\mathrm{TeV}^2}{M_{Pl}}\right)^2 \sim \mathrm{mm}^{-2}. \tag{6.2.5}$$

Hence, an interesting general consequence of such logarithmic potentials is that the mass of the radion is naturally in the millimeter range for supersymmetry breaking and fundamental scales $\Lambda \sim M_* \sim$TeV. This order of magnitude result is important for mm range gravity experiments, because the Weyl rescaling introduces a gravitational strength coupling of the radion to the Standard Model fields, so that radion exchange modifies the Newtonian potential to

$$V(r) = -\frac{G_N m_1 m_2}{r}\,(1 + 2e^{-m_R r}). \tag{6.2.6}$$

For a radion which determines the size of an n dimensional bulk, the coefficient of the exponential is $4n/(n+4)$, so that an observation of a coefficient corresponding to $n = 2$ would be a dramatic signal of our mechanism.

It might be argued that, since M_* is larger than 50–100 TeV for $n = 2$ from astrophysics and cosmology ([12, 13]), m_R will be sufficiently large that the range of the radion-mediated force will be considerably less than than a mm, making an experimental discovery extremely difficult. This conclusion is incorrect, for several reasons:

- The astrophysical and cosmological limits are derived from graviton emission and hence constrain the gravitational scale, which may be somewhat larger than the fundamental scale, M_*.

- It is the scale of supersymmetry breaking on the branes, Λ, which determines m_R, and this may be less than M_*, reducing m_R and making the range of the Yukawa potential larger.

- The radion mass may be reduced from the order of magnitude estimate $m_R \approx \Lambda^2/M_{Pl}$ by powers of $\log R$, depending on the function f which appears in the potential (1), as occurs in the theory described in the next section.

- Finally, the cosmological and astrophysical limits on the fundamental scale are unimportant in the case that the bulk contains more than one 2D subspace, but as discussed in section 4, the radions still have masses $\sim \text{mm}^{-1}$.

6.3 Explicit model

In this section we present a specific effective theory that stabilizes two large extra dimensions, without relying on input parameters with particularly large (> 10) ratios. The framework for our model is as follows. Supersymmetry in the bulk guarantees a vanishing bulk cosmological constant. Embedded in the 6D spacetime is a set of parallel three-branes that can be regarded as non-supersymmetric defects. Following closely the example of [4], the tensions of these three-branes themselves compactify the extra dimensions. We take the bulk bosonic degrees of freedom to be those of the supergravity multiplet, namely, the graviton g_{AB} and the anti-self-dual 2-form A_{AB}. The 2-form A_{AB} does not couple to any of the three-branes and can be set to zero in our case. We can also have a set of massless bulk scalars ϕ_i contained in hypermultiplets. The relevant part of the Bosonic action is then

$$S = S_{Bulk} + S_{Brane} \tag{6.3.7}$$

where

$$S_{Bulk} = \int d^4x d^2y \sqrt{-G} \left(-2M^4 R + \sum_i (\partial \phi_i)^2 + \cdots \right) \tag{6.3.8}$$

is the bulk action and

$$S_{Branes} = \int d^4x \sum_i \sqrt{-g_i} \left(-f_i^4 + \sum_a \mathcal{L}_a(\psi_a, \phi|_a) + \cdots \right) \tag{6.3.9}$$

is the action for the branes [14]. Here the f_i^4 are the brane tensions and $\mathcal{L}_i$ are Lagrangians for fields ψ_i that may live on the branes, which can also depend on the value of bulk fields ϕ evaluated on the brane $\phi|_a$. G is the 6d metric, g_i is the induced metric on the i'th brane, and we have set the bulk cosmological constant to zero.

Note that while S_{Bulk} must be accompanied by all the extra fermionic terms to have SUSY in the bulk, the brane actions do not have to linearly realize SUSY at all, although they may realize SUSY non-linearly. In particular, there need not be any trace of superpartners on the brane where the Standard Model fields reside. The only reason we need SUSY in the bulk is to protect against the generation of a bulk cosmological constant Λ_{bulk}, which would make a contribution $\sim \Lambda_{bulk} A$ to the potential for the area modulus and spoil our picture with logarithmic potentials.

Our model has three 3-branes, two of which couple to scalars ϕ and ϕ'. The dynamics on the brane impose boundary conditions on the bulk scalar fields. In particular, imagine that the the brane defects create brane-localized potentials for ϕ, which want ϕ to take on the value v_1^2 on one brane and v_2^2 on the other. This will lead to a repulsive contribution to the potential for the area. The same two branes will be taken to have equal and opposite magnetic charges for the scalar ϕ', setting up a vortex-antivortex configuration for ϕ' which will lead to an attractive potential. The balance between these contributions provides a specific realization of how competing dependences on $\log R$ can lead to an exponentially large radius without very large or small input parameters.

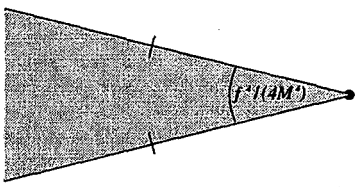

Figure 6.1: The two transverse dimensions in the presence of a three-brane with tension f^4. The shaded region is excluded, and the two borders of the excluded region are to be identified.

We begin by reviewing how the brane tensions can compactify the two extra dimensions [14, 15]. Suppose we ignore for the time being the branes' couplings to bulk scalars, in which case the relevant terms in the action in the low-energy limit are

$$S = -\int d^4x \sum_i \sqrt{-g_i} f_i{}^4 - 2M^4 \int d^4x d^2y \sqrt{-G} R. \tag{6.3.10}$$

For the case in which only a single brane is present, the static solution to Einstein's equations is

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu + \mathcal{G}_{mn}(y) dy^m dy^n, \tag{6.3.11}$$

where $\mathcal{G}_{mn}$ is the 2D Euclidean metric everywhere but at the position of the three-brane, where it has a conical singularity with deficit angle

$$\delta = \frac{f^4}{4M^4}. \tag{6.3.12}$$

As expected, this is in exact correspondence with the metric around point masses in 2+1 dimensional gravity [15]. As shown in Figure 6.1, the spatial dimensions transverse to the brane are represented by the Cartesian plane with a wedge of angle δ removed. Adding a second brane removes a further portion of the Cartesian plane. In fact, if $\frac{f_1^4}{4M^4} + \frac{f_2^4}{4M^4} > 2\pi$, then the excluded region surrounds the allowed portion, as in Figure 6.2. In this case Einstein's equations have a static solution that features a compact space with spherical

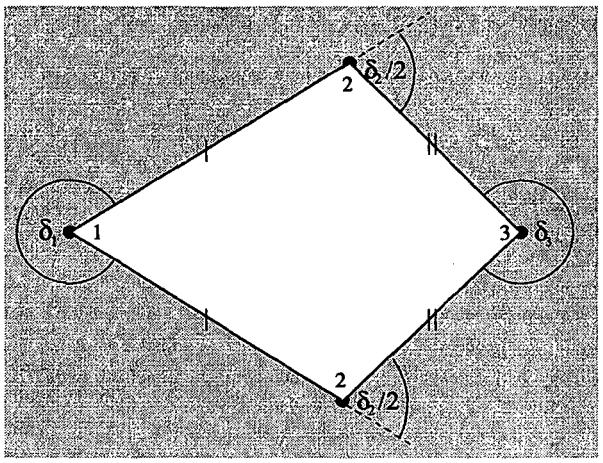

Figure 6.2: A compact space can be obtained given three branes whose corresponding deficit angles δ_i add up to 4π. Identifications to be made are indicated by hash marks. Note that in contrast to the brane in Figure 6.1, branes 1 and 3 in this figure have tensions larger than $4\pi M^4$.

topology, *provided* that a third brane of tension $f_3^4 = 16\pi M^4 - f_1^4 - f_2^4$ is placed at the intersecting lines of exclusion. In general, a set of three-branes has a static solution with spherical topology if

$$\sum_i \frac{f_i^4}{4M^4} = 4\pi, \tag{6.3.13}$$

that is, the deficit angles must add up to 4π.

If a set of branes compactifies the space in this manner, then the 4D effective theory is given by including in the action of (6.3.10) the massless excitations about the classical metric. Thus we replace $\eta_{\mu\nu} \to \overline{g}_{\mu\nu}(x)$ and allow $\mathcal{G}_{mn}(y)$ to fluctuate about δ_{mn} in the bulk. The induced metric on a given brane will differ from $\overline{g}_{\mu\nu}(x)$ by terms involving the fields associated with the brane separations, which we temporarily ignore. The curvature breaks up into two pieces $R^{(4)}$ and $R^{(2)}$, the Ricci scalars built out of $\overline{g}_{\mu\nu}(x)$ and $\mathcal{G}_{mn}(y)$,

respectively. Then, using the Gauss-Bonnet Theorem for spherical topology,

$$\int d^2y\sqrt{\mathcal{G}}R^{(2)} = -8\pi, \tag{6.3.14}$$

along with the fact that $R^{(4)}$ has no y dependence, we can integrate over the extra dimensions to obtain

$$S = -\int d^4x\sqrt{-\overline{g}}\left(\sum_i f_i^4 - 16\pi M^4 + 2(\int d^2y\sqrt{\mathcal{G}})M^4R^{(4)}\right). \tag{6.3.15}$$

In this action it is explicit that adjusting the deficit angles to add up to 4π is equivalent to tuning the 4D cosmological constant to zero.

To develop our specific model we consider the case of three three-branes on a space of spherical topology. Then the "shape" of the extra dimensions is fixed by the branes' deficit angles, or equivalently, by their tensions. However, the size of the extra dimensions,

$$A = \int d^2y\sqrt{\mathcal{G}}, \tag{6.3.16}$$

is completely undetermined. Moreover, the scalar associated with fluctuations of A, the radion, is massless and mediates phenomenologically unacceptable long-range forces. To stabilize the volume of the extra dimensions and give the radion a mass, we couple bulk scalar fields to two of the branes, which, for simplicity, we assume have equal tensions f. The scalar profiles will generate a potential $V_\phi(A)$ that is minimized for a certain value $\overline{A}$ of the volume of the compactified space. Adding the scalar action to (6.3.15) yields a total potential

$$V(A) = V_\phi(A) + \sum_i f_i^4 - 16\pi M^4. \tag{6.3.17}$$

The effective cosmological constant,

$$\Lambda_{eff} = V(\overline{A}), \tag{6.3.18}$$

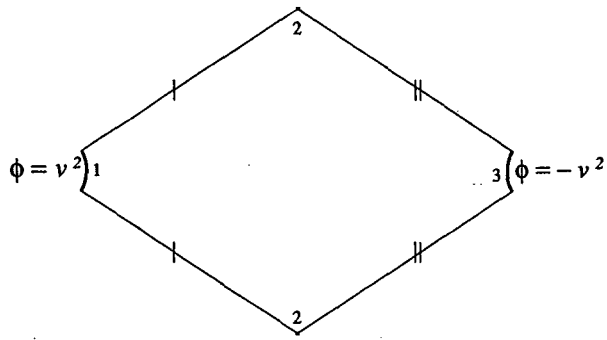

Figure 6.3: The boundary conditions on ϕ. Identifications to be made are indicated by hash marks.

can then be made to vanish by a single fine tuning of fundamental parameters. The back-reaction on the spatial geometry that is induced by the scalars is discussed below.

We work with two massless bulk scalars, ϕ and ϕ', which induce repulsive and attractive forces, respectively. In treating the scalar fields, we will for simplicity ignore their back-reaction on the metric and assume that they propagate in the flat background with conical singularities set up by the branes. It is easy to see that the effect of back-reaction can be made parametrically small if the scalar energy scales are somewhat smaller than M_*, and none of our conclusions are affected.

Suppose that on branes 1 and 3 of Figure 6.3, ϕ is forced to take on unequal values v_1^2 and v_3^2, respectively. This can for instance be enforced if the non-SUSY brane defects generate a potential for ϕ on the branes, analogous to what was considered in [3]. Because ϕ is massless in the bulk, we are free to perform a constant field redefinition and take $v_1^2 = -v_3^2 \equiv v^2$. We account for the brane thicknesses by enforcing these values for ϕ to hold along arcs of finite radius $r_* \sim 1/M_*$, and not just at individual points. The field configuration in the bulk is then given by solving Laplace's equation with these boundary conditions.

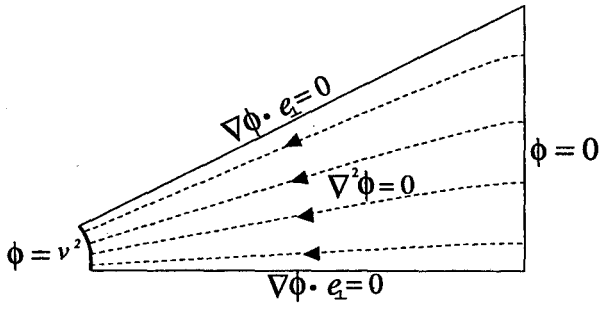

Figure 6.4: A boundary-value problem that determines ϕ. Here $\mathbf{e}_\perp$ refers to the unit vector normal to the relevant boundary, and lines of $\nabla\phi$ are shown dashed. The solution for the full space of Figure 6.3 is given by first evenly reflecting across the bottom horizontal line, and then performing an odd reflection (i.e., $\phi \to -\phi$) across the vertical line where $\phi = 0$.

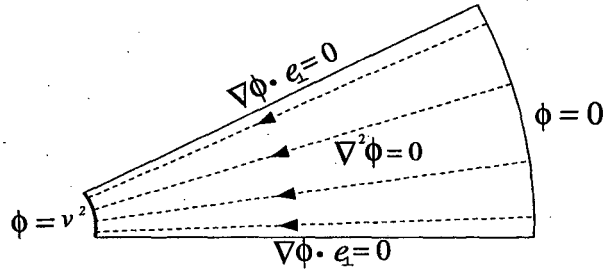

Figure 6.5: The simplified boundary-value problem for ϕ.

Keeping in mind the identifications to be made between the various edges of the space in Figure 6.3, the symmetry of the problem tells us that the field configuration is found by solving the problem depicted in Figure 6.4, and then reflecting that solution appropriately. For simplicity we consider instead a slightly different problem which, unlike that shown in Figure 6.4, is trivially solved. As indicated in Figure 6.5, we take the boundary at which $\phi = 0$ holds to be an arc of radius R, rather than a straight line, so that the solution in this region is immediately found to be

$$\phi = v^2 \frac{\log(R/r)}{\log(R/r_*)}, \tag{6.3.19}$$

where r measures the distance from the (missing) left vertex of the pie slice. The total

energy of this configuration is

$$4\int d\theta \int_{r_*}^{R} dr r \frac{(\nabla\phi)^2}{2} = \theta_0 \frac{v^4}{\log(R/r_*)}, \tag{6.3.20}$$

where $\theta_0 = 2\pi - \frac{f^4}{4M^4}$. Thus, ϕ sets up a $1/\log R$ repulsive potential. It is not difficult to prove using simple variational arguments that the same conclusion is reached when one solves the "real" problem involving the triangle rather than the pie slice.

Now suppose that the same two branes that couple to ϕ carry topological charge under a derivatively coupled field ϕ'. That is, under any closed loop containing a brane we have

$$\int \mathbf{dl} \cdot \nabla\phi' = n\theta_0 w^2, \tag{6.3.21}$$

where w is a fixed parameter with unit mass dimension and n is an integer. Non-zero charge $n \neq 0$ is only possible if we make the identification

$$\phi \sim \phi + \theta_0 w^2. \tag{6.3.22}$$

In order to be able to solve Laplace's equation on a compact space, the branes must carry equal and opposite charges, which we take to correspond to $n = \pm 1$. The configuration for ϕ' is then found by solving Laplace's equation with $\nabla\phi' = \pm\frac{w^2}{r_*}\mathbf{e}_{||}$ on the branes (the gradient runs clockwise on one brane and counterclockwise on the other). This sets up the the vortex-antivortex field configuration for ϕ' shown in Figure 6. For simplicity, in order to calculate the energy in this configuration we once again work on a pie slice (Figure 6.7) rather than a triangle, and it is easily proved that this modification does not affect the essential scaling of the energy with the area. With this simplification the solution is

$$\phi' = w^2\theta + C, \tag{6.3.23}$$

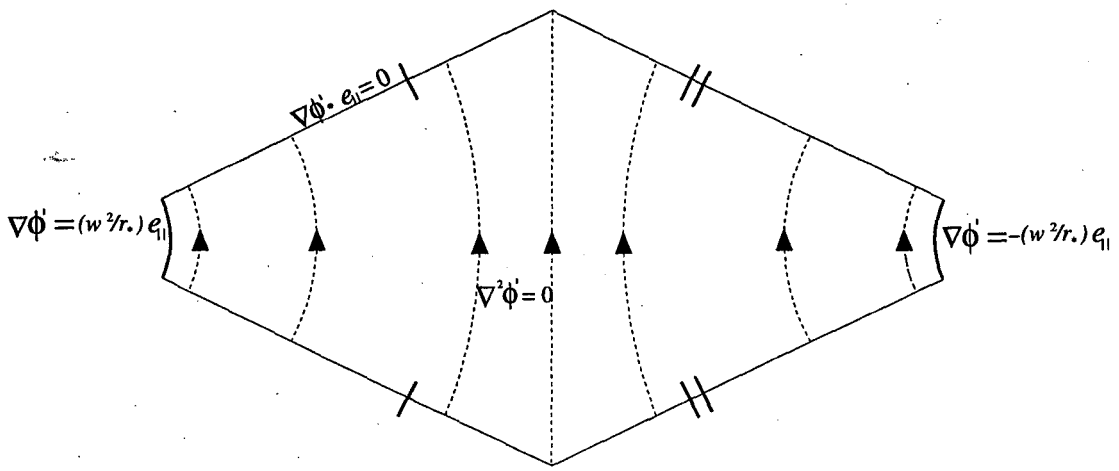

Figure 6.6: The configuration of ϕ'. Each brane carries a topological charge, which generates an attractive potential.

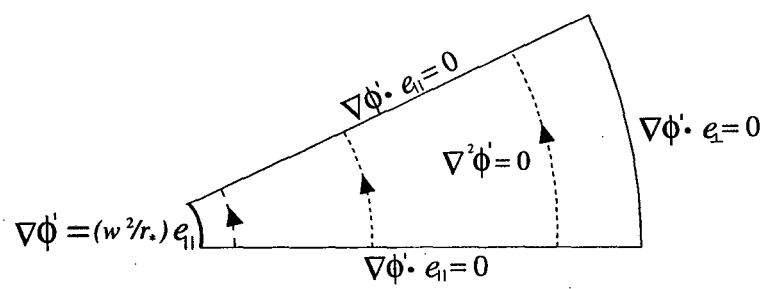

Figure 6.7: The simplified boundary-value problem for ϕ' . Here $\mathbf{e}_{\perp}$ and $\mathbf{e}_{\parallel}$ are the unit vectors normal and parallel, respectively, to the relevant boundary.

where θ is the angular coordinate and C is an undetermined, irrelevant constant. The energy of the configuration is then found to be

$$4\int d\theta \int_{r_*}^{R} dr r \frac{(\nabla\phi')^2}{2} = \theta_0 w^4 \log(R/r_*). \tag{6.3.24}$$

so we have found an attractive potential that will balance the repulsive contribution of (6.3.20). ¿From (6.3.20) and (6.3.24), we see that the full potential is

$$V(R) = \theta_0 \frac{v^4}{\log(R/r_*)} + \theta_0 w^4 \log(R/r_*) + \sum_i f_i^4 - 16\pi M^4, \tag{6.3.25}$$

which is minimized when

$$R = \bar{R} = r_* e^{v^2/w^2}. \tag{6.3.26}$$

Even a mild ratio $v/w \sim 6$ yields an exponentially large radius $\bar{R} \sim 10^{16} r_*$. The effective cosmological constant,

$$\Lambda_{eff} = V(\bar{R}) = \sum_i f_i^4 - 16\pi M^4 + 2\theta_0 v^2 w^2, \tag{6.3.27}$$

can be made to vanish by a single tuning of v, w, and the brane tensions.

Note that we can now see explicitly that the presence of the non-supersymmetric brane defects can not generate a bulk cosmological constant. The presence of the branes leads to logarithmic variation for the bulk fields, which does indeed break SUSY and generate a potential for the area modulus. However, since any *constant* field configuration preserves SUSY, the SUSY breaking in the bulk must be proportional to the *gradient* of the bulk scalar fields, which drops as $1/r$ with distance r away from the branes. Therefore, it is impossible to induce a cosmological constant, since this would amount to an constant amount of SUSY breaking throughout the bulk. In fact, a very simple power-counting argument shows that all corrections to the energy are logarithmic functions of the area.

Given a specific form for the logarithmic potential (6.3.25), we can work out the mass of the area modulus, which is

$$m_R^2 \sim \frac{\bar{R}^2 V''(\bar{R})}{M_{Pl}^2} \sim \frac{v^4}{M_{Pl}^2 \log^3(\bar{R}/r_*)}. \tag{6.3.28}$$

Interestingly, m_R is suppressed by $(\log(\bar{R}/r_*))^{3/2}$ compared to the naive estimate M_*^2/M_{Pl}. Hence even for $v \sim M_*$ as large as 100 TeV, the range of the radion-mediated Yukawa potential is 0.1 mm – accessible to planned experiments.

6.4 Four and Six Extra Dimensions

Since the logarithmic form of the propagator occurs only in two dimensions, one may worry that the ideas in this paper are only applicable to the case of two large dimensions. This is the case most severely constrained by astrophysical and cosmological constraints [1, 12, 13], which demand the 6D Planck scale $M_* > 50$ TeV, seemingly too large to truly solve the hierarchy problem. One possibility is that the true Planck scale of the ten dimensional theory could be $\sim$ O(TeV), and the 6D Planck scale of $\sim$ 50 TeV could arise if the remaining four dimensions are a reasonable factor $O(10)$ bigger than a $(\text{TeV})^{-1}$. But we don't have to resort to this option. As pointed out in [7, 10], the presence of two-dimensional subspaces where massless fields can live is sufficient to generate logarithms. Take the case of four extra dimensions. Imagine one set of parallel 5-branes filling out the 12345 directions, and another set filling out the 12367 directions. They will intersect on 3-dimensional spaces where 3-branes can live. These 3-branes can act as sources for fields living on each of the 5-branes, which effectively propagate in two sets of orthogonal 2D subspaces. Once again, bulk SUSY can guarantee a vanishing "cosmological constant" for

each of the 2D subspaces. The SUSY breaking at the intersections can set up logarithmically varying field configurations on the 5-branes that leads to a potential of the form $V(\log A_1, \log A_2)$ for the areas A_1, A_2 of the 2D subspaces. Minimizing the potential, each radius can be exponentially large, and the ratio of the radii will also be exponential, but the value of M_{Pl} will require the largest radius to be very much smaller than a mm. It would be interesting to build an explicit model along these lines.

Even without an explicit model, however, we can see that the scale of the radion masses is unchanged. The logarithmic potential still gives $m_{Ri} \approx \Lambda^2/M_{Pl} \approx \text{mm}^{-1}$, for $\Lambda \approx 1$ TeV. After Weyl rescaling, each radion couples with gravitational strength to the Standard Model and should show up in the sub-millimeter measurements of gravity.

6.5 Other ideas

There is an alternative way in which theories with two transverse dimensions can generate effectively exponentially large radii. The logarithmic variation of bulk fields can force the theory into a strong-coupling region exponentially far away from some branes, and interesting physics can happen there. This is the bulk analog of the dimensional transmutation of non-Abelian gauge theories, which generate scales exponentially far beneath the fundamental scale and trigger interesting physics, such as e.g. dynamical supersymmetry breaking [10]. It is tempting to speculate that such strong-coupling behavior might effectively compactify the transverse two dimensions. Recently, Cohen and Kaplan have found an explicit example realizing this idea [16]. They consider a massless scalar field with non-trivial winding in two transverse dimensions: a global cosmic string. Since the total energy of the string diverges logarithmically with distance away from the core of the vortex,

we expect gravity to become strongly coupled at exponentially large distances. Indeed, Cohen and Kaplan find that the metric develops a singularity at a finite proper distance from the vortex core, but argue that the singularity is mild enough to be rendered harmless. What they are left with is a non-compact transverse space, with gravity trapped to an exponentially large area

$$\bar{A}M_*^2 = e^{M_*^4/f_\pi^4} \tag{6.5.29}$$

where f_π is the decay constant of the string. A ratio of $M_*/f_\pi \sim 2.5$ is all that is needed to solve the hierarchy problem in this case. This model is a natural implementation of the ideas of [1], to solve the hierarchy problem with large dimensions, together with the idea of trapping gravity in non-compact extra dimensions as in [17]. Unlike [2], however, the bulk geometry is not highly curved everywhere, but only near the singularity. Thus, gravity has essentially been trapped to a flat "box" of area A in the transverse dimensions, and the phenomenology of this scenario is essentially the same as that of [1]. An attractive aspect of this scenario is that, unlike both our proposal in this paper and those of [2, 3], no modulus needs to be stabilized in order to solve the hierarchy problem. This also points to a phenomenological difference between our proposal and that of [16]. While both schemes generate an exponentially large area for two transverse dimensions, there is no light radion mode in [16] whereas we have a light radion with $\sim$mm^{-1} mass.

6.6 Conclusions

In this paper, we have shown how to stabilize exponentially large compact dimensions, providing a true solution to the hierarchy problem along the lines of [1] which is on the same footing as technicolor and dynamical SUSY breaking. Of course, there are many

mysteries other than the hierarchy problem, and the conventional picture of beyond the Standard Model physics given by SUSY and the great desert had a number of successes. So why do we bother pursuing alternatives? Are we to think that the old successes are just an accident?

A remarkable feature of theories with large extra dimensions is that the phenomena that used to be understood inside the energy desert can also be interpreted as arising from the space in the extra dimensions. Certainly all the qualitative successes of the old desert, such as explaining neutrino masses and proton stability, can be exactly reproduced with the help of the bulk [1, 19, 20, 18], in such a way that e.g. the success of the see-saw mechanism in explaining the scale of neutrino masses is not an accident. As we have mentioned, there is even hope that the one quantitative triumph of the supersymmetric desert, logarithmic gauge coupling unification, can be exactly reproduced so that the old success is again not accidental. We find it encouraging that it is precisely the same sorts of models-with two dimensional subspaces, SUSY in the bulk broken only on branes- which allows us to generate exponentially large dimensions. Hopefully, in the next decade experiment will tell us whether any of these ideas are relevant to describing the real world.

Bibliography

[1] N. Arkani-Hamed, S. Dimopoulos and G. Dvali, *Phys. Lett.* **B429** 263 (1998), hep-ph/9803315; I. Antoniadis, N. Arkani-Hamed, S. Dimopoulos and G. Dvali, *Phys. Lett.* **B436** 257 (1998), hep-ph/9804398; N. Arkani-Hamed, S. Dimopoulos and G. Dvali, *Phys. Rev.* **D59** 086004 (1999), hep-ph/9807344.

[2] L. Randall and R. Sundrum, *Phys. Rev. Lett.* **83**, 3370 (1999).

[3] W.D. Goldberger and M.B. Wise, *Phys.Rev.Lett.* **83**, 4922 (1999).

[4] R. Sundrum, *Phys. Rev.* **D59**, 085010 (1999), hep-ph /9807348.

[5] N. Arkani-Hamed, S. Dimopoulos and J. March-Russell, hep-th/9809124.

[6] C. Bachas, hep-ph/9807415.

[7] I. Antoniadis and C. Bachas, hep-th/9812093.

[8] I. Antoniadis, C. Bachas, E. Dudas, hep-th/9906039.

[9] L. Ibanez, hep-ph/9905349.

[10] N. Arkani-Hamed, S. Dimopoulos and J. March-Russell, hep-th/9908146.

[11] G. Dvali, hep-ph/9905204.

[12] S. Cullen and M. Perelstein, *Phys. Rev. Lett.* **83** (1999) 268, hep-ph/9903422.

[13] L. Hall and D. Smith, *Phys. Rev.* **D60** (1999) 085008, 1999, hep-ph/9904267.

[14] R. Sundrum, hep-ph/9805471.

[15] S. Deser, R. Jackiw, and G. 't Hooft, *Ann. Phys.* **152**, 220 (1984).

[16] A. G. Cohen and D.B. Kaplan, hep-th/9910132.

[17] L. Randall and R. Sundrum, hep-th/9906064.

[18] N. Arkani-Hamed and S. Dimopoulos, hep-ph/9811353.

[19] N. Arkani-Hamed, S. Dimopoulos, G. Dvali and J. March-Russell, hep-ph/9811448.

[20] N. Arkani-Hamed and M. Schmaltz, hep-ph/9903417.

Chapter 7

TeV Theories of Flavor and Large Extra Dimensions

7.1 Introduction

The extreme weakness of gravity is usually attributed to the fundamental mass scale of gravity being very much larger than that of the strong and electroweak interactions. The standard model provides no understanding of how this enormous difference in scales is stabilized against radiative corrections. Despite this gauge hierarchy problem, the need for extraordinarily large physical mass scales has been accepted as a central feature in theories of physics beyond the standard model. The unification of the gauge coupling constants at 10^{16} GeV strengthens this view. Furthermore, the absence of flavor and CP violating phenomena, beyond that explained by the weak interactions, has made it all but impossible to construct theories of flavor at accessible energies, and suggests that the fundamental mass scale for flavor physics is very far above the TeV scale.

Over the last two decades, the usual approach to addressing the gauge hierarchy problem has been to modify particle physics between the weak and Planck scales. However, there is another possibility: gravity can be modified at and beneath the TeV scale, as was

realized in [1, 2, 3, 4]. In this scenario, the fundamental mass scale of gravity can be brought far beneath the conventional Planck scale, perhaps as low as a TeV, in the presence of sub-millimeter sized new spatial dimensions serving to dilute the strength of gravity at long distances. These dimensions have not been detected since the Standard Model fields are localized on a three-dimensional wall, or "3-brane", in the higher dimensional space. Such a scenario can naturally be accommodated in string theory [2], where the wall on which the SM fields live can be a D-brane.

Remarkably, despite the profound modifications of physics both at sub-millimeter and TeV scales, this scenario is not excluded by any known lab, astrophysical or cosmological constraints [3]. This realization opens up the possibility that there may be a number of experimentally viable approaches to addressing the hierarchy problem which involve the basic ingredients of modifying gravity at or beneath the TeV scale, and localizing matter fields to branes in extra dimensions. An interesting modification of gravity has been proposed recently [5] where the gravitational metric describing the 4 usual coordinates of spacetime depends on the location in the extra dimensions. Such metrics result from solving Einstein's equations in the presence of brane configurations, and lead to spatial localization of the graviton zero mode in the extra dimensions. Various schemes for solving the hierarchy problem have been based on this [5, 6, 7, 8]. All these schemes, and the original scheme with large extra dimensions, share a common feature: we live on a 3-brane located in the extra dimensions in which gravity propagates. From the viewpoint of our 3-brane, the fundamental mass scale is the TeV scale, and this is the scale at which quantum gravity gets strong.

Lowering the fundamental cutoff close to the TeV scale obliterates the usual ultraviolet desert in energy scales. On the theoretical side, this seems to destroy the attractive picture of gauge coupling unification. More pressingly, there are in principle dangerous effects from higher dimension operators now only suppressed by the TeV scale giving e.g. disastrously large rates for proton decay.

However, it has been realized that the space in the extra dimensions replaces the old desert as the new arena in which such issues can be addressed. For instance, the old picture of logarithmic gauge coupling unification close to the Planck scale may be mimicked by the logarithmic variation of classical fields in sets of two large dimensions [9]. * Furthermore, the difficulties associated with higher-dimensional operators can also find a natural resolution using higher-dimensional locality. Indeed intrinsically higher-dimensional ways of suppressing proton decay were proposed in [3, 1, 11]. † After proton decay, the most serious issue is that of flavor-changing neutral currents. Dimensional analysis suggests that the flavor scale should be above 10^4 TeV from the K_L—K_S mass difference, and greater than 10^5 TeV from CP mixing in the neutral kaons. While this naive estimate can be avoided, it has proved extraordinarily hard to construct theories at the TeV scale which provide an explanation for the small flavor parameters. It is natural to ask whether extra dimensions offer any new possibilities for evading these problems. In [13], a higher-dimensional mechanism was proposed for generating the fermion mass hierarchy, and preliminary arguments were given to suggest that the FCNC problem could also be avoided. It is our purpose in

*Another approach to gauge coupling unification bases on power-law running of higher-dimensional gauge couplings has been discussed in [10].

†In a different context, higher-dimensional locality has been used to ameliorate the SUSY flavor problem in anomaly-mediated models [12].

this paper to extend and generalize these ideas to realistic and elegant theories of flavor at the TeV scale which are safe from FCNC effects. As we will see, *it is the physics of extra dimensions that allows us to naturally bring flavor physics down to the TeV scale.*

In this paper we therefore study effective theories of flavor with a low fundamental mass scale Λ. What do we mean by "effective theories of flavor"? At scale Λ there is a fundamental theory, presumably string theory, which has some low energy effective theory. It is conceivable that this is just the standard model; with entries in the Yukawa matrices somehow set to the required hierarchical values, and with all higher dimensional operators absent. We consider this unlikely, but, since we do not know the low energy limit of string theory, we must make some assumptions about the form of the low energy effective theory.

We assume that the effective theory beneath Λ is based on some symmetry group G and has an effective Lagrangian

$$\mathcal{L}_{eff} = \sum_i \frac{c_i}{\Lambda^{p-4}} \mathcal{O}_i^p \supset \left(\frac{\varphi}{\Lambda}\right)^n f f^c H + ... \qquad (7.1.1)$$

where i runs over all G invariant operators, $\mathcal{O}_i^p$, p labels the dimension of the operator and c_i are unknown dimensionless couplings of order unity. An example of an operator which leads to a small Yukawa coupling for the fermion f to the Higgs H is shown. This assumption implies that the small dimensionless parameters of flavor physics must arise spontaneously from $\langle\varphi\rangle/\Lambda$, where φ is a field of the low energy theory. We call φ a flavon field: the effective theory must explain why it has a vev small compared to the fundamental scale. In an effective theory of flavor, given the symmetry group and the field content, and in extra dimensions the brane configuration, flavor can be understood in the low energy theory itself.

In section 2 we discuss several difficulties encountered in building theories of flavor at the TeV scale in 4 dimensions. Higher dimension operators lead to flavor-changing and CP violating effects which are hard to tame, even with a flavor symmetry. An Abelian symmetry cannot prevent enormous $K\bar{K}$ mixing, and a non-Abelian symmetry results in disastrous flavor-changing (pseudo-) Goldstone bosons. Even the maximal flavor symmetry, $U(3)^5$, is not quite sufficient to protect against large electric dipole moments for the electron and neutron. Finally, the flavon quanta are themselves very light, and the exchange of these particles in the low energy theory also generates disastrous four fermion operators. These difficulties are illustrated with a $U(2)$ flavor symmetry group.

In section 2.6 we discuss the minimal $U(3)^5$ flavor structure in 4 dimensions, in which the three Yukawa matrices are each promoted to a single flavon field. This structure has been used to argue that flavor physics can occur at a low scale [14]. We show that CP must also be spontaneously broken for the fundamental scale to be under 10 TeV. However, as it stands, this minimal $U(3)^5$ structure is not an effective theory of flavor. The flavon fields contain a hierarchy of vevs which are simply imposed by hand and not derived from the low energy effective theory. This could be remedied by introducing a hierarchy of symmetry breakings at a sequence of scales beneath Λ, as proposed by Froggatt and Nielson [15]. However, this would produce flavons at each of these scales, and some would be very light indeed, and their exchange would induce disastrous flavor and CP violating interactions. This flavon exchange problem appears generic to effective theories of flavor with low Λ in 4d.

With extra dimensions, there is a new possible origin for the small flavor param-

eters: symmetries which are broken strongly on some source brane may be only weakly broken on our brane because the source brane is distant from us [13]. This idea is explored in section 3. It is most easily implemented by making the flavon field, φ, a bulk field, which is coupled to a source on the distant source brane so that it has a small vev on our brane. The obstacles to constructing effective theories of flavor encountered in 4d are immediately removed: there is now an origin for the small parameters in the flavon vev – we have a real theory for the small parameters – and yet this is done without introducing light flavons, solving the flavon exchange problem. Furthermore, with order unity breaking of the discrete flavor group on the distant branes, the pseudo-Goldstone masses can be raised to the TeV scale. [‡]

The origin of this success is to understand flavor from a hierarchy of distances in the extra dimensions, and not from a hierarchy of mass scales in our 4d world. In section 3 we also discuss another phenomenon which is generic in theories of flavor in extra dimensions when the bulk flavon field possesses non-linear interactions. This means that the flavor breaking felt on our brane is sensitive to the value of the flavon field in the bulk, not just to its value on our brane. The "sniffing" of flavor breaking in the bulk can lead to interesting phenomena. For example, in section 6, we study a $U(2)^5$ theory, which incorporates features of the $U(2)$ 4d theory, and in which sniffing plays a crucial role in symmetry breaking.

In section 4 we construct a complete realistic theory of flavor in extra dimensions with Λ in the region of 5 – 10 TeV. The flavor group is maximal, $U(3)^5 \times CP$, and the minimal set of flavons propagate in the bulk, taking classical values which result from shining

[‡]An alternative way to make the pseudo-Goldstones massive is to gauge the flavor symmetry in the bulk [16]. Although there are horizontal gauge bosons, they are less dangerous than usual since they propagate in the bulk.

from just three source branes. Each source brane breaks a discrete subgroup of each $U(3)$ using only triplet vevs, and the three source branes may be identical to each other.

In section 5 we consider a particularly simple brane configuration for realizing this $U(3)^5$ theory: our 3-brane and the three source 3-branes are located on a 4-brane, so that shining occurs in 1 dimension. This makes the calculation of the Yukawa matrix, and the additional flavor changing effects from the bulk, remarkably simple.

In section 6 we study the possibility of a smaller non-Abelian flavor symmetry, and introduce a variety of bulk flavons in a way motivated by the observed quark spectrum. This theory illustrates some of the possibilities opened up for flavor physics in extra dimensions. For example, a new mechanism for suppressing the neutron electric dipole moment is proposed.

Predictive theories of fermion masses can result if the source branes have a symmetrical geometrical configuration, as would be expected in a dynamical theory of brane stabilization. In section 7 we study theories in which the 9 quark masses and mixing angles are given quite successfully in terms of just 5 free parameters. These theories are inherently extra-dimensional, with the precise predictions reflecting the geometry of the brane configuration, and the location of our three-brane.

7.2 Challenges to a low flavor scale

In this section we summarize the major challenges to lowering the scale of flavor physics close to the TeV scale in theories with four spacetime dimensions. The difficulties are mostly well-known (see e.g. [17]), but it is useful to have them collected in one place.

Process	O	Bound on Λ (TeV)
ϵ_K	$(ds^c)(\bar{s}\bar{d}^c)$	10^5
	$(\bar{s}\bar{\sigma}^\mu d)^2$	10^4
Δm_K	$(ds^c)(\bar{s}\bar{d}^c)$	10^4
	$(\bar{s}\bar{\sigma}^\mu d)^2$	10^3
Δm_D	Analogous to above	10^3
	"	5×10^2
Δm_B	Analogous to above	5×10^2
	"	5×10^2

Table 7.1: Bounds on Λ from $\Delta F = 2$ processes.

7.2.1 Dimensional analysis

The most serious obstacle to lowering the flavor scale Λ comes from flavor-changing neutral currents, most severely from the kaon system. With the coefficients c_i in eqn.(7.1.1) taken to be of unit magnitude with large phases, the bounds on Λ coming from the operators contributing to $\Delta F = 2$ processes (ϵ_K and $\Delta m_K, \Delta m_D, \Delta m_B$) are presented in Table 7.1. The bounds on Λ from the left-right operators for $\Delta m_K, \Delta m_D$ are enhanced by a factor of ~ 3 due to the QCD enhancement in running from Λ down to the hadronic scale. There are also bounds on Λ, far above the TeV scale, coming from $\Delta F = 1$ processes such as $\mu \to e\gamma$ and $K_L \to \mu e$.

7.2.2 Implications for model-building

While this may suggest that the scale of flavor should be above $\sim 10^4 - 10^5$ TeV, it is also possible that whatever is responsible for suppressing the Yukawa couplings of the light generations also adequately suppresses the flavor-changing operators. For instance, if a weakly broken flavor symmetry G_F is responsible for the fermion mass hierarchy, the same G_F could suppress the dangerous operators.

It is easy to see that even this idea fails for generic flavor symmetries. The reason is that the most dangerous effects arise not directly from operators that *violate* G_F, but rather from G_F invariant operators which violate flavor when rotated into the mass basis. Suppose for instance that G_F is Abelian with different charges for the first and second generations. Then, the flavor symmetry allows the higher dimension operators

$$\frac{a}{\Lambda^2}(Q_1 D_1^c)(\bar{Q}_1 \bar{D}^c{}_1) + \frac{b}{\Lambda^2}(Q_2 D_2^c)(\bar{Q}_2 \bar{D}^c{}_2) \tag{7.2.2}$$

with $a, b \sim O(1)$ whereas operators of the form $(Q_2 D_1^c)(\bar{Q}_1 \bar{D}^c{}_2)$ will have suppressed coefficients. Nevertheless, when we rotate the fields to go to the mass eigenstate basis, we generate an operator

$$\sim \frac{(a-b)\theta_c^2}{\Lambda^2}(Q_2 D_1^c)(\bar{Q}_1 \bar{D}^c{}_2) \tag{7.2.3}$$

where we have assumed that the Cabbibo angle dominantly comes from the down sector. Note that unless $a = b$ to high accuracy, this still forces $\Lambda > 10^3 - 10^4$ TeV. Having the Cabbibo angle come dominantly from the up sector helps, but still requires $\Lambda > 10^2$ TeV. The only way out is for $a = b$, however an Abelian flavor symmetry is not enough to enforce equality. As claimed, we see that the central challenge is to ensure that G_F *invariant* operators remain harmless when rotated to the mass eigenstate basis, and this requires G_F to be non-Abelian. If we ignored this issue, in other words if we assume that for some reason these invariant higher-dimensions operators are absent or have their coefficients magically tuned to equality, then even Abelian symmetries are enough to adequately suppress FCNC effects from "directly" flavor-violating operators [18]. For instance, for any Abelian flavor symmetry we expect to have the operator

$$\frac{(\lambda_s \theta_c)^2}{\Lambda^2}(Q_2 D_1^c)(\bar{Q}_1 \bar{D}^c{}_2), \tag{7.2.4}$$

and even assuming maximal phase this requires $\Lambda > 7\,\mathrm{TeV}$ if we take m_s at the lower end of its range, $\sim 90\,\mathrm{MeV}$, as is currently favored from the lattice [19]. Notice, however, that in a two-Higgs doublet theory this bound turns into $\Lambda > 7 \tan\beta\,\mathrm{TeV}$, so we can not tolerate large $\tan\beta$.

In the SM, such higher dimension operators are generated by integrating out $W's$ at the weak scale, but enormous FCNC's are not generated. This is because the SM gauge interactions respect the $U(3)^5$ flavor symmetry acting separately on each of the (Q, U, D, L, E) fields, explicitly broken only by the Yukawa matrices. In the $U(3)^5$ symmetric limit, all the operators are generated automatically with equal coefficients; this maximal flavor symmetry is strong enough to ensure that *flavor symmetric* operators are harmless when rotated to the mass basis. It is then natural to explore the possibility that the true flavor symmetry is the maximal one $G_F = U(3)^5$, and we will consider this possibility both in the context of four dimensions and in extra dimensions. We find that while it is difficult to believe in a real theory based on $U(3)^5$ in 4D, it is easy to construct elegant theories based on $U(3)^5$ in extra dimensions.

However, there is a strong constraint, even on theories based on a $U(3)^5$ flavor symmetry, coming from electric dipole moments of the electron and neutron. Any flavor symmetry would allow an operator of the form e.g.

$$\frac{e^{i\varphi}\lambda_d}{\Lambda^2}(Q_1 H)\sigma^{\mu\nu} D_1^c, \tag{7.2.5}$$

and if the phase φ is $O(1)$, this requires $\Lambda > 40\,\mathrm{TeV}$ from the neutron edm [21]. A similar operator in the lepton sector gives $\Lambda > 100\,\mathrm{TeV}$ from the electron edm. While it may be possible to lower Λ below the $10^4 - 10^5\,\mathrm{TeV}$ barrier by imposing powerful enough flavor

symmetries, we cannot lower it past 40 TeV without making further assumptions about CP violation. We must assume that CP is primordially a good symmetry, and is broken by the same fields breaking G_F. This gives a hope that the phases in the mass and edm operators are the same and therefore in the mass eigenstate basis there is no phase in the edm operator.

7.2.3 Flavor-changing goldstone bosons

We have argued that the flavor group, G_F, cannot be Abelian: controlling flavor changing effects from higher dimension operators points to a large non-Abelian flavor symmetry group. If G_F is continuous the spontaneous breaking produces familons – flavor-changing Goldstone bosons – leading to the very stringent bound $\Lambda > 10^{12}$ GeV. Gauging the flavor symmetry allows the familons to be eaten, but the weakness of the breaking then tells us that there will be horizontal gauge bosons with masses much smaller than the fundamental scale, whose exchange leads to flavor changing problems. The only option is to have G_F be a large, *discrete*, non-Abelian symmetry. However, even this case typically is excluded by the accidental occurence of pseudo-Goldstone bosons. At the renormalizable level, the potential for the flavon fields only contains a few G_F invariant interactions, and this typically gives an accidental continuous symmetry, reintroducing Goldstone bosons. Higher order operators that respect only the discrete symmetry will give masses to these pseudos which are suppressed by ratios of flavon VEVs to the fundamental scale. Moreover, these ratios will be raised to high powers, given the large size of the discrete group, and we are thus left with extremely light pseudo-goldstone bosons that can be produced, for instance, in K decays.

7.2.4 The flavon exchange problem

In the low energy effective theory, beneath the fundamental scale Λ, the Yukawa couplings are generated by operators of the form

$$\mathcal{L} \sim \left(\frac{\varphi}{\Lambda}\right)^n f f^c H \tag{7.2.6}$$

If we set all but one of the φ's and the Higgs to its vev, we have an effective coupling to φ

$$\left(\frac{\langle\varphi\rangle}{\Lambda}\right)^{n-1} \frac{\langle H\rangle}{\Lambda} f f^c \varphi \sim \frac{m_{ij}}{\langle\varphi\rangle} f_i f_j^c \varphi \tag{7.2.7}$$

Tree-level flavon exchange then generically generates flavor-changing 4-fermion operators that are suppressed only by the *flavon* mass and not by the scale Λ. Unless the flavon potentials are fine-tuned, we expect that the flavon masses m_φ are of the same magnitude as the vev $\langle\varphi\rangle$, which must be smaller than Λ in order to produce small Yukawa couplings. Using the same interactions which generate the Yukawa couplings, tree-level flavon exchange can generate dangerous 4-fermion operators. Of course, if the flavon masses dominantly respect the flavor symmetry, the induced operators will be flavor-symmetric, and if the flavor symmetry is powerful enough these operators may be harmless. However, since the flavon vevs themselves break the flavor symmetry at a scale $\sim \langle\varphi\rangle \sim m_\varphi$, we expect generically that the flavon masses will have $O(1)$ flavor breaking. The generated 4-fermi operators are clearly most dangerous if the light generation Yukawa couplings are generated by a single flavon coupling. For instance, suppose that the 12 element of the down mass matrix is produced by the vev of a flavon φ_{12}. then, tree-level flavon exhange can induce $\Delta S = 2$ operators

$$\frac{1}{\Lambda^2 \langle\varphi_{12}\rangle^2} (Q_2 D_1^c H)(\bar{Q}_1 \bar{D}_2^c H^\dagger) \tag{7.2.8}$$

which forces Λ back above $\sim 10^3 - 10^4$ TeV from Δm_K and ϵ_K.

7.2.5 An example: $G_F = U(2)$

There is a very simple theory, with $G_F = U(2)$ [20], which gives highly successful quark mass matrices, and alleviates the FCNC problem in supesymmetric theories. However, this theory has been studied for the case of very large Λ – what happens when Λ is reduced towards the TeV scale?

We study just the two light generations, which transform as $U(2)$ doublets ψ^a, where $a = 1, 2$. The flavons are in a doublet φ_a and an antisymmetric tensor A_{ab}. The structure of the Yukawa matrices follows from

$$\mathcal{L}_{Yuk} \sim h_1 \varphi_a \varphi_b \psi^a \psi^{cb} H + h_2 A_{ab} \psi^a \psi^{cb} H. \tag{7.2.9}$$

Whatever triggers a vev for φ and A, we can always choose a basis so that $\varphi \propto \begin{pmatrix} 0 \\ \sqrt{\epsilon} \end{pmatrix}$ and $A_{ab} \propto \epsilon' \epsilon_{ab}$, yielding the interesting structure

$$\lambda \sim \begin{pmatrix} 0 & \epsilon' \\ -\epsilon' & \epsilon \end{pmatrix}. \tag{7.2.10}$$

While placing the first two generations in $U(2)$ doublets goes a long way in erasing dangerous flavor-changing effects, there are higher dimension $U(2)$ invariant operators, in particular

$$\mathcal{O}_{bad} = \frac{1}{\Lambda^2} Q_a D_b^c \bar{D}^{ca} \bar{Q}^b, \tag{7.2.11}$$

which give disastrously large contributions to Δm_K and ϵ_K, forcing $M_F > 10^5$ TeV. This suggests that more complex models are needed with more than one $U(2)$ factor. Before discussing such possibilities, however, let us proceed by assuming that for some reason this

dangerous higher-dimensional operator is *not* present in the theory [§]. We now show that, even making this assumption, a 4d $U(2)$ theory still requires $\Lambda > 10^3$ TeV.

The difficulty for the 4d theory is that there are physical states lighter than M_F charged under flavor, the flavons themselves. As an illustration, suppose that we generate vevs for φ and A via independent mexican-hat potentials

$$V(\varphi, A) \sim |\varphi^* \varphi - p^2|^2 + |A^* A - a^2|^2. \tag{7.2.12}$$

Of course, most disastrously, we get goldstone modes from the breaking of the global $U(2)$, and $K \to \pi +$ familon would force all the scales above $\sim 10^{12}$ GeV. We should really be considering a large discrete subgroup of $U(2)$, and there will be other terms in the potential that can lift the familon masses. Even if this is done, however, we are still left with light flavons of mass $\sim p, a$. The tree-exchange of A in particular generates

$$|h_2|^2 \frac{H^* H (\epsilon^{ab} Q_a D_b^c)(\epsilon_{ij} \bar{Q}^i \bar{D}^{cj})}{\Lambda^2 a^2}, \tag{7.2.13}$$

which contains the dangerous $(Q_1 D_2^c)\overline{Q}_2 \overline{D}_1^c$ operator. Note that the mass of A is not $U(2)$ violating, and so we have generated a $U(2)$ invariant operator which is nevertheless dangerous. The coefficient of the operator is real so there is no contribution to ϵ_K, but there is a strong constraint from Δm_K: to produce the small 12 entries of the Yukawa matrices we need

$$\frac{a}{\Lambda} \sim \lambda_s \theta_c, \tag{7.2.14}$$

leading to $\Lambda > 10^3$ TeV.

[§] In a Froggatt-Nielsen theory, for instance, as long as the only coupling between SM fields and the heavy Froggatt-Nielsen fields involve flavons, the coefficient of such an operator can be suppressed by many loop factors.

7.2.6 Minimal $U(3)^5$ in 4D

As already mentioned, the largest symmetry group of the standard model Lagrangian in the limit of vanishing Yukawa couplings is $U(3)_Q \times U(3)_{U^c} \times U(3)_{D^c} \times U(3)_L \times U(3)_{E^c}$. Imposing $U(3)^5$ on the underlying theory therefore gives the strongest possible symmetry suppression of flavor violating processes in the effective theory. In the simplest realization of $U(3)^5$ (which we illustrate for the quark sector alone), the symmetry is broken by the VEVs of a single χ_u and a single χ_d, transforming as $(\overline{3}, \overline{3})$ under $U(3)_Q \times U(3)_{U^c}$ and $U(3)_Q \times U(3)_{D^c}$, respectively. The effective Lagrangian has the form of equation (7.1.1), where the flavor and gauge invariant $\mathcal{O}_i^p$ are constructed from χ_u, χ_d, and standard model fields. Fermion masses, for example, come from the operators

$$\frac{1}{\Lambda} Q \chi_u U^c \tilde{H} \qquad \text{and} \qquad \frac{1}{\Lambda} Q \chi_d D^c H. \tag{7.2.15}$$

Having too low a flavor scale Λ leads to conflict with experiment. Strong bounds come from flavor conserving operators such as

$$|H^\dagger D_\mu H|^2 \qquad \text{and} \qquad (H^\dagger D_\mu H) \overline{Q} \gamma^\mu Q, \tag{7.2.16}$$

which give anomalous contributions to the ρ parameter and to fermion couplings to the Z boson, and require $\Lambda > 6$ and 7 TeV, respectively[¶]. Other dimension 6 operators that lead to similar precision electroweak limits are listed in [24]. Atomic parity violation experiments and direct searches at LEPII for 4-lepton couplings place only slightly milder bounds of $\Lambda >$ 3 TeV.

Provided these requirements due to flavor conserving phenomena are met, operators that arise due to flavor breaking are relatively safe. For example, it is impossible

[¶]Here and below, we set the relevant $c_i = 1$ to obtain bounds on Λ.

to construct higher dimensional operators that induce $K - \overline{K}$ mixing using only the VEV of χ_d, because χ_d is diagonal in the down quark mass basis[‖]. One must instead consider operators that involve χ_u, such as

$$\frac{c}{\Lambda^2}(\overline{Q}\overline{\sigma}^\mu \frac{\chi_u^\dagger \chi_u}{\Lambda^2} Q)^2, \tag{7.2.17}$$

which in the mass basis contains

$$\frac{c}{\Lambda^2}(\overline{D}\overline{\sigma}^\mu V_{CKM}^\dagger \overline{\lambda}_U^2 V_{CKM} D)^2, \tag{7.2.18}$$

where $\overline{\lambda}_U^2 = \mathrm{Diag}(\lambda_u^2, \lambda_c^2, \lambda_t^2)$. This gives the $\Delta S = 2$ piece

$$\frac{c}{\Lambda^2}\lambda_t^4 (V_{td}^* V_{ts})^2 (\overline{d}\overline{\sigma}^\mu s)^2, \tag{7.2.19}$$

which leads to bounds from Δm_K and ϵ_K of $\Lambda > .5$ and 5 TeV, provided the phase of $c(V_{td}^* V_{ts})^2$ is of order one. $\Delta S = 1$ processes give weaker bounds.

As mentioned in section 2.2, the most stringent bound on Λ arises because, a priori, there is no reason to expect any relations between the phases of the c_i that appear in equation (7.1.1). The Yukawa interaction

$$Q\frac{\chi_d}{\Lambda} D^c H \tag{7.2.20}$$

and the electric dipole moment operator

$$\frac{1}{\Lambda^2} F_{\mu\nu} Q \frac{\chi_d}{\Lambda} \sigma^{\mu\nu} D^c H \tag{7.2.21}$$

can simultaneously be made real and diagonal. However, since these operators' coefficients have independent phases, we should expect that the coefficient in front of the EDM operator

[‖]This is not exactly true, as equation (7.2.15) gives only the leading order pieces in the Yukawa interactions, and leaves out operators like $Q\chi_u\chi_u^\dagger\chi_d D^c H$, for instance. However, in spite of the large top Yukawa coupling, we find that these additional contributions are not dangerous, and we omit them from our discussion.

will be complex in the mass basis, and generically, we get a contribution to the neutron EDM that is too large unless $\Lambda > 40$ TeV. To evade this bound we must require that CP is a symmetry of the underlying theory, broken spontaneously by χ vevs. Because the same flavon, χ_d, gives rise to both the Yukawa interaction and the EDM operator, spontaneous CP violation guarantees that there is no contribution to the neutron EDM at leading order, and the bound on Λ disappears**.

Provided that the scale Λ is larger than roughly 7 TeV, and that CP is broken spontaneously, minimal $U(3)^5$ sufficiently suppresses all dangerous operators that arise in a spurion analysis. Ideally, though, a flavor symmetry should do more than simply control higher dimensional operators; it should also accomodate a simple understanding of fermion mass hierarchies and mixing angles. Unfortunately, if we insist on a low flavor scale, addressing masses and mixings in the context of $U(3)^5$ becomes problematic. One might attempt to explain mass hierarchies by introducing a few sets of χ's that acquire VEVs at very different scales. However, the $U(3)^5$ mechanism for suppressing dangerous operators requires that only a single χ_d and a single χ_u exist. If instead there were, say, two of each, then a combination

$$a_1\chi_d^1 + a_2\chi_d^2 \tag{7.2.22}$$

would appear in the down quark Yukawa interaction, while a different combination

$$b_1\chi_d^1 + b_2\chi_d^2 \tag{7.2.23}$$

would appear in the down quark EDM operator. There is no generic reason for the second combination to be real in the basis that make makes the first combination real and diagonal

**More precisely, after taking into account higher order contributions to both Yukawa and EDM interactions, the bound is reduced to $\Lambda > 500$ GeV.

(although it is reasonable to assume that corresponding entries of the two combinations are of the same order of magnitude), so the EDM bound $\Lambda > 40$ TeV returns. Similarly, the operator

$$\frac{c}{\Lambda^2}\left(Q(c_1\chi_d^1 + c_2\chi_d^2)D^c\right)\left(Q(f_1\chi_d^1 + f_2\chi_d^2)D^c\right)^\dagger \tag{7.2.24}$$

leads to a stringent bound, $\Lambda > 7$ TeV, coming from the CP violating parameter ϵ_K. Thus we are led to work with only a single set of χ's, whose hierarchical VEVs, we might imagine, arise due to a sequential breaking of the flavor group at widely separated scales. But by adopting this view we encounter the flavon exchange problem problem of section 2.4: one expects the masses of the various flavons that compose the χ's to be of the same order of magnitude as their VEVs, and thus the masses of the lightest flavons to be much smaller than Λ. Unless Λ is quite large, these light flavons mediate flavor changing and CP violating processes at unacceptable levels.

7.3 Small parameters from extra dimensions

What is the origin of the small dimensionless flavor parameters of the standard model? All attempts at understanding these numbers have been based on two ideas. One idea is that these parameters vanish at tree level and are generated radiatively, and that the loop factor is small. In a perturbative theory, with coupling parameters of order unity, the loop factor is of order $1/16\pi^2$. The second idea is that the small fermion mass ratios and mixing angles arise as a ratio of mass scales of the theory, presumably generated dynamically. Such is the situation in Froggatt-Nielsen type theories and in extended technicolor models.

In theories with extra dimensions, however, another attractive possibility arises. Suppose there are flavor symmetries that are primordially exact on our brane, but which

are strongly broken on a distant brane. If bulk fields charged under these symmetries are present, this symmetry breaking is "shined" from the distant branes [13], and there is a new origin for the small parameters, namely, the large volume in the extra dimensions. The fermion mass ratios and mixing angles are small not because of small breaking on the distant branes, but rather due to the flavor breaking messenger's propagation over large distances across the bulk. Fundamentally, the origin is again one of a ratio of mass scales. However, these are set by the distances in the brane configuration, and result in completely new physics possibilities different from other scenarios.

Effects of this shining can be grouped into two categories: spurion effects arising from the free classical theory, and classical and quantum "sniffing" effects, arising from nonlinearities in the Lagrangian.

7.3.1 Free, classical shining

The basic shining effect can be understood as the classical, free propagation of the flavon field through the bulk. From the viewpoint of physics on our brane, the flavor breaking is at this level equivalent to classical spurion effects. We assume that there is some flavor symmetry group G_F, which acts on the matter fields Q_i, U_j, D_k, L_m and E_n, where $\{i, j, k, m, n\}$ specify the representations under which the fields transform. We further assume that G_F is broken at order one on some distant brane by a source J[††]. If J couples to some bulk field, then that field can mediate the flavor symmetry breaking to our wall.

[††]In what follows, we have taken the dimensionality of J to be that of a scalar field living on the symmetry breaking brane. Later, for simplicity, we will set J=1.

For example, suppose the Lagrangian is [‡‡]

$$\mathcal{L} \supset \int d^4x\, dy_{\|}^m dy_{\perp}^n \frac{J_{kl}\chi^{kl}}{M_*^{\frac{n-4}{2}}} \delta^n(y_\perp - y_0) + \int d^4x\, dy^{m+n} \frac{\chi^{kl}}{M_*^{\frac{m+n+2}{2}}} L_k E_l H \delta^{m+n}(y), \quad (7.3.25)$$

where the symmetry is broken on a $(3+m)$-brane at location y_0 in the extra dimensions, and where M_* is the fundamental scale . Since the source brane is an extensive object, it acts as a point source for a Yukawa potential in n dimensions. This is completely analagous to a charged plate in 3 dimensions being described as a point source in 1 dimension. Knowing this, it is simple to write down what the profile of the χ field is as a function of y (neglecting nonlinear interactions),

$$\chi = J\Delta(m_\chi; y) = \frac{J}{M_*^{\frac{n-4}{2}}(2\pi)^{\frac{n}{2}}} \left(\frac{m}{|y|}\right)^{\frac{n-2}{2}} K_{\frac{n-2}{2}}(m|y|) \quad (7.3.26)$$

where $|y|$ is the distance from the source brane to the point in question, K_n is the modified Bessel function, and n is the codimension of the source of χ in the space in which χ propagates. For $m_\chi|y| \ll 1$, this takes the asymptotic form

$$\chi \approx \frac{JM_*}{2\pi} \log(m_\chi|y|) \qquad (n=2) \quad (7.3.27)$$

$$\approx \frac{J\,\Gamma(\frac{n-2}{2})}{4\pi^{\frac{n}{2}} M_*^{\frac{n-4}{2}} |y|^{n-2}} \qquad (n>2) \quad (7.3.28)$$

and for $m_\chi|y| \gg 1$,

$$\chi \approx \frac{J\, m_\chi^{\frac{n-3}{2}}}{2(2\pi)^{\frac{n-1}{2}} M_*^{\frac{n-4}{2}}} \frac{e^{-m_\chi|y|}}{|y|^{\frac{n-1}{2}}} \quad (7.3.29)$$

In this example, the lepton Yukawa matrix will be

$$\lambda_l^{mn} = \frac{1}{M_*^{(n+m+2)/2}} \chi^{mn}. \quad (7.3.30)$$

[‡‡]Here we have allowed that χ transform as a reducible multiplet under G_F for the sake of generality. In an actual model there may be many χ fields, each transforming as an irreducible multiplet.

The flavor symmetry breaking parameters are then either power or exponentially suppressed functions of the distances between branes. If the different elements of χ are generated on different branes, we can, at least in principle, generate a fully general texture, and likewise for the quarks. For example, with $G_F = U(3)^5$ the flavon fields $\chi_{u,d,e}$ appear as single multiplets on our wall, and yet the various entries can have values which are hierarchical.

7.3.2 Classical and quantum sniffing

Flavor breaking from extra dimensions is much more interesting than simply taking the values of χ and its derivatives on our brane at $y = 0$ and using them in a spurion analysis. Non-linearities in the bulk Lagrangian can induce a wide variety of effects which probe flavor breaking at non-zero y.

The simplest examples of this are classical non-linear effects. One of the most significant is the generation of a vev for a bulk field without a direct source brane. Consider a situation with two source branes, with sources J_m^1 and J_n^2 and two bulk fields which have vevs generated on these branes, χ_1^m and χ_2^n. If there is, in addition, another bulk field φ^{mn} transforming as a product representation of the 1 and 2 representations, we naturally have a term in our Lagrangian $\varphi_{mn}^* \chi_1^m \chi_2^n$. As a consequence, φ will also take on a vev in the bulk, and hence on our wall as well.

This is a very familiar situation, even in four dimensions. However, in four dimensions, we typically expect a value $\varphi \propto \chi_1 \chi_2$. In extra dimensions, the vev can typically be much larger. As shown in Fig. 7.1, the fact that the source for φ is spread throughout the bulk means that the dominant contribution to its shining can come from a region distant from our brane, defeating our four-dimensional intuition.

Formally, we want to sum the contributions to φ on our wall from every point in space. Given the assumed coupling, if we take our wall at $y = 0$, and the sources for χ_1 and χ_2 at y_1 and y_2, respectively, we can then calculate

$$\varphi(0) = \int d^n y \Delta(m_\varphi; |y|)\chi_1(y)\chi_2(y). \tag{7.3.31}$$

If the propagators are dominantly exponentials, some small region will dominate this integral, and we can take

$$\varphi(0) \approx \Delta V \Delta(m_\varphi; |\overline{y}|)\chi_1(\overline{y})\chi_2(\overline{y}), \tag{7.3.32}$$

where $\overline{y}$ is some representative point within the volume ΔV where the integrand is appreciable. In cases where the particles are light, ΔV can be very large. Even in cases where the particles are heavy, if they are even an order one factor lighter than the fundamental scale, the volume will typically be larger than one by a factor $1/m_\chi^n$.

We illustrate this with the following example:

Consider a brane configuration with our brane localized at $(0, 0, 0)$ in three extra dimensions, while one source brane is at $y_1 = (10, 0, 0)$ and the other at $y_2 = (10, 3, 0)$ in units where $M_* = 1$. Further, take the masses to be $m_{\chi_1} \approx m_{\chi_2} \approx m_\varphi \approx 1/3$. We can calculate the vev of φ numerically, and find on our brane we have $\chi_1 \approx 2 \times 10^{-4}$, $\chi_2 \approx 2 \times 10^{-4}$, and $\varphi \approx 3 \times 10^{-6}$. We can understand the larger value of φ as also being enhanced by a volume factor ΔV of (7.3.32) being larger than one, and we show this graphically in Fig. 7.2.

If we further give the fields moderately different masses, $m_{\chi_1} \approx 1/3$, $m_{\chi_2} \approx 1/2$, $m_\varphi \approx 1/5$, we find $\chi_1 \approx 2 \times 10^{-4}$, $\chi_2 \approx 5 \times 10^{-5}$, but $\varphi \approx 10^{-5}$, much larger than the naive expectation if all other masses are order the fundamental scale. This is very sensitive to the

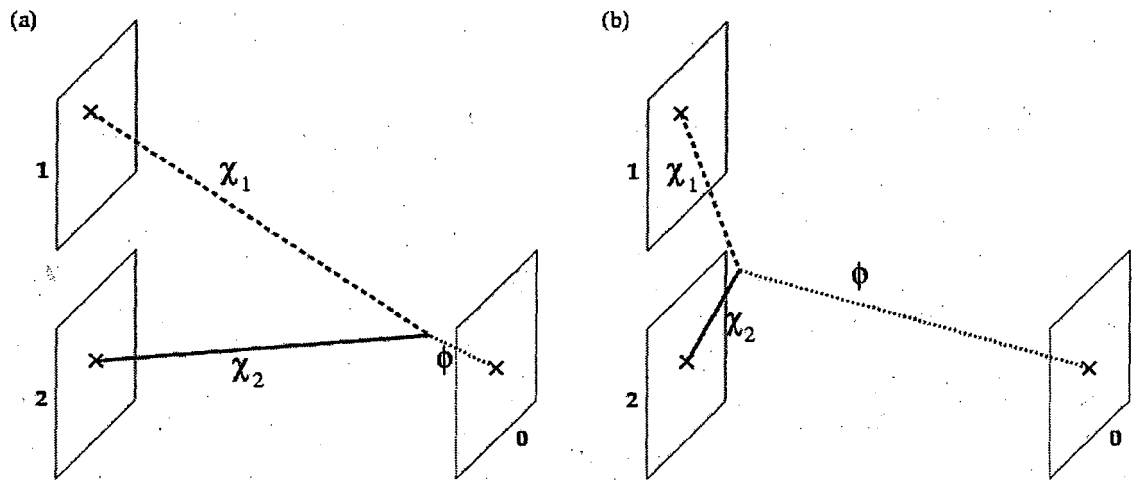

Figure 7.1: Contributions to the vev of the field φ. Our brane is designated 0, while the source branes for χ_1 and χ_2 are 1 and 2, respectively. (a) is supressed by two propagators while (b) is suppressed by only one. As a consequence, (b) will typically dominate. If the mass of φ is even an order one factor lighter than either of the other fields, the difference can be further amplified.

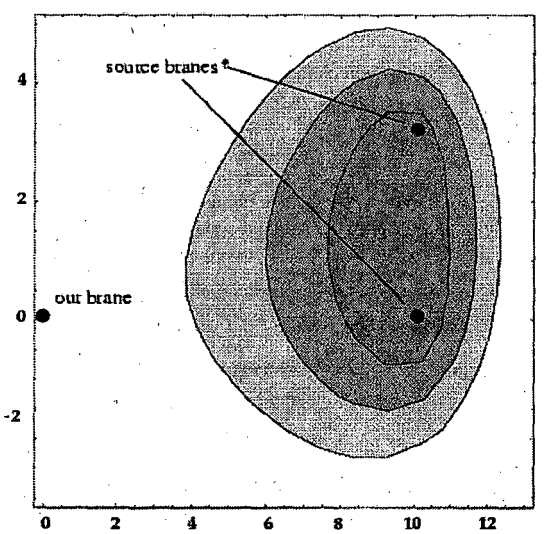

Figure 7.2: We plot here a $z = 0$ slice of the first, second and third efolds of the integrand of (7.3.31) from its maximum for the given example. Notice that the region contributing to the integral is both large ($\Delta V > 1$) and far from our brane.

brane geometry and the masses of the bulk fields, of course, and the predictivity suffers as a consequence. However, it illustrates one very important difference between flavor model building in extra dimensions versus that in the usual four.

Once a bulk vev for a field exists, it can act to further regenerate another field that may fall off more quickly. In the previous example, if we reach a region where $\chi_1\varphi > \chi_2$, then this region can act as a further source for χ_2, dominating for some regions of y over the source brane. This can be understood rather simply: in situations where the vev profile is dominantly exponential, (i.e., when $m_{\chi_2} y \gg 1$), and if further $m_\varphi < m_{\chi_2}$, it can be advantageous to exploit the presence of χ_1, for instance, and propagate as a φ, as we illustrate in 7.3(a). Of course, if the regenerated value of χ_2 is sufficiently large, it can again regenerate φ in certain circumstances and the nonlinearities can dominate the entire problem. It is important to be aware of this when employing non-linear effects in model building.

All of the effects discussed so far arise from classical field theory in the bulk. They could be obtained, in principle, by solving the non-linear classical fields equations.* Quantum effects in the bulk may be just as, if not more, important. For example, if $m_\varphi + m_{\chi_1} < m_{\chi_2}$, it may be advantageous to pay the price of a loop factor and propogate as a $\varphi - \chi_1$ loop, which is shown in the "Tie-Fighter" diagram of figure 7.3(b).

As another example of non-linear bulk interactions changing the physics on our brane at $y = 0$, consider a bulk field φ with a φ^4 interaction and a source brane at y_s. Ignoring the non-linearity, an operator involving φ^p on our wall would have classically a

*There is one remaining classical effect, namely the generation of local operators through bulk non-linearities, which we will discuss at the conclusion of this section.

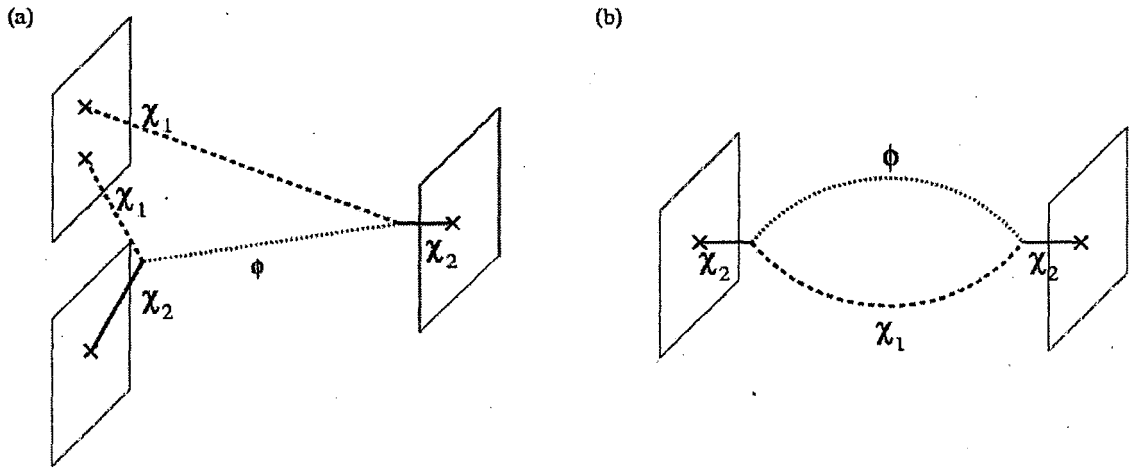

Figure 7.3: In addition to the standard direct propagation of χ_2, there can be other contributions. In the presence of an external field χ_2, it can be advantageous to propagate as a φ, as shown in (a). Even absent such an external field, the propagation through a loop of lighter particles may be dominant compared to the direct propagator (b).

coefficient of $\Delta(m_\varphi; y_s)^p$. This is illustrated in 7.4(a) for $p = 3$. However, if $m_\varphi |y_s| \gg 1$, there will be a large exponential suppression of this contribution, so that the dominant effect may instead come from the loop diagram of 7.4(b), which gives a contribution to $<\varphi(0)^3>$ of

$$L^2 \int d^n y \, \Delta(y)^3 \varphi(y), \tag{7.3.33}$$

where L is a loop factor [†]. This quantum effect can lead to a very large deviation of $<\varphi(0)^3>$ from $<\varphi(0)>^3$, since it involves only one power of $e^{-m_\varphi |y_s|}$.

For the case of $p = 2$ and an operator on our brane involving φ^2, in addition to the tree contribution there is the 1-loop contribution shown in Fig. 7.5a, equal to

[†]Actually, the exact expression involves an integration both over 4 and higher dimensional momenta. The result can be re-expressed (upon wick rotating to Euclidean space) as

$$\int \frac{d^4 k_4}{(2\pi)^4} \frac{d^4 \bar{k}_4}{(2\pi)^4} \int d^n y \Delta(\sqrt{m_\varphi^2 + k_4^2}, y) \Delta(\sqrt{m_\varphi^2 + \bar{k}_4^2}, y) \Delta(\sqrt{m_\varphi^2 + (k_4 + \bar{k}_4)^2}, y) \varphi(y) \tag{7.3.34}$$

In our expression in the text, we are *overestimating* this effect by replacing $\Delta(\sqrt{m_\varphi^2 + k_4^2}, y)$ with $\Delta(m_\varphi, y)$. Henceforth, we will often make similar approximations in discussions of sniffing.

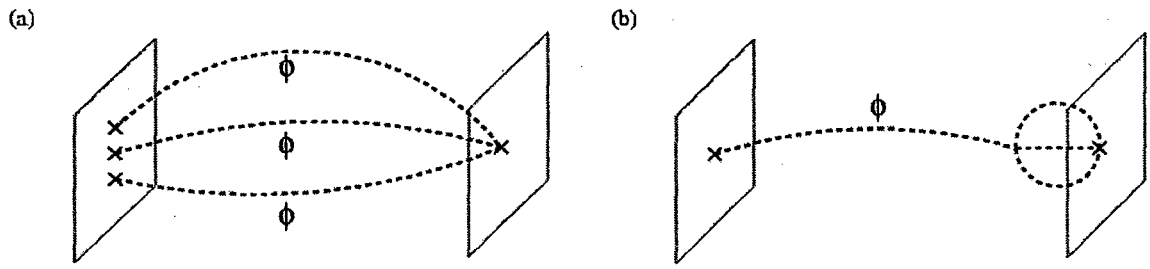

Figure 7.4: Two contributions to the value of $<\varphi>^3$ on our brane. The contribution in (a) is the classical spurion contribution. The contribution in (b) is due to sniffing and can often be larger than that of (a).

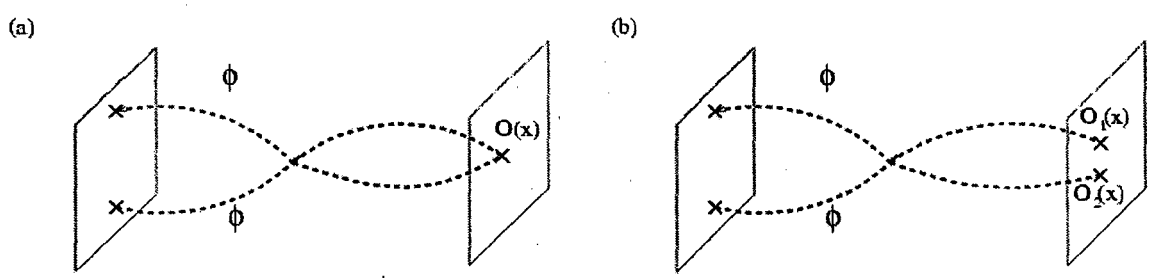

Figure 7.5: Two examples of "sniffed" contributions to an operator on our wall. In (a) a quantum loop corrects the value of $<\chi^2>$ on our wall. In (b), a φ field can interact in the bulk and generate a local operator.

$$L \int d^n y \, \Delta(y)^2 \varphi(y)^2. \tag{7.3.35}$$

In this case, two propagators traverse the whole space from $y = y_s$ to $y = 0$ in both tree and loop diagrams, so it may appear that the loop correction is small and unimportant. However, in theories of flavor, we will find that the loop diagram contains flavor breaking not present at tree level. For instance, if φ transforms as a multiplet of G_F, it may point in different directions in flavor space at different y; this is a situation which arises when multiple source branes are present. At tree level the only flavor breaking is given by $\varphi(0)$, whereas at the loop level the flavor breaking of $\varphi(y)$ is also probed. We say that additional flavor breaking is "sniffed" in the bulk from points $y \neq 0$.

In addition to quantum corrections to $<\varphi^p>$, there can be classical corrections as well. By integrating out a φ field which interacts in the bulk, we generate local operators as we illustrate in figure 7.5(b). In this particular example, given operators $\varphi(x_1)O_1(x_1)$ and $\varphi(x_2)O_2(x_2)$, we generate a local operator $\varphi^2(x)O_1O_2(x)$. Absent bulk corrections, this operator would have a coefficient $<\varphi>^2$, but sniffing contributions can change this. Although both φ-legs on our wall are evaluated at the same point in spacetime, this is not a quantum effect and does not receive the same loop suppression as in Fig. 7.5(a).

7.3.3 Spatial derivatives of the flavon field

The bulk flavon fields have y dependent profiles, and, because Lorentz invariance is violated in directions perpendicular to our brane, one might imagine that wherever a χ field appears in the Lagrangian of our brane, we could just as easily write $(a_n\partial_n + a_{mn}\partial_m\partial_n + ...)\chi$, with no need to contract indices of extra dimensional derivatives [21]. Unless the χ mass

were much smaller than the scale Λ, the derivative terms would not be strongly suppressed. Nor would χ and its derivatives necessarily be proportional to each other, since the field can have sources on several branes, potentially leading to a variety of troublesome flavor changing effects. However, if Lorentz invariance is broken *spontaneously* (as is the case if standard model fields are localized on a D-brane, for instance), only certain derivative terms are allowed. In the low energy effective theory, we simply have SM fields localized to our brane, the bulk χ field, and the goldstones of spontaneously broken translational invariance Y^m, which give the position of our brane in the extra dimensions [22]. Thus, all terms involving derivatives of a single power of χ must feature either $\Box^j\chi$ for some integer j, or a brane tension-suppressed coupling to the goldstone. For instance, we can have terms like

$$\partial_\mu Y^m(\overline{Q}\,\partial_m\chi\,\overline{\sigma}^\mu Q) \qquad \text{or} \qquad Q\Box\chi D^c H, \tag{7.3.36}$$

but not something like

$$Q\,\partial_m\chi\,D^c H, \tag{7.3.37}$$

because extra-dimensional derivatives with uncontracted indices amount to *explicit* breaking of Lorentz invariance.

Of course, the localized fields have finite profiles in the bulk, and we can contract derivatives of χ with derivatives of wall fields in the full extra-dimensional theory. However, the effective field theory argument just given indicates that in our wall's low energy theory only terms involving $\Box^j\chi$ will be generated. It is straightforward to see how this comes about explicitly from a microscopic description. Let us label the $(4+n)$ dimensional spacetime coordinates as (x^μ, y^m), where x^μ are the 4D coordinates on our brane, and y^m are the coordinates of the extra n dimensions. We will use an index $K = (\mu, m)$ that runs over all

$(4+n)$ dimensions. Consider the Lorentz invariant term

$$\frac{1}{\Lambda^{n+3}} \int d^4x \int d^n y \, (\partial_K Q^{(4+n)})(\partial^K \chi^{(4+n)}) D^{c(4+n)} H^{(4+n)}. \tag{7.3.38}$$

We assign the standard model fields $\psi^{(4+n)}$ Gaussian profiles in the extra dimensions, so that their relation to the canonically normalized fields in 4D is

$$\psi^{(4+n)}(x,y) = \left(\frac{2\Lambda^2}{\pi}\right)^{n/4} \psi^{(4)}(x) e^{-\Lambda^2 |y|^2}, \tag{7.3.39}$$

while the x-independent bulk flavon VEV is given by

$$\chi^{(4+n)}(x,y) = \Lambda^{n/2} \chi^{(4)}(y). \tag{7.3.40}$$

In terms of the canonically normalized fields, (7.3.38) becomes

$$-\Lambda^{n-1} \int d^4x \, Q^{(4)}(x) \left(\int d^n y \, e^{-3\Lambda |y|^2} y^m \, \partial^m \chi^{(4)}(y) \right) D^{c(4)}(x) H^{(4)}(x), \tag{7.3.41}$$

where we have neglected factors of π and 2. After integration by parts, the piece involving χ becomes

$$\int d^n y \, (6\Lambda^2 |y|^2 - n) \chi^{(4)}(y) e^{-3\Lambda |y|^2}. \tag{7.3.42}$$

This is of the form $\int d^n y \, f(y) e^{-\alpha |y|^2}$, which is equivalent to

$$\int d^n q \, \tilde{f}(q) e^{-|q|^2/\alpha} = (e^{-\Box/\alpha} f)\Big|_{y=0}, \tag{7.3.43}$$

where $\tilde{f}(q)$ is the Fourier transform of $f(y)$. In our case, we have $f(y) = \chi^{(4)}(y)(a + b|y|^2)$, with a and b real, which satisfies

$$e^{-\Box/\alpha} (\chi^{(4)}(a + b|y|^2))\Big|_{y=0} = \sum_{j=0}^{\infty} c_j \, \Box^j \chi^{(4)}\Big|_{y=0}, \tag{7.3.44}$$

with real coefficients c_j. Although there can be many χ derivatives that appear in the Yukawa interaction, we see that they are all proportional to $\Box^j\chi$ for some integer j and thus, by the equations of motion, all proportional to χ itself.

In contrast, an operator like

$$\partial_m\chi\partial^m\chi O_{SM}(x), \tag{7.3.45}$$

where O_{SM} is an operator of standard model fields, cannot be brought into a form involving only $\Box^j\chi_d$. The presence of these operators has model dependent effects which we will discuss in later sections.

7.3.4 Harmless flavon exchange

As discussed in section 2.4, breaking flavor symmetries at low Λ in four dimensions generates harmful flavor-changing operators through the exchange of flavons, due to the smallness of the flavor-breaking scales relative to the fundamental scale. In sharp contrast, there is no reason in extra dimensional theories to expect that the bulk flavon masses are closely related to the sizes of the Yukawa couplings, as these small parameters are no longer ratios of mass scales. However, even if the bulk fields *were* very light, the harmful operators still receive no subsequent enhancement. This is due to the IR softness of bulk propagators in extra dimensions. Returning to the example of section 2.4, let us suppose that φ_{12} lives in p extra dimensions. Then, the coefficient of the induced 4-fermi operator is (working in units with $M_* = 1$)

$$v^2 \int \frac{d^p\kappa}{(2\pi)^p} \frac{1}{\kappa^2 + m_\varphi^2}. \tag{7.3.46}$$

Note that we integrate over the extra dimensional momenta κ since this momentum is not conserved. The important point is that for $p > 2$, this integral is dominated in the UV and

is insensitive to m_φ! Therefore, the generated operator is *not* enhanced by $1/m_\varphi$ factor. In fact, the dominant contribution is m_φ independent and generates a flavor-symmetric operator suppressed by powers of M_* and additional loop factors. Sub-dominant contributions need *not* respect the flavor symmetry, but are small enough; the leading corrections go as

$$L(p)\log m_\varphi^2,\ p=2; \qquad L(p)m_\varphi,\ p=3; \qquad L(p)m_\varphi^2,\ p=4,5,\cdots \tag{7.3.47}$$

where

$$L(p)=\frac{1}{p2^{p-1}\pi^{p/2}\Gamma(p/2)} \tag{7.3.48}$$

is a loop factor.

We can also return to the example of our toy $U(2)$ theory discussed in section 2.5. The tree-level exchange of a *bulk* A field produces the operator

$$L(n)\frac{H^*H(Q_aD_b^c)(\bar{Q}^a\bar{D}^{cb})}{M_*^4}. \tag{7.3.49}$$

There is no inverse dependence on the A mass at all, and we can therefore tolerate $M_* \sim 1-10$ TeV, roughly three orders of magnitude below the bound on Λ in the 4D case.

7.4 A $U(3)^5$ theory in extra dimensions with 3 source branes

Having introduced the shining of flavor breaking from distant branes, we now discuss the construction of $U(3)^5$ models in extra dimensions. In the models we will describe, χ_u and χ_d (which in this context are bulk fields) will be the only flavons that couple directly to the standard model fields of our 4D universe, just as in minimal $U(3)^5$ in 4D. The authors of [13] have applied the shining framework described in section 3 to the case of $U(3)^5$. In their picture, χ_u (χ_d) couples to nine source fields $\varphi^{u,ij}$ ($\varphi^{d,ij}$). Each of these source fields

transforms as χ_u (χ_d) and is localized on its own distinct brane. The sources acquire VEVs of the form

$$< \varphi^{u,ij}_{kl} > \sim < \varphi^{d,ij}_{kl} > \sim \Lambda \delta^i_k \delta^j_l, \tag{7.4.50}$$

that is, each of the nine φ^d sources essentially shines a single element of the down quark mass matrix, and similarly for the up sector. The magnitude of each source is taken to be roughly Λ, but large fermion mass ratios are still possible by requiring some source branes to be closer to our brane than others. This represents a significant improvement over the minimal case in 4D: only a single χ_d and a single χ_u appear in the Yukawa interactions, and yet a simple explanation for the hierarchical nature of the fermion masses is achieved.

This picture is far from complete, however. The most serious deficiency is that no understanding is provided of why $V_{CKM} \sim I$. Related to this is the fact that the VEVs of (7.4.50) do not comprise a justifiable starting point, as we will now argue. To avoid problems with goldstone bosons, we work with a large discrete subgroup of $U(3)^5$ rather than with $U(3)^5$ itself (because the breaking is order unity, we avoid the light pseudo-goldstone bosons that appear in the 4D case). The directions of the eighteen φ^u and φ^d VEVs are thus fixed in various directions that do not depend on bulk dynamics. The important point is that there is no reason for the direction of a source on one brane to be related in any particular way to the direction of a source on another. This reasoning argues against the arrangement of VEVs in (7.4.50), and more generally, it tells us that we should expect order unity CKM mixing angles if all the sources are on separate branes. Suppose, for example, that a φ^u localized on one nearby brane breaks $U(3)_Q \times U(3)_{u^c} \to U(2)_Q \times U(2)_{u^c}$, while on another brane, a φ^d independently breaks $U(3)_Q \times U(3)_{d^c} \to U(2)_Q \times U(2)_{d^c}$. We imagine that

these branes are the ones nearest us, so that these φ's shine the leading order contributions to the quark mass matrices. The $U(3)_Q \times U(3)_{d^c}$ symmetry allows us to take

$$\varphi^d = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & v \end{pmatrix}, \tag{7.4.51}$$

and then, using the $U(3)_{u^c}$ symmetry, we can write write

$$\varphi^u = M \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & v' \end{pmatrix}, \tag{7.4.52}$$

where the form of $M \in U(3)$ is fixed by the explicit breaking. The point is that there is no reason for φ^u and φ^d to choose the same unbroken $U(2)_Q$, and there is not in general a basis in which both φ^u and φ^d are diagonal, because the $U(3)_Q$ freedom is used up entirely in diagonalizing either one or the other. Generically, we expect the (23) entry of φ^u to be roughly as large as its (33) entry[‡], and since the leading order form of $\chi_{u(d)}$ on our brane is simply proportional to φ^u (φ^d), we should expect a large CKM mixing angle, contrary to what is observed.

7.4.1 A complete $U(3)^5$ model

We now describe a model that retains the successes of the picture just described, but which in addition predicts small mixing angles. The model is remarkably simple. We assume the existence of a series of source branes, each of which has localized on it a triplet under $U(3)_Q$, a triplet under $U(3)_{u^c}$, and a triplet under $U(3)_{d^c}$. Nothing special distinguishes any of the source branes - we will even assume for simplicity that they are identical

[‡]Using the residual $U(2)_Q$ symmetry respected by φ^d, the (13) entry can be made to vanish.

copies of each other - except that they are located at different distances from our brane. The three triplets on each brane acquire VEVs near the fundamental scale and act as sources for bulk flavons χ_u and χ_d. We again regard the true flavor group as a large discrete subgroup of $U(3)^5$ to avoid goldstones, so the potential for each triplet features a discrete series, rather than a continuum, of minima. Each triplet's VEV is stuck at one of these minima, unable to tunnel from one to another. Moreover, the directions chosen by the sources on one brane are not related in any particular way to the directions chosen on a different brane. What we have, effectively, is explicit breaking on each source brane, with the triplets getting fixed, complex VEVs that point in uncorrelated directions.

Let us work out the implications of this simple scenario. On the brane nearest ours, the triplet sources acquire VEVs that, if we exploit our $U(3)^3$ freedom, we can write as

$$T_Q^1 = \begin{pmatrix} 0 \\ 0 \\ v_Q \end{pmatrix}, \qquad T_u^1 = \begin{pmatrix} 0 \\ 0 \\ v_u \end{pmatrix}, \text{ and } T_d^1 = \begin{pmatrix} 0 \\ 0 \\ v_d \end{pmatrix}, \tag{7.4.53}$$

with v_Q, v_u, and v_d real and not much smaller than M_*. In fact, these sources could even be localized on *our* brane. The bulk flavons are shined by the triplet sources due to the brane interactions[§]

$$T_Q^1 \chi_u T_u^1 \qquad \text{and} \qquad T_Q^1 \chi_d T_d^1. \tag{7.4.54}$$

Consider, for the moment, the extreme case in which this brane is by far the closest one to our ours - the closest by so much that, on our wall, we can ignore contributions to flavon VEVs coming from all other sources. In contrast to the example that led to the alignment

[§]If the sources are on our brane, the third generation quarks acquire mass from direct couplings to the triplets.

problem of equations (7.4.51) and (7.4.52), both χ_u and χ_d are shined from the same nearby brane, and simultaneously take the form

$$\chi_{u,d} \propto \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & A_{u,d} \end{pmatrix}. \tag{7.4.55}$$

Now let us consider the additional contributions to $\chi_{u,d}$ that are shined from more distant sources. The VEVs of (7.4.53) respect a residual $U(2)^3$ symmetry that can be used to write the sources on the second nearest brane as

$$T_Q^2 = v_Q \begin{pmatrix} 0 \\ \sin\theta_Q \\ \cos\theta_Q e^{i\alpha_Q} \end{pmatrix}, \quad T_u^2 = v_u \begin{pmatrix} 0 \\ \sin\theta_u \\ \cos\theta_u e^{i\alpha_u} \end{pmatrix}, \text{ and } T_d^2 = v_d \begin{pmatrix} 0 \\ \sin\theta_d \\ \cos\theta_d e^{i\alpha_d} \end{pmatrix}, \tag{7.4.56}$$

where we assume for simplicity that $T^\dagger T$ is the same in the various discrete minima. After we include the effects of the shining interactions

$$T_Q^2 \chi_u T_u^2 \qquad \text{and} \qquad T_Q^2 \chi_d T_d^2, \tag{7.4.57}$$

the flavon VEVs on our brane take the form

$$\chi_{u,d} \propto \begin{pmatrix} 0 & 0 & 0 \\ 0 & \epsilon & \epsilon \\ 0 & \epsilon & A \end{pmatrix}_{u,d}, \tag{7.4.58}$$

with $\epsilon_{u,d} \ll A_{u,d}$. At this stage, the VEVs in (7.4.53) and (7.4.56) still admit a $U(1)^3$ symmetry that can be used to write the sources on the third brane as

$$T_Q^3 = v_Q \begin{pmatrix} \mathrm{s}\varphi_Q \\ \mathrm{c}\varphi_Q\, \mathrm{s}\rho_Q\, e^{i\beta_Q} \\ \mathrm{c}\varphi_Q\, \mathrm{c}\rho_Q\, e^{i\gamma_Q} \end{pmatrix}, \quad T_u^3 = v_u \begin{pmatrix} \mathrm{s}\varphi_u \\ \mathrm{c}\varphi_u\, \mathrm{s}\rho_u\, e^{i\beta_u} \\ \mathrm{c}\varphi_u\, \mathrm{c}\rho_u\, e^{i\gamma_u} \end{pmatrix}, \quad T_d^3 = v_d \begin{pmatrix} \mathrm{s}\varphi_d \\ \mathrm{c}\varphi_d\, \mathrm{s}\rho_d\, e^{i\beta_d} \\ \mathrm{c}\varphi_d\, \mathrm{c}\rho_d\, e^{i\gamma_d} \end{pmatrix}. \tag{7.4.59}$$

Including the effects of this brane, we arrive at the Yukawa texture

$$\lambda_{u,d} \sim \begin{pmatrix} \epsilon' & \epsilon' & \epsilon' \\ \epsilon' & \epsilon & \epsilon \\ \epsilon' & \epsilon & A \end{pmatrix}_{u,d} , \tag{7.4.60}$$

which features both a hierarchy of eigenvalues and small mixing angles.

This model features a simple symmetry breaking pattern. If we include only the nearest brane, its sources break $U(3)^3 \to U(2)^3$. Bringing the second brane into the picture then breaks $U(2)^3 \to U(1)^3$. Finally, moving the third brane into place breaks $U(1)^3 \to$ nothing. Note that this breaking pattern is not put in by hand, but rather follows inevitably from the fact that the sources transform as triplets and acquire fixed VEVs pointing in random different directions. Note also that we work with three source branes only because this is the minimal set required to break $U(3)^3$ entirely. Given that at least three exist, the success of our picture is insensitive to how many branes there are in all. Additional branes, being further away, will give small contributions to the Yukawa couplings, leaving the texture of equation (7.4.60) unchanged.

The Yukawa texture suggests the approximate relations $|V_{ij}| \sim m_{d_i}/m_{d_j}$, which work reasonably well for all mixing angles except for θ_c. The fact that θ_c naively comes out to small is not a serious problem, because the entries of (7.4.60) come with unknown coefficients of order 1 due to the unknown angles θ, φ, *etc.*, that appear in equations (7.4.56) and (7.4.59): by taking $\tan\varphi_d \sim \frac{1}{5}\sin\rho_d$ we obtain the correct size $\theta_c \sim 1/4$. Somewhat surprisingly, the more closely *aligned* the source triplets on the second and third branes are, the *larger* θ_c is. In sections 6 and 7 we consider different models that accomodate θ_c more easily.

It is advantageous to give the flavons slightly different masses, with $m_{\chi_u} > m_{\chi_d}$, to explain why the mass hierarchies are stronger for the up quarks than for the down quarks. Doing so makes all the more pressing the question of why $m_t > m_b$, given that these masses are essentially shined from the same brane. Having two Higgs doublets with large $\tan\beta$ leads to flavor changing problems, as we will see below. A simple alternative that leads to no phenomenological difficulties is to have a $\sim 1/60$ suppression of v_d relative to v_u. Even irrespective of flavor changing issues, this approach may be more appealing than a non-SUSY large $\tan\beta$ scenario, because here the different-sized VEVs are given to two fields, T_u and T_d, that transform entirely differently under the flavor symmetry. In contrast, if we have two Higgs doublets, H_d and $\tilde{H}_u$ transform identically under the gauge symmetry and are both flavor singlets, so it is especially difficult to understand how one is chosen to have a much larger VEV than the other. A more interesting approach to understanding m_b/m_t will be described in section 7. The details of how the present $U(3)^5$ model can give realistic fermion masses and mixings are important but should not obscure the central point: having source triplets with uncorrelated VEVs leads *automatically* to a CKM matrix with small mixing angles.

Another attractive feature of this model is that it violates CP spontaneously, as the VEVs of equations (7.4.56) and (7.4.59) give the off-diagonal elements of λ order unity phases. If we impose CP as a symmetry of the underlying theory, this model exhibits the same solution to the EDM problem described in section 2.6 in the context of $U(3)^5$ in 4D.

7.4.2 Flavor-changing from the bulk

The only question is whether there are additional challenges in suppressing dangerous operators, now that we are working in extra dimensions. As discussed in section 3, physics in our 4D universe can be sensitive not only to the values of χ_u and χ_d on our brane, but also to their values away from our brane, due to "sniffing" effects. For instance, if we have bulk couplings $\int d^n y\, Tr(\chi_u^\dagger \chi_u \chi_u^\dagger \chi_u)$, and $\int d^n y\, (Tr(\chi_u^\dagger \chi_u))^2$, then in the up quark EDM operator we can replace the matrix $\langle \chi_u(y=0)\rangle_{ij}$ with

$$\frac{\langle \chi_u^\dagger(0)\rangle_{mn}}{16\pi^2} \int d^n y\, \langle \chi_{uim} \chi_{unj}(y)\rangle \Delta^2(y), \tag{7.4.61}$$

or with

$$\frac{1}{(16\pi^2)^2} \int d^n y\, \langle \chi(y)\rangle_{ij} \Delta^3(y), \tag{7.4.62}$$

where we have included loop factors from integrating over 4D momenta. Diagrams representing these contributions are shown in Figs. 7.5a and 7.4b, respectively. Because they do not have the same flavor structure as χ_u, we need to check that these contributions are not problematic. The largest contribution to the 1-1 entry comes from the piece of (7.4.61) proportional to λ_t,

$$\frac{\lambda_t}{16\pi^2} \int d^n y\, \langle \chi_{u13} \chi_{u31}(y)\rangle \Delta^2(y). \tag{7.4.63}$$

$\Delta(y)$ is largest near our brane ($y \sim 0$), but in this region $\langle \chi_{u13}\chi_{u31}(y)\rangle$ nearly vanishes in the mass diagonal basis. On the other hand,$\langle \chi_{u13}\chi_{u31}(y)\rangle$ is largest near the third most distant brane, located at y_3. Using the short distance form for the propagator to evaluate $\langle \chi_{u13}\chi_{u31}(y)\rangle$ in this region, we get a contribution of roughly

$$\frac{\lambda_t}{16\pi^2} \lambda_u^2 \frac{1}{dS_n} f(m_\chi), \tag{7.4.64}$$

where $f(m_\chi) \sim (\frac{1}{m_\chi})^{(4-n)}$, $\mathrm{Log}(m_\chi)$, and 1 for $n = 2$ or 3, $n = 4$, and $n > 4$ respectively, and where dS_n is the surface area of the unit sphere in n dimensions. The potential mild enhancement from $f(m_\chi)$ cannot nearly compensate for the extra factor of $\frac{\lambda_u}{16\pi^2}$ relative to what we have for $\langle \chi_u(0) \rangle_{11}$, so this contribution is harmless. Sniffing contributions to the down quark EDM are similarly suppressed.

Sniffed versions of

$$(Q\chi_d D^c)(Q\chi_d D^c)^\dagger \qquad \text{and} \qquad (Q\chi_d D^c H)(Q\chi_d D^c H)^\dagger \tag{7.4.65}$$

yield $\Delta S = 2$ operators with coefficients of approximate size

$$\frac{1}{16\pi^2} \int d^n y \, \langle \chi_{d12} \chi^*_{d21}(y) \rangle \Delta^2(y), \tag{7.4.66}$$

and

$$v^2 \int d^n y \, \langle \chi_{d12} \chi^*_{d21}(y) \rangle \Delta^2(y), \tag{7.4.67}$$

respectively. Again concentrating on the region around y_3, we estimate the integrals as roughly $(\lambda_s \theta_c)^2 \frac{1}{dS_n} f(m_\chi)$. If the coefficient were simply $(\lambda_s \theta_c)^2$, ϵ_K would require $\Lambda > 7$ TeV. Since these contributions are further suppressed by either a loop factor or by v^2, they are safe.

A different challenge posed by the extra dimensions involves the bulk flavon derivatives described in section 3.3. If we allowed all flavor invariant terms with extra dimensional derivatives acting on χ_d, then in the basis that diagonalized and made real the down quark Yukawa interaction

$$Q\left((1 + a_n \partial_n + a_{mn} \partial_m \partial_n + ...)\chi_d\right) D^c H, \tag{7.4.68}$$

the EDM operator

$$F_{\mu\nu}Q\left((1+b_n\partial_n+b_{mn}\partial_m\partial_n+...)\chi_d\right)\sigma^{\mu\nu}D^cH \tag{7.4.69}$$

would in general be complex, leading to the familiar EDM bound $\Lambda > 40$ TeV. However, as discussed in section 3.3, the only drivatives of χ_d allowed in equations (7.4.68) and (7.4.69) are those of the form $\Box^j\chi_d = m^{2j}\chi_d$, provided Lorentz invariance is broken spontaneously. In this case derivative terms are harmless as far as EDM's are concerned. On the other hand, the operator

$$\frac{1}{\Lambda^2}\left(Q\frac{\partial_K\chi_d}{\Lambda^2}D^c\right)\left(Q\frac{\partial^K\chi_d}{\Lambda^2}D^c\right)^\dagger \tag{7.4.70}$$

cannot be brought into a form involving only $\Box^j\chi_d$. Because $\partial_m\chi_d$ is not in general proportional to χ_d, the most pessimistic view is then that ϵ_K forces us to take $\Lambda > 7$ TeV. (In fact, at this point it becomes clear why using large $\tan\beta$ to explain $m_t > m_b$ is disasterous: the $\Delta S = 2$ piece of (7.4.70) has a coefficient that is proportional to $\tan^2\beta$). Note, however, that the $\Lambda > 7$ TeV interpretation assumes that the derivative terms are entirely unsuppressed: if $m_\chi = \Lambda/S$, for instance, then the bound is reduced by a factor of S. Also, it is conceivable that $\partial_a\chi_d$ *is* nearly proportional to χ_d. For example, if the source branes lie along along the same direction from ours, then in the case of three extra dimensions, the derivative contributions that are not proportional to χ_d are suppressed by factors of $1/(\Lambda r_i)$ relative to the leading non-derivative contribution, where r_i are the various distances of the source branes from our brane. $K-\overline{K}$ mixing is most sensitive to contributions shined from branes responsible for the light quark masses. Taking $(mr)\sim 5$ for these branes, and $m\sim\Lambda/3$, we find that the bound on Λ is reduced by a factor ~ 15. The general point is that bounds derived by considering terms involving flavon derivatives are softer than those

obtained from operators without derivatives, as they are more sensitive to the flavon mass, and to the details of the brane configuration.

We have seen that $U(3)^5 \times$CP models with triplet sources are generically safe for $\Lambda \approx 5$ TeV, provided that Lorentz invariance is broken spontaneously. Next we will show that specific models can be safe at this scale without qualification. In particular, we present what we consider the simplest specific realization of our $U(3)^5 \times$CP scenario, and find that regardless of how Lorentz invariance is violated, and regardless of whether m_χ is suppressed relative to Λ, both flavon derivative and "sniffing" effects are harmless.

7.5 A concrete realization of $U(3)^5$

Here, we will consider a concrete arrangement of branes in our $U(3)^5$ scenario. The arrangement is very simple and furthermore allows analytic calculation of FCNC effects. We will see that the potential flavor-changing effects are very suppressed by this particular set-up; for instance all χ derivatives are exactly aligned with χ.

We imagine that even though there are $n \geq 2$ extra spatial dimensions, flavor is associated with only one of them, which we parametrize by y. Our 3-brane and several source 3-branes are taken to lie in a 4-brane described by $(x,\ 0 \leq y \leq L)$, where we compactify on an interval $[0, L]$ (rather than a circle) of moderately large size, $L \sim 10 M_*^{-1}$. The 3-branes are spread out roughly evenly in the space available to them and are then naturally spaced between $\sim 1 - 10$ times M_*^{-1}, so that the question of what determines the sizes of the inter-brane separations is to some extent obviated. The flavon χ is taken to propogate only on this 4-brane.

In one infinite extra dimension, the χ propagator is just $e^{-m|y-y'|}$. When the

dimension is compactified, the propagator depends on the boundary conditions. We will impose the conditions $\chi = 0$ at the boundaries of the interval, so that the propagator on the strip from y_1 to y_2 is

$$\Delta(y_1, y_2) = \left(\theta(y_1 - y_2)\frac{\sinh[m(L-y_1)]}{\sinh[mL]}\sinh[my_2] + \theta(y_2 - y_1)\frac{\sinh[my_1]}{\sinh[mL]}\sinh[m(L-y_2)]\right) \tag{7.5.71}$$

Note that this goes to $e^{-m|y_1-y_2|}$ when $L \to \infty$, as it should. The classical profile for χ is then

$$\chi_{cl}(y) = \sum_i \chi_i \Delta(y_i, y), \tag{7.5.72}$$

where

$$\chi_{i(a\alpha)} = T_{Lia} T_{Ri\alpha} \tag{7.5.73}$$

is the source for χ shone from the $i'th$ wall. Suppose that our 3-brane, located at $y = y_*$, is positioned to the left of all the other branes on the strip, i.e. $y_* < y_3$ where y_3 is the location of the nearest wall. Then, for all $y < y_3$ the profile of χ is

$$\chi_{cl}(y) = \bar{\chi}\sinh[my], \tag{7.5.74}$$

where

$$\bar{\chi} = \sum_i \chi_i \frac{\sinh[m(L-y_i)]}{\sinh[mL]}. \tag{7.5.75}$$

Note that the $\chi_{cl}(y)$ at different values of $y < y_3$, are all proportional to the same matrix, and so all derivatives of χ_{cl} evaluated on our wall are diagonal in the same basis as χ_{cl} itself. Therefore, even allowing for heavy $m_\chi \sim \Lambda$ and explicit Lorentz violation in the extra

dimension, there are no problems with derivative terms. In the absence of interactions in the bulk, the FCNC analysis is identical to the standard $U(3)^5$ spurion analysis.

In four dimensions, $U(3)^5$ with minimal flavons leads to exact lepton flavor conservation. As we have discussed, in higher dimensions, derivative operators have the potential to induce some level of flavor changing, but in this simple realization of $U(3)^5$, where $\partial_m \chi \propto \chi$, one might expect again that flavor changing is absent. This intuition is incorrect, as sniffing effects give us sensitivity to the value of χ in the bulk, and thus to regions where it is not diagonal in the mass basis. However, it is easy to see that these effects are highly suppressed.

We can illustrate this by considering the process $\mu \to 3e$, which occurs due to the presence of the operator

$$E^c_k \chi^{kl} \chi^\dagger_{lm} \overline{E}^{c\,m} E^c_n \overline{E}^{c\,n}. \tag{7.5.76}$$

In addition to the spurion contribution, which is diagonal in the flavor basis, we have the sniffed contribution

$$< \chi\chi^\dagger >= \int dy d^4x \chi_{cl} \chi^\dagger_{cl}(y) \Delta(y,x)^2. \tag{7.5.77}$$

If we assume a brane geometry where our brane is at $y = 0$, and the μ and e branes are at positions $0 < y_\mu < y_e$, the sniffed contributions from the region $y < y_\mu$ will all be proportional to χ. Thus, the first flavor changing piece comes in the region $y_\mu < y < y_e$. We calculate the coefficient of the operator to be

$$< \chi\chi^\dagger >^{12}_{FC} \simeq \frac{\lambda_e \lambda_\mu m^3}{32\pi^2 (\mathrm{Log}\lambda_\mu)^2} \approx 10^{-13}, \tag{7.5.78}$$

which gives a completely unobservable rate for $\mu \to 3e$. Thus, while sniffing does allow for

flavor violations even in a $U(3)^5$ theory (where naively lepton flavor is conserved!), they are easily small enough to be harmless.

7.6 Smaller flavor symmetries

In the previous two sections we developed extra-dimensional models of flavor with low Λ and $G_F = U(3)^5$. One might wonder how difficult it is to work instead with smaller flavor groups. For instance, we have already seen that taking the symmetry to be $U(2)$ is problematic: the extra dimensions alleviate the flavon exchange problem, but the $U(2)$-invariant, non-renormalizable operator of equation (7.2.11) forces $\Lambda > 10^5$ TeV. Of course, we could hope that this operator is simply not generated by the underlying theory, but if one wants to assume that all invariant operators are present, then we need a larger symmetry. Here we adopt the group $U(2)^3$ and consider the quark sector alone. We again take CP to be a symmetry of the underlying theory in hopes of evading the EDM bound.

With this choice of flavor symmetry both $Q_3 D_3^c H$ and $Q_3 U^c \tilde{H}$ are flavor singlets, so to explain $m_b \gg m_t$, we might require two Higgs doublets with large $\tan\beta$. Unfortunately, as we have already seen, unless a single flavon multiplet is responsible for the elements of λ_d involving the light generations (which will not be the case for $U(2)^3$), then we expect to get the operator

$$\frac{(\lambda_s \theta_c \tan\beta)^2}{\Lambda^2}(Q_2 D_1^c)(\overline{Q}_1 \overline{D}_2^c), \tag{7.6.79}$$

forcing $\Lambda > 400$ TeV. Thus, we instead use a single Higgs but enlarge the flavor symmetry to include an extra $U(1)$ factor under which only D_3^c is charged. We introduce a bulk flavon θ, whose charge under this $U(1)$ is opposite that of D_3^c, and take the VEV of θ on our wall to be $\sim 1/60$.

Next we introduce another bulk flavon, φ, a doublet under $U(2)_Q$. We can choose a basis in which its VEV on our wall is

$$\varphi = \begin{pmatrix} 0 \\ v \end{pmatrix}, \tag{7.6.80}$$

with v real, and to yield a reasonable V_{cb} we take $v \sim 1/30$. Our picture for the symmetry breaking that gives masses to the light generations is designed to preserve certain features of the standard $U(2)$ fermion mass texture, in particular the relation $\theta_c \approx \sqrt{m_d/m_s}$. We imagine that on a distant brane the subgroup $U(2)_Q \times U(2)_{d^c}$ is broken down to $U(2)$ by a source that shines the bulk flavon χ_d, transforming as $(2, \overline{2})$ under this subgroup. Similarly, from a different brane we have $U(2)_Q \times U(2)_{u^c} \to U(2)$ breaking transmitted by χ_u, a $(2, \overline{2})$ under *this* subgroup (note that we do not expect the two walls to preserve the same $U(2)$). Finally, we imagine that on our brane both $U(2)_{d^c}$ and $U(2)_{u^c}$ are broken *primordially* to their $SU(2)$ subgroups.

What does this assortment of breakings say about the Yukawa matrices? The flavor symmetry allows us the freedom to choose convenient forms for the χ_u and χ_d VEVs, but because of the primordial breaking on our brane, and because of the freedom already used to fix the form of φ, we are only allowed arbitrary $SU(2)_{u^c} \times SU(2)_{d^c} \times U(1)_Q$ transformations, where the $U(1)_Q$ acts on Q_1 alone. These transformations allow us to take

$$\chi_d = v_d \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad \chi_u = v_u \begin{pmatrix} e^{-i\delta} & 0 \\ 0 & 1 \end{pmatrix} \tag{7.6.81}$$

on our wall, with both v_u and v_d real. We are stuck with a phase in χ_u; this is the origin of CP violation in the model. The leading order couplings of these flavons to the quarks are

$$Q_L \overline{\chi}_d{}^L{}_l \epsilon^{lm} D^c{}_m \qquad \text{and} \qquad Q_L \overline{\chi}_u{}^L{}_l \epsilon^{lm} U^c{}_m, \tag{7.6.82}$$

so that at this stage the Yukawa textures are

$$\lambda_D = \begin{pmatrix} 0 & v_d & 0 \\ -v_d & 0 & \theta v \\ 0 & 0 & \theta \end{pmatrix} \qquad \text{and} \qquad \lambda_U = \begin{pmatrix} 0 & v_u e^{i\delta} & 0 \\ -v_u & 0 & v \\ 0 & 0 & 1 \end{pmatrix}. \tag{7.6.83}$$

To give masses to the charm and strange quarks, we assume the presence of two additional bulk fields, $\xi_{u,d}$, that transform as 2×2 under $U(2)_Q \times U(2)_{u^c,d^c}$[¶]. However, these flavons are not shined from distant branes, but rather have VEVs induced in the bulk by the interactions

$$\mathcal{L} \supset \varphi_L \overline{\chi}^L{}_l \varphi_M \overline{\xi}^{lM}. \tag{7.6.84}$$

Due to its sensitivity to the flavon masses and to the brane geometry, the size of the sniffed ξ is essentially a free parameter. Note, however, that the orientation and phase of ξ is determined entirely by the orientation and phase of φ and χ, so that we have

$$\xi_{u,d} = \begin{pmatrix} 0 & 0 \\ 0 & v'_{u,d} \end{pmatrix}, \tag{7.6.85}$$

where v'_u and v'_d are both real.

Including the leading order coupling of all flavons to the quarks yields the textures

$$\lambda_D = \begin{pmatrix} 0 & v_d & 0 \\ -v_d & v'_d & \theta v \\ 0 & 0 & \theta \end{pmatrix} \qquad \text{and} \qquad \lambda_U = \begin{pmatrix} 0 & v_u e^{i\delta} & 0 \\ -v_u & v'_u & v \\ 0 & 0 & 1 \end{pmatrix}, \tag{7.6.86}$$

and leads to the approximate relations

$$\theta_c \sim \sqrt{\frac{m_d}{m_s}} \sim \frac{v_d}{v'_d}, \qquad \left|\frac{V_{ub}}{V_{cb}}\right| \sim \sqrt{\frac{m_u}{m_c}} \sim \frac{v_u}{v'_u}, \tag{7.6.87}$$

[¶]It would be problematic to instead introduce doublets under $U(2)_{u^c}$ and $U(2)_{d^c}$ for this purpose, because the relation $\theta_c \approx \sqrt{m_d/m_s}$ would be spoiled by the Yukawa term $Q_L \overline{\varphi}_Q{}^L \varphi_{dl} \epsilon^{lm} D^c{}_m$. Moreover, the bulk coupling $\overline{\varphi}_Q \chi_d \varphi_d$ would regenerate χ_d in the vicinity of our wall, and would also disrupt the texture (we might expect, for instance $\lambda_{d21} \sim \lambda_{d22}$).

$$\frac{m_s}{m_b} \sim \frac{v'_d}{\theta}, \qquad \frac{m_c}{m_t} \sim v'_u, \qquad V_{cb} \sim v. \tag{7.6.88}$$

We get reasonable values for all observables by taking $v'_d \sim 3 \times 10^{-4}$, $v'_u \sim 3 \times 10^{-3}$, $v_d \sim 6 \times 10^{-5}$, $v_u \sim 2 \times 10^{-4}$, $v \sim 1/30$, and $\theta \sim 1/60$. Note that the CKM matrix is of the form

$$V_{CKM} = \begin{pmatrix} R^u_{12} & & \\ & 1 & \end{pmatrix} \begin{pmatrix} e^{i\delta} & & \\ & R_{23} & \end{pmatrix} \begin{pmatrix} R^d_{12} & & \\ & 1 & \end{pmatrix}, \tag{7.6.89}$$

so that the unitarity triangle relations for this model will simply be those of standard $U(2)$.

How safe is this model? In the mass basis, we expect to have the operator of equation (7.2.11) generated with coefficient $\sim (\lambda_s \theta_c)^2$, so that the bound from ϵ_K is reduced to $\Lambda > 7$ TeV. The issue of the neutron EDM is more subtle. Despite the fact that CP is broken spontaneously, one might expect this model to have an EDM problem because the mass and EDM matrices are produced by several flavons, rather than by a single multiplet as in $U(3)^5$. However, an attractive feature of this model is that, in the mass basis, the phase δ appears only in the CKM matrix and *not* in the leading order EDM matrices. This is clear from (7.6.86): rotating $U_1 \to e^{-i\alpha} U_1$ makes the mass and EDM matrices completely real, because the 1-2 entry of the up quark EDM matrix started out with precisely the same phase as λ_{U12}. Higher order contributions to these matrices disrupt this cancellation. We find that the order of magnitude of the contribution to the neutron EDM is determined by the coupling

$$Q_L \overline{\chi}_u{}^L{}_l \epsilon^{lm} \overline{\chi}_u{}^M{}_m \xi_{dnM} \epsilon^{np} D^c{}_p. \tag{7.6.90}$$

Setting $p = L = 1$ gives 1-1 entries in both the down quark EDM and mass matrices of roughly $v_u^2 v'_d e^{i\delta}$ in the flavor basis. In the mass basis, the 1-1 element of EDM matrix has approximate size v_d^2/v'_d, so the phase of that element will be roughly $\frac{v_u^2 v'^2_d}{v_d^2} \sim 10^{-6} - 10^{-5}$,

suggesting that the bound on Λ is reduced to well below 1 TeV.

7.7 Predictive Theories

We find the picture for generating flavor of sections 4 and 5, based on the maximal $U(3)^5$ flavor group and involving only three identical flavor-breaking branes with triplet sources pointing in random directions, to be both elegant and plausible. Nevertheless, it does not provide a *predictive* theory of flavor. Furthermore, there is no explanation of the b/t hierarchy. There are two points that need to be addressed in order to build more predictive theories based on $U(3)^5$:

- The brane geometry must be more constrained.
- The directions in $U(3)^5$ space shone by the triplets must be more constrained.

In this section, we will present some examples of more predictive theories along these lines. In order to deal with the second point, we will assume that the dynamics on the branes is such that the triplet sources have identical strengths and can only shine in three orthogonal directions, which we take to be $(0,0,1), (0,1,0)$, and $(1,0,0)$. If we continue to work with just three parallel source branes, as in the previous sections, we would be stuck with $V_{CKM} = 1$. Therefore we consider other brane configurations. In particular, we imagine that the triplets T_u, T_d, and T_Q are all localized on different defects, which we label as U^c, D^c, and Q branes. We take all the Q branes to be parallel and equally spaced, and similarly for the U^c and D^c branes. However, the Q branes intersect at right angles with both the U^c and D^c branes, and at the junctions χ_u and χ_d have sources. The Yukawa matrices are thus shined from the points of intersection on a grid of flavor breaking branes.

7.7.1 Simple grid models

We will describe three simple and predictive grid models. In the first, the triplet sources are localized on three sets of three parallel five-branes of infinite extent. Labeling the extra dimensions by the numbers $1, 2, ..., n$, we take the Q, U^c, and D^c branes to fill extra dimensions 1 and 2, 2 and 3, and 1 and 3, respectively. At the four-dimensional intersection of a Q brane with a U^c (D^c) brane, the T_Q and T_u (T_d) triplets shine the bulk flavon χ_u (χ_d). We could further imagaine the existence of an additional bulk field $\chi_{u,d}$ that transforms as $(\overline{3}, \overline{3})$ under $U(3)_{D^c} \times U(3)_{U^c}$. This flavon would not induce dangerous operators on our brane, and would simply make the picture more symmetric. We take the masses of the three flavons, as well as the spacings between the Q, U^c, and D^c branes, to be identical. Until we have a theory that determines the inter-brane separations, we can't justify the regularity of the grid, however the symmetry of the system ensures that the configuration is at least a local extremum of the potential. In an attempt to understand why $m_t \gg m_b$, we imagine that our three-brane is located at one of the Q - U^c intersections, but is not in contact with a D^c brane. Below we will find it necessary to make the offset, the shortest distance from our three-brane to the nearest Q - D^c intersection, much smaller than the brane spacing. The configuration is represented in Figs. 7.6a and 7.6b. In Fig. 7.6a we project onto the 1-3 plane passing through our three-brane, so that the Q - U^c intersections appear as points. In Fig. 7.6b we do the same for the 2-3 plane passing through our brane, so that the Q - D^c intersections appear as points. We have placed our universe near a corner of the grid, where it is easiest to attain hierarchical quark masses for all three generations.

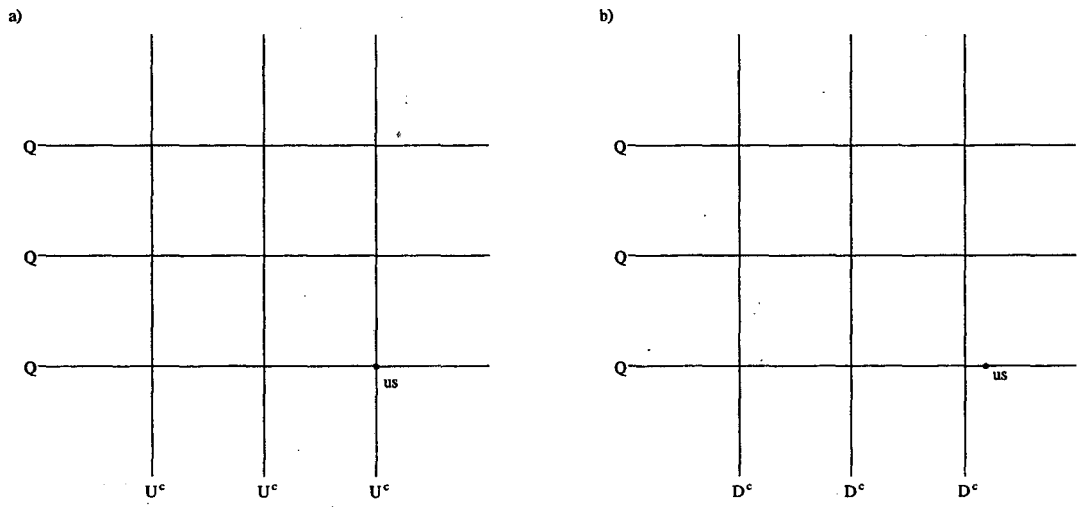

Figure 7.6: The brane configuration for the first grid model. We project onto a plane parallel to the D^c branes for (a), and parallel to the U^c branes for (b).

We stress that the starting point for this theory is a remarkably symmetrical configuration of source branes. In the absence of our 3 brane, and of spontaneous breakings, the configuration is completely symmetrical with respect to interchanging any pair of the extra dimensions 1,2 and 3. The labels Q, U^c, D^c are just labels of identical sets of branes. The lack of symmetry only occurs by virtue of the position of our own 3 brane and the gauging on it. Flavor symmetry is built into the large scale structure of the bulk, and is explicitly broken only at a point defect.

Already we can see that the resulting mass matrices will have an interesting structure. For instance, keeping only the exponential dependence of the χ propagators, the up mass matrix has the form

$$\lambda^U \sim \begin{pmatrix} \epsilon^{2\sqrt{2}} & \epsilon^{\sqrt{5}} & \epsilon^2 \\ \epsilon^{\sqrt{5}} & \epsilon^{\sqrt{2}} & \epsilon \\ \epsilon^2 & \epsilon & 1 \end{pmatrix} \tag{7.7.91}$$

where $\epsilon \sim \exp -(m_\chi S)$ is the suppression factor with S the interbrane spacing. This sort of pattern is not expected in 4d theories of flavor, where usually only integer powers of a

small parameters appear.

We now proceed to a quantitative analysis. As mentioned above, we take the magnitudes of each source VEV to be precisely the same, and allow the triplets to point in three orthogonal directions. We can choose a basis where the triplets on the Q, D^c,and U^c branes nearest us all point in the 3 direction; for conciseness we label these branes as $Q3$, U^c3, and D^c3. The D^c3 - $Q3$ intersection shines the Yukawa coupling

$$\lambda^D_{33} = \alpha_\chi \, \Delta(m_\chi; y^D_{33}), \tag{7.7.92}$$

where α_χ absorbs the source strengths and their couplings to the bulk flavons (we take these to be the same for χ_u and χ_d), and y^D_{33} is the distance from our brane to the D^c3-Q^c3 intersection. Note that if there are k extra dimensions, the propagator is given by equation (7.3.26) with $n = k - 1$, because the shining is from the 4D intersection of two five-branes. Meanwhile, from the U^c3-$Q3$ intersection we get the Yukawa coupling

$$\lambda^U_{33} = \alpha_\chi \, \Delta(m_\chi; y_{cutoff}) + \alpha_T, \tag{7.7.93}$$

where the second term comes from the direct coupling of the triplet sources to standard model fields, and the propagator has been cutoff at a distance $y_{cutoff} \sim 1/\Lambda$. Of course, with α_T, α_χ, m_χ, and the offset y^D_{33}, we have more than enough freedom to fit the top and bottom quark masses; the hope is that these four free parameters plus the brane spacing S can be simultaneously chosen to give reasonable CKM mixing angles and mass ratios for the other quarks as well. An additional hope is that the parameter values for a successful fit not be too far from unity. In this case, the smallness of m_b/m_t is not put in by hand by simply choosing $\alpha_\chi \ll \alpha_T$, but is instead a consequence of our location at a U^c-Q intersecion, away

from D^c branes. Note that even in the limit that we are very near a $D^c - Q$ interesection, m_b/m_t is still suppressed by a factor if $\Gamma(\frac{n-2}{2})/4\pi^{n/2}$.

The model we have described is quite constrained. For a given configuration of triplet VEVs, the quantities y^D_{33}/S and Sm_χ specify all ratios of Yukawa matrix elements except those involving λ^U_{33}; by further fixing $\alpha_\chi m_\chi^{k-3}$, where k is the number of extra dimensions, we determine the magnitudes of all Yukawa matrix elements except λ^U_{33}, which is given only once we choose α_T. Thus there are four free paramaters to predict six masses and three mixing angles[‖]. The predictions turn out to be wrong. A qualitative reason for this can be understood by considering only the two nearest Q, U^c, and D^c branes. We must be able to choose the source VEVs on these branes as ($Q3$, $Q2$), (U^c3, U^c2), and (D^c3, D^c2) - if there were a repetition in any of the VEV directions, then three sets of three branes would not be sufficient to give masses to all of the quarks[**]. The size of m_s/m_b is approximately the ratio of the contributions to χ_d from $Q2$ - D^c2 and $Q3$ - D^c3 shining. The distance from our brane to $Q2$ - D^c2 is longer than that to $Q3$ - D^c3 by at least S, regardless of y^D_{33}/S; if we make Sm_χ larger than roughly 2 or 3, then m_s/m_b automatically comes out too small. Meanwhile the dominant contribution to V_{cb} comes from the ratio of the contributions to χ_d from $Q2$ - D^c3 and $Q3$ - D^c3 shining. For the moderate values of of Sm_χ needed for m_s/m_b, getting $V_{cb} \ll 1$ requires the offset y^D_{33} to be substantially smaller than the spacing S (by roughly a factor of 3 or more). With y^D_{33}/S constrained in this way, the ratio of the charm mass, which arises dominantly from $Q2$ - U^c2 shining, to the bottom mass, comes out too small[††]. Of course, this problem can be avoided if we introduce

[‖]CP violation is discussed below.

[**]We could have (U^c3, U^c3), and a massless up quark, but this makes the problem described below only more severe.

[††]In the case of four extra dimensions, for example, if we require $.036 < V_{cb} < .042$ and $1/24 < m_s/m_b <$

an additional free parameter, for instance by letting χ_u and χ_d have different masses, by allowing the spacing between the U^c and D^c branes to be different, or by giving χ_u and χ_d different couplings to the triplet sources. However, even if adjustments like these are made to accomodate m_c/m_b, we find that it is not possible to simultaneously obtain accurate predictions for all other mass ratios and mixing angles.

A simple modification of the brane grid just described is to eliminate the U^c branes and place the T_u sources on the same branes as the T_Q's. To make the picture more symmetric, we could imagine that on the D^c branes we have additional triplet sources $T_{Q'}$ that transform under yet another $U(3)$, under which all standard model fields are singlets. In the original grid model, our brane needed to be located at an intersection of different branes to get the additional contribution to λ^U_{33} from the direct coupling of triplet sources to standard model fields; here, the direct coupling is automatic provided only that we reside on a Q/U^c brane. Two other important differences distinguish this grid from the previous one: first, there are now only three independent sources that shine χ_u; second, the classical profiles for χ_u and χ_d shining are no longer identical - if the source branes are co-dimension l objects, then the χ_u profile is determined using equation (7.3.26) with $n = l$, while for the χ_d profile one should use $n = l + 1$ (as the intersections of D^c and Q branes have co-dimension $l + 1$). One fortunate effect of the latter difference is to increase m_c/m_b from what the previous grid gave, for given choices of y^D_{33}/S, S, and m_χ.

To determine how well this grid can fit quark masses and mixings, we need to specify a VEV configuration. For all six quarks to acquire mass, there must not be repetitions of VEV orientations (for example, we need one each of D^c3, D^c2, and D^c1). Unfortunately,

1/75, then we obtain $m_c/m_b < 1/23$, with the value of m_c/m_b optimized when $y^D_{33}/S \sim .08$ and $Sm_\chi \sim .7$.

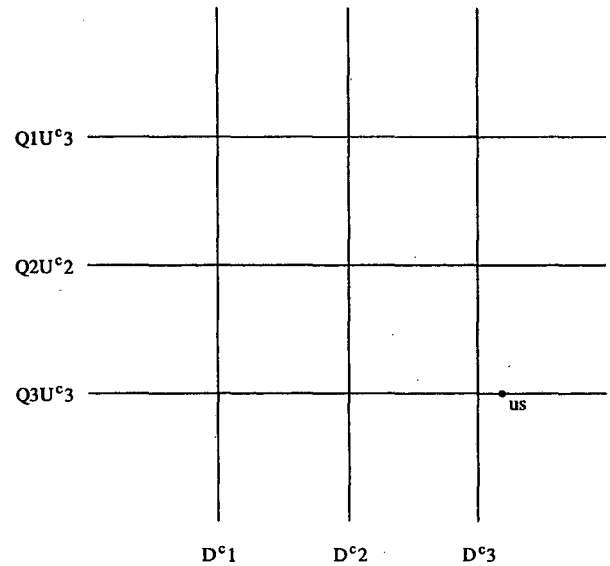

Figure 7.7: The brane configuration for the second grid model. The numbers indicate the directions of the triplet source VEVs.

giving a 3-2-1 pattern to all three sets of triplet VEVs leads automatically to an up quark mass that is too large compared to m_d. In light of this we choose *not* to make the VEVs on the most distant Q/U^c brane $(Q1, U^c1)$, but instead choose them to be $(Q1, U^c3)$, as shown in Fig. 7.7 (choosing $(Q1, U^c2)$ leads to too large a contribution to θ_c coming from the up sector). With this VEV configuration, the up quark Yukawa matrix has the texture

$$\lambda^U \sim \begin{pmatrix} 0 & 0 & \epsilon' \\ 0 & \epsilon & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{7.7.94}$$

In particular, we have $m_u = 0$, which is allowed at second order in chiral perturbation theory[25].

Interestingly, this grid model is slightly less constrained than the previous one: because the χ_u and χ_d sources have different dimensionality, the ratio m_c/m_b depends on m_χ and S independently, so that there are five free parameters (m_χ, S, y_{33}^D, α_χ, α_T). With

these chosen to be (.43, 3.3, .43, .1, ~ 1)[‡‡], and with source branes of co-dimension two, we obtain a mixing matrix with elements of magnitude

$$V_{CKM} = \begin{pmatrix} .975 & .223 & .0040 \\ .223 & .974 & .037 \\ .0045 & .037 & .999 \end{pmatrix}, \tag{7.7.95}$$

and find masses

$$m_d = 1.9\,\text{MeV} \qquad m_u = 0 \tag{7.7.96}$$

$$m_s = 70\,\text{MeV} \qquad m_c = 1.4\,\text{GeV} \tag{7.7.97}$$

$$m_b = 4.2\,\text{GeV} \qquad m_t = 174\,\text{GeV}, \tag{7.7.98}$$

where we have included RGE running. The mass ratios and the magnitudes of the CKM matrix elements are consistent with those inferred from data, except that $m_s/m_d = 37$ is too high by $\sim 50\%$, and $|V_{ub}/V_{cb}| = .11$ is too large by $\sim 10\%$. Note that we are partially successful in understanding the smallness of m_b/m_t: the dimensionless parameter required to fit this mass ratio, α_χ, is $\sim 1/10$ rather than $\sim 1/60$. The most serious problem with the model as presented so far is that there is no CP violation. This is easily remedied: if we allow the triplet VEVs to be complex, then we are left with an irremovable phase in the CKM matrix, provided the phases of the T_u VEVs on the $(Q3, U^c3)$ and $(Q1, U^c3)$ branes are different.

We find it encouraging that the simple, regular grid shown in Fig. 7.7 can describe quark masses and mixings so well. Perhaps the strangest, least desirable feature of this grid is our peculiar location relative to it. This motivates the grid shown in Fig. 7.8, in which our brane is located at the precise center. Given that we are in the middle, located on a

[‡‡]The precise value of α_T is fixed by fitting the top mass.

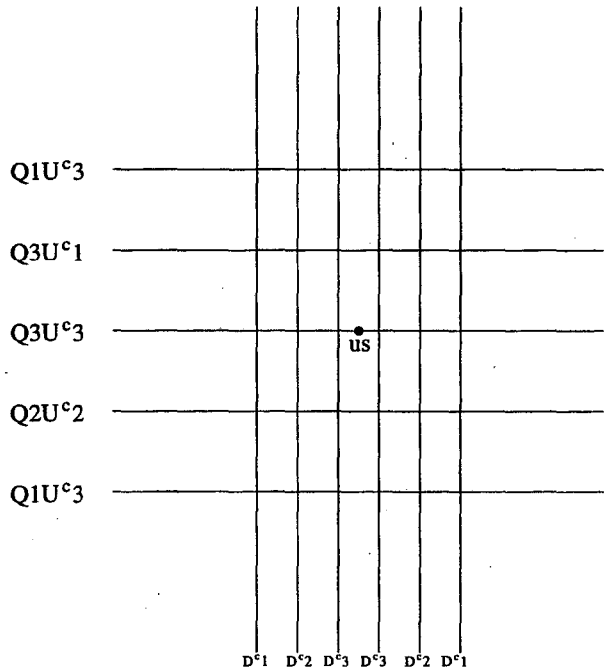

Figure 7.8: The configuration of branes and source VEVs in the third grid model.

Q/U^c brane and in between two D^c branes, this construction features the minimum number of branes required to give masses to all three down-type quarks. We have eliminated the free parameter y^D_{33}; in its place we allow the D^c brane spacing to differ from the spacing of the Q/U^c branes. With the orientation of source VEVs shown in Fig. 7.8, the up quark Yukawa matrix is

$$.\lambda^U \sim \begin{pmatrix} 0 & 0 & \epsilon' \\ 0 & \epsilon & 0 \\ \epsilon & 0 & 1 \end{pmatrix}, \tag{7.7.99}$$

so that now the up quark obtains a small mass proportional to $\epsilon\epsilon'$. If we choose our free parameters (m_χ, S_{D^c}, S_{Q/U^c}, α_χ, α_T) to be (.23, 2.75, 8, .2, ~ 1), then we obtain

$$V_{CKM} = \begin{pmatrix} .976 & .219 & .0057 \\ .219 & .975 & .039 \\ .0030 & .039 & .999 \end{pmatrix}, \tag{7.7.100}$$

and

$$m_d = 4.3\,\text{MeV} \qquad m_u = 1.3\,\text{MeV} \tag{7.7.101}$$

$$m_s = 130\,\text{MeV} \qquad m_c = 1.3\,\text{GeV} \tag{7.7.102}$$

$$m_b = 4.3\,\text{GeV} \qquad m_t = 174\,\text{GeV}. \tag{7.7.103}$$

All masses and mixings agree with experiment at the 50% level: $|V_{td}|$ and m_d/m_s are both too small by $\sim 25\%$, while $|V_{ub}|/|V_{cb}|$ is too large by $\sim 50\%$. Again, we find it intriguing that the symmetric grid of Fig. 7.8, with our 3 brane at its center, can account rather well for the pattern of quark masses and mixings.

7.8 Conclusions

The gauge hierarchy problem has motivated several directions for constructing theories beyond the standard model. Each of these has presented certain challenges and opportunities for making progress on the flavor problem. Constructing realistic theories of fermion masses in technicolor theories without fundamental scalars proved to be very difficult - especially incorporating the heavy top quark. In the simplest supersymmetric theories, the Yukawa couplings of the standard model are simply copied as superpotential interactions. As in the standard model there is an economical description which provides no understanding of the origin of flavor. The ideas for understanding the origin of small dimensionless Yukawa couplings are the same as for theories without supersymmetry: perturbative loops or the Froggatt Nielsen mechanism using hierarchies of mass scales. While there are new twists on these old ideas – superpartners can be in the loop and the dynamics of many strongly interacting supersymmetric theories are understood – at the end of the

day one is tempted to say that supersymmetry did not allow much progress in understanding flavor. In fact, for supersymmetric theories the question has been how to avoid taking a step in the wrong direction: there are severe constraints from flavor-changing and CP violating processes on the form of the soft supersymmetry breaking interactions involving squarks and sleptons. While the answer motivated some flavor groups, it may be that these constraints are telling us more about how supersymmetry is broken than about how flavor is broken.

In contrast, if we live on a three brane at some location in the bulk, with the fundamental scale from our viewpoint of order a TeV, then the constraints on theories of flavor are radically altered, and a whole new world of flavor models is opened up. At first sight it again appears that we are heading in the wrong direction: how could disastrous flavor changing effects be avoided from operators generated at such a low scale, from familons and from light flavon exchange? We have argued that all three objections are immediately removed by having a discrete non-Abelian flavor group spontaneously broken on source branes in the bulk. The fundamental scale of flavor breaking on these source branes is order unity, but the breaking effects on our 3-brane are small because the source branes are distant from us. The origin of the flavor parameters is now a convolution of two effects: the geometrical configuration of the source branes in the bulk and our location relative to them, and the random relative orientations of the flavor breaking vevs on the various source branes. Phenomenology places some constraints on these effects, and energetics suggest that the brane configuration will be highly symmetrical. We find that the convolution of these two effects has sufficient complexity to lead to the collection of mystifying flavor numbers

nature has given us, while still originating from a very simple and elegant symmetrical structure. An interesting aspect of this picture is that flavor symmetry is a crucial feature of the extra dimensions and is important in determining the brane configurations in the bulk. On the other hand, our gauge interactions are restricted to our 3-brane, and are unimportant from the viewpoint of the bulk.

It is important to stress that these theories really are new, and cannot be mimicked by 4 dimensional theories. For example, the relative size of entries in the Yukawa matrices are governed by distances to sources from which flavor breaking is shone. These relative distances involve factors, such as $\sqrt{2}$ and $\sqrt{5}$ as shown in eqn. [7.7.91], which are characteristic of the spatial geometry. Furthermore, these theories of flavor can occur whether the extra dimensions are large, small or infinite, and whether the background geometry is flat or curved. In addition, there are new ideas for a qualitative explanation of features of the fermion mass spectrum. In grid theories a fermion mass hierarchy is *inevitable* – our 3-brane must be located closer to some source branes than to others. The uniquely heavy top quark is explained by having our 3-brane located on source branes which break flavor symmetry in the up sector. The origin of m_t/m_b may be a brane configuration such as the one shown in Figure 8, where our 3-brane lies equidistant between two source branes for breaking flavor in the down sector.

As well as a new structure for flavor, and new ideas for qualitative features of the fermion mass spectrum, theories in extra dimension offer a completely new mechanism for obtaining precise flavor parameter predictions. It is striking that there have been only a few theoretical ideas which lead to relations amongst the flavor observables, such as $\theta_c \approx$

$\sqrt{m_d/m_s}$ and $m_b \approx 3m_\tau$. Texture zeros, symmetry properties of the Yukawa matrices, and grand unified relations between up, down and lepton sectors have been the most important tools. Extra dimensions offer a completely new possibility: the Yukawa matrices on our 3-brane at y_0 are given by $\chi(y_0) = \sum_i \chi_i \Delta(y_0 - y_i)$ where y_i is the location of source i which shines the flavor breaking χ_i via the propagator $\Delta(y_0 - y_i)$. The positions y_i depend only on the lattice spacing, and there may be few possible spontaneous choices for the orientation of χ_i. This basic idea can be implemented in a wide range of models.

There are clearly very many source brane structures to be considered, even concentrating on those with high symmetry, and one may question whether such constructions are plausible origins for the quark and lepton mass matrices. We find little reason to prefer the alternative picture of multiple Frogatt-Nielsen fields and flavons with masses enormously high compared to the TeV scale. Rather than debate the relative merits, it seems worth exploring this new class of theories in which there is a spatial geometry of flavor.

Bibliography

[1] N. Arkani-Hamed, S. Dimopoulos and G. Dvali, Phys. Lett. **B429** (1998) 263, hep-ph/9803315.

[2] I. Antoniadis, N. Arkani-Hamed, S. Dimopoulos and G. Dvali, Phys. Lett. **B436** (1998) 257, hep-ph/9804398.

[3] N. Arkani-Hamed, S. Dimopoulos and G. Dvali, Phys. Rev. **D59** (1999) 086004, hep-ph/9807344.

[4] R. Sundrum, hep-ph/9805471; N. Arkani-Hamed, S. Dimopoulos and J. March-Russell, hep-th/9809124.

[5] L. Randall and R. Sundrum, hep-th/9905221.

[6] L. Randall and R. Sundrum, hep-th/9906064.

[7] N. Arkani-Hamed, S. Dimopoulos, G. Dvali and N. Kaloper, hep-th/9907209.

[8] J. Lykken and L. Randall, hep-th/9908076.

[9] I. Antoniadis and C. Bachas, hep-th/9812093; N. Arkani-Hamed, S. Dimopoulos and J. March-Russell, hep-th/9908146.

[10] K. Dienes, E. Dudas and T. Gherghetta, Nucl. Phys. **B537** (1999) 47, Phys. Lett. **B436** (1998) 55.

[11] N. Arkani-Hamed and M. Schmaltz, hep-ph/9903417.

[12] L. Randall and R. Sundrum, hep-th/9810155.

[13] N. Arkani-Hamed and S. Dimopoulos, hep-ph/9811353.

[14] L.J. Hall and L. Randall, Phys. Rev. Lett. **65** (1990) 2939.

[15] C.D. Froggatt and H.B. Nielsen, Nucl. Phys. **B147** (1979) 277.

[16] Z. Berezhiani and G. Dvali, Phys. Lett. **B450** (1999) 24.

[17] N. Arkani-Hamed, C.D. Carone, L.J. Hall and H. Murayama, Phys. Rev. **D54** (1996) 7032.

[18] A. Antaramian, L.J. Hall and A. Rasin, Phys. Rev. Lett. **69** (1992) 1871; L.J. Hall and S. Weinberg, Phys. Rev. **D48** (1993) 979.

[19] ALPHA and UKQCD Collaboration (Joyce Garden et al.), hep-lat/9906013.

[20] R. Barbieri, G. Dvali and L. Hall, Phys. Lett. **B377** (1996) 76.

[21] T. Banks, M. Dine and A. Nelson, hep-th/9903019.

[22] R. Sundrum, Phys. Rev. **D59** (1999) 085009.

[23] G. Giudice, R. Rattazzi and J. Wells hep-ph/9811291; E. Mirabelli, M. Perelstein and M. Peskin hep-ph/9704448.

[24] R. Barbieri and A. Strumia, **hep-ph/9905281**.

[25] D.B. Kaplan and A.V. Manohar, Phys. Rev. Lett. **56** (1986) 2004.

Chapter 8

Supersymmetry Breaking from Extra Dimensions

8.1 Introduction

The four forces of nature are each characterized by a mass scale: $\sqrt{1/G_N} = M_P \approx 10^{19}$ GeV for gravity, $\Lambda_W \approx 10^3$ GeV for the weak interaction, $\Lambda_{QCD} \approx 0.1$ GeV for the strong interaction and $m_\gamma = 0$ for the electromagnetic interaction. What is the origin of these diverse scales? Over the last 25 years a single dominant viewpoint has developed: the largest scale, that of gravity, is fundamental, and the other scales are generated by a quantum effect in gauge theories known as dimensional transmutation. If the coupling strengths of the other forces have values $\alpha_P \approx 1/30$ at the fundamental scale, then a logarithmic evolution of these coupling strengths with energy leads, in non-Abelian theories, to the generation of a new mass scale

$$\Lambda \approx M_P \, e^{-1/\alpha_P} \tag{8.1.1}$$

where the interaction becomes non-perturbative. On the other hand, Abelian theories, like QED, remain perturbative to arbitrarily low scales. For strong and electromagnetic interactions this viewpoint is immediately successful; but for the weak interaction the success

is less clear, since the weak interactions are highly perturbative at the scale Λ_W. If Λ_W is generated by a dimensional transmutation, it must happen indirectly by some new force getting strong and triggering the breakdown of electroweak symmetry. There have been different ideas about how this might occur: the simplest idea is technicolor, a scaled up version of the strong force[1]; another possibility has the new strong force first triggering supersymmetry breaking which in turn triggers electroweak symmetry breaking[2]. For our purposes the crucial thing about these very different schemes is that they have a common mechanism underlying the origin of Λ_W: a dimensional transmutation, caused by the logarithmic energy evolution of a gauge coupling constant, generates the exponential hierarchy of (8.1.1).

In this letter, we propose an alternative mechanism for generating Λ_W exponentially smaller than the fundamental scale. Our scheme requires two essential ingredients beyond the standard model: supersymmetry, and compact extra dimensions of space. The known gauge interactions reside on a 3-brane, and physics of the surrounding bulk plays a crucial role in generating an exponentially small scale of supersymmetry breaking.

Our mechanism is based on the idea of "shining" [3]. A bulk scalar field, φ, of mass m, is coupled to a classical source, J, on a brane at location $y = 0$ in the bulk, thereby acquiring an exponential profile $\varphi \propto Je^{-m|y|}$ in all regions of the bulk distant from the source, $m|y| \gg 1$. If our brane is distant from the source, then this small exponential, arising from the propagation of the heavy scalar across the bulk, can provide an origin for very small dimensionless numbers on our brane, in particular for supersymmetry and

electroweak symmetry breaking

$$\Lambda_W \propto M_* \, e^{-mR} \tag{8.1.2}$$

where R is the distance scale of our brane from the source brane, and M_* is the fundamental scale of the theory. The possibility of such a supersymmetry-breaking mechanism has been noted before qualitatively [3]. If some of the extra dimensions are very large, M_* can be significantly below M_P, and could even be of order Λ_W, providing an alternative viewpoint on the mass scales of the four fources of nature [4]. We are concerned with the case of $M_* \gg \Lambda_W$, although M_* need not be as large as M_P. In this letter we give an explicit construction of shining which preserves 4-dimensional supersymmetry, but triggers an exponentially small amount of supersymmetry breaking due to the presence of our brane. A possible worry is that R might run to infinity, thus minimizing the vacuum energy and restoring supersymmetry. We exhibit simple mechanisms, based on the same supersymmetric shining, which stabilize the extra dimensions with finite radius.

8.2 Shining of Chiral Superfields

We begin by constructing a 5d theory, with a source brane shining an exponential profile for a bulk scalar, such that the equivalent 4d theory is exactly supersymmetric. The 5d theory possesses N=1 supersymmetry in a representation containing two scalar fields, φ and φ^c, together with a four-component spinor $\Psi = (\psi, \psi^c)$. The equivalent 4d theory has two families of chiral superfields $\Phi(y) = \varphi(y) + \theta\psi(y) + \theta^2 F(y)$ and $\Phi^c(y) = \varphi^c(y) + \theta\psi^c(y) + \theta^2 F^c(y)$. In the 4d theory, y can be viewed as a parameter labelling the families of chiral superfields.

Using this 4d chiral superfield notation, we write the bulk action as

$$S_B = \int d^4x\, dy \quad \Big(\int d^4\theta (\Phi^\dagger \Phi + \Phi^{c\dagger} \Phi^c) + \int d^2\theta \Phi^c (m + \partial_y) \Phi \Big) \tag{8.2.3}$$

Viewed as a 4d theory, we have manifest supersymmetry, with the y integral summing over the family of chiral superfields. The form of the superpotential appears somewhat unusual; however, on eliminating the auxiliary fields, the action in terms of component fields describes a free Dirac fermion and two complex scalar fields in 5d. The 5d Lorentz invariance is not manifest in (8.2.3), but this form is useful to us, since it makes the 4d supersymmetry manifest.

Next we locate a 3-brane at $y = 0$, and require that it provides a source, J, for a chiral superfield in a way which preserves 4d supersymmetry:

$$W_S = \int dy \delta(y)\, J \Phi^c, \tag{8.2.4}$$

where we choose units so that the fundamental scale of the theory $M_* = 1$. The conditions that this source shines scalar fields into the bulk such that supersymmetry is not spontaneously broken are

$$F(y) = (m - \partial_y)\varphi^c = 0 \tag{8.2.5}$$

$$F^c(y) = J\delta(y) + (m + \partial_y)\varphi = 0 \tag{8.2.6}$$

The first of these does not have any non-trivial solutions that do not blow up at infinity, or which are well-defined on a circle. The second, however, has the solution

$$\varphi(y) = -\theta(y) J e^{-my}, \tag{8.2.7}$$

in infinite flat space and

$$\varphi(y) = \frac{-Je^{-my}}{1-e^{-2\pi mR}} \quad y \in [0, 2\pi R), \tag{8.2.8}$$

on a circle. Thus we see that φ has taken on a non-zero profile in the bulk, but in a way that the energy of the system remains zero and one supersymmetry remains unbroken. Interestingly, this is not the profile that occurs with non-supersymmetric shining, but is asymmetric, shining in only one direction. One may have thought that the gradient energy for any profile of a bulk scalar field would neccessarily break supersymmetry, but our example shows this is not the case. The $|F^c|^2$ contribution to the vacuum energy includes the $|\partial_y\varphi|^2 + |m\varphi|^2$ terms as expected, but these are cancelled by $\varphi^*\partial_y\varphi$ terms, and at $y = 0$ by terms which arise because J is coupled to the combination $(m + \partial_y)\varphi(0)$. Note that if we had written a linear term for Φ instead of Φ^c, we would have shined a profile for φ^c in the opposite direction. Likewise, if we had chosen a negative value for m, we would shine φ in the opposite direction, since the 5d theory is invariant under $m \to -m, \;\; y \to -y$.

8.3 Supersymmetry breaking

Having learned how to shine a chiral superfield from a source brane across the bulk, we now investigate whether a probe brane, located far from the source at $y = \overline{y}$, can sample the small value of $\varphi(\overline{y})$ to break supersymmetry by an exponentially small amount on the probe brane. In addition to superfields which contain the standard model fields, the probe brane contains a standard model singlet chiral superfield X, and has a superpotential

$$W_P = \int dy\, \delta(y - \overline{y})(W_{MSSM} + \Phi X) \tag{8.3.9}$$

where W_{MSSM} is the superpotential of the minimal supersymmetric standard model. This superpotential has F-flatness conditions

$$F^c(y) = J\delta(y) + (m + \partial_y)\varphi = 0 \tag{8.3.10}$$

$$F(y) = \delta(y - \overline{y})x + (m - \partial_y)\varphi^c \tag{8.3.11}$$

$$F_X = \varphi(\overline{y}). \tag{8.3.12}$$

The first equation can only be satisfied by having a shined value for $\varphi(\overline{y}) \neq 0$. Clearly, the first and third equations cannot be simultaneously satisfied: we have an O'Raifeartaigh theory, and supersymmetry is spontaneously broken. As always in an O'Raifeartaigh theory, at tree level there is a flat direction: the value for x is undetermined, and if it is non-zero it acts as a source shining φ^c. It is simple to understand what is going on. In the presence of the source brane, the field φ is shined from the source brane, generating an exponentially small linear term for X on the probe brane. After we have integrated out the heavy fields φ and φ^c we are simply left with the superpotential on the probe brane

$$W_P \sim Je^{-m\overline{y}}X, \tag{8.3.13}$$

which generates a nonzero $F_X \sim Je^{-m\overline{y}}$.

This is not a precise equality, as the probe brane resists a non-zero $\varphi(\overline{y})$, and provides a back reaction on the bulk. It is simple to show that this effect is qualitatively insignificant.

If the fifth dimension is a circle, then we can imagine that the probe brane is stabilized at some location on the circle, or that it will drift such that it is immediately next to the source brane where the resulting supersymmetry breaking is smallest, as in

figure 8.1. In either case, we generate an exponentially small supersymmetry breaking scale F_X.

Notice that this is *not* in the same spirit as recent works that use bulk dynamics to transmit distantly broken supersymmetry[5]. Rather, in our case, in the absence of either source or probe brane, supersymmetry remains unbroken. It is the simultaneous presence of *both* branes that leads to the exponentially small supersymmetry breaking. A simple option for mediating the supersymmetry breaking from F_X to the standard model superpartners is to add non-renormalizable operators to the probe brane

$$\begin{aligned}\Delta S_P = \int d^4x dy \delta(y - \overline{y})(\int d^4\theta(\frac{1}{M_*^2} X^\dagger X Q^\dagger Q + ...) \\ + \int d^2\theta(\frac{1}{M_*} X W^\alpha W_\alpha + ...))\end{aligned} \tag{8.3.14}$$

where Q is a quark superfield and W^α a standard model gauge field strength superfield. We have inserted M_* explicitly, so that the soft masses of the standard model superpartners and x are $\widetilde{m} \sim F_X/M_* \sim (J/M_*)e^{-m\overline{y}}$. Until now we have not specified the values for J and m; the most natural values are $J \approx M_*^2$ and $m \approx M_*$.

Our entire theory is remarkably simple, and is specified by the bulk action S_B of (8.2.3), the source brane superpotential W_S of (8.2.4), and the interactions of (8.3.9) and (8.3.14) on our brane.

8.4 Radius Stabilization

Mechanisms for dynamical supersymmetry breaking by dimensional transmutation[6] typically suffer from the "dilaton runaway problem" when embedded in string theory[7]: since the coupling constant α_P is a dynamical field, the vacuum energy is minimized as

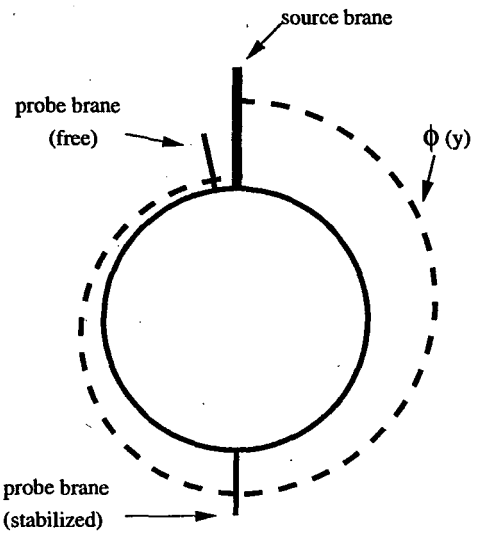

Figure 8.1: The schematic profile of φ in the extra dimension. Whether our brane is stabilized at some position or free to move under the given forces, we can achieve an exponentially small value for φ and hence exponentially suppressed supersymmetry breaking.

$\alpha_P \to 0$, where the theory becomes free. In our case, it appears there is an analogous problem. Taking the supersymmetry-breaking brane to be free to drift, the vacuum energy of the theory is

$$E \sim J^2 e^{-4\pi R m}, \tag{8.4.15}$$

so it is energetically favorable for the radius to grow to infinity. However, in contrast with dynamical supersymmetry breaking scenarios, where one must simply assume that the dilaton vev is somehow prevented from running to infinity, stabilizing R turns out to be quite simple.

Consider adding to the model of the previous section a second bulk multiplet (Φ', Φ'^c), of mass m', with interactions

$$W' = \int dy \, \left[\delta(y) J' \Phi'^c + \delta(y - \overline{y}) X' \left(\Phi' + A\right)\right] \tag{8.4.16}$$

where A and J' are constants and X' is a chiral superfield. The terms in this superpotential

are nearly identical to those of (8.3.9) and (8.2.4), except for the presence of the constant A on the probe brane. We assume that both A and J' are real. In complete analogy with the shining of φ, the scalar φ' acquires a profile

$$\varphi'(y) = -J'\theta(y)e^{-m'y}. \tag{8.4.17}$$

Writing $\overline{y} = \theta R$, the F-flatness condition for X' becomes

$$m'R\theta = \log \frac{J'}{A(1 - e^{-2\pi Rm'})}, \tag{8.4.18}$$

which defines a real function $R(\theta)$ provided that $J'/A > 0$. We assume m' is less than m (by a factor of roughly 30, for very large M_*), so that, for a given value of θ, the radius is essentially determined by the condition $F_{X'} = 0$, with a small correction $\frac{\Delta R}{R} \sim \frac{m}{m'}e^{-m/m'}$ coming from the $|F_X|^2$ contribution to the potential. However, we have already seen that the vacuum energy is minimized when the probe brane drifts completely around the circle. The value of R is thus immediately fixed by equation (8.4.18), with $\theta = 2\pi$. Its precise value depends on A and J', but if we take their ratio to be of order unity, then we find $2\pi Rm' \sim 1$. The supersymmetry breaking F-term is then $F_X \sim Je^{-2\pi mR} \sim Je^{-m/m'}$, so that the higher dimension interactions of (8.3.14) give superpartner masses

$$\widetilde{m} \sim e^{-m/m'}M_*. \tag{8.4.19}$$

In this model the mass of the radion, the field associated with fluctuations of the size of the circle, is $m_{radion} \sim F_X/M_P \sim 1$ TeV (M_*/M_P).

Alternatively one can stabilize R in an entirely supersymmetric fashion. Here we describe just one of a number of ways in which this can be done. Imagine supplementing

the "clockwise" shining of φ' due to W' with "counterclockwise" shining of a different scalar $\widetilde{\varphi}^c$ of comparable mass, $\widetilde{m}$, through the added superpotential terms

$$\widetilde{W} = \int dy \left[\delta(y)\widetilde{J}\widetilde{\Phi} + \delta(y-\overline{y})\widetilde{X}\left(\widetilde{\Phi}^c + B\right)\right]. \tag{8.4.20}$$

Note that because $\widetilde{\Phi}$ (rather than $\widetilde{\Phi}^c$) couples to the source, the shining is in the opposite direction as that of φ'. The F-flatness condition for $\widetilde{X}$,

$$\widetilde{m}R(2\pi - \theta) = \log \frac{B}{\widetilde{J}(1 - e^{-2\pi R\widetilde{m}})}, \tag{8.4.21}$$

and the F-flatness condition for X' independently determine R as a function of θ, and for broad ranges of parameters the combined constraints are satisfied by unique values of θ and R. This supersymmetric stabilization of the radius yields $m_{radion} \sim M_*^2/M_P$, far above the TeV scale.

8.5 Gauge-Mediated Supersymmetry Breaking

We have presented a complete model in which exponentially small supersymmetry breaking is generated as a bulk effect and communicated to the standard model via higher-dimension operators. It is straightforward to modify the model so that the supersymmetry breaking is mediated instead by gauge interactions[8].

Consider the O'Raifeartaigh superpotential

$$W = X(Y^2 - \mu^2) + mZY. \tag{8.5.22}$$

At tree level x is a flat direction, but provided $\mu^2 < m^2/2$, radiative effects stabilize x at the origin and give $m_x^2 \sim \mu^2/16\pi^2$. Supersymmetry is broken by $F_X = -\mu^2$. Models using an O'Raifeartaigh superpotential to achieve low-energy supersymmetry breaking have been

constructed in the past, but have required a small value for μ^2 to be input by hand. Instead, we use supersymmetric shining as an origin for the parameters μ^2 and m by coupling the brane superfields X, Y, and Z to the shone Φ according to

$$W_{hidden} = \lambda_1 X(Y^2 - \Phi(\overline{y})^2) + \lambda_2 \Phi(\overline{y}) ZY, \tag{8.5.23}$$

where λ_1 and λ_2 are both of order unity and $\lambda_1 < \lambda_2^2/2$. Next we introduce couplings to messenger fields Q and $\overline{Q}$ transforming under the standard model gauge group* ,

$$W_{messenger} = \alpha_1 XQ\overline{Q} + \alpha_2 \Phi(\overline{y}) Q\overline{Q}. \tag{8.5.24}$$

By taking $\alpha_2^2 > \alpha_1 \lambda_1$ we ensure that the messenger scalars do not acquire vevs. These superpotentials give Q and $\overline{Q}$ supersymmetric masses and supersymmetry-breaking mass splittings of comparable order, $M \sim \sqrt{F} \sim \varphi(\overline{y})$. The messengers then feed the supersymmetry breaking into the standard model in the usual way, yielding soft supersymmetry-breaking parameters of order $\widetilde{m} \sim \frac{1}{16\pi^2}\varphi(\overline{y})$. Fixing the radius R by either of the mechanisms already described then leads to $\widetilde{m} \sim \frac{M_*}{16\pi^2} e^{-m/m'}$. Note that this is truly a model of low-energy supersymmetry breaking, with $\sqrt{F} \sim 16\pi^2 \widetilde{m} \sim 100$ TeV, allowing for decays of the NLSP within a detector length. Moreover, this small value for $\sqrt{F}$ is favored by cosmology in that it suppresses the gravitino energy density[9].

While there is typically a severe μ problem in gauge-mediated theories [10], it is easily solved with our mechanism by shining μ in the superpotential with a term

$$W \supset \lambda \varphi(\overline{y}) H_1 H_2. \tag{8.5.25}$$

*The superpotentials of (8.5.23) and (8.5.24) are not justified by symmetries. However, it is not difficult to modify things, for instance by shining both Φ *and* Φ^c, in such a way that symmetries select superpotentials that give the same essential results.

With $\lambda \sim 1/30$, problems of naturalness are much less severe than in theories where supersymmetry is broken dynamically. If $B\mu = 0$ at tree level, radiative effects can generate a small $B\mu$ and large $\tan\beta$ [11]. Likewise, in gravity mediated theories, a shined term $\int d^2\theta\Phi(\overline{y})H_1H_2$ can also generate an appropriate value for μ, while $\int d^4X^\dagger XH_1H_2$ generates $B\mu$. Although φ is related to supersymmetry breaking, this is distinct from the Giudice-Masiero mechanism. Absent the superfield X, supersymmetry is preserved, but the value of μ is unchanged.

Depending on whether supersymmetric or supersymmetry breaking stabilization of the radius is employed, the radion mass is either $m_{radion} \sim M_*^2/M_P$ or $m_{radion} \sim \sqrt{F}/M_P \sim$ 1 eV (M_*/M_P). Even the latter case is safe, since the limit on the radion mass is on the mm^{-1} scale, at the limits of experimental probes of gravity at short distances.

8.6 Conclusions

Dimensional transmutation, (8.1.1), and shining, (8.1.2), are alternative mechanisms for taking a dimensionless input of order 30 and generating an exponentially small mass hierarchy. These mass hierarchies can explain the scales of symmetry breaking, for instance of a global flavor symmetry, or of supersymmetry, as we have discussed. While dimensional transmutation is a quantum effect requiring an initial coupling which is highly perturbative, $1/\alpha_P \approx 30$, shining is classical and requires a bulk distance scale of size $R \approx 30M_*^{-1}$. Such a radius can in turn be stabilized in a simple way. We presented two standard ways of communicating this exponentially small supersymmetry breaking, through higher-dimensional operators or via standard model gauge interactions. It is clearly possible to employ other mechanisms, such as those discussed in [5]. Our theories are remarkably

simple, using only free classical dynamics in one extra dimension. Extensions to more dimensions should be straightforward. While we have concentrated on constructing effective theories with exponentially small global supersymmetry breaking, it will be interesting to embed these models in a consistent local supergravity. It will also be interesting to explore whether any of these mechanisms can be realized in the D-brane construction of non-BPS states in string theory.

Bibliography

[1] S. Weinberg, *Phys. Rev.* **D19** (1979) 1277; L. Susskind, *Phys. Rev.* **D20** (1979) 2619.

[2] E. Witten, *Nucl. Phys.* **B188** (1981) 513.

[3] N. Arkani-Hamed and S. Dimopoulos, hep-ph/9811353.

[4] N. Arkani-Hamed, S. Dimopoulos and G. Dvali, *Phys. Lett.* **B429** 263 (1998), hep-ph/9803315; *Phys. Rev.* **D59** 086004 (1999), hep-ph/9807344; I. Antoniadis, N. Arkani-Hamed, S. Dimopoulos and G. Dvali, *Phys. Lett.* **B436** 257 (1998), hep-ph/9804398; N. Arkani-Hamed, S. Dimopoulos and J. March-Russell, hep-th/9809124.

[5] L. Randall, R. Sundrum, hep-th/9810155; E.A. Mirabelli, M.E. Peskin, *Phys. Rev.* **D58** (1998) 065002, hep-th/9712214; D.E. Kaplan, G.D. Kribs, M. Schmaltz, hep-ph/9911293; Z. Chacko, M.A. Luty, A.E. Nelson, E. Ponton, hep-ph/9911323.

[6] I. Affleck, M. Dine and N. Seiberg, *Nucl. Phys.* **B256** (1985) 557; for a review see: E. Poppitz and S. Trivedi, hep-th/9803107.

[7] M. Dine and N. Seiberg, *Phys. Lett* **B162** (1985) 299.

[8] L. Alvarez-Gaume, M. Claudson, M.B. Wise, *Nucl. Phys.* **B207** (1982) 96; M. Dine, A.E. Nelson Y. Nir and Y. Shirman, *Phys. Rev.* **D53** (1996) 2658.

[9] A. de Gouvea, T. Moroi and H. Murayama, *Phys. Rev.* **D56** (1997) 1281.

[10] G. Dvali, G.F. Giudice, A. Pomarol, *Nucl. Phys.* **B478** (1996) 31-45, hep-ph/9603238.

[11] R. Rattazzi, U. Sarid, *Nucl. Phys.* **B501** (1997) 297, hep-ph/9612464.

Chapter 9

Conclusion

With the Large Hadron Collider at CERN scheduled to come on line in 2006 or 2007, we have hope of finally studying fully the energy regime we have known since Fermi first wrote down his theory of weak interactions. Likewise, BaBar and Bell will give us tremendous data into the nature of CP violation, while SNO, KAMLand, Minos, K2K and other will give us a plethora of data on neutrino masses.

We have seen that we already have tremendous information on neutrino masses and CP violation. We have great hopes for signals from supersymmetry, or the presence of extra dimensions, and already have made great strides in preparing for the flood of information that these experiments should yield.

However, given these future data, it is imperative that we do not simple bide our time waiting for the results. Given the many possibilities that we have already developed, the particular signals, even if within one of the frameworks already thought of, could be something we have not considered. We have the hope of understanding better the nature of matter, of pushing our understanding of the history of the universe back farther, and of answering innumerable questions posed over the last seventy years. In doing so, hopefully,

and in all likelihood, we will generate volumes more.

Appendix A

The Stability Index

A.1 "Formalism" of the Stability Index

It is difficult to establish a formalism for the stability index, because it involves an inherently ill-defined quantity, namely, what constitutes an order one quantity. However, the potential instability of various predictions to variations in these order one parameters makes some attempt to quantify this necessary. Such a quantification should be relatively insensitive to what precisely constitutes an "order one quantity".

Therefore, we demand the following quantities of the index:

• An "order one" quantity should be defined as a quantity x with some probability distribution $P(x)$ to occur in an interval about 1. For reasons that will become clear later, it will be useful to consider instead the quantity $\overline{P}(y)$, where $x = 10^y$.

- This distribution should be "sensible", namely

  1. $P(x)$ should be an even function in $Log(x)$, that is, $\overline{P}(y)$ is even in y.

  2. $\overline{P}(y)$ should achieve its maximum value at 0.

3. $\overline{P}(y)$ should have a spread characterized by its variance,

$$v^2 = \int_{-\infty}^{\infty} y^2 \overline{P}(y)$$

the variance then quantifying what "order one" is numerically.

4. A product of two sensible distributions, correlated or uncorrelated, should be sensible.

- The index should have similar implications regardless of $\overline{P}(y)$, so long as it is sensible.
- The definition of $\overline{P}(y)$ should be the only necessary input.

We shall explore the motivation for these assumptions and will shortly see that the presented index nearly meets the requirements, and with minor modifications can meet them enitrely.

We assume that the expectation value of x, and of any products of x, is unity. It follows immediately that $\overline{P}(y)$ should be even in y. We do not have strong arguments in favor of this assumption, and if it were relaxed, the formalism could be suitably modified.

For instance, consider the seemingly sensible distribution

$$P(x) = \begin{cases} \frac{3}{8}, & \text{if } \frac{1}{3} \leq x \leq 3; \\ 0, & \text{otherwise.} \end{cases}$$

which has been normalized to give total probability 1. The expectation value of a product of n uncorrelated variables with such a distribution would be

$$< \overline{X} >= \int d^n x \prod_i P(x_i) x_i = (\frac{10}{3})^n. \qquad \text{(A.1.1)}$$

Such a numerical pile-up of the central value of a product of order unity coefficients is excluded by our assumption.

What constitutes a "sensible" distribution is, of course, a judgement call. Examples of what we consider sensible distributions would be

- Flat distributions taking on the value $1/a$ from $-a/2$ to $a/2$
- Exponential distributions with standard deviation σ
- Linearly decreasing distributions of the form

$$\overline{P}(y) = \begin{cases} (\frac{1}{ab})(-\frac{b}{a}|y| + b), & \text{if } -a \leq y \leq a; \\ 0, & \text{otherwise.} \end{cases}$$

In fact, it can be shown that the last case it just the product of two uncorrelated quantities of the first type.

In all of these cases, the next moment (x^4) is irrelevant in quantifying the likelihood of the variable being within a particular region about zero. Requirement 3 is then simply a statement that a sensible distribution should simply have one quantity, its variance, to determine how confident we are that the variable is within that region. This will then allow us to be more confident in deducing the significance of the variance of some product.

This being stated, we can actually go about constructing some approximation of confidence intervals. The ability to describe the distribution of one variable by its variance is useful in allowing us to calculate the variances for higher products. We begin by writing the formal expression for the probability distribution of n uncorrelated variables $x_i = 10^{y_i}$ with probability distributions $\prod_i \overline{P}(y_i)$. We have

$$\overline{P}(z) = \int d^n y (\prod_i \overline{P}_i(y_i)) \delta(z - \sum_i y_i) \tag{A.1.2}$$

This expression is tedious to calculate for given $\overline{P}(y)$, particularly for large n.

However, its variance is a relatively simply calculation.

$$v_z^2 = \int dz \overline{P}(z) z^2 = \int dz\, d^n y (\sum_i y_i)^2 (\prod_i \overline{P}_i(y_i)) \delta(z - \sum_i y_i) \tag{A.1.3}$$

Expanding the squared term we find terms

$$\int dz\, d^n y\, y_i\, y_j (\prod_i \overline{P}_i(y_i)) \delta(z - \sum_i y_i) = \begin{cases} 0, & \text{for } i \neq j; \\ v_i^2, & \text{if } i = j, \end{cases} \tag{A.1.4}$$

giving

$$v_{z\,uncorr}^2 = \sum_i v_i^2 \tag{A.1.5}$$

For n correlated variables, a similar calculations yields

$$v_{z\,corr}^2 = n^2 v_0^2 \tag{A.1.6}$$

where v_0^2 is the variance of the original variable.

Thus, a product of n correlated order one quantities is far more unstable than a product of n uncorrelated order one quantities. Simply counting the total number of order one coefficients is not sufficient. Thus we will refer to a product of the form

$$\prod_i x_i^{n_i} \tag{A.1.7}$$

as having index $(\sum_i n_i)$ of type $(n_1, n_2, ..., n_m)$. If some of the n_i are repeated, we use the shorthand of writing n^j, if n is repeated j times. We assume all order one quantities have the same distribution. A product of type $(n_1, n_2, ..., n_m)$, has variance

$$v_{(n_i)}^2 = v_0^2 \sum_i n_i^2 \tag{A.1.8}$$

This works extremely well for products of order one coefficients. However, a sum of order one coefficients is not necessarily order one. In these cases, it is usually best to perform a Monte Carlo to determine the distribution.

A.2 Sensible distributions

To characterize the probability of a general product to be within a certain region about 1, it is necessary to explore the particular forms of various distributions. We consider three reasonable distributions to be i) the flat distribution, ii) the Gaussian distribution, and iii) the linearly decreasing distribution.

A product of two equal width flat distributions yields a linear distribution, so we need only consider the flat and Gaussian cases. Gaussian distributions are well understood: products of variables with Gaussian $\overline{P}(y)$ functions are again Gaussian, allowing standard statistical techniques to be applied.

Products of flat distributions very quickly become characterized by Gaussian distributions. We have performed explicit Monte Carlos for $n = 1, 2, 3, 5, 7, 9$ uncorrelated variables. Even by $n = 2$ the Gaussian approximation is good, and for $n \geq 3$ it is very good. We thus believe it is reasonable to simply use Gaussian distributions, making a statistical interpretation of the variances simple.

For a standard, we propose using a distribution with variance $v = \sqrt{\frac{1}{12}}$, which corresponds to the variance of a flat distribution for $-\frac{1}{2} \leq y \leq \frac{1}{2}$. Changing the width of such a distribution from 1 to a would amount to multiplying this variance by a. Such generally mild sensitivity of the index to variations in the initial distribution is one its desirable qualities. We can then take "1-v" and "2-v" regions with $|y| \leq v$ and $|y| \leq 2v$,

respectively. As should be clear, these should not be interpreted as the precise 67% and 95% 1 and 2-σ regions, because σ is not precisely defined. They are simply regions of medium and strong confidence, respectively.

As an example, consider a prediction with an unknown coefficient of order one quantities of the form $x_1^2 x_2 x_3$. We say this has index $2+1+1=4$ of type $(2,1,1)$, which we will write in shorthand as $(2,1^2)$. Assuming the standard variance given above, this coefficient has variance $v = \sqrt{\frac{2^2+1^2+1^2}{12}} = \sqrt{\frac{1}{2}}$. Thus, we can have medium confidence that the prediction for x is known within a factor of $10^v = 5$, and strong confidence the the prediction is within a factor of $10^{2v} = 25$.

We can also see that this reduces to the expected prediction in the case of a variable of index 1 of type (1). It will have variance $v = \sqrt{\frac{1}{12}}$ which gives medium confidence that the prediction is known within a factor of 1.9, and strong confidence it is known within a factor of 3.8. This is a good consistency check that the index predicts what we would expect in the case of a single order one coefficient.

A.3 Reassessing the uncertainties in the $U(2)$ neutrino model

In lieu of the preceding analysis, we address the index type of the predictions already presented, and thus assess strong and medium confidence regions of each prediciton. We list all uncertainties for the general theory in table A.1.

In the general theory with the S-field, the atmospheric mixing angle is completely stable, while the solar angle is of approximate type $(1^4) = (1,1,1,1)$. However, it involves a sum of order one coefficients, motivating the use of Monte Carlos. Since a sum is involved, the relative sign of the order one quantities becomes relevant, and we list those

quantity	type (*=approx type)	Sign convention	Range of Med Confidence	Range of Strong Confidence
θ_{atm}	0	n/a	$\frac{\pi}{4}$ exact	exact
$\theta_{\odot}$	4*	+	$(0.002, 0.03)$	$(0.0006, 0.1)$
$\theta_{\odot}$	4*	-	$(0.005, 0.012)$	$(0.0001, 0.06)$
m_{ν_L}	$(2, 1^2)$	none	$(2, 50)$eV	$(0.4, 250)$eV
m_{N_L}	$(4, 3, 2^2, 1)$	none	(17MeV, 40GeV)	(300keV, 1.9TeV)
m_{N_H}	$(4, 2^4)$	none	(470MeV, 870GeV)	(11MeV, 37TeV)

Table A.1: General Theory: uncertainties in predicitions. The regions listed here are simply for the uncertainty due to order one coefficients. Additional error due to uncertainty in input quantities, in particular in m_ν, can also be significant.

cases seperately. These Monte Carlos allow us to claim that we have medium confidence that $\theta_{\odot}$lies within $(0.002, 0.03)$ and strong confidence that it lies within $(0.0001, 0.1)$, giving large overlap of the BP98 region.

The mass of the pseudo-Dirac neutrino has stability index 5 of type $(2, 1^3)$, giving a medium confidence to know this within a factor of 5, and strong confidence within a factor of 25. Given the uncertainty in δm^2_{atm} and $\delta m^2_{\odot}$, which determine the prediction, m_ν could conceivably be as low as 0.1eV.

The masses of the right-handed states are not known so well. The mass prediction is, for the heavier state, of type $(4, 2^4)$, and, for the lighter state, of type $(4, 3, 2^2, 1)$. This would give medium confidence to know the masses at factors of 43 and 48, and strong confidence at factors of 1800 and 2300, respectively. The cosmological implications of these neutrinos are very uncertain, given that the lighter could be well over a TeV in mass.

Without the S field, certain uncertainties change. The precise nature of the changes

depends on which splitting operators are included and what sign convention is taken. Because of the large number of permutations, we list only the basic results. The atmospheric angle is, as expected, completely certain. The solar angle becomes slightly more uncertain, but still overlaps BP98 well. The heaviest two righthanded masses typically become less certain by a factor of roughly 100, but the uncertainty is so large that the phenomenological predictions remain the same. The only dramatic difference in the variant theory is that ν_μ has a medium confidence region on its mass of $(24\text{eV}, 4\text{keV})$, and a strong confidence region of $(2\text{eV}, 48\text{keV})$. Including the uncertainties in the input quantities, the mass could be as low as $0.4eV$, which escapes the KARMEN bound, although narrowly.

Ernest Orlando Lawrence Berkeley National Laboratory
One Cyclotron Road | Berkeley, California 94720

Prepared for the U.S. Department of Energy under Contract No. DE-AC03-76SF00098