

Fermilab Library  
QC721 .W89 1983 010101 000  
Workshop on Grand Unified  
Proceedings. Edited by K. Odak



0 1159 0005383 2



FERMILAB

OCT 8 1984

KEK Report 84-12  
September 1984  
H

LIBRARY

PROCEEDINGS OF  
WORKSHOP ON GRAND UNIFIED THEORIES  
AND COSMOLOGY

KEK, Tsukuba, Japan  
December, 7-10, 1983

Edited by

K. ODAKA and A. SUGAMOTO

NATIONAL LABORATORY FOR  
HIGH ENERGY PHYSICS

FERMI  
QC721  
.W89  
1983



KUNIO NAGATANI

(23.1.1936 - 18.3.1984<sup>†</sup>)

FERMILAB  
LIBRARY

Proceedings of  
Workshop on Grand Unified Theories  
and Cosmology  
KEK, Tsukuba, Japan  
December, 7-10, 1983

Edited by  
K. Odaka and A. Sugamoto

## Organizers of the Workshop

J. Arafune (Cosmic Ray Research, University of Tokyo)  
M. Fukugita (Research Institute of Fundamental Physics, Kyoto University)  
H. Sato (Research Institute of Fundamental Physics, Kyoto University)  
H. Sugawara (National Laboratory for High Energy Physics)  
M. Yoshimura (National Laboratory for High Energy Physics)

Foreword	
Proton Decay in Grand Unified Theories	
Z. Hioki .....	1
32 Kton Water Cerenkov Detector (JACK)	
M. Koshihara .....	24
Nucleon Decay Experiment at Kolar Gold Field	
S. Miyake .....	32
Astrophysical Constraints on the Monopole Abundance	
M. Fukugita .....	38
Issues in Monopole-Catalyzed Proton Decay	
Y. Kazama .....	61
Present Status of Monopole Search	
F. Kajino .....	78
Underground Searches for Anomalous Penetrating Particles	
S. Orito .....	88
A Plan of Monopole Search Experiment Using 100m <sup>2</sup> Calorimeter of Akeno Air Shower Array	
T. Hara .....	95
Present Status of Monopole-Search with Superconducting Induction Coils	
T. Watanabe .....	110
Recent Observation of the Galactic Center	
J. Inatani .....	122
Infrared Observation of the Early Universe	
T. Matsumoto .....	131
Current Status of Missing Mass Problem	
F. Takahara .....	148
Some Topics in Cosmic Rays	
J. Arafune .....	169

Observation of Ultra High Energy Cosmic Rays ( $\geq 10^{10}$ GeV)		Quantum Effects in Generalized Kaluza-Klein Theories	
M. Nagano .....	177	O. Yasuda .....	418
Point Source of Very High Energy Gamma Ray		Stability of Solutions in Kaluza-Klein Theory	
T. Kifune .....	201	M. Yamasaki .....	428
A. Brief Review on Neutrino Mass Measurements		Fermion in a Modified Kaluza-Klein Model	
T. Yasuda and T. Ohshima .....	211	T. Muta .....	444
Measurement of the Mass of the Electron Neutrino Using Electron Capture in $^{163}\text{Ho}$		Composite Quasi Nambu-Goldstone Fermions	
S. Yasumi .....	240	T. Yanagida .....	452
Present Status of the Solar Neutrino Problem			
N. Itoh .....	250		
Nuclear Physics in the Solar Neutrino Problem			
K. Nagatani .....	251		
Scenarios of Galaxy Formation			
H. Sato .....	275		
Galaxy Formation Theory and Large Scale Structure			
S. Ikeuchi .....	287		
Inflationary Universe Model			
K. Sato .....	303		
Problems in the Perturbation Analysis of a Universe Dominated by a Coherent Scalar Field			
M. Sasaki .....	349		
Comments on the Chaotic Inflation			
H. Kodama .....	359		
Proton Decay in Supersymmetric Models			
N. Sakai .....	376		
Mass of Superpartners in SGUT			
H. Komatsu .....	381		
On Kaluza-Klein Theories			
Y. Fujii .....	388		

#### FOREWORD

The workshop on Grand Unified Theories and Cosmology was held at KEK, on December, 7-10, 1983. This workshop attracted about 100 participants, who covered a wide range of research fields, and as a result stimulated many fruitful discussions. We asked the speakers of the workshop to submit typewritten manuscripts of their talk. Those manuscripts are presented in these proceedings. We wish to thank these who contributed to this volume. Shortly after this workshop was finished, one of the speakers Professor Kunio Nagatani died on March 18, 1984 at the age of 48. We sincerely regret his untimely death. However, with the help of Masataka Fukugita, we could include in this volume his picture and the copies of his transparencies presented at this workshop. The workshop was partially supported by Grant in Aid for Scientific Research and also by Physics Department of KEK.

#### Editors

K. Odaka

A. Sugamoto

## Proton Decay in Grand Unified Theories

Zenrō HIOKI

Department of Physics, Kyoto University, Kyoto 606, JAPAN

### Abstract

A review on the recent status of proton decay studies in the SU(5) and the SO(10) grand unified theories is presented.

- § 1. Introduction
- § 2. SU(5) GUT and low energy parameters
- § 3. Proton decay I : SU(5) GUT
- § 4. Proton decay II : SO(10) GUT
- § 5. Summary

### § 1. Introduction

At the present, we have quite successful theories of the weak, electromagnetic and strong interactions, i.e., the standard SU(2) × U(1) electroweak theory<sup>1)</sup> and the Quantum Chromodynamics (QCD) based on the color SU(3) symmetry.<sup>2)</sup> In fact, various phenomena at energies attainable by present experiments are consistently described by these theories.

However, we have to recognize that these theories include some unsatisfactory points. For example, the Weinberg angle,  $\theta_W$ , in the electroweak theory is a free parameter since there is no theoretical relation between the SU(2) and the U(1) coupling constants,  $g$  and  $g'$  ( $\tan \theta_W \equiv g'/g$ ). We also do not know the reason for the charge quantization. These are expected to be solved in a more fundamental theory.

One of the most attractive approaches to such a theory is the grand unification, which aims at describing the electroweak and the strong interactions by a gauge theory based on a larger (simple) group including SU(3) × SU(2) × U(1).

In this scenario, we have two predictions which can be readily checked by low energy data: the values of  $\sin^2 \theta_W$  and  $m_b$  (the b-quark mass). However, what is more striking is the prediction that proton (and bound neutron) is not stable!

Among various theories (models)<sup>3)</sup> proposed so far, the one based on SU(5) is the simplest<sup>4)</sup>, and its predictions for  $\sin^2 \theta_W$ <sup>5)-9)</sup> and  $m_b$ <sup>6),10)</sup> are in good agreement with experimental

data. Therefore, calculations of the proton lifetime,  $\tau_p$ , have also been made mainly in this SU(5) grand unified theory (GUT).<sup>7),9),11)</sup>

Recently, however, experimental data on  $\tau_p$  which seem to contradict this theory have been reported (IMB collaboration).<sup>12)</sup> Experiment by another collaboration, KAMIOKANDE, has also started data taking.<sup>13)</sup> We are now in a very exciting situation.

In this article, I will review the recent status of the proton decay studies in the SU(5) GUT, and also the SO(10) GUT,<sup>14)</sup> which is another strong candidate for the realistic model. (Proton decay in supersymmetric GUTs is also interesting but I do not mention these theories. See, e.g., Ref.15.) In §2, taking the SU(5) GUT as an example, I explain how to determine the values of parameters of the theory, and show its successful predictions on  $\sin^2 \theta_W$  and  $m_b$ . Then, the proton decay in the SU(5) theory is discussed in §3. There I will also discuss possible ways<sup>16) -21)</sup> to rescue this theory from the contradiction with the data. The SO(10) GUT is described in §4, where we see that this theory allows the proton lifetime longer than that of the SU(5) GUT. The final section is devoted to summary.

## § 2. SU(5) GUT and low energy parameters

A model based on SU(5) symmetry, the prototype of grand unified theories, was first proposed by Georgi and Glashow.<sup>4)</sup>

In this model, quarks and leptons of one generation are assigned to  $\mathbf{5}^*$  and  $\mathbf{10}$  representations of SU(5). Under such assignment, this model satisfies the anomaly-free condition. We have 24 gauge bosons. Among them, 12 correspond to the well-known light bosons, and another 12 are superheavy ones,  $X_i^\pm, Y_i^\pm$  ( $i = 1 \sim 3$ : color index,  $m_X = m_Y$ ).

In addition, at least two Higgs multiplets are necessary in order to break SU(5) spontaneously to SU(3)  $\times$  SU(2)  $\times$  U(1) and further to SU(3)  $\times$  U(1)<sub>EM</sub>. Usually, **24** and **5** Higgs multiplets are taken for the former and the latter breakings respectively.

This scheme is the most compact one, and is called "minimal SU(5) theory (model)". For the time being, we restrict our discussions to this minimal SU(5) GUT. (Of course, we can incorporate a different set of Higgs multiplets, but this increases the number of free parameters and the predictive power of the theory is diminished.)

Let us first consider the Weinberg angle.  $\sin^2 \theta_W$  is no longer a free parameter in the GUTs. The gauge principle demands only one coupling constant, and the U(1) generator is identified with a combination of the SU(5) generator  $T^a$  ( $a = 1 \sim 24$ ). Therefore, the structure of the weak neutral current, and consequently the value of  $\sin^2 \theta_W$  is determined only by the group structure. As is well-known,  $\sin^2 \theta_W = 3/8$  in the SU(5) theory.<sup>4)</sup>

Accidentally, early experimental data were  $\sin^2 \theta_W^{\text{exp}} \approx 0.37 - 0.38$  ( $\approx 3/8$ ), but the above prediction is inconsistent with present more refined data,<sup>22)</sup>  $\sin^2 \theta_W^{\text{exp}} \approx 0.21 - 0.22$ . This seeming disagreement vanishes if we consider the higher order effects as was first pointed out by Georgi, Quinn and Weinberg.<sup>5)</sup>

Generally, a large radiative correction like  $\sim g^2 \ln(M/m)$  appears if two extremely different mass scales  $M$  and  $m$  are included in a calculation. The data on  $\sin^2 \theta_W$  are obtained in low energy (at most  $m \sim M_{W,Z}$ ) experiments on the neutral current interactions. On the other hand, the value  $\sin^2 \theta_W = 3/8$  is considered to be the one obtained from the neutral current renormalized at  $M \sim M_X$  (grand unification scale).  $M_X$  must be very large ( $\gtrsim 10^{14-15}$  GeV) so that proton does not decay too fast. Therefore, the higher order effects cannot be neglected in order to calculate  $\theta_W(M_W)$ , the Weinberg angle which can be directly compared with low energy data.

These effects are expressed in terms of the running coupling constants governed by the following equations.

$$\mu \frac{dg_i}{d\mu} = \beta_i(g_i) \quad , \quad (2.1)$$

where  $i = 1, 2$  and  $3$  correspond to  $U(1)$ ,  $SU(2)$  and  $SU(3)$  respectively. We know the value of  $g_3(\mu)$  (QCD coupling) at low energy ( $\mu \lesssim M_W$ )<sup>23)</sup>. Further,  $g_2(M_W)$  and  $g_1(M_W)$  are expressed by  $\alpha(M_W)$  (QED coupling) and  $\sin \theta_W(M_W)$ , and the value of  $\alpha(M_W)$

is calculable from  $\alpha(0) = 1/137.036$ .

Using these values and the well-known  $SU(3)$ ,  $SU(2)$  and  $U(1)$   $\beta$  functions, and assuming that  $SU(5)$  symmetry is restored at  $\mu = M_X$ , i.e.,  $g_i(M_X) = g_{SU(5)}$ , Georgi et al.<sup>5)</sup> obtained the values of  $\sin^2 \theta_W(M_W)$  and  $M_X$  as functions of  $\alpha_{\text{QCD}}$  and  $\alpha$ . Subsequently, Buras et al.<sup>6)</sup> made more detailed computations of  $\sin^2 \theta_W(M_W)$  and also the value of  $m_D/m_T$  (fermion mass terms made by 5 Higgs satisfy the relation  $m_D = m_T$  at tree level).

Strictly speaking, however, we must not neglect the mass effects of  $W^\pm$ ,  $Z$  and  $X^\pm$ ,  $Y^\pm$  bosons in  $\beta$  functions even at  $\mu < M_{W,Z}$  and  $\mu < M_{X,Y}$  respectively, and  $SU(5)$  symmetry must not be restored at  $\mu = M_X$  (i.e., the values of  $g_i(M_X)$  do not coincide with each other). In the above calculations, such points are not taken into account. Therefore, in order to determine the values of parameters more accurately, we have to use the  $\beta$  functions including these effects (threshold effects).

First systematic study of those effects was done by Ross,<sup>24)</sup> who showed that the value of  $M_X$  thereby decreases by a factor of 6. Subsequently, several authors<sup>7) - 10)</sup> have made computations of parameters, and we have now the following results (at the level of two-loop renormalization group equations) for  $\sin^2 \theta_W(M_W)$ ,  $m_D$  and  $M_X$ . (I have shown the values referred in the review by Marciano<sup>25)</sup>.)



Input data

$$\begin{aligned} \text{QED coupling : } \alpha(0) &= 1/137.036 , \\ \text{QCD coupling : } \Lambda_{\overline{\text{MS}}} &= 0.16_{-0.08}^{+0.10} \text{ GeV}^{23} , \end{aligned} \quad (2.2)$$

$$\left( \alpha_{\text{QCD}}(\mu) = \frac{1.51}{\ln(\mu^2/\Lambda_{\overline{\text{MS}}}^2)} \left\{ 1 - 0.74 \frac{\ln(\ln(\mu^2/\Lambda_{\overline{\text{MS}}}^2))}{\ln(\mu^2/\Lambda_{\overline{\text{MS}}}^2)} \right\} \right) ,$$

Output

$$\begin{aligned} \sin^2 \theta_W(M_W) &= 0.214_{-0.003}^{+0.004} , \\ m_b/m_\tau &= 2.9 \pm 0.2 , \\ M_X &= (2.1_{-1.2}^{+1.7}) \times 10^{14} \text{ GeV} , \end{aligned} \quad (2.3)$$

( If we have a fourth generation,  $m_b/m_\tau$  ratio increases by  $\approx 0.25$  and disagrees with the data . )

Experimental data<sup>26)</sup>

$$\begin{aligned} \sin^2 \theta_W^{\text{exp}}(M_W) &= 0.215 \pm 0.014 , \\ ( m_b/m_\tau )^{\text{exp}} &= 2.6 - 2.9 . \end{aligned} \quad (2.4)$$

The agreement of  $\sin^2 \theta_W$  with its data is excellent, and a severe constraint on the number of fermion generations obtained from the value of  $m_b/m_\tau$  is noteworthy.

### § 3. Proton decay I: SU(5) GUT

In order to evaluate the proton lifetime  $\tau_p$ , there are further two necessary steps:

- i) calculations of enhancement factor (i.e., inclusion of radiative corrections to the tree amplitude), and
- ii) reasonable estimate of the hadronic matrix element.

The enhancement factor is usually investigated in the framework of operator analysis combined with the renormalization group method.<sup>27)</sup> These studies were done by several authors.<sup>6),28)-30)</sup> The value of this factor is about 2.9 in the amplitude.

As for the estimate of the matrix element, we have various standard techniques developed in the studies of ordinary hadron decays: the bag model, the relativistic composite model, ... . By using these tools, the matrix element (and the proton lifetime) has been computed by many authors.<sup>6),7),11)</sup>

The average values of  $\tau_p$  and partial life  $\tau_p^{e^+\pi^0}$  for the mode  $p \rightarrow e^+\pi^0$  (which is considered as the main decay mode in the SU(5) GUT with branching ratio  $\sim 40\%$ ) are

$$\begin{aligned} \tau_p &\approx 2 \times 10^{29 \pm 1.7} \text{ yr} , \\ \tau_p^{e^+\pi^0} &\approx 4.5 \times 10^{29 \pm 1.7} \text{ yr} . \end{aligned} \quad (3.1)$$

On the other hand, the recent IMB experiment<sup>12)</sup> has given

$$\tau_p^{e^+\pi^0} > 1 \times 10^{32} \text{ yr} , \quad (3.2)$$

which is much longer than the above SU(5) predictions. (I was informed during this Workshop that the lower bound (3.2) changed as  $\tau_p^{e^+\pi^0} > 1.65 \times 10^{32} \text{ yr.}$ )<sup>31)</sup>

Since, there is a proportionality  $\tau_p \propto M_X^4 \propto \Lambda_{\overline{MS}}^4$ , consistency seems to recover somehow if  $\Lambda_{\overline{MS}} > 0.5 - 0.6 \text{ GeV}$ . However, this would produce another difficulty. As  $\Lambda_{\overline{MS}}$  becomes larger, the predicted  $m_b$  increases and becomes larger<sup>6),10)</sup> than the appropriate value shown in Eq.(2.3). I show here the  $\Lambda_{\overline{MS}}$  dependence of  $m_b$ .

$\Lambda_{\overline{MS}}$ ( GeV )	$m_b$ ( GeV )	
0.08	4.7	
0.16	5.1	(3.3)
0.26	5.5	

(Results by Oliensis and Fischler<sup>10)</sup> for  $m_t = 20 \text{ GeV}$ .)

Therefore, the IMB data drive the minimal SU(5) GUT into a corner.

Can we rescue this theory with slight modifications? In the following, I show some possible ways<sup>16)-21)</sup> to make  $\tau_p$  longer.

i) Inclusion of a **45** Higgs multiplet

This is based on the discussion by Jarlskog.<sup>16)</sup> In a theory

with more than two generations, particle mixing among them should be taken in the most general way. This is applied to the SU(5) GUT with 3 generations. Then, if there remain some new parameters (i.e., parameters which do not appear in the well-known charged current sector) in the proton-decay-inducing X and Y boson interactions, we may be able to adjust them freely and even rotate the proton decay away.

Although this Jarlskog's argument does not work in the minimal scheme<sup>29),32)</sup>, her idea was shown to be valid if a **45** Higgs multiplet is included.<sup>17),32)</sup> Subsequently, concrete calculations in such extended scheme were performed by several authors, and the above results were shown not to be affected by the inclusion of radiative corrections.<sup>18)</sup> (However, it should be noted that the inclusion of **45** Higgs multiplet changes the calculation of  $m_b/m_\tau$  since the tree mass relation given by this Higgs multiplet is  $m_b/m_\tau = 1/3$ .)

ii) Inclusion of mass-split multiplets

Another approach is to make  $M_X$  larger by adding some new particles which give different contributions to each of  $\beta_i$  functions.

One way is to use some new fermion multiplets. Frampton and Glashow<sup>20)</sup> extended the fermion sector  $5^*+10$  to  $(5^*+5)$  and  $(10+10^*)$ , and examined how  $\tau_p$  changes for various cases in which some parts of the multiplets are made light ( $\sim M_W$ ; masses of others are  $\sim M_X$ ). They found that  $\tau_p$  can be increased by a

factor  $\sim 2000$ . (In this case, the predicted Weinberg angle,  $\sin^2 \theta_W$ , changes  $\sim +0.009$ .)

Second one is to use Higgs multiplets instead of fermions. Hagiwara et al.<sup>21)</sup> studied this possibility, and found that adding a split **24** Higgs multiplet can increase  $\tau_p$  by a factor of 200 without changing the value of  $\sin^2 \theta_W$ .

Besides these modifications, papers<sup>33)</sup> appeared, which assert that minimal SU(5) prediction for  $\tau_p$  is not inconsistent with the IMB data. The authors obtained this results from the calculations of the matrix elements by the B-S equation technique (the estimated  $\tau_p$  is larger than the usual one by a factor of about 100). However, I cannot understand why their results are so different from those by many other authors.

#### § 4. Proton decay II: SO(10) GUT

SO(10) grand unified theory<sup>14)</sup> is also interesting and attractive though its particle content is complicated slightly compared to the SU(5) GUT. (We have 45 gauge bosons, and several Higgs multiplets are necessary to realize the desirable symmetry breaking<sup>34)</sup>.) This theory is automatically anomaly-free due to the group property. Moreover, all fermions in one generation can be assigned to a **16** representation. (As was mentioned in § 2, **5\*** and **10** are both necessary in the case of SU(5) GUT.)

As is well-known, there are two representative symmetry breaking patterns:

$$\begin{aligned} \text{(I)} \quad & \text{SO}(10) \rightarrow \text{SU}(5) \rightarrow \text{SU}(3) \times \text{SU}(2) \times \text{U}(1) , \\ \text{(II)} \quad & \text{SO}(10) \rightarrow \text{SU}(3) \times \text{SU}(2)_L \times \text{SU}(2)_R \times \text{U}(1)_{B-L} \\ & \rightarrow \text{SU}(3) \times \text{SU}(2) \times \text{U}(1) . \end{aligned} \quad (4.1)$$

(Strictly speaking, each pattern has further several branches, some of which will be discussed later.)

Concerning the proton decay, the pattern (I) is similar to the SU(5) GUT. Detailed studies in this pattern were done,<sup>35)</sup> and  $\tau_p$  was found to be almost same or shorter than  $\tau_p^{\text{SU}(5)}$ . So, I concentrate my talk on the pattern (II).

Pattern (II) gives a possible answer to the question "Why is the nature left-handed?" Furthermore, if a solution which allows relatively light  $W_R$  (right-handed gauge boson) is possible, this scheme is quite interesting not only from theoretical (aesthetic) but also from phenomenological point of view.

$\tau_p$  in this case becomes longer as  $W_R$  becomes lighter. Tosa et al.<sup>36)</sup> gave a simple formula connecting  $M_X$  and  $M_{W_R}$  as

$$M_X = M_X^{\text{SU}(5)} \sqrt{\frac{M_X^{\text{SU}(5)}}{M_{W_R}}} , \quad (4.2)$$

where  $M_X^{\text{SU}(5)}$  means the X boson mass in the SU(5) GUT.

Of course, we are not allowed to make  $M_{W_R}$  too light since it may affect low energy phenomenology. Rizzo and Senjanovic<sup>37)</sup>

made low-energy re-analysis and found the following solutions.

(i) Standard picture

$$\sin^2 \theta_W(M_W) \approx 0.22-0.23, \quad M_{W_R} \gtrsim 10^{9-10} \text{ GeV}, \quad (4.3)$$

$$\tau_P^{SU(5)} \lesssim \tau_P \lesssim 10^{33} \text{ yr}.$$

(ii) Weakly broken parity picture

$$\sin^2 \theta_W(M_W) \approx 0.27-0.28, \quad M_{W_R} \approx 150-240 \text{ GeV}, \quad (4.4)$$

$$\tau_P \gtrsim 10^{38} \text{ yr}.$$

The solution (ii) seems quite interesting. Although we must abandon the possibility to detect the proton decay, we will be able to observe the restoration of the left-right symmetry by future accelerators. Unfortunately, however, the measured  $W^\pm$  and Z boson masses<sup>38),39)</sup> seem to be inconsistent with the predicted values in the solution (ii);

$$M_W \approx 70-72 \text{ GeV}, \quad M_Z \approx 80-84 \text{ GeV}, \quad (4.5)$$

(Predictions in the solution (ii))<sup>37)</sup>

$$\begin{aligned} m_W^{\text{exp}} &= 80.9 \pm 1.5 \pm 2.4 \text{ GeV (UA1)} \\ &= 81.0 \pm 2.5 \pm 1.3 \text{ GeV (UA2)}, \\ m_Z^{\text{exp}} &= 95.6 \pm 1.5 \pm 2.9 \text{ GeV (UA1)} \\ &= 91.9 \pm 1.3 \pm 1.4 \text{ GeV (UA2)}. \end{aligned} \quad (4.6)$$

Let us consider the solution (i). In this case, it becomes completely impossible for us to see the L-R symmetry restoration. Then, can't we expect any new phenomena at relatively low energy region? Fortunately, there still remains such a possibility.

The breaking pattern (II) can have another intermediate energy scale as

$$(II') \quad SU(3) \times SU(2)_L \times SU(2)_R \times U(1)_{B-L}$$

$$\xrightarrow{M_{W_R}} SU(3) \times SU(2) \times U(1)_{B-L} \times U(1)_R \quad (4.7)$$

$$\xrightarrow{M'} SU(3) \times SU(2) \times U(1).$$

The equation (4.2) is still valid in this case. That is, the size of  $M'$  (mass of the second neutral gauge boson  $Z'$ ) is not constrained.  $Z'$  can be light as long as any phenomenological difficulties do not appear. In fact, it was shown possible<sup>40)</sup> to have  $Z'$  with a mass as low as 200 GeV. Using two-loop renormalization group equations, Ma and Whisnant<sup>41)</sup> showed that  $\tau_P$  in this case becomes much longer (as high as  $10^{39}$  yr).

The SO(10) GUTs are now quite attractive. However, all calculations except for Ref.35) do not include the threshold effects. As was mentioned in § 2, there may occur non-negligible changes in the values of the parameters by these effects. Therefore, further detailed studies are desired.

§ 5. Summary

I have reviewed the proton decay studies in the conventional SU(5) and SO(10) grand unified theories.\*) I have shown necessary steps for the determination of the values of parameters in the basic Lagrangian, successful predictions for  $\sin^2 \theta_W(M_W)$  and  $m_p$ , and the resultant proton lifetime.

The minimal SU(5) GUT is now in a quite difficult situation because of the IMB experiment. In order to rescue this theory, i.e., to get longer  $\tau_p$ , several authors have suggested some modifications.

On the other hand, the SO(10) GUTs with breaking through the left-right symmetric stage are consistent with various data (including  $\tau_p^{\text{exp}}$ ) provided  $M_{WR} \gtrsim 10^{9-10}$  GeV. Moreover, we may be able to see new phenomena related to the second neutral gauge boson at energies attainable by future accelerators. Strictly speaking, however, these results were obtained from the calculations without the threshold effects and therefore more systematic studies are necessary.

In addition, we should pay further effort to eliminate various ambiguities in the theoretical computations. The largest one comes from the value of the QCD coupling constant,

---

\*) I did not mention supersymmetric grand unified theories. The lifetime and the decay mode of proton are quite different in such theories.<sup>15)</sup>

i.e., the size of  $\Lambda_{\overline{MS}}$ . If we can have more refined data on  $\sin^2 \theta_W$  and  $M_{W,Z}$ , we can use them as new inputs instead of  $\Lambda_{\overline{MS}}$  for the determination of  $M_X$ , and consequently  $\tau_p$ . This will be possible in the near future by new high energy accelerators TRISTAN, LEP, ... .

At the present, we do not know whether the grand unification scenario is correct or not. The proton decay can give us a significant key to this important problem. More refined experimental data from IMB, KAMIOKANDE and other collaborations are highly desired.

Acknowledgement

I would like to thank Takeo Inami for careful reading of the manuscript.

References

- 1) S.L. Glashow, Nucl. Phys. 22 (1961), 579.  
S. Weinberg, Phys. Rev. Letters 19 (1967), 1264.  
A. Salam, Elementary Particle Theory, ed. by N. Svartholm (Almqvist and Wiksell, Stockholm, 1968), p.367.  
S.L. Glashow, J. Iliopoulos and L. Maiani, Phys. Rev. D2 (1970), 1285.  
M. Kobayashi and T. Maskawa, Prog. Theor. Phys. 49 (1973), 652.

- 2) H. Fritzsch, M. Gell-Mann and H. Leutwyler, Phys. Letters 47B (1973), 365.  
S. Weinberg, Phys. Rev. Letters 31 (1973), 494; Phys. Rev. D8 (1973), 4482.  
H.D. Politzer, Phys. Rev. Letters 30 (1973), 1346.  
D.J. Gross and F. Wilczek, Phys. Rev. Letters 30 (1973), 1343; Phys. Rev. D8 (1973), 3633.
- 3) As reviews:  
J. Ellis, Lectures presented at the 21st Scottish Universities Summer School in Physics, 10-30 August 1980, St Andrews, Scotland (CERN-preprint Ref. TH. 2942).  
P. Langacker, Phys. Reports 72 (1981), 185.
- 4) H. Georgi and S.L. Glashow, Phys. Rev. Letters 32 (1974), 438.
- 5) H. Georgi, H.R. Quinn and S. Weinberg, Phys. Rev. Letters 33 (1974), 451.
- 6) A.J. Buras, J. Ellis, M.K. Gaillard and D.V. Nanopoulos, Nucl. Phys. B135 (1978), 66.
- 7) T. Goldman and D.A. Ross, Nucl. Phys. B171 (1980), 273.
- 8) W.J. Marciano, Phys. Rev. D20 (1979), 274.  
E.A. Paschos, Nucl. Phys. B159 (1979), 285.  
W.J. Marciano and A. Sirlin, Phys. Rev. Letters 46 (1981), 163.  
J. Kubo and S. Sakakibara, Phys. Letters 103B (1981), 39.
- 9) J. Ellis, M.K. Gaillard, D.V. Nanopoulos and S. Rudaz, Nucl. Phys. B176 (1980), 61.  
C.H. Llewellyn Smith, G.G. Ross and J.F. Wheeler, Nucl. Phys. B177 (1981), 263.  
L. Hall, Nucl. Phys. B178 (1981), 75.  
P. Binétruy and T. Schücker, Nucl. Phys. B178 (1981), 293; 307.  
W.J. Marciano, Field Theory in Elementary Particles, ed. by B. Kursunoglu and A. Perlmutter (Plenum, NY, 1982), p.71.  
I. Antoniadis, C. Kounnas and C. Roiesnel, Nucl. Phys. B198 (1982), 317.
- 10) D.V. Nanopoulos and D.A. Ross, Nucl. Phys. B157 (1979), 273; Phys. Letters 108B (1982), 351.  
J. Oliensis and M. Fischler, Phys. Rev. D28 (1983), 194.  
M. Fischler and J. Oliensis, Phys. Rev. D28 (1983), 2027.  
See also:  
K. Inoue, A. Kakuto and Y. Nakano, Prog. Theor. Phys. 62 (1979), 307.
- 11) M.B. Gavela, A. Le Yaouanc, L. Oliver, O. Pène and J.C. Raynal, Phys. Rev. D23 (1981), 1580.  
Y. Tomozawa, Phys. Rev. Letters 46 (1981), 463.  
J.M.F. Labastida and F.J. Ynduráin, Phys. Rev. Letters 47, (1981), 1101.

- V.S. Berezinsky, B.L. Ioffe and Ya.I. Kogan, Phys. Letters 105B (1981), 33.
- H. Arisue, Prog. Theor. Phys. 66 (1981), 1395.
- J.F. Donoghue and E. Golowich, Phys. Rev. D26 (1982), 3092.
- N. Isgur and M.B. Wise, Phys. Letters 117B (1982), 179.
- A.W. Thomas and B.H.J. McKellar, Nucl. Phys. B227 (1983), 206.
- W. Lucha, Phys. Letters 122B (1983), 381.
- T. Okazaki and K. Fujii, Phys. Rev. D27 (1983), 188.
- S. Meljanac, D. Palle and D. Tadić, Nucl. Phys. B228 (1983), 56.
- Y. Dupont, T.N. Pham and Tran N. Truong, Nucl. Phys. B228 (1983), 77.
- S.J. Brodsky, J. Ellis, J.S. Hagelin and C.T. Sachrajda, Preprint SLAC-PUB-3141 (1983).
- 12) IMB Collaboration: R.M. Bionta et al., Phys. Rev. Letters 51 (1983), 27.
- 13) K. Arisaka and T. Kajita, Talk at this Workshop.
- 14) H. Fritzsch and P. Minkowski, Ann. of Phys. 93 (1975), 193.  
M.S. Chanowitz, J. Ellis and M.K. Gaillard, Nucl. Phys. B128 (1977), 506.  
H. Georgi and D.V. Nanopoulos, Nucl. Phys. B155 (1979), 52.
- 15) N. Sakai, Preprint TIT/HEP-78, August 1983.  
N. Sakai, Talk at this Workshop.  
J. Ellis, Talk presented at the 1983 International Symposium on Lepton and Photon Interactions at High Energies, Cornell Univ., 4-9 August 1983 (CERN-preprint Ref. TH. 3718) and references therein.
- 16) C. Jarlskog, Phys. Letters 82B (1979), 401.
- 17) S. Nandi, A. Stern and E.C.G. Sudarshan, Phys. Letters 113B (1982), 165.
- 18) V.S. Berezinsky and A.Yu. Smirnov, Phys. Letters 97B (1980), 371.  
D. Altschüler, P. Eckert and T. Schücker, Phys. Letters 119B (1982), 351.
- See also:  
G.P. Cook, K.T. Mahanthappa and M.A. Sher, Phys. Letters 90B (1980), 398.
- 19) G. Segrè and H.A. Weldon, Phys. Rev. Letters 44 (1980), 1737.
- 20) P.H. Frampton and S.L. Glashow, Phys. Letters 131B (1983), 340 ; 135B (1984), 515 Erratum.
- 21) K. Hagiwara, F. Halzen and K. Hikasa, Wisconsin-preprint MAD/PH/142, November 1983.
- 22) J.E. Kim, P. Langacker, M. Levine and H.H. Williams, Rev. Mod. Phys. 53 (1981), 211.  
F. Dydak, Phil. Trans. R. Soc. of London A304 (1982), 43.
- 23) A.J. Buras, Proceedings of the 1981 International Symposium on Lepton and Photon Interactions at High Energies, Univ. of Bonn, 24-29 August 1981, ed. by W. Pfeil, p.636.

- 24) D.A. Ross, Nucl. Phys. B140 (1978), 1.
- 25) W.J. Marciano, Talk presented at the Fourth Workshop on Grand Unification, Univ. of Pennsylvania, April 21-23, 1983 (Preprint BNL-33415).
- 26) This value of  $\sin^2 \theta_W^{\text{exp}}(M_W)$  was derived from the actual experimental data by taking the radiative corrections into account:  
 A. Sirlin and W.J. Marciano, Nucl. Phys. B189 (1981), 442.  
 See also:  
 S. Sakakibara, Prog. Theor. Phys. 64 (1980), 619; Phys. Rev. D24 (1981), 1149.  
 C.H. Llewellyn Smith and J.F. Wheeler, Phys. Letters 105B (1981), 486.  
 E.A. Paschos and M. Wirbel, Nucl. Phys. B194 (1982), 189.  
 J.F. Wheeler and C.H. Llewellyn Smith, Nucl. Phys. B208 (1982), 27; B226 (1983), 547 Erratum.
- 27) K.G. Wilson, Phys. Rev. 179 (1969), 1499.  
 M.K. Gaillard and B.W. Lee, Phys. Rev. Letters 33 (1974), 108.  
 G. Altarelli and L. Maiani, Phys. Letters 52B (1974), 351.  
 Theoretical foundation for the application of this technique to the GUT was given in  
 Y. Kazama and Y-P. Yao, Phys. Rev. D21 (1980), 1116; D21 (1980), 1138; D25 (1982), 1605.
- 28) F. Wilczek and A. Zee, Phys. Rev. Letters 43 (1979), 1571.  
 L.F. Abbott and M.B. Wise, Phys. Rev. D22 (1980), 2208.
- 29) J. Ellis, M.K. Gaillard and D.V. Nanopoulos, Phys. Letters 88B (1979), 320.
- 30) M. Daniel and J.A. Peñarrocha, Rutherford-preprint RL-83-042.
- 31) IMB Collaboration, Talk at Fermilab (1983).
- 32) R.N. Mohapatra, Phys. Rev. Letters 43 (1979), 893.
- 33) A.N. Mitra and R. Ramanathan, Phys. Letters 128B (1983), 381; Trieste-preprint IC/83/109.
- 34) Symmetry breaking patterns of SO(10) were studied, e.g., in  
 S. Rajpoot, Phys. Rev. D22 (1980), 2244.  
 M. Yasuë, Prog. Theor. Phys. 65 (1981), 708; Phys. Rev. D24 (1981), 1005.
- 35) T. Matsuki and N. Yamamoto, Prog. Theor. Phys. 69 (1983), 1505.  
 Y. Hara, Prog. Theor. Phys. 68 (1982), 642; Nucl. Phys. B214 (1983), 167.  
 Y. Hara, T. Matsuki and N. Yamamoto, Prog. Theor. Phys. 68 (1982), 652.
- 36) Y. Tosa, G.C. Branco and R.E. Marshak, Phys. Rev. D28 (1983), 1731.
- 37) T.G. Rizzo and G. Senjanović, Phys. Rev. D24 (1981), 704; D25 (1982), 235.  
 See also:



M. Yasuë and K. Matumoto, Prog. Theor. Phys. 67 (1982), 1899.

W.E. Caswell, J. Milutinovič and G. Senjanovič, Phys. Rev. D26 (1982), 161.

38) UA1 Collaboration: G. Arnison et al., Phys. Letters 122B (1983), 103; 126B (1983), 398; 134B (1984), 469.

39) UA2 Collaboration: M. Banner et al., Phys. Letters 122B (1983), 476; P. Bagnaia et al., Phys. Letters 129B (1983), 130.

40) V. Barger, E. Ma and K. Whisnant, Phys. Rev. D28 (1983), 1618.

Several authors have studied the  $SU(2) \times U(1) \times U(1)$  electro-weak theory in the  $SO(10)$  GUT. See the recent preprint by C.N. Leung and J.L. Rosner (FERMILAB-Pub-83/90-THY) and references therein.

41) E. Ma and K. Whisnant, Phys. Letters 131B (1983), 343.

#### Note added

Recently, Berezinsky and Smirnov have proposed a new mechanism to make  $\tau_p$  much longer (at least three orders of magnitude) in the minimal  $SU(5)$  grand unified theory, in which non-renormalizable terms with a coupling constant governed by a mass  $M(> M_X)$  is introduced:

V.S. Berezinsky and A.Yu. Smirnov, CERN-preprint Ref. TH. 3786.

### 32 Kton Water Cerenkov Detector(JACK)

A proposal for detailed studies of nucleon decays  
and for low energy neutrino detection

KAMIOKANDE collaboration

M.KOSHIBA

#### 1. Purposes

- 1) Determination of branching ratios of nucleon decays in the  $10^{32}$  year life time range.
- 2) Rubakov effect in the water detector
- 3)  $n-\bar{n}$  oscillation
- 4) Solar  $B^8$  neutrino
- 5) Rubakov effect neutrinos from the Sun
- 6) Neutrinos resulting from the gravitational collapse of star.
- 7) Search for very heavy stable particles, like fractional charge bare quarks and/or monopoles.
- 8) Multiple muons

#### 2. Detector

The proposed detector is shown schematically in Figs. 1-a, and -b. Besides the inner volume viewed by  $\sim 11,000$  PMT's of  $20''\phi$ , 40% photocathode coverage of the entire surface, it has an anti-coincidence layer of 1.5m thickness, which is a necessity if one is to work on the low energy neutrino of several MeV.

All the phototubes will be equipped with FADC's or with ADC + TDC.

#### 3. Expected performances

With the 40% coverage by photocathode,  $2 \times 20''\phi$  PMT/m<sup>2</sup>, of the entire surface one can with good efficiency detect electrons of  $\sim 6$  MeV.

Monte Carlo simulations of low energy-neutrino events are shown in Figs. 2.

The expected low energy neutrino event rates are summarized

in Table I. Note that, not only the respectable event rates, we can determine the source direction with reasonable accuracy.

As to the nucleon decay, even with a 30% probability for successful reconstruction,  $10^{32}$  years life time will yield 40 clear cases of nucleon decay per year in the fiducial 22,000 ton of water, containing 2500 ton of free protons.

It has been shown in the present KAMIOKA experiment that showering and non-showering tracks can be clearly distinguished.

With the still better, by a factor of 2, light collection combined with the good timing, it is expected that the separations of  $e$  and  $\gamma$  as well as of  $\mu$  and  $\pi$  among the nucleon decay products are within the reach.

#### 4. Cost and Feasibility

The approximate cost of this experiment is;

Excavation and water tank	$\text{¥}2.0 \times 10^9$
20''PMT @200,000 x 120,000	$\text{¥}2.4 \times 10^9$
Electronics	$\text{¥}1.0 \times 10^9$
Contingency	$\text{¥}0.5 \times 10^9$
Total	$\text{¥}5.9 \times 10^9$

( $\sim 24$ M\$)

It is possible to excavate  $40\text{m}\phi \times 40\text{m}$ h cave at  $\sim 1000$  m underground in KAMIOKA. Another possibility is at Gran Sasso in Italy.

The present KAMIOKANDE collaboration is eager to work on this experiment.

Table I

Expected Low Energy Neutrino Events

(4 $\pi$  Anti-coincidence shield mandatory)

## Parameters for 32000 ton Water Cerenkov Detector

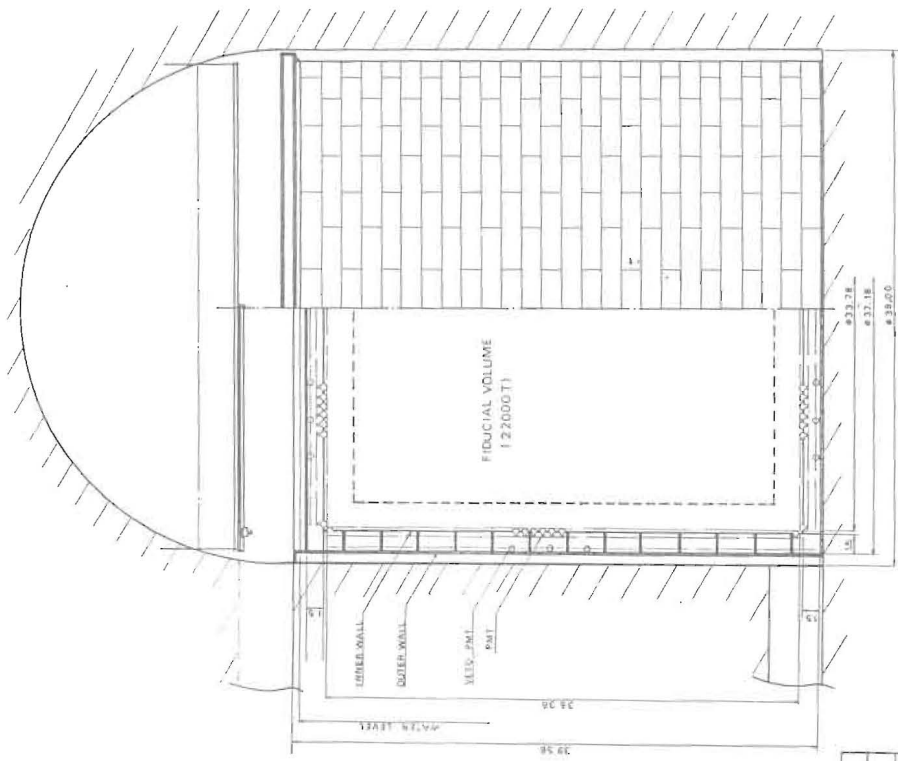
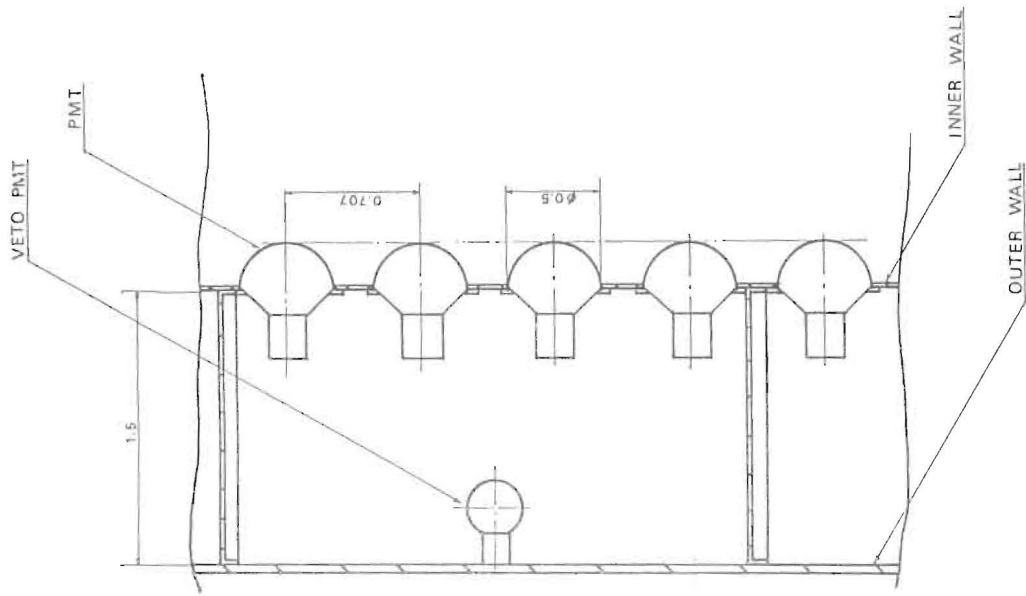
1. Cylindrical tank of 37.18m $\phi$  x 39.76mh
2. Total mass = 32000 t  
Fiducial mass = 22000 t(29.78m $\phi$  x 31.36mh)
3. Number of phototubes(20") = 11000  
Extra number for veto(20") = 500
4. Average light transparency  $\geq$  35m
5. Expected photoelectrons/MeV = 5.8 p.e./MeV
6. Expected event rate estimated from present KAMIOKANDE
 

	Total	Fiducial
Single ring events	7.4/day	5.1
multi ring events	3.3	2.3

Source of $\nu$	Solar B <sup>8</sup>	Solar Rubakov effect	Gravitational collapse at Galactic center
$E_{\nu_e}$	$E_{\nu_e} \leq 14$ MeV $\langle E_{\nu_e} \rangle > 10$ MeV	$E_{\nu_e} \leq 53$ MeV $\langle E_{\nu_e} \rangle > 35$ MeV	$\langle E_{\nu_e} \rangle \sim 16$ MeV
$\nu_e$ flux at Earth	$5.6 \times 10^6 \nu_e$ * /cm <sup>2</sup> /sec.	$< 1.2 \times 10^4 \nu_e$ ** /cm <sup>2</sup> /sec.	$7.7 \times 10^{10} \nu_e$ /m.sec. *** for a few m.sec. duration
KAMIOKANDE 880 ton H <sub>2</sub> O	$6.0(\nu_e e + \nu_e \bar{e})$ per day. ( $E_e \geq 6$ MeV)	$35 \nu_e e + \nu_e \bar{e}$ $53 \nu_e 0 + e \bar{F}$ per year. ( $E_e \geq 6$ MeV)	$2.3(\nu_e e + \nu_e \bar{e})$ per m.sec.burst ( $E_e \geq 6$ MeV)
JACK 22000 ton H <sub>2</sub> O	$150(\nu_e e + \nu_e \bar{e})$ per day. ( $E_e \geq 6$ MeV)  $\bar{\theta}_{\nu_e e} \leq 19^\circ$  $\bar{\theta}_{\text{scatt}}(8 \rightarrow 4 \text{ MeV})$ $= 100^\circ$ Source direction $\leq 8^\circ$ per day.	$2.4(\nu_e e + \nu_e \bar{e})$ $3.6(\nu_e 0 + e \bar{F})$ per day. ( $E_e \geq 6$ MeV)  $\bar{\theta}_{\nu_e e} \leq 10^\circ$  $\bar{\theta}_{\text{scatt}}(16 \rightarrow 8 \text{ MeV})$ $= 50^\circ$ Source direction $\leq 6^\circ$ per month.	$58(\nu_e e + \nu_e \bar{e})$ per m.sec.burst ( $E_e \geq 6$ MeV)  $\bar{\theta}_{\nu_e e} \leq 14^\circ$  $\bar{\theta}_{\text{scatt}}(12 \rightarrow 6 \text{ MeV})$ $= 65^\circ$ Source direction $\leq 2^\circ$ per m.sec.burst

\* Bahcall, J.N., et al; Rev.Mod.Phys., 54, (1982), 767\*\* Arafune, J. et al; P.R.L., 50, (1983), 1901 and P.L.B., (1983)

\*\*\* Mazurek, T.J., et al; DUMAND, Hawaii (1980)



SUPERKAMIOKANDE	
1/200	UNIT : METER
SKETCH	
30.11.83	
UNIV. TOKYO Y.T.	

Fig. 1-a

1/20

Fig. 1-b

SUPERKAMIOKANDE	
1/200	UNIT : METER
SKETCH	
30.11.83	
UNIV. TOKYO	

Electron (5MeV)

<< SUPER NDE >>

- DEVELOPMENT -

NUM	:	37
RUN	:	0
EVENT	:	36
TIME	:	29/11/83
		20.11.12

< NUMBER OF PHOTO ELECTRONS >

TOTAL P.E.	:	21
MAX P.E.	:	2
THRES P.E.	:	0.2
NORM	:	0

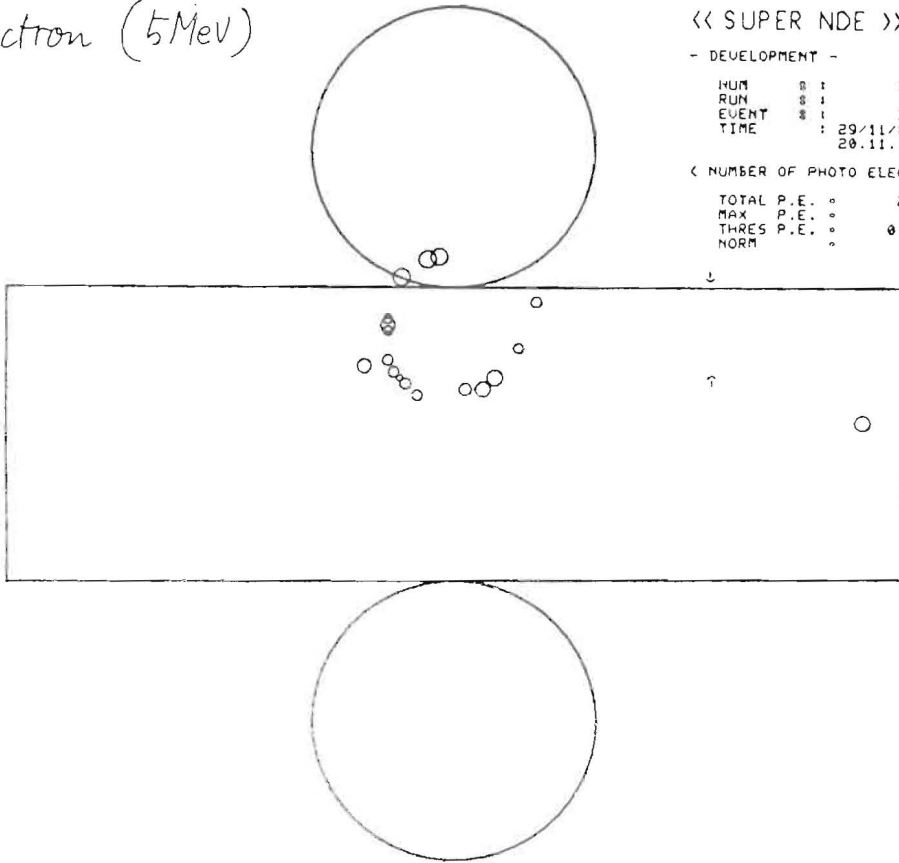


Fig. 2-a

Electron (10MeV)

<< SUPER NDE >>

- DEVELOPMENT -

NUM	:	18
RUN	:	0
EVENT	:	17
TIME	:	29/11/83
		19.10.29

< NUMBER OF PHOTO ELECTRONS >

TOTAL P.E.	:	54
MAX P.E.	:	2
THRES P.E.	:	0.4
NORM	:	0

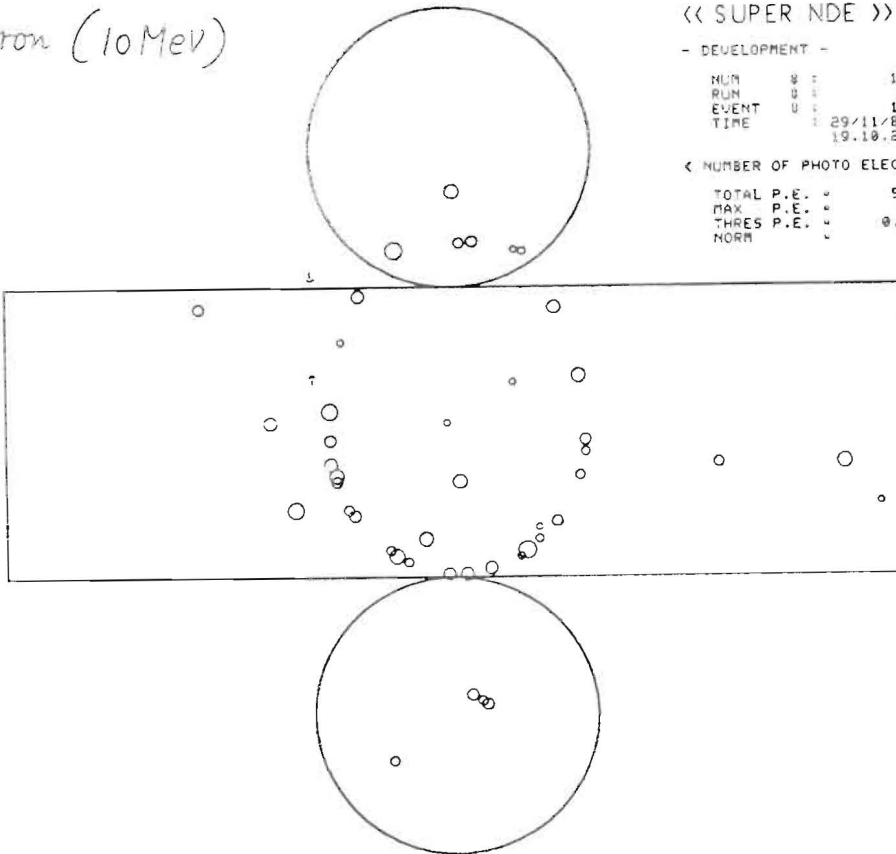


Fig. 2

## NUCLEON DECAY EXPERIMENT AT KOLAR GOLD FIELD

M.R.Krishnaswamy, M.G.K.Menon, N.K.Mondal, V.S.Narasimham and B.V.Sreekantan  
Tata Institute of Fundamental Research, Bombay, India  
Y.Hayashi, N.Ito, S.Kawakami and T.Nakamura  
Osaka City University, Osaka, Japan  
S.Miyake  
Institute for Cosmic Ray Research, University of Tokyo, Tokyo, Japan

### Abstract

Five events have been observed in the K.G.F. nucleon decay experiment with tracks fully confined to the detector volume. It is shown that their characteristics are in conformity with the decay of bound nucleons and that the background due to neutrino interactions within the detector volume is small. Based on these data, mean life time for nucleons bound in iron nuclei is estimated as about  $1.1 \times 10^{31}$  years.

### 1. Introduction

The Kolar Gold Field nucleon decay experiment has been in operation since October 1980 at the depth of 2,300 m, equivalent to 7,600 m.w.e. of standard rock.

The detector, of floor area 6 m x 4 m and 3.7 m in height, is composed of 34 layers of proportional counters, with 1.2 cm thick iron plates between the layers. The counters have a cross section of  $10 \times 10 \text{ cm}^2$  with a thickness of 2.3 mm iron and are in two lengths of 4 m and 6 m. The alternate layers of counters are arranged orthogonally to obtain a three dimensional view of tracks. The total weight of the detector is about 140 tons. The trigger is a 5 layer coincidence in any of 11 consecutive layers. There is also an additional trigger whereby tracks crossing any 2 counters in 2 layers in 3 consecutive layers are also recorded. In every trigger, it records the position of hit counters and information on the ionisation deposited in each counter.

During effective running time of 2.5 years, about 1900 events have been recorded. These have been classified into various categories  
1. atmospheric muons about 1,700, 2. neutrino interactions in rock about 50  
3. neutrino interactions inside detector 29, and 4. nucleon decay candidates  
5 confined cases.

### 2. Nucleon Decay Candidates

Five events have been classified as candidates for nucleon decay which are fully confined in the detector and constitute rather strong evidence for nucleon decay.

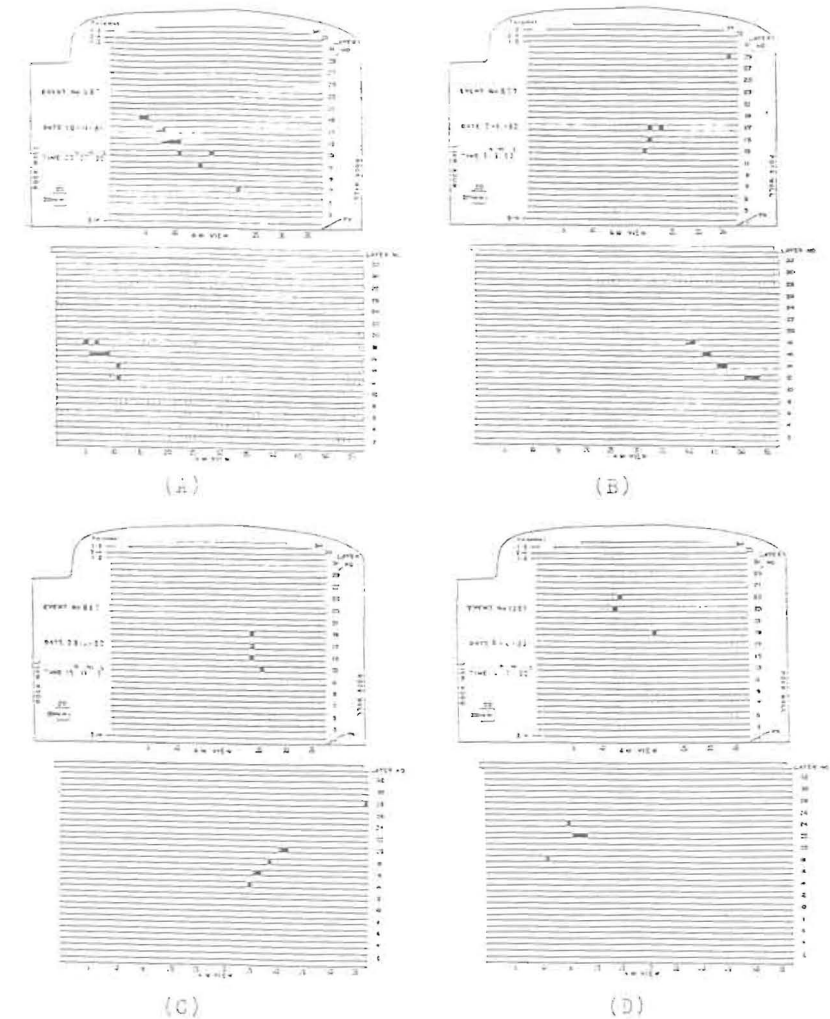
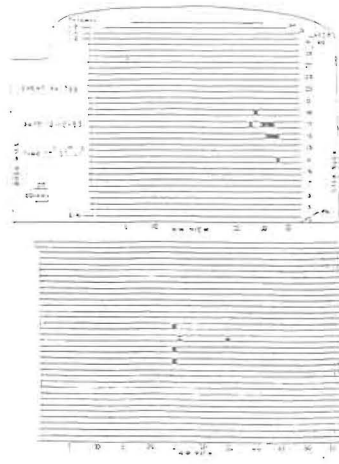


Fig. 1



Two orthogonal over views of the fully confined events are shown in Fig. 1. The odd and even numbered layers correspond to the 6 m and 4 m counters. The black squares are the counters which show ionization above the threshold.

Detailed plots of ionisation of the hit counters, converted to the equivalent number of minimum ionising particles are given separately for the two orthogonal views in Fig.2,3,4,5 and 6.

(E)

Fig.1 (E)

Event No. 587 ; The pattern of hit counters and ionisation in Fig. 2 is typical of electromagnetic cascades, with the absence of any clear penetrating tracks. Where the main axis has the angular co-ordinates  $\theta = 58^\circ$  and  $\phi = 35^\circ$  (from North to South clockwise). The total range of the shower measured along the axis is  $\sim 20$  radiation lengths and the total ionisation corresponds to 42.6 particles.

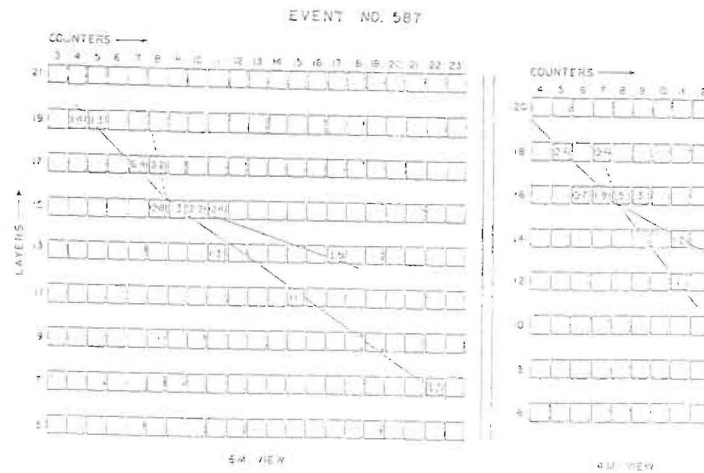


Fig. 2

Based on the conventional track length integral method, we estimate the total energy of the event as 980 MeV with an uncertainty of  $\pm 20\%$ . The profiles of the event are easily understood as that the event is composed of separate showers originating from a point in the layer 15th (top) and emitted in back to back configuration. A plausible interpretation of the event in terms of proton decay is a decay into a positron (upwards) and neutral pion (downwards).

Event No 867 ; This event (Fig.3), besides being a fully confined one, has the following distinguishing features ; 1) A kink at the point B with the angle of deflection of about  $40^\circ$ . 2) Normal ionisation along the path BC. 3) Increased ionisation at the end point 'A'. These features suggest the creation of a particle at the point 'C' which slowed down to point 'B' and produced a decay particle 'B to C'. From the identification of the particles, using range and ionisation, this event can be interpreted as  $P \rightarrow \nu + K^+$  and  $K^+$  decayed into muon and neutrino, in view of the fact that the measured value of muon is in good agreement with that of kaon decay. Dotted line in the figure is showing other possible cases, however, it is less probable because of the difficulty to understand high ionisation in the counter of 15th layer.

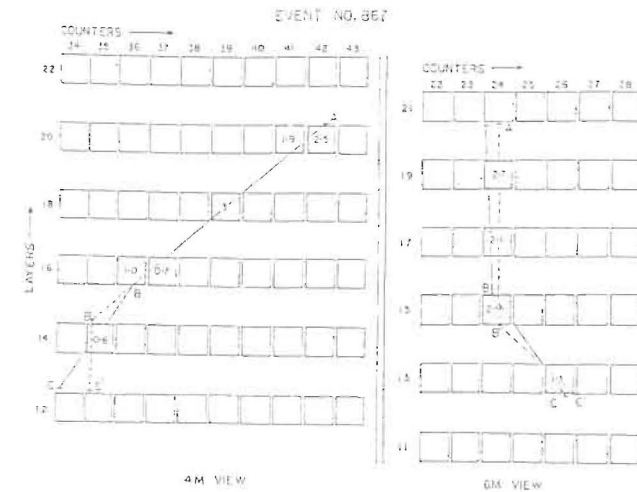


Fig. 3

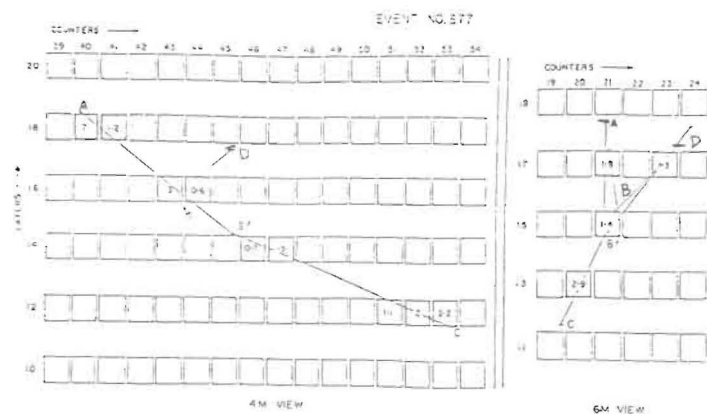


Fig. 4

Event No. 877 ; The main characteristics of this event (Fig.4) are ;

- 1) A non showering particle pion or muon, traversing the path 'BC' of  $150 \text{ g/cm}^2$  before stopping.
- 2) Two slow particles 'BA' and 'BD', both pion or muon, are emitted to upward with an opening angle of about  $100^\circ$ .
- 3) Total energy of the event is close to  $1 \text{ GeV}$ .

A detailed analysis of the event , and in particular the ionisation in the hit counters and their disposition along the path suggests a reasonable fit to the decay mode of  $P \rightarrow \mu^+ + K_S^0$  and  $K_S^0 \rightarrow \pi^+ + \pi^-$ . The total energy of pions and their opening angle are consistent with decay of a slowly moving  $K^0$  of momentum  $\sim 300 \text{ MeV/c}$ .

Event No. 1465 ; This event spreads just six layers and has 3 tracks with large opening angles. From the ionisation of tracks, 'PA' and 'PB' can be single tracks and remaining one is gamma ray with  $\emptyset$  missing layers. 'PA' and 'PB' pass through materials of about  $55$  and  $35 \text{ g/cm}^2$  each and if these tracks are pions, their energy can be estimated as  $270$  and  $220 \text{ MeV}$  respectively. Therefore, this event can be interpreted as  $N \rightarrow \nu + \eta$  and  $\eta \rightarrow \pi^+ + \pi^- + \pi^0$ , in view of the fact that the measured value of  $E_\pi$

is in good agreement with that to be expected in such a process. For a low energy neutrino interaction, with more than two pions production, is suppressed by a factor of  $10$  compared to single pion process.

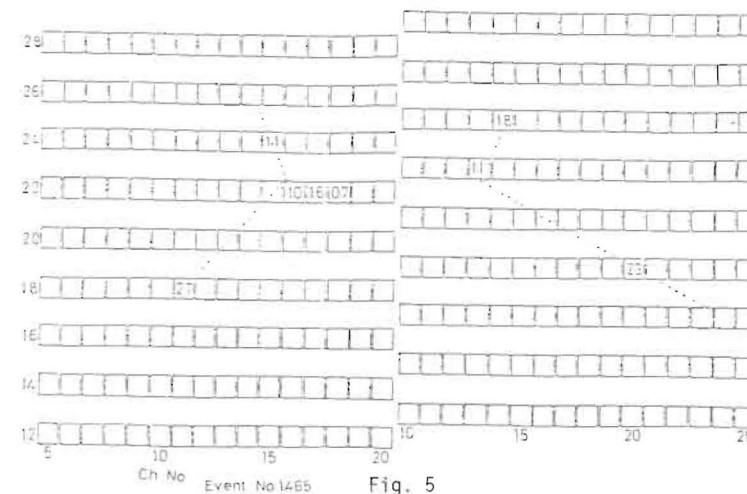


Fig. 5

Event No. 1766 ; The event has been found recently but not fully analysed yet. From the ionisation deposited to the detector, visible energy can be estimated as about  $600 \text{ MeV}$  excluding particle mass. The event can be interpreted as  $P \rightarrow e^+ + K^0$ , and  $K^0 \rightarrow \pi^+ + \pi^-$  and total energy is estimated as about  $900 \text{ MeV}$  in this case.

### 3. Estimate on the Lifetime of Bound Nucleons

Based on 5 candidate events within a fiducial weight of  $60$  tons during a livetime of  $2.5$  years, and also the estimated detection efficiency of  $0.5$  inclusive of hadron absorption in the nucleus, the lifetime for nucleons is estimated as about  $1.1 \times 10^{31}$  years.

For some more details and discussion of background, see References ;  
Krishnaswamy M.R. et al ; Phys. Lett, B 106,339, 1981

Phys. Lett. B 115,349, 1982

Proc. Int. Colloq. on Baryon Nonconservation  
1982 in Bombay, 1983 in Frascati

Pramana Suppl. 115 1982

Pramana 19,525, 1982



Astrophysical Constraints on the Monopole Abundance

M. Fukugita

Research Institute for Fundamental Physics  
Kyoto University, Kyoto, 606 Japan

Abstract

Astrophysical constraints on the monopole abundance are surveyed. Various constraints derived from the monopole catalysed proton decay are discussed more in detail.

1. Introduction

The search for magnetic monopole has been greatly attracting our interest. One of the reason for this is that the monopole is predicted naturally to appear in grand unified theories when a simple group is spontaneously broken into a group that contains a U(1) symmetry<sup>[1]</sup>. Another motivation that has triggered our interest is an observation of a candidate event which might be interpreted as a Dirac monopole passing through an induction coil<sup>[2]</sup>.

The mass of GUT monopole ( $M_m$ ) is of the order of  $M/\alpha$  with  $M$  the mass scale of the symmetry breaking and  $\alpha$  the U(1) charge, and therefore we expect  $M_m \sim 10^{16}$  GeV. [In SU(5) grand unification the monopole mass<sup>[3]</sup> is restricted to be  $M_X/(8\alpha_{em}/3) \leq M_m \leq 1.79 M_X/(8\alpha_{em}/3)$ .] Since such a heavy monopole can be produced only in the early universe with a temperature  $T \sim M$ , one can explore, by searching for the monopole, the hot universe as early as  $t \sim 10^{-38}$  sec, which may be compared with  $t \sim 10^5$  yr for the microwave background radiation or  $t \sim 1$  sec for the He synthesis. The monopoles produced in the early universe will survive up to the present epoch<sup>[4,5]</sup>, and are subject to a variety of astrophysical constraints as well as those obtained in laboratories.

We think a standard velocity of this heavy monopole being as slow as  $\beta = v/c \sim 10^{-3}$ . If monopoles cluster with the galaxy or the local supercluster, their velocity must be of the order of the virial velocity,  $10^{-3}$  for the galaxy and  $3 \times 10^{-3}$  for the supercluster. If relic monopoles are distributed uniformly throughout the universe, the velocity dispersion characterised by their temperature is very small, and their velocity relative to our galaxy is again of the order  $10^{-3}$  as a consequence of a proper motion of our galaxy. The presence of galactic magnetic field, of course, might change the situation.

An important aspect which would be borne by the GUT monopole is that it would catalyse the nucleon decay with the cross section typical of strong interactions (Rubakov effect)<sup>[6]</sup>, although there are some subtle points yet to be clarified before giving a full credit to this prediction<sup>[7]</sup>. The Rubakov effect, if it happens, leads to quite significant consequences not only in particle physics but also in astrophysical environments. A constraint on cosmic monopole flux derived therefrom<sup>[8,9]</sup> could be so strong that it would discourage any efforts to look for monopoles in laboratories.

In this talk I shall present a survey on astrophysical constraints on the monopole abundance, with some emphasis on that derived from the Rubakov effect by the present author and collaborators, and also by others. Another interesting problem, the problem of monopole production in the early universe<sup>[4,5,10]</sup>, is not mentioned in the present talk.

2. Constraints on the monopole abundance derived without resorting to the Rubakov effect.

The experimental limit on the cosmic monopole flux varies depending on the velocity of monopoles. The limit so far obtained<sup>[11]</sup> is summarized as

$$F_m \leq 5 \times 10^{-15} (\beta/10^{-2})^{-1.65} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1} \quad \text{for } 10^{-4} \leq \beta \leq 10^{-2}$$

$$F_m \leq 5 \times 10^{-15} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1} \quad \text{for } 10^{-2} \leq \beta \leq 1 . \quad (1)$$

For  $\beta \leq 10^{-4}$  the energy loss of a monopole is very small and the monopole escapes from detection in energy-loss experiments. The limit for such a slowly-moving monopole comes only from induction experiments<sup>[2,12]</sup>:  $F_m \leq 3.7 \times 10^{-11} \text{ cm}^{-2} \text{ sr}^{-1} \text{ sec}^{-1}$ .

*Constraints from the galactic magnetic field.* Astrophysical arguments which lead to a limit on monopole flux may be classified into two. The first uses the fact that an excessive presence of monopoles could have exhausted the astrophysical magnetic field<sup>[13]</sup>, and the second is from the cosmological mass density. The monopole with density  $n_m$  depletes the magnetic-field energy at the rate

$$-\frac{d}{dt} \frac{B^2}{8\pi} = g v_m n_m B . \quad (2)$$

The condition that the galactic magnetic field should not decay within the typical field generation time ( $\tau$ ) leads to the Parker's limit,

$$F_m = \frac{n_m v_m}{4\pi} < \frac{1}{4\pi} \frac{B}{8\pi g \tau} . \quad (3)$$

For  $B=3\mu$  Gauss and  $\tau=10^8$  yr from the dynamo theory for galactic magnetic field we obtain<sup>[13,14]</sup>

$$F_m < 2 \times 10^{-16} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} .$$

The galactic magnetic field is not uniform but it fluctuates over a typical length scale of  $\lambda=100-150$  pc<sup>[15]</sup>, and a monopole traversing the galaxy does not always gain energy from the magnetic field, but sometimes loses its energy. Therefore the loss of field energy is smaller by a factor of  $(r_g/\lambda)^{1/2}$ , giving<sup>[16]</sup>

$$F_m \leq 10^{-16} (r_g/\lambda)^{1/2} \sim 10^{-15} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} . \quad (4)$$

When a monopole moves very fastly, the energy loss is not as effective as discussed here, and the limit on  $F_m$  is loosened (e.g., for  $M_m=10^{16}$  GeV,  $F_m \leq 10^{-15} [\beta/10^{-2.5}]^2$  for  $\beta \geq 10^{-2.5}$ )<sup>[16]</sup>. See fig.1 below.

Let us note here that this limit is obtained without taking account of the evolution of monopole numbers in the galaxy, i.e., it is assumed that monopole flux does not change over  $10^8$  yr. If we consider a monopole-antimonopole plasma, we expect magnetic field and monopole kinetic energy fluctuation interconverting periodically, rather than the irreversible damping of the magnetic field<sup>[17,18]</sup>. The authors in ref.17 obtained a bound looser than Parker's, requiring that the period of oscillation be longer than the period, over which we expect an approximate constant magnetic field (the alignment time for interstellar grains  $\tau \geq 2.4 \times 10^5$  yr is taken in ref.17). It is argued in ref.18 that the magnetic field undergoes the decay only resonantly and the decay does not occur for a large monopole flux. Nevertheless the Parker's limit on the cosmic monopole flux fascinates experimentalists who are searching for monopoles because of the simplicity of the argument, and perhaps of the fact that the limit comes just at a value which they can reach with a reasonable effort.

*Constraints on the intracluster monopole flux.* A similar argument, albeit with more uncertainties, also applies to the monopole flux in the intracluster space. Using  $B=0.01-0.1\mu$  Gauss and the dynamical time  $\tau=10^9$  yr for the cluster, one may obtain<sup>[19]</sup>

$$F_m (\text{intracluster}) \leq 10^{-18} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} . \quad (5)$$

*Constraints from the mass density.* The condition that monopoles should not give a cosmological mass density larger than the critical density of the universe  $\rho_c = 3H_0^2/8\pi G = 10.5 h^2 \text{ keV/cm}^3$  ( $h$  is the Hubble constant  $H_0$  in the unit of  $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ) leads to the limit,

$$F_m \leq 5 \times 10^{-15} (M_m/10^{16} \text{ GeV})^{-1} (\beta/10^{-3}) h^{-2} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} , \quad (6)$$

when monopoles are uniformly distributed in the universe. If  $\beta < 10^{-3}$ , it is likely that monopoles are clumped in the galaxy, and the bound on the galactic monopole abundance is loosened. In this case we may use the constraint that the monopole mass should not exceed the dynamical mass of the galaxy  $\sim 10^{12} M_\odot$ , giving

$$F_m \lesssim 3 \times 10^{-10} (M_m / 10^{-16} \text{ GeV})^{-1} (\beta / 10^{-3}) \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} . \quad (7)$$

It is interesting to note that if  $F_m$  is in the range  $(0.6-3) \times 10^{-10} (M_m / 10^{-16} \text{ GeV})^{-1} (\beta / 10^{-3})$ , the galactic halo<sup>[20]</sup> may be dominated by monopoles. Cabrera noted that this might happen for  $M_m \sim 10^{-16} \text{ GeV}$  taking his candidate event as a monopole<sup>[2]</sup>. This, however, largely violates the Parker limit. The possibility opens only when  $M_m \geq 10^{19} \text{ GeV} = M_{pl}$ . (Authors in ref.17 and 18 argued the possibility of the monopole halo for  $M_m > 10^{17} - 10^{18} \text{ GeV}$ .)

*Constraints on monopole abundance in stars.* We may also apply Parker's argument to the magnetic field in stars. With a supplementary assumption that the monopole has a "virialised" velocity  $v \sim (\frac{1}{2} M_m v^2 = GM_m / R_0 \approx g B R_0)$ , we are led to a constraint on the monopole abundance in stars, and also to that on the monopole flux if the monopole-trapping power of stars (see Sec.6) is known. The results thus derived are:

	B	$\tau$	$n_M (\text{gr}^{-1})$
the sun <sup>[21]</sup>	$10^2 - 10^3$ Gauss	10 yr	$< 1/2 \times 10^7$
the Earth <sup>[21]</sup>	$\sim 10^2$ Gauss	$10^4$ yr	$< 1/6 \times 10^9$
$A_p$ stars <sup>[22]</sup>	$10^3 - 10^4$ Gauss	$\sim 5 \times 10^8$ yr	$< 1/1.5 \times 10^{13}$ .

Here  $A_p$  (peculiar A star) is a star, in which the direction of magnetic field is opposite to that of rotation of the star, and hence the magnetic field is supposed to be frozen when the star is formed<sup>[23]</sup>. (The author in ref.22 made a more elaborate argument.)

For monopoles captured in the earth there is also a constraint derived in a different way<sup>[24]</sup>: It is known that the terrestrial magnetic field reverses irregularly with epochs lasting typically a  $10^6$  years. In the time of reversal ( $\sim 10^3$  yr) a monopole and an antimonopole annihilate each other producing a heat. The constraint that this should not heat the earth too much gives  $n_M \lesssim 5 \times 10^{-5} \text{ gr}^{-1}$ . This limit, however, is considerably weaker than that discussed above.

Finally we mention a recent experiment<sup>[25]</sup> using a geochemical method searching for a monopole track in a rock<sup>[26]</sup>, by assuming that a some nucleus (Al, Mn etc.) will be trapped by a monopole while its traversing the earth's crust [see also Sect.3]. If this indeed happens, the nucleus-monopole

composite undergoes nuclear collision, while passing through a rock, resulting in the formation of a trail of lattice defects. The authors claimed that the search for tracks in  $4.6 \times 10^8$  yr-old mica (muscovite) places an upper limit of  $F_m < 10^{-17} \cdot 10^{-16} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$  for monopole with velocities around  $\beta \approx 3 \times 10^{-4} - 10^{-3}$ . For monopoles moving more slowly, the diamagnetic repulsion due to atomic electrons greatly suppresses the probability of forming a bound state with a nucleus. If a monopole would have captured a proton or electron before reaching to the earth, the capture cross section of a nucleus could be different from that used to estimate the flux.

### 3. Rubakov effect in matter

To discuss the Rubakov effect<sup>[6]</sup> in laboratories and in astrophysical environments, we have to know the behaviour of a monopole in matter. The Rubakov effect occurs when the monopole comes sufficiently close to the nucleon. In matter, however, the probability of a slowly-moving monopole coming close to the nucleus is greatly suppressed by repulsive forces with two different origins<sup>[27]</sup>. The first is due to the fact that the monopole-nucleus system carries an extra angular momentum  $q = (eg/4\pi)Z = Z \times (\pm 1/2, \pm 1, \dots)$  ( $Z$ =charge of the nucleus). The wave function for the lowest angular momentum state behaves as  $\sim (\beta/\beta_0)^v$  near the origin with  $v = -1/2 + (1/4 + |q|)^{1/2}$  for a spinless nucleus. For an  $s \neq 0$  nucleus, it follows that

$$v = -\frac{1}{2} + \left[ \frac{1}{4} + |q|(1-2(1+\kappa)s) \right]^{1/2} , \quad (9)$$

where  $\kappa$  is the anomalous magnetic moment of the nucleus,

$$\kappa = (\mu/\mu_B)(A/Z)(1/2s) - 1$$

( $\mu_B$ =Bohr magneton). Therefore the cross section for monopole-nucleus scattering receives a velocity-dependent factor

$$F(\beta) \sim (\beta/\beta_0)^{2\text{Re } v} , \quad \beta_0 \sim 1/[r_0^A m_A^{1/3}]$$

with  $r_0 \sim 1.2 \text{ fm}$ . Taking the  $\beta$ -dependence for the exothermic reaction  $\bar{\sigma} \sim 1/\beta$ , we obtain for the Rubakov process,

$$\sigma_R \sim F(\beta) \bar{\sigma} \sim \bar{\sigma} (\beta/\beta_0)^{2\text{Re } v} \sim \beta^{-2\text{Re } v - 1} . \quad (10)$$

When  $\nu < 0$ , as is for the most of familiar nuclei, we expect a suppression. Some examples are:

				F( $\beta$ )		
	Z	2 ReV	$\beta_0$	$\beta=10^{-3}$	$\beta=5 \times 10^{-4}$	$\beta=10^{-4}$
$^4\text{He}$	2	1.236	0.0275	0.017	0.0071	$9.7 \times 10^{-4}$
$^{16}\text{O}$	8	3.123	0.00434	0.0098	0.0012	$7.7 \times 10^{-6}$
$^{28}\text{Si}$	14	4.385	0.00206	0.042	0.0020	$1.7 \times 10^{-6}$
$^{40}\text{A}$	18	5.082	0.00128	—	0.0084	$2.4 \times 10^{-6}$
$^{56}\text{Fe}$	26	6.280	0.00082	—	0.046	$1.9 \times 10^{-6}$

The Rubakov effect is largely suppressed in matter with heavy elements for the slowly-moving monopole. For example, an iron detector is not sensitive to the Rubakov effect for  $\beta < 5 \times 10^{-4}$ . In a water detector<sup>[28,29]</sup> only the hydrogen component (2/18) is sensitive to a monopole with  $\beta < 10^{-3}$ . Namely multiple catalysed proton decays are expected in a water detector with a 15 m depth when  $\sigma_R$  (monopole + H)  $\geq 10$  mb rather than  $\geq 1$  mb.

The nuclei with  $\kappa > 0$  are rather exceptional. Of them familiar examples are  $^{27}\text{Al}$  (abundance ~1.5% in the earth),  $^{19}\text{F}$ ,  $^{55}\text{Mn}$  and hydrogen. For these nuclei  $\text{Re } \nu = -1/2$ , which leads to an "enhancement" factor for monopole-nucleus scattering. The cross section reads

$$\sigma_R \sim 1/\beta^2. \quad (10)$$

We also have  $\text{Re } \nu = -1/2$  for a neutron, and the cross section takes this form. For monopole-hydrogen scattering this is the  $\beta$ -dependence obtained in the relativistic calculation<sup>[30]</sup>. If the Rubakov effect does not happen at all, the nucleus with  $\kappa > 0$  may form a stable bound state with a monopole<sup>[30,31]</sup> by an attractive force, once the monopole approaches sufficiently close to the nucleus.

Even in the case when the attractive force is present between the monopole and nucleus, there is yet another suppression for a slowly-moving monopole approaching the atom from infinity: The effect of the monopole magnetic field on the atomic electrons induces a repulsive force between the monopole and the atom. The repulsion potential is  $\Delta E = nZ^2 \text{Ry}$  with  $\text{Ry}$  the Rydberg constant and  $n$  a fractional number that depends on the element<sup>[32]</sup>. For helium atom, for

instance,  $\Delta E_{\text{He}} \sim 16 \text{ eV}$ <sup>[33]</sup>, and hence a monopole with  $\beta < 10^{-4}$  can hardly approach the helium nucleus. This threshold velocity increases as  $\sim Z/A^{1/2}$  for heavier atoms, and this repulsion force greatly suppresses the Rubakov effect to happen for  $\beta < 10^{-4}$  even for the element with  $\kappa > 0$ .

An exception is the case for hydrogen atom. There exists the ground state, the energy of which is not affected when a monopole approaches the proton<sup>[33]</sup>, and hence there appears no barrier factor.

#### 4. Rubakov effect in ordinary stars

*Monopoles in the Earth.* The velocity of monopoles which would have been captured in the Earth is  $\beta < 3 \times 10^{-5} (M_m/10^{16} \text{ GeV})^{-1}$  (see Sect.6). The thermal velocity of nuclei at the centre of the Earth is  $\beta_{\text{th}} \sim 2.7 \times 10^{-5}$  ( $T = 4 \times 10^3 \text{ K}$ ). Thus monopoles trapped in the Earth do not approach nuclei, and the Rubakov process is strongly suppressed. Therefore, contrary to the claim made in ref.34, we do not obtain any bound on monopole abundance from the heat flow<sup>[27]</sup>. [The strongest upper bound so far obtained for the Earth remains to be  $n_m < 1/6 \times 10^9 \text{ gr}$ <sup>[21]</sup>. See (8).]

*Monopoles in the sun.* In hydrogen we expect the Rubakov process to happen effectively with the cross section

$$\sigma \sim \sigma_0 \beta^{-2} \quad (11)$$

with  $\sigma_0$  the high energy cross section of the order typical of strong interactions. ( $\sigma_0$  could be suppressed to the order of  $(100 \text{ GeV})^{-2}$ . This suppression will not alter the arguments given below, however.) The cross section (11) may apply until it grows up for a small  $\beta$  to  $\sigma = \pi d^2$  ( $d$ =mean distance between atoms in matter), beyond which many-body effects will be important and cut-off effects may start to work.

The prime candidate for stars rich in hydrogen is the sun<sup>[35]</sup>. The frequency of the Rubakov effect in the sun is given by

$$\begin{aligned} f &= \int n_m n_H (\sigma_R v_{\text{rel}}) d^3x \\ &= c \sigma_0 \left( \int n_m \frac{\rho_H}{\beta} d^3x \right) \cdot (6 \times 10^{23} / \text{cm}^3) \end{aligned} \quad (12)$$

where  $\beta_{\text{th}} = (1-1.7) \times 10^{-3}$  ( $T = 10^6 - 1.6 \times 10^7 \text{ K}$ ). We see in a typical calculation of the solar interior<sup>[36]</sup> that  $\rho_H/\beta$  depends only weakly on the position in the sun for  $0 < M_r < 0.5$  ( $M_r$  is the fractional mass in the unit of  $M_\odot$ ) and it

varies at most by a factor 3 in this range of  $M_L$ . Therefore we have

$$f = (6 \times 10^{23} / \text{cm}^3) \cdot N_m \langle \sigma_0 \langle \rho_H / B \rangle \rangle, \quad (13)$$

with  $N_m$  the total number of monopoles  $\int d^3x n_m$ .

The strongest constraint on the monopole abundance in the sun is derived in a consideration of the neutrino flux<sup>[35]</sup>. Since the SU(5) monopole appears in the  $(d_3^c, e^-)$  SU(2) sector, the monopole catalysis  $m+p \rightarrow m+\bar{\nu}+\pi^+$  is forbidden<sup>[37]</sup>. The allowed  $\Delta B \neq 0$  processes are:  $m+p \rightarrow m+e^+$ ,  $m+p \rightarrow m+e^++\pi^0$ ,  $m+p \rightarrow m+e^++(\rho^0, \eta, \omega)$  etc. In the third process  $\rho^0$ ,  $\eta$  and  $\omega$  immediately decay into pions. Then  $\pi^+$  is absorbed by hydrogen before the decay, but  $\pi^+$  thus produced eventually decays into  $\mu^+$ , and  $\mu^+$  in its turn decays into  $\bar{\nu}_\mu + e^+ + \nu_e$  mostly after it is stopped. (Near the centre the absorption of  $\pi^+$  by helium reduces the  $\mu^+$  by a factor of 2.) We therefore expect the  $\nu_e$  flux with the average energy  $\langle E_\nu \rangle = 35$  MeV eventually arising from the Rubakov process. (The  $\nu_\mu$  and  $\bar{\nu}_\mu$  flux with  $\langle E_\nu \rangle = 30-35$  MeV also arises, but no  $\bar{\nu}_e$  flux is anticipated.) There is yet another process which gives  $\nu_e$ ,  $m+p \rightarrow m+\mu^++K^0$ ,  $K^0 \rightarrow (K_L^0, K_S^0)$ , and  $K_L^0 \rightarrow \pi^- + e^+ + \nu_e$ . The neutrino that appears in this process has a higher energy, but the expected flux is very small because of a strong  $K_S^0$  regeneration effect in the sun. This process is not important, unless the  $\mu^+ K^0$  mode dominates the proton decay.

Using the result of solar neutrino experiment with a  $C_2Cl_4$  detector observing  $\nu_e + {}^{37}\text{Cl} \rightarrow e^- + {}^{37}\text{A}^*$ <sup>[38]</sup>, we may obtain a limit on the  $\nu_e$  flux that might arise in the Rubakov process, and hence on the monopole abundance in the sun. The capture cross section increases as  $E_\nu$ ,  $\sigma_{\text{cap}} \propto \sum_i (E_\nu - \epsilon_i)^2 |M_i|^2$  and is computed to be

$$\sigma_{\text{cap}}(\nu_e + {}^{37}\text{Cl}) = 8.5 \times 10^{-41} \text{ cm}^2 \quad (14)$$

for  $\nu_e$  from the  $\mu^+$  decay using the Bahcall's matrix element<sup>[39]</sup>. [A shell model calculation gives  $\sigma_{\text{cap}} \approx 9.9 \times 10^{-41} \text{ cm}^2$ <sup>[40]</sup>.] Allowing for the excess capture rate  $1\text{SNU} = 10^{-36}$  caps/atom·s, we obtain a limit on the Rubakov neutrino flux on the earth

$$J_{\nu_e} \leq 1.2 \times 10^6 / \text{cm}^2 \cdot \text{s}, \quad (15)$$

and hence,

$$n_m \leq (1/4 - 8 \times 10^{12} \text{ gr}) \cdot (\sigma_0 / 10^{-27} \text{ cm}^2)^{-1} (B_L / 0.5)^{-1}, \quad (16)$$

$B_r$  being the average number of  $\pi^+$  in the catalysed proton decay (we expect in SU(5)  $B_r = 0.5$  in the proton decay without a monopole. The presence of  $m+p \rightarrow m+e^+$  may suppress  $B_r$  by a factor of 2<sup>[41]</sup>). Corresponding to this constraint the luminosity excess due to monopoles does not exceed  $L_m \leq 10^{-5} L_\odot$ , and the disturbance of magnetic field at the surface  $\Delta B|_{R=R_\odot} \leq 10^{-6}$  Gauss. These values are small enough and do not cause any effects on the evolutionary scenario of the sun. We notice that this bound (16) is stronger than any other bounds (see (8)), unless  $\sigma_0 < 10^{-7} \times (10^{-27} \text{ cm}^2)$ .

We now discuss a possibility to detect a Rubakov neutrino, or at least to improve the upper limit using a massive detector in the underground experiment<sup>[42,43]</sup>. The Rubakov process in the sun corresponding to the limit (15) and (16) produces 120  $e^-$  events/1000 ton·yr in a water detector, as

$$\begin{aligned} 70 \text{ events} & \text{ for } \nu_e + {}^{16}\text{O} \rightarrow e^- + {}^{17}\text{F} \quad (\text{ref.44}) \\ 40 \text{ events} & \text{ for } \nu_e + e^- \rightarrow e^- + \nu_e \\ 10 \text{ events} & \text{ for } \begin{pmatrix} - \\ \nu \end{pmatrix} + e^- \rightarrow e^- + \begin{pmatrix} - \\ \nu \end{pmatrix} \end{aligned}$$

On the other hand, atmospheric neutrino events ( $E_\nu \leq 50$  MeV) (mainly  $\bar{\nu}_e + p \rightarrow e^+ + n$ ) do not exceed 0.2 events/1000 ton·yr at Kamioka, and 1.5 events/1000 ton·yr at Ohio<sup>[45]</sup>. Then, in principle, it seems possible to measure the monopole abundance down to

$$n_m \approx (1/1 \times 10^{15} \text{ gr}) (\sigma_0 / 10^{-27} \text{ cm}^2)^{-1} (B_r / 0.5)^{-1}. \quad (17)$$

In an iron detector, the large capture cross section ( $\sigma_{\text{cap}} = (0.17-0.69) \times 10^{-39} \text{ cm}^2$ <sup>[46]</sup>), increases the sensitivity, and 700-2800 events/1000 ton·year are expected corresponding to (16). In a water detector we may increase the sensitivity by making use of NaCl solution, because of a large capture cross section of Cl due to the presence of the analogue state in A [20 times larger than  $\sigma(\nu_e + {}^{16}\text{O})$ , see (14)]. With 35% NaCl solution we expect three times more events in the same detector.

*Monopoles in the Jovian planets.*<sup>[42]</sup> Another candidate for stars rich in hydrogen is the Jovian planets (Jupiter, Saturn, Uranus and Neptune). It has long been known from the infrared observation and the Bond albedo measurement<sup>[47,48]</sup> that there is an intrinsic heat generation in these planets. The magnitude of this heat for Jupiter and Saturn is now measured more precisely through the nearby-flight of Pioneer 10/11<sup>[49]</sup> and Voyager I<sup>[50]</sup>.

The intrinsic heat for Jupiter and Saturn is  $(1.5-1.8) \times 10^{-6}$  erg/g·sec, almost two orders of magnitude larger than the heat generation in the earth  $\approx 5 \times 10^{-8}$  erg/g·sec, which is usually attributed to the activity of Th, U and  $^{39}\text{K}$ . The current view<sup>[48]</sup> ascribes the source of this intrinsic heat to gradual release of the gravitational energy liberated at the birth of those planets. This scenario, however, fails to explain coherently the magnitude of the excess heat in each planet and must invoke additional stories; downwards migration of helium in Saturn and upwards convective transport of heavy elements in Neptune etc<sup>[48]</sup>.

Here let us ask whether there is a possibility that this excess heat (or a part of it) could be attributed to the Rubakov process taking place in the Jovian planets. If the monopole distribution is reasonably uniform, eq.(13) also applies to this case. In Jupiter thermal momentum is still larger than the Fermi momentum or zero point oscillation momentum of the lattice. We then use  $\beta \approx \beta_{th}$ , and we see that

$$n_m \approx (1/3-8 \times 10^{14} \text{ gr})(\sigma_0/10^{-27} \text{ cm}^2)^{-1} \quad (18)$$

is required to account for the whole intrinsic heat of Jupiter. Contrary to this, if monopoles behave like an ideal gas in the gravitational field of the planet, the monopoles, in the absence of internal magnetic field, are concentrated in the core. The core is supposed to consist of rock and ice (mainly  $\text{H}_2\text{O}$ ,  $\text{CH}_4$  and  $\text{NH}_3$ ). Using an estimate ice/rock  $\approx 14/5$ <sup>[51]</sup>, we suppose that 10% of the core consists of hydrogen. Therefore monopole density required in this case is

$$n_M \approx 1/(1-2 \times 10^{14} \text{ gr})(\sigma_0/10^{-27} \text{ cm}^2)^{-1} \quad (19)$$

In the presence of a strong magnetic field ( $\sim 1000$  Gauss) the required density is well in the middle of these two typical cases.

For Saturn the argument follows similarly and numbers similar to (18) and (19) are obtained. For Uranus and Neptune the intrinsic heat generation per unit mass is an order of magnitude smaller than that for Jupiter and Saturn ( $\sim 0.2 \times 10^7$  erg/g·s for Neptune and  $\leq 0.2 \times 10^7$  erg/g·s for Uranus<sup>[48]</sup>). A principal part of their interiors consists of an ionic ocean of  $\text{H}_3\text{O}^+\text{OH}^-$  (with dissolved  $\text{NH}_3$ )<sup>[52]</sup>, and the hydrogen component is  $\sim 4/36$ . We then obtain the monopole density  $n_m \approx (1/1-2 \times 10^{15} \text{ gr})(\sigma_0/10^{-27} \text{ cm}^2)^{-1}$  for these planets.

The monopole densities required to heat up the Jovian planets are consistent with each others. We stress that they are about two orders of magnitude

smaller than the limit (16), the strongest limit so far obtained for the object in the solar system. In the last section we show that the monopole density in Jupiter may not be very much different from that in the sun (they differ at most by an order of magnitude; see (27) below). Therefore the possibility that the Rubakov effect heats up the Jovian planets is not yet reprehensible. Furthermore if monopoles are present in the sun with the density (18)-(19), the Rubakov neutrino from the sun would be detectable in the underground experiment [see (17)].

So far we have made an argument, by assuming that the whole intrinsic heat is to be ascribed to the monopole, for simplicity. On the other hand, the monopole heat does not cause a significant effect on the thermal evolution of the Jovian planets, if the monopole density is slightly less. This is easily seen by employing the adiabatic-convective cooling model<sup>[48,53]</sup>. The thermal evolution equation with monopole heat is given by

$$4\pi R^2 \sigma (T_s^4 - T_m^4) = - \int dm \left( \frac{dE}{dt} - \frac{P}{2} \frac{d\rho}{dt} \right)$$

where  $4\pi R^2 \sigma T_s^4 = (1 - \text{Bond albedo}) \times (\text{solar energy flux})$  is the solar heat absorbed in Jupiter and  $4\pi R^2 \sigma T_m^4 = L_m$  is the monopole heat. Using the equation of the state for Jupiter, the evolution equation reads

$$dt = -\alpha T^{-3.757} [1 - (T_s^4 + T_m^4)/T^4]^{-1} dT \quad .$$

After integration we see that the cooling time of Jupiter after inclusion of the monopole heat does not differ more than ten percent from that without the monopole heat, if the monopole density is by a factor of 3 less.

In table 1 we summarise, for convenience, the limit on the monopole density for various objects in the solar system. Direct comparison of the result, of course, is not meaningful.

#### 5. Rubakov effect in neutron stars and the limit on the cosmic monopole flux

Strong bounds have been derived on the cosmic monopole flux from the excess limit on the X-rays from neutron stars<sup>[8,9,54,55]</sup>. In the neutron star the Rubakov process releases energy at a rate

$$L = m_N n \sigma_{NR} v_{rel} N_m \quad (20)$$

where  $v_{rel} \sim v_F = 0.37(\rho_N/3 \times 10^{14} \text{ gr})^{1/3}$  is large and the velocity dependence of the cross section is not important. (Nevertheless we still use (11), i.e.,  $\sigma v = c\sigma_0/\beta_F = 2.7c\sigma_0$  for consistency of our notation.)

When the monopole hits the neutron star, the monopole readily loses the energy because of a high electron density<sup>[56]</sup>  $[dE/dx \sim 4\pi^2 n_e (e\beta)^2 p_e^{-1} \beta^{-11} (\text{GeV/cm})\beta]$  and it will easily be captured by the neutron star. The total number of monopoles accumulated in the neutron star in the period  $\tau$  is given by

$$N_m = 4\pi F_m \tau R^2 \left[ 1 + \left( \frac{\beta_{esc}}{\beta} \right)^2 \right] \cdot \tau, \quad (21)$$

provided that the monopole-antimonopole annihilation does not take place significantly. For  $R \approx 10 \text{ km}$  and  $M \approx M_\odot$ ,  $N_m \approx 4 \times 10^{36} (F_m/\text{cm}^2 \cdot \text{s} \cdot \text{sr}) \cdot (\beta/10^{-3})^{-2}$ . In the presence of sufficient numbers of monopoles neutron stars will be X-ray sources as a result of monopole catalysed nucleon decay. An upper limit on the monopole flux is obtained, if one finds a limit on the total luminosity  $L$ . Important uncertainties arise not only in the procedure to find the source X-ray luminosity  $L_Y$ , but also in the relation between  $L$  and  $L_Y$  when  $L_Y$  exceeds  $\sim 10^{31} \text{ erg} \cdot \text{s}^{-1}$ <sup>[57]</sup>: The relation depends much on the equation of state used for the neutron star. (If the pion condensation does not occur,  $L \approx L_Y$  holds up to  $L_Y \leq 10^{33} \text{ erg} \cdot \text{s}^{-1}$ <sup>[57]</sup>.) Let us summarise the arguments that were used to give a limit on  $F_m$  and some critiques on them (we hereafter assume  $\beta = 10^{-3}$ ): (1) The negative result of Einstein Observatory serendipitous searches for discrete X-ray sources<sup>[58]</sup>, when combined with the expected number density of neutron stars,  $n_{NS} \geq 4 \times 10^{-3} \text{ pc}^{-3}$ , leads to the luminosity limit  $L_{dis}^Y \leq 10^{-31} \text{ erg} \cdot \text{s}^{-1}$ . Taking the neutrino luminosity into account,  $L \leq 10^{-33} \text{ erg} \cdot \text{s}^{-1}$  leads to<sup>[8]</sup>

$$F_m \leq 1 \times 10^{-23} (\sigma_0/10^{-27} \text{ cm}^2)^{-1} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}. \quad (22)$$

A critique<sup>[54]</sup> to this argument is that the significant absorption for soft X-rays in the interstellar matter [absorption length:  $\lambda_{abs} \approx (6 \text{ pc}) [E/0.1 \text{ keV}]^3 (n_H/\text{cm}^{-3})^{-1}$ <sup>[54]</sup>] reduces the sight to 100 pc (without absorption it is 1 kpc) and the limit would be evaded if the number density of neutron stars ( $n_{NS}$ ) is an order of magnitude less.

(2) The X- and UV-rays emitted by neutron stars contribute to the diffuse X- and UV-ray background. Dimopoulos et al.<sup>[9]</sup> obtained a constraint ( $L_Y < 3 \times 10^{30} \text{ erg} \cdot \text{s}^{-1}$ ) from the measured diffuse X-ray background<sup>[59]</sup>, assuming  $n_{NS}$  similar to the above value. Using our  $\sigma_0$ , their bound is

$$F_m \leq 2 \times 10^{-26} (\sigma_0/10^{-27} \text{ cm}^2)^{-1} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}. \quad (23)$$

Kolb et al.<sup>[8]</sup> estimated the luminosity limit  $L_Y \leq 2 \times 10^{32} \text{ erg} \cdot \text{s}^{-1}$  from the limit on the total power radiated during the lifetime of the neutron star<sup>[60]</sup>, assuming also the birth rate ( $\sim 1/100 \text{ yr}$ ). Allowing for the neutrino luminosity,  $L \leq 10^{-36} \text{ erg} \cdot \text{s}^{-1}$  leads to

$$F_m \leq 1 \times 10^{-20} (\sigma_0/10^{-27} \text{ cm}^2)^{-1} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}. \quad (24)$$

The absorption of X-(and UV-)rays by the interstellar matter will considerably increase the upper bound on  $L_Y$ . This is particularly true for the former analysis<sup>[9]</sup>, while the cut-off of soft X-rays ( $E_Y > 0.2 \text{ keV}$ )<sup>[60]</sup> makes the absorption correction small for the latter<sup>[8]</sup>. On the other hand a decrease of  $n_{NS}$  causes an amplified luminosity ( $L$ ) increase of  $\sim n_{NS}^{-3}$  (not  $\sim n_{NS}^{-1}$ ). If  $L_Y \sim 10^{33} \text{ erg} \cdot \text{s}^{-1}$  would be allowed (a factor 1.7 less in  $n_{NS}$ , say), the bound would be loosened to

$$F_m \leq 1 \times 10^{-18} (\sigma_0/10^{-27} \text{ cm}^2)^{-1} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}. \quad (24a)$$

(3) Freese et al.<sup>[54]</sup> obtained a limit on the excess luminosity using an observed (young) pulsar PSR1929+10 ( $\tau = 3 \times 10^6 \text{ yr}$ ). From the estimate of photon luminosity  $L_Y = 2.6 \times 10^{30} (R/15 \text{ km})^2 \text{ erg} \cdot \text{s}^{-1}$ , the limit on the monopole flux is

$$F_m \leq 3 \times 10^{-23} (\sigma_0/10^{-27} \text{ cm}^2)^{-1} \left( \frac{R}{15 \text{ km}} \right)^4 \left( \frac{M}{1.4 M_\odot} \right)^{-2} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}. \quad (25)$$

The recent parallax measurement of PSR1929+10 suggests, however, that the distance to this pulsar is as far as  $\approx 250 \text{ pc}$ <sup>[61]</sup> rather than  $\approx 60 \text{ pc}$ <sup>[62]</sup> they used. If we adopt this new distance,  $L_Y$  is more than  $5 \times 10^{31} (R/15 \text{ km})^2$  and the neutrino luminosity would dominate ( $L \sim 10^{34} \text{ erg} \cdot \text{s}^{-1}$  is allowed for  $R = 15 \text{ km}$ ). The bound will then be loosened and it is at most

$$F_m \leq 6 \times 10^{-19} (\sigma_0/10^{-27} \text{ cm}^2)^{-1} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}. \quad (25a)$$

In summary, although most of the reported upper limits on the monopole flux is as stringent as  $F_m < 10^{-22} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$  for  $\sigma_0 = 10^{-27} \text{ cm}^2$ , the safe limit obtained after allowing for various uncertainties is much loose and it is of the order of

$$F_m \leq 10^{-18} (\sigma_0/10^{-27} \text{ cm}^2) \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} . \quad (26)$$

With the present knowledge on the neutron star interior, it seems difficult to conclude whether the monopole-antimonopole annihilation would significantly reduce the monopole number captured in the neutron star. We should bare in mind that monopoles could well be concentrated in the small region near the centre under the very strong gravitational field, if the neutron star is normal conducting, and the annihilation could reduce the number<sup>[63]</sup>.

## 6. Local monopole flux

There is a suggestion that the local monopole flux might result from a diffuse cloud of monopoles which are in newtonian orbits about the sun, and the local flux may be significantly enhanced<sup>[21]</sup>. A model calculation shows that an enhancement up to  $\sim 0[50(M_m/10^{16} \text{ GeV})^2]$  is possible<sup>[64]</sup> near the sun. In any case the local monopole flux may well be different from the average flux in the galaxy.

For ordinary stars whether the monopole which hits the star will stop or not depends on the velocity and mass of monopole. The Eddy current energy loss for the sun is<sup>[65]</sup>

$$\frac{dE}{dx} = -(10 \sim 100) \text{ GeV cm}^2 \text{ g}^{-1} \beta \rho$$

depending on the position in the sun. A monopole with the velocity  $\beta \sim 10^{-3}$  that particularly concerns us will be stopped in the sun. Assuming a similar energy loss for Jupiter, it is shown that monopoles moving more slowly than  $\beta \sim 10^{-3} (M_m/10^{16} \text{ GeV})^{-1}$  will stop even in Jupiter<sup>[34]</sup>. A more elaborate calculation for the deceleration of monopoles in a plasma<sup>[66]</sup> shows that a monopole will stop in the sun in a distance  $\approx 0.01 R_\odot (M_m/10^{16} \text{ GeV}) (\beta/10^{-3})$ . So far as the monopole stops, eq.(21) applies for the total number of monopoles captured in stars. In the ordinary star the gravitational field is not strong, and a care about the monopole-antimonopole annihilation is not necessary.

Thus the ratio of monopole density captured in the sun and Jupiter is

$$\frac{n_m(\odot)}{n_m(\text{J})} \sim 0.5 \quad (27)$$

for  $M_m \sim 10^{16} \text{ GeV}$  and  $\beta \sim 10^{-3}$ . We expect that the monopole densities for these stars are not very much different.

One can also obtain a limit on the local monopole flux from (16),

$$F_m \leq 1.2 \times 10^{-21} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} (\sigma_0/10^{-27} \text{ cm}^2)^{-1} (\text{Br}/0.5)^{-1} , \quad (28)$$

which is comparable to the limit from the excess luminosity limit of neutron stars. The monopole flux which provides the monopole abundance required to account for the excess heat in Jupiter is<sup>[42]</sup>,

$$F_m \sim 1 \times 10^{-23} (\sigma_0/10^{-27} \text{ cm}^2)^{-1} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1} , \quad (29)$$

or less when there are some monopoles in rocks which would be captured at the time of formation of Jupiter. It is most likely that this flux is smaller than the limit obtained from neutron stars.

We also note here that the search for the solar Rubakov neutrino could explore the monopole flux down to<sup>[42,43]</sup>

$$F_m \sim 6 \times 10^{-24} (\sigma_0/10^{-27} \text{ cm}^2)^{-1} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1} , \quad (30)$$

corresponding to (17). It has been thought that any direct search for monopoles on the earth is by no means possible, if the monopole flux is as small as that originally derived from neutron stars. Searching for the neutrino flux from the sun, albeit not direct, will provide us with a unique method, by using the sun as a collector, to search for monopoles at a prohibitively small flux.

A situation for the limit on the monopole flux is summarised in fig.1.

I have benefited from conversations with J. Arafune, K. Hayashi, H. Sato, H. Suda and S. Yanagita on the subjects in this talk. I am also grateful to Professor D.J. Stevenson for his useful correspondence.



## References

1. G. 't Hooft, Nucl. Phys. B79, 276 (1974);  
A.M. Polyakov, Zh. Eksp. Teor. Fiz. Pis'ma 20, 430 (1974) [JETP Lett. 20, 194 (1974)].
2. B. Cabrera, Phys. Rev. Lett. 48, 1378 (1982).
3. C.P. Dokos and T.N. Tomaras, Phys. Rev. D21, 2940 (1980);  
A.N. Schellekens and C.K. Zakos, Phys. Rev. Lett. 50, 1242 (1983).
4. Ya.B. Zel'dovich and M.Yu. Khlopov, Phys. Lett. 79B, 239 (1978).
5. J.P. Preskill, Phys. Rev. Lett. 43, 1365 (1979).
6. V.A. Rubakov, Inst. Nucl. Research (Moscow) preprint P-0211 (1981);  
Zh. Eksp. Teor. Fiz. Pis'ma 33 (1981) 658 [JETP Lett. 33 (1981) 644];  
Nucl. Phys. B203, 311 (1982);  
C.G. Callan Jr., Phys. Rev. D25, 2141 (1982); *ibid.* D26, 2058 (1982).
7. Y. Kazama, in this Proceedings.
8. E.W. Kolb, S.A. Colgate and J.A. Harvey, Phys. Rev. Lett. 49, 1373 (1982).
9. S. Dimopoulos, J. Preskill and F. Wilczek, Phys. Lett. 119B, 320 (1982).
10. e.g., A.H. Guth and S.-H.H. Tye, Phys. Rev. Lett. 44, 631 (1980);  
M.B. Einhorn, D.L. Stein and D. Toussaint, Phys. Rev. D21, 3295 (1980);  
P. Langacker and S.Y. Pi, Phys. Rev. Lett. 45, 1 (1980);  
M.B. Einhorn and K. Sato, Nucl. Phys. B180, 385 (1981);  
M. Izawa and K. Sato, Prog. Theor. Phys. 68, 1574 (1982).
11. E.N. Alexeyev et al., Lett. Nuovo Cim. 35, 413 (1982); in Proc. 18th Int. Cosmic Ray Conf., Bangalore 1983 vol.5;  
M.R. Krishnaswamy et al., *ibid.*;  
D.E. Groom et al., Phys. Rev. Lett. 50, 573 (1983);  
J. Bartelt et al., Phys. Rev. Lett. 50, 655 (1983);  
R. Bonarelli et al., Phys. Lett. 126B, 137 (1983);  
T. Mashimo et al., Phys. Lett. 128B, 327 (1983);  
T. Doke et al., Phys. Lett. 129B, 370 (1983);  
G. Tarlé, S.P. Ahlen and T.M. Liss, Phys. Rev. Lett. 52, 90 (1984);  
F. Kajino et al., Phys. Rev. Lett. 52, 1373 (1984).
12. B. Cabrera, M. Taber, R. Gardner and J. Bourg, Phys. Rev. Lett. 51, 1933 (1983).
13. E.N. Parker, Ap. J. 160, 383 (1970).
14. S.A. Bludman and M.A. Ruderman, Phys. Rev. Lett. 36, 840 (1976);  
G. Lazarides, Q. Shafi and T.F. Walsh, Phys. Lett. 100B, 21 (1981).
15. J.R. Jokipii and E.N. Parker, Ap. J. 155, 799 (1969);  
J.R. Jokipii, I. Lerche and R.A. Schonmer, Ap. J. 157, L119 (1969).
16. M.S. Turner, E.N. Parker and T.J. Bogdan, Phys. Rev. D26, 1296 (1982).
17. E.E. Salpeter, S.L. Shapiro and I. Wasserman, Phys. Rev. Lett. 49, 1114 (1982).
18. J. Arons and R.D. Blandford, Phys. Rev. Lett. 50, 544 (1983).
19. Y. Raphaeli and M.S. Turner, Phys. Lett. 121B, 115 (1983).
20. J.P. Ostriker and P.J.E. Peebles, Ap. J. 186, 467 (1973);  
S.M. Faber and J.S. Gallagher, Ann. Rev. Astron. Astrophys. 17, 135 (1979);  
D.J. Hegyi and K.A. Olive, Phys. Lett. 126B, 28 (1983).
21. S. Dimopoulos, S.L. Glashow, E.M. Purcell and F. Wilczek, Nature 298, (1982) 824.
22. D.M. Ritson, SLAC preprint PUB-2977 (1982).
23. G.W. Preston, Ap. J. 164, 309 (1971);  
J.D. Landstreet, E.F. Borra, J.R.P. Angel, R.M.E. Illing, Ap. J. 201, 624 (1975).
24. R.A. Carrigan Jr., Nature 288, 348 (1980).
25. P.B. Price, S. Guo, S.P. Ahlen and R.L. Fleischer, Phys. Rev. Lett. 52, 1265 (1984).
26. R.L. Fleischer, P.B. Price, R.T. Woods, Phys. Rev. 184, 1398 (1969).
27. J. Arafune and M. Fukugita, Phys. Rev. Lett. 50, 1901 (1983).
28. S. Errede et al., Phys. Rev. Lett. 51, 245 (1983);
29. T. Kajita, in this Proceedings.
30. W.J. Marciano and I.J. Muzinich, Phys. Rev. Lett. 50, 1035 (1983);  
See also Y. Kazama, C.N. Yang and A.S. Goldhaber, Phys. Rev. D15, 2287 (1977); V.A. Rubakov and M.S. Serebryakov, Nucl. Phys. B218, 240 (1983).
31. L. Bracci and G. Fiorentini, Phys. Lett. 124B, 29 (1983);  
J. Makino, M. Maruyama and O. Miyamura, Prog. Theor. Phys. 69, 1042 (1983);  
C.J. Goebel, Madison preprint MAD/TH-146 (1983);  
K. Olaussen, H.A. Olsen, I. Øverbø and P. Osland, Phys. Rev. Lett. 52, 325 (1984).
32. W.V.R. Malkus, Phys. Rev. 83, 899 (1951).
33. S.D. Drell, N.M. Kroll, M.T. Mueller, S.J. Parke and M.A. Ruderman, Phys. Rev. Lett. 50, 644 (1983).
34. M.S. Turner, Nature 302, 804 (1983).
35. J. Arafune and M. Fukugita, Phys. Lett. 133B, 380 (1983).
36. e.g., D.D. Clayton, *Principles of Stellar Evolution and Nucleosynthesis*, McGraw Hill, New York (1968).
37. J. Ellis, D.V. Nanopoulos and K.A. Olive, Phys. Lett. 116B, 127 (1982);  
F.A. Bais, J. Ellis, D.V. Nanopoulos and K.A. Olive, Nucl. Phys. B219, 189 (1983).

38. R. Davis Jr. et al., Phys. Rev. Lett. 20, 1205 (1968);  
J.N. Bahcall and R. Davis Jr., Science 191, 264 (1976);  
J.K. Rowley et al., BNL preprint 27190 (1980).
39. J.N. Bahcall, Phys. Rev. 135, B137 (1964).
40. T.W. Donnelly and W.C. Haxton, Nucl. Phys. A287, 506 (1977).
41. C. Schmid, Phys. Rev. D28, 1802 (1983).
42. J. Arafune, M. Fukugita and S. Yanagita, Kyoto University, RIFP preprint 531 (1983).
43. J. Arafune and M. Fukugita, to be published.
44. T.W. Donnelly, Phys. Lett. 43B, 93 (1973).
45. E.C.M. Young, in *Cosmic Ray at Ground Level*, The Institute of Physics, London (1973);  
T.K. Gaisser et al., Phys. Rev. Lett. 51, 223 (1983);  
A. Dar, Phys. Rev. Lett. 51, 227 (1983).
46. J.N. Bahcall and S. Frautschi, Phys. Rev. 136, B1547 (1964);  
N. Itoh and Y. Kohyama, Nucl. Phys. A306, 527 (1978).
47. D.H. Menzel, W.W. Coblenz and C.O. Lampland, Ap. J. 73, 177 (1926);  
H. Jeffreys, MNRAS 84, 534 (1924).
48. e.g., W.B. Hubbard, Rev. Geophys. Space Phys. 18, 1 (1980).
49. A.P. Ingersoll et al., Science 207, 439 (1980); in *Jupiter*, Univ. Arizona Press, Tucson (1976).
50. R. Hanel et al., Icarus 53, 262 (1983); J. Geophys. Res. 86, 8705 (1981).
51. W.B. Hubbard et al., J. Geophys. Res. 85, 5909 (1980).
52. e.g., D.J. Stevenson, Ann. Rev. Earth Planet. Sci. 10, 257 (1982).
53. W.C. De Marcus et al., Ap. J. 63, 2 (1958);  
H.C. Graboske et al., Ap. J. 199, 265 (1975);  
W.B. Hubbard et al., Icarus 30, 305 (1977).
54. K. Freese, M.S. Turner and D.N. Schramm, Phys. Rev. Lett. 51, 1625 (1983).
55. A.K. Drukier and G.G. Raffelt, Max-Planck-Institut preprint MPI-PAE/PTh 57/83 (1983).
56. V. Martem'yanov and S. Khakimov, Zh. Eksp. Teor. Fiz. 62, 35 (1972) [JETP 35, 20 (1972)].
57. K.A. van Riper and D.Q. Lamb, Ap. J. 244, L13 (1981).
58. F.A. Córdova, K.O. Mason and J.E. Nelson, Ap. J. 245, 609 (1981);
59. F. Paresce and P. Jakobsen, Nature 288, 119 (1980).
60. J. Silk, Ann. Rev. Astron. Astrophys. 11, 269 (1973).
61. D.C. Backer and R.A. Sramek, Ap. J. 260, 512 (1982).
62. M.J. Salter, A.G. Lyne and B. Anderson, Nature 280, 478 (1979).
63. V.A. Kuzmin and V.A. Rubakov, Phys. Lett. 125B, 372 (1983).
64. K. Freese and M.S. Turner, Phys. Lett. 123B, 293 (1983).
65. S.P. Ahlen and K. Kinoshita, Phys. Rev. D26, 2347 (1982).
66. A.J.S. Hamilton and C.L. Sarazin, Ap. J. 274, 399 (1983).
67. P.H. Eberhard, R.R. Ross, L.W. Alvarez and R.D. Watt, Phys. Rev. D4, 3260 (1971).
68. R.R. Ross, P.H. Eberhard, L.W. Alvarez and W.D. Watt, Phys. Rev. D8, 698 (1973).
69. T. Ebisu and T. Watanabe, Kobe Univ. preprint 83-11 (1983).

Table 1. Limits on the monopole density for various objects in the solar system.

	$n_m$	Refs.
meteorites (induction expt.)	$\leq 5 \times 10^{-4}/g$	67
lunar materials (induction expt.)	$\leq 1.7 \times 10^{-4}/g$	68
the earth (monopole-antimonopole annihilation heat)	$\leq 5 \times 10^{-5}/g$	24
iron ore (induction expt.; heating the sample above Curie temp.)	$\leq 2.3 \times 10^{-6}/g$	69
the sun (lifetime of magnetic field)	$\leq 5 \times 10^{-8}/g$	21
the earth (lifetime of magnetic field)	$\leq 2 \times 10^{-10}/g$	21
the sun (Rubakov neutrino, Davis' expt.)	$\leq 2 \times 10^{-13}/g$ $\times (\sigma_0/10^{-27} \text{ cm}^2)^{-1} (\text{Br}/0.5)^{-1}$	35
Jovian planets (Rubakov heat)	$\approx (1-2) \times 10^{-15}/g$ $\times (\sigma_0/10^{-27} \text{ cm}^2)^{-1}$	42

Figure captions

Fig.1. Limits on the cosmic monopole flux. The region below the curve is "forbidden". The notation shown are as follows: (1)  $dE/dX$ , counter experiments<sup>[11]</sup>; (2)  $I$ , induction experiments<sup>[12]</sup>; (3)  $G$ , geochemical measurements searching for monopole-nucleus composites<sup>[25]</sup>; (4) PARKER; lifetime of the galactic magnetic field<sup>[16]</sup>; (5)  $MD_1$ , cosmological mass density for unclumped monopoles; (6)  $MD_2$ , dynamical mass of galaxy (for clumped monopoles); (7)  $NS_1$ , excess X-ray luminosity of neutron stars (original limit)<sup>[8,9]</sup>; (8)  $NS_2$ , the "safe" limit from excess X-ray luminosity of neutron stars; (9) SOLAR  $\nu$ , Rubakov neutrinos from the sun<sup>[35]</sup> (using the Davis' expt.); (10)  $J$ , intrinsic heat generation in Jupiter<sup>[34,42]</sup>; (11) SOLAR  $\nu$  (future), an expected limit which may be accessible in underground experiments searching for Rubakov neutrinos<sup>[42,43]</sup>.

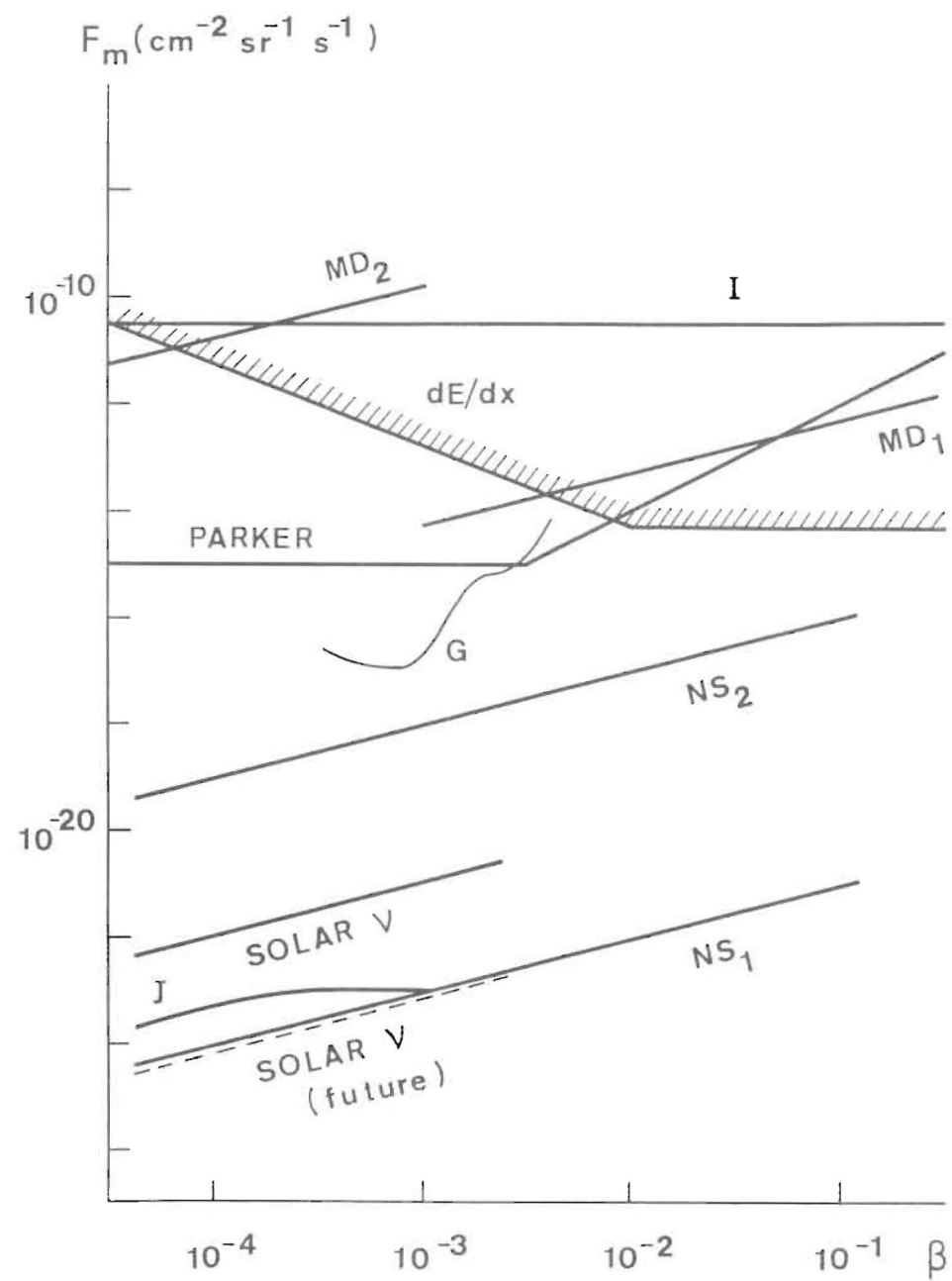


Fig. 1  
- 60 -

Issues in Monopole-Catalyzed Proton Decay\*

Yoichi Kazama

National Laboratory for High Energy Physics(KEK),  
Tsukuba, Ibaraki, 305 Japan

ABSTRACT

Some of the unsettled issues in monopole-catalyzed proton decay are discussed, following a short pedagogical review for non-experts.

\* Invited talk given at the Workshop on Grand Unification and Cosmology, Dec. 7-10, 1983, KEK, Tsukuba, Japan

It has already been about two years since Rubakov [1] and Callan [2] pointed out the fascinating possibility of monopole-catalyzed proton decay. Since then, numerous investigations have been made [3] and our understanding of the phenomenon has certainly been deepened. This, however, does not mean that we have the complete grasp of this intricate subject. There still remain a couple of important issues, clarification of which is needed in order to speak confidently of the "fast" decay of a proton in the presence of a monopole.

In this talk, I wish to discuss some of these issues, while recapitulating what we have learned in the past two years. To make the discussions understandable to non-experts, I shall first present the essence of why a GUT monopole may induce proton decay, with minimum of mathematics. I shall then go on to a critical reassessment of this "standard" story and discuss some of the unsettled issues. They include; the question of the boundary condition at the monopole core, the role of the weak interaction scale, and the effects of higher partial waves. The effects of the strong interaction, which is obviously the most difficult one to deal with, will not be discussed in this talk.

2. WHY CAN A GUT MONOPOLE CATALYZE PROTON DECAY?

Let me first describe the essence of why a GUT monopole may induce baryon number violating processes without suppression. Consider, for simplicity, a 't Hooft-Polyakov type SU(2) monopole [4], which is embedded in a peculiar way in the standard SU(5) GUT. Namely the generators,  $\vec{T}$ , of the SU(2) group (to be denoted by SU(2)<sub>M</sub>) are chosen to be

$$\vec{T} = \begin{pmatrix} 0 & & & \\ & 0 & \frac{\vec{\tau}}{2} & \\ & & & 0 \end{pmatrix}, \quad (1)$$

where  $\vec{\tau}$  are Pauli matrices. For one generation of fermions in the 5 and the 10 representations, there are four Weyl doublets for this SU(2) group:

$$\begin{pmatrix} d_3 \\ e^+ \end{pmatrix}_{R,L}, \quad \begin{pmatrix} u_c \\ u_1 \end{pmatrix}_{R,L} \quad (2)$$

(1,2,3 are color indices)

Obviously one can see that the upper and the lower members of these doublets carry different quantum numbers with respect to the color hypercharge, the electric charge, the weak hypercharge, as well as the baryon number. Thus it is

not unreasonable to anticipate that something exotic might happen around such a monopole. In fact, by rather simple arguments, we can list three features which strongly indicate that indeed peculiar physics should take place.

(i) As is well-known, an SU(5) monopole has a tiny core of size  $1/M_X$  ( $M_X$  is the X boson mass) inside of which the relevant SU(2)<sub>M</sub> Higgs fields, in this case a part of 24 of Higgs, practically vanish. This means that SU(2)<sub>M</sub> is a good symmetry inside the core and it requires little energy to cause transition between the upper and the lower members of the SU(2)<sub>M</sub> doublets. Thus we should expect to find, e.g.,  $d_3$  and  $e^+$  with equal probability;

$$|d_{3R}\rangle = |e_{+R}\rangle, \text{ etc.} \quad (3)$$

This means in particular, the baryon number is indeterminate inside the core.

(ii) Now indeterminacy of some quantum number is not sufficient to cause a reaction which actually changes these quantum numbers. In this regard, the peculiar nature of the special partial wave, namely the one with J(angular momentum) = 0, plays a crucial role. As was recognized long ago, there exists an extra angular momentum in the system of a monopole and a charged particle. In the case of SU(2) monopole, it coincides with the "isospin"  $\vec{T}$ . In the regular gauge, where all the fields are smooth and the Higgs field points along the radial direction, the expression for the total angular momentum J takes the form

$$\vec{J} = \vec{L} + \vec{S} + \vec{T}, \quad \vec{L} \equiv \vec{r} \times \vec{p}, \quad (4)$$

where  $\vec{S}$  is the ordinary spin and  $\vec{T}$  represents the extra "spin". For the isospinor, spin 1/2 fermions of our interest,  $S = \vec{\sigma}/2$ , and  $T = \vec{\tau}/2$ , where  $\vec{\sigma}$  and  $\vec{\tau}$  are two independent sets of Pauli matrices. Because of this extra spin, one can combine  $\vec{L}, \vec{S}$ , and  $\vec{T}$  to form a state with zero angular momentum. What is so special about this state? To see this, form a scalar product of  $\vec{J}$  with  $\hat{r}$ , the unit radial vector. One obtains

$$\vec{J} \cdot \hat{r} + \vec{\tau} \cdot \hat{r} = 0. \quad (5)$$

Since the Higgs field points along  $\hat{r}$  in the isospace,  $\vec{\tau}/2 \cdot \hat{r}$  is nothing but the unbroken U(1) generator, which we shall hereafter call  $T_3$ -charge. ( $\vec{\tau} \cdot \hat{r}$  becomes  $T_3$  in the 'physical' gauge, where the Higgs points along the third direction). Now in the asymptotic region, the momentum  $\vec{p}$  for the outgoing (incoming) wave is in the same (opposite) direction as  $\hat{r}$ . Thus  $\vec{\sigma} \cdot \hat{r}$  is (minus) the helicity for the outgoing (incoming) wave.

We now see that in the J=0 sector, the helicity and the "charge" are

intimately connected. Eq.(5) says that there are two cases: (a) If the "charge" is unchanged during a process, so is  $\vec{\sigma} \cdot \hat{r}$ . But this means that the helicity of the incoming and the outgoing waves must be different. This is the helicity-flip charge-non-flip case. (b) If, on the other hand, the charge is not conserved,  $\vec{\sigma} \cdot \hat{r}$  must also change and this leads to the helicity-non-flip charge-flip process. What is important is that either (a) or (b) must occur for J=0.

In the case of an abelian (Dirac) point monopole, the gauge field cannot carry charge and the case (b) should not occur. In fact this was explicitly checked some time ago [5]. Due to the point singularity of the abelian monopole, one must restrict the fermion wave functions to those which satisfy a certain boundary condition at the monopole position in order to ensure the hermiticity of the Hamiltonian of the system. This boundary condition is such that it connects the right-handed and the left-handed components and effects the helicity flip scattering.

For a non-abelian monopole, the situation is quite the contrary. First, it is a smooth object with no singularity and hence no such boundary condition exists. This means that for massless fermions, which we shall consider hereafter, there is no force which mixes right and left handed components. (To be precise, of course there is the axial anomaly, which violates the chiral symmetry. But as it will be pointed out later, this turned out to be inessential in understanding the gist of the catalysis process.) Secondly, non-abelian monopole can carry charge; it can make transitions to and from the so called dyon states. Thus these features say that the case (b) should occur for the non-abelian system. (iii) Another important feature of the J=0 sector is that such a partial wave does not feel any centrifugal barrier. This means that the baryon number violating process under discussion will not be suppressed. More precisely speaking, this is a feature characteristic of the sector with minimum possible angular momentum, not just for the "S wave" case. Since we are talking about the feature outside the monopole core, it suffices to discuss the case of an abelian monopole. The extra angular momentum in that case is of the form  $-eg\hat{r}$ , where e and g are, respectively, the electric charge of the fermion and the magnetic charge of the monopole. Due to the famous Dirac quantization condition, the product eg, which we shall denote by q, is either an integer or a half integer. Then the minimum value of the orbital angular momentum (denoted by  $\hat{L}$ ), including now  $-eg\hat{r}$  piece, is |q| and, when coupled with the spin of the fermi field, the minimum total angular momentum  $J_{\min}$  is |q| - 1/2. The one particle wave function of the fermion satisfies the Dirac equation,

$$(\not{\partial} + ie A_{\text{MON}})\psi \equiv \not{D}\psi = 0 \quad (6)$$

Operating  $\not{D}$  once more brings this to the form

$$\left(\partial_t^2 - \frac{1}{r^2}\partial_r r^2 \partial_r + \frac{L^2 - q^2}{r^2} - \vec{\sigma} \cdot \hat{r} \frac{q}{r^2}\right)\psi = 0 \quad (7)$$

Now for  $J=J_{\text{min}}$ , we substitute

$$L^2 = |q|(|q|+1) \quad (8)$$

$$\vec{\sigma} \cdot \hat{r} = q/|q| \quad ,$$

where the second equation is nothing but Eq.(5) for the abelian case. We then get,

$$\left(\partial_t^2 - \frac{1}{r^2}\partial_r r^2 \partial_r + \frac{|q|}{r^2} - \frac{|q|}{r^2}\right)\psi_{J_{\text{min}}} = 0 \quad (9)$$

We see that, regardless of the value of  $q$ , the centrifugal barrier due to the extra angular momentum is precisely cancelled by the coupling of the Dirac spin to the magnetic field.

Putting together the three points we have listed, it should now be clear why a GUT monopole is expected to induce baryon number violating processes with unsuppressed (hence probably strong interaction) rate.

### 3. STANDARD ANALYSIS

In the previous section, the essence of why a GUT monopole can induce baryon number violating processes was described in a pedagogical manner without using much mathematics. Now to gain deeper understanding of more dynamical aspects of the phenomenon and to discuss some intricate and important issues, we must give a bit more detailed account of the analysis performed by Rubakov, Callan, and others. In this section we shall review what may now be called the "standard" analysis and leave its critical examination to the next section.

Let us begin by summarizing the strategy of Rubakov and Callan. Obviously to solve the entire quantum-field-theoretical system of a non-abelian monopole and fermions is hopelessly difficult. We must make a sensible approximation, and the following simplified picture is expected to give a good account of the process. Since the non-abelian nature of the problem is effective only within the tiny monopole core, one first replaces the dynamics inside by a set of effective boundary conditions on the fermi fields at the core, which is obtained

by solving the Dirac equation in the background of a non-abelian monopole field. Outside the core, the monopole field is of abelian type, and, by making use of the spherical symmetry of the  $J=0$  partial wave sector, the system turned out to be reducible to a Schwinger-like two dimensional field theory. This was exactly solved by Rubakov and Callan (by different methods), and baryon number violating fermionic condensates were found to form.

Let us be a little more specific. Consider, for simplicity, a model  $SU(2)$  gauge theory with a doublet of Dirac fields

$$\psi = \begin{pmatrix} \psi_R \\ \psi_L \end{pmatrix} \quad (10)$$

Here and hereafter  $R$  and  $L$  denote chirality. In the  $J=0$  sector, the Weyl doublet  $\psi_R$  can be written in the form

$$\psi_R = \frac{1}{\sqrt{4\pi}} \frac{1}{r} (R_+(r,t) \eta_+^L + R_-(r,t) \eta_-^L \eta_+^T) \quad , \quad (11)$$

where  $\eta_{\pm}^L$  and  $\eta_{\pm}^T$  are, respectively, the two component Lorentz and  $SU(2)$  spinors which satisfy

$$\begin{aligned} \vec{\sigma} \cdot \hat{r} \eta_{\pm}^L &= \pm \eta_{\pm}^L \\ \vec{\tau} \cdot \hat{r} \eta_{\pm}^T &= \pm \eta_{\pm}^T \end{aligned} \quad (12)$$

Of course we have a similar expression for the left handed component. As for the gauge field, we shall take into account only the  $S$  wave (Coulomb) excitation, which will be represented by  $\lambda$ . It is related to the radial component of the gauge field by

$$A_r = -\lambda' \equiv -\partial_r \lambda, \quad E_r = \dot{\lambda}' \equiv \partial_t \partial_r \lambda \quad (13)$$

Then with a bit of algebra, the original four dimensional action,

$$I = \int d^4x \left[ -\frac{1}{4e^2} F_{\mu\nu}^a F^{\mu\nu a} + \bar{\psi} i \gamma^\mu (\partial_\mu + (A_\mu^a + A_{\mu \text{MON}}^a) \frac{\tau^a}{2i}) \psi \right] \quad (14)$$

can be reduced into the following form:

$$\begin{aligned} I &= \int_{-\infty}^{\infty} dt \int_0^{\infty} dr \left( \frac{2\pi r^2}{e^2} (\dot{\lambda}')^2 + \frac{4\pi}{e^2} K^2 \lambda'^2 \right. \\ &\quad + \bar{R}_N i \not{\partial} R_N + \bar{L}_N i \not{\partial} L_N \\ &\quad \left. - \frac{\dot{\lambda}'}{2} (\bar{L}_N \gamma^0 \gamma^5 L_N - \bar{R}_N \gamma^0 \gamma^5 R_N) \right) \end{aligned}$$

$$+ \frac{K}{r} (\bar{R}_N i\gamma_5 R_N + \bar{L}_N i\gamma_5 L_N) \quad (15)$$

We need to explain various symbols in the above expression: The function  $K(r)$  describes the extended nature of the non-abelian monopole and has the profile depicted in Fig.1.  $K$  is equal to one at the origin and it falls exponentially like  $\exp(-M_X r)$  for large  $r$ . The gamma matrices in (15) are the two dimensional ones and our convention is  $\gamma^0 = \sigma_2$ ,  $\gamma^1 = i\sigma_1$  and  $\gamma^5 = \sigma_3$ . They are defined in the isospace.  $R_N$  and  $L_N$  are two component, charge-neutral spinors which are defined by

$$\begin{cases} R_N = e^{i\frac{\lambda}{2}\gamma_5} R \\ L_N = e^{-i\frac{\lambda}{2}\gamma_5} L \end{cases}, \quad \begin{cases} R = \begin{pmatrix} R_+ \\ -iR_- \end{pmatrix} \\ L = \begin{pmatrix} L_- \\ -iL_+ \end{pmatrix} \end{cases} \quad (16)$$

Let us look at the second term of Eq.(15). As we saw, inside the core the  $K$  function is of order one. So if the gauge coupling is sufficiently small,  $4\pi K^2/e^2$  is large and the fluctuation of  $\hat{\lambda}$  should be negligible. This means that  $\lambda$  and the fermi field practically decouple and we may solve the one particle Dirac equation inside the core:

$$i\hat{\lambda}R_N + \frac{K(r)}{r} i\gamma_5 R_N = 0 \quad (17)$$

(Similarly for  $L_N$ )

Upon solving this one finds that at the core, i.e.,  $r_0 = 1/M_X$ , for the energy  $E \ll M_X$ , the fermion fields satisfy the condition (hereafter called RC boundary condition).

$$\begin{cases} R_{N+} = -R_{N-} \\ L_{N+} = -L_{N-} \end{cases} \quad \text{at } r = r_0 \quad (18)$$

This corresponds, in the case of the SU(5) theory, the baryon number violating relation discussed in Eq.(3) of Section 2.

Outside the core, we can set the  $K$  function to zero and we need to solve the following simplified theory, supplemented by the boundary condition Eq.(18):

$$I_{\text{outside}} = \int_{-\infty}^{\infty} dt \int_0^{\infty} dr \left( \frac{2\pi r^2}{e^2} (\hat{\lambda}')^2 + \bar{R}_N i\hat{\lambda}R_N + \bar{L}_N i\hat{\lambda}L_N - \frac{\hat{\lambda}}{2} (\bar{L}_N \gamma^0 \gamma_5 L_N - \bar{R}_N \gamma^0 \gamma_5 R_N) \right) \quad (19)$$

This system can be solved exactly either by the path integral method or by the bosonization technique. We shall not go into that in this talk. Instead, we will summarize the results of such analyses below.[1,2,3]

(i) Charge-neutral baryon number violating four-body condensates form with the magnitude  $\sim 1/r^6$ . As far as the chirality is concerned, these condensates include both chirality-violating and chirality-preserving ones. Existence of the latter type clearly shows that the primary cause of the condensation is the effective boundary condition at the core, not the chiral anomaly. The role of the anomaly is to allow the former type of condensates to exist.

(ii) Although the one particle boundary condition seems to violate the electric and the color charges as well, actually, they are happily conserved: Charged condensates are suppressed by the factor

$$\frac{2\pi}{\alpha} \left( \frac{r_0}{r} \right)^1 + \frac{\alpha}{2} \ln 2 + O(\alpha^2) \quad \left( \text{for } \frac{\alpha}{2\pi} \frac{r}{r_0} \gg 1 \right), \quad (20)$$

which vanishes as the size of the core,  $r_0$ , goes to zero. This is of course due to the large Coulomb energy required to form such a configuration. The fact that the suppression factor is not exactly zero simply describes the polarization phenomenon illustrated in Fig.2.

Thus we see that, as far as the leading approximation described above is concerned, the kinematical arguments of Section 2 are well-supported.

#### 4. CRITICAL EXAMINATION OF THE STANDARD ANALYSIS

We are now ready to examine the validity of the standard analysis, which we have just reviewed.

1. The first question to ask is if the effective boundary condition adopted in the standard analysis is truly correct. As the catalysis process crucially depended upon this boundary condition, it is of extreme importance. Let us recall that the boundary condition was derived by solving the one particle Dirac equation inside the core with the "justification" that, for small  $e^2/4\pi$ ,  $\hat{\lambda}$  is practically frozen there. This "justification" can be challenged. We know that the concept of the monopole core is only an approximate one and the true situation is like that depicted in Fig.3. There must inevitably be a region (region C) where  $K$  function is neither close to 1 nor to zero. In this region,  $\hat{\lambda}$ (Coulomb) fluctuation is not negligible and it couples to the charge deposited in the core, which is dictated by the RC boundary condition. This should cost Coulomb energy of order  $e^2 M_X$  and one may wonder if it does not spoil the boundary



condition itself.

Yamagishi [6] studied this question in some detail by means of a variational calculation, where parameters are introduced to represent various boundary conditions. His conclusion was that as the ratio of  $M_X$  to the fermion mass goes to infinity, the helicity flip, not the charge flip, is favored.

Does this result then tell us that the RC boundary condition should be modified? Not necessarily so. Even if the boundary condition is such that it may seem to cost a lot of energy for one particle motion, it is obvious that a group of two or more particles with zero net charge, if they are close together in space and in time, can move freely without costing much energy. This is in accordance with the result described in the previous section that neutral condensates form without suppression while those with net non-zero charge are very much suppressed. Since this discussion applies to any charge coupled to a "long range" force, we may summarize as follows: As long as we can arrange particles such that the process does not leave behind net gauge charges in a "small" region of space around the monopole, then the "charge-violating" one particle boundary condition is perfectly OK, leading to unsuppressed amplitudes.

II. The preceding discussion in turn suggests that we should examine the balance of all the gauge charges in SU(5) theory. For this purpose we have listed all the gauge quantum numbers carried by the  $SU(2)_M$  doublets of fermions in Table 1. Let us now try to write down the most energetically favorable baryon number violating processes, namely the ones for which no net gauge charges are deposited in the region around the monopole. One such process is

$$u_L^1 + u_R^2 \longrightarrow d_{3R}^c + e_L^+ \quad (21)$$

(This can proceed via X boson exchange.) This reaction, however, won't go due to the helicity-charge constraint, i.e., Eq.(5). For example, consider  $e_L^+$ . For this field,  $\vec{\sigma} \cdot \vec{r} = -2T_3 = 1$ , while  $\vec{\sigma} \cdot \hat{p} = -1$ . Thus  $\hat{p} = -\hat{r}$ , i.e. it can only be incoming, and cannot appear on the right hand side of Eq.(21). In fact this is not a special example; we can easily prove the following.

1. Regardless of the boundary condition (and the number of fermion generations), every process, involving J=0 fermions only, must leave behind a net non-zero sum for at least one of the gauge charges. Proof: First consider the weak isospin  $T_3^W$ . It is clear from the table that for every particle (and anti-particle),

$$\Delta T_3^W = T_{3out}^W - T_{3in}^W \leq 0 \quad (22)$$

where the equal sign holds only for the right handed ones. Thus to ensure

$\Delta T_3^W = 0$ , all the particles participating in the process must be right handed. But for every right handed particle,  $\Delta Q_{EM} > 0$ . Thus  $\Delta T_3^W = 0$  and  $\Delta Q_{EM} = 0$  are incompatible. q.e.d.

2. For the  $T_3$  violating RC boundary condition, color and the electromagnetic charges are violated or preserved together. (Proof is easy and omitted)

By looking at Table 1., one can immediately exhaust all the possible baryon number violating processes which preserve color and  $Q_{EM}$ . There are only twelve of them.

$$\begin{aligned} u_{1L} + u_{2L} &\longrightarrow d_{3R}^c + e_R^+ \\ e_L^+ + d_{3L}^c &\longrightarrow u_{1R} + u_{2R} \\ u_{2L} + e_R^- &\longrightarrow u_{1L}^c + d_{3R}^c \\ u_{2R}^c + e_L^+ &\longrightarrow u_{1R} + d_{3L} \\ u_{2L} + d_{3R} &\longrightarrow u_{1L}^c + e_R^+ \\ u_{2R}^c + d_{3L}^c &\longrightarrow u_{1R} + e_L^- \\ u_{1L} + e_R^- &\longrightarrow u_{2L}^c + d_{3R}^c \\ u_{1R}^c + e_L^+ &\longrightarrow u_{2R} + d_{3L} \\ u_{1L} + d_{3R} &\longrightarrow u_{2L}^c + e_R^+ \\ u_{1R}^c + d_{3L}^c &\longrightarrow u_{2R} + e_L^- \\ u_{2R}^c + u_{1R}^c &\longrightarrow d_{3L} + e_L^- \\ e_R^- + d_{3R} &\longrightarrow u_{1L}^c + u_{2L}^c \end{aligned} \quad (23)$$

As was shown in 1, for all these processes, the weak charge is not balanced. Now the question is; how much energy would it cost because of this?

This is not an easy one to answer. The reason is that, as it will be explained, both  $M_W$  and  $M_X$  are involved in this business. Since  $\Delta T_3^W \neq 0$  does not mean much in the region where weak SU(2) is broken, we must first look for the region where it is not broken. The part of  $\phi_5$ , the  $\underline{5}$  of Higgs, which controls the breaking of  $SU(2)_W$ , is a singlet with respect to the  $SU(2)_M$ . That is, it feels the presence of the monopole only indirectly through its coupling to the SU(2) non-singlet part of the  $\underline{24}$  of Higgs. Since this occurs only within the core, we expect  $\langle \phi_5 \rangle$  to be equal to  $eM_W$ , the vacuum value, all the way down to the core. Whether weak SU(2) is restored inside the core is a difficult dynamical question, which has not been studied in the literature. If it is restored, then  $T_3^W$  must be balanced there by excitation of weak-charge-carrying fields and one

must evaluate the excitation energy to see if it is large.

III. The third and the final topic I would like to discuss is the role of the higher partial waves, which, through radiative corrections, can affect the physics of the  $J=0$  sector we have been discussing. One might say that because of the smallness of the fine structure constant  $\alpha$  the radiative effects must be a tiny correction. This, however, need not be so in the strong magnetic field of the monopole. There are two effects which come to our mind.

(a) The first is the effect of the anomalous magnetic moment of the fermi field, which we have been neglecting. It is known [5] that if a constant extra magnetic moment,  $\alpha/2\pi$ , is put in by hand, the wave function of the fermion is very much suppressed near the monopole position. The suppression factor at  $r \sim 1/M_X$  is of the order  $\exp(-\alpha M_X/(4\pi m))$  which is extremely small. Of course in reality we must take into account the effect of the form factor. Although such a calculation has not been done, the works of reference [7], which dealt with the case of the strong uniform magnetic field background, give a good hint. If we denote such a magnetic field by  $H$ , these authors found that the usual expression  $(\alpha/2\pi)(e/2m)H$  for the anomalous term is replaced by the expression  $(m\alpha/4\pi)(\ell n(2eH/m^2))^2$  in the case of large  $H$ . Because of the slow logarithmic growth, this indicates that the suppression of the wave function near the monopole core due to the anomalous moment should be immaterial.

(b) The second issue is the decay of the  $J=0$  wave by bremsstrahlung into higher  $J$  states. If this happens at a sufficient rate, the centrifugal barrier present for the higher partial waves would suppress the catalysis process. The differential rate for the process shown in Fig.4 is given by the expression

$$\begin{aligned} \frac{dW_E}{dp} &= \sum_{\sigma=\pm 1,1} \sum_{j=\pm 1}^{\infty} \frac{\alpha}{2} (2j+1) \frac{E+m}{E'+m} \left| \int_0^{\infty} dr k \frac{k'}{k} [J_{\sqrt{j(j+1)+\frac{\alpha}{2}}}(k'r) \sin(kr+\delta) \right. \\ &\quad \left. - \sigma \frac{E'+m}{E+m} J_{\sqrt{j(j+1)-\frac{\alpha}{2}}}(k'r) \cos(kr+\delta)] J_{j+1/2}(pr) \right|^2 \\ \frac{dW_M}{dp} &= \sum_{\sigma=\pm 1,1} \sum_{j=\pm 1}^{\infty} \frac{\alpha}{2} (2j+1) \frac{E+m}{E'+m} \left| \int_0^{\infty} dr k \frac{1}{2j+1} \frac{k'}{k} [J_{\sqrt{j(j+1)+\frac{\alpha}{2}}}(k'r) \sin(kr+\delta) \right. \\ &\quad \times [J_{j+3/2}(pr)(\sigma j + \sqrt{j(j+1)}) + J_{j-1/2}(pr)(-\sigma(j+1) + \sqrt{j(j+1)})] \\ &\quad \left. - \sigma \frac{E'+m}{E+m} J_{\sqrt{j(j+1)-\frac{\alpha}{2}}}(k'r) \cos(kr+\delta) \right. \\ &\quad \left. \times [J_{j+3/2}(pr)(\sigma j - \sqrt{j(j+1)}) - J_{j-1/2}(pr)(\sigma(j+1) + \sqrt{j(j+1)})] \right|^2, \end{aligned} \quad (24)$$

where  $J_\nu(Z)$  is the Bessel function of order  $\nu$ ,  $E'=E-P$ ,  $k=\sqrt{E^2-m^2}$ ,  $k'=\sqrt{E'^2-m^2}$ , and  $\delta=\frac{1}{2}\cos^{-1}(m/E)$ . The subscripts E and M stand for electric and magnetic photon emission respectively. Analytical evaluation of these expressions seems quite difficult and numerical study is currently being carried out. [8] (In the case of a quark, gluon emission rate can also be computed in a similar manner.) One can also ask about the effect of virtual emissions. This is more complicated due to the necessity of performing the renormalization, which is cumbersome in the angular momentum basis we are forced to stick to. We hope to be able to report on this subject in the near future.

## 5. EPILOGUE

Although it is clearly not possible, in a talk, to cover all the aspects of the fascinating subject of the monopole-catalyzed proton decay, we hope that the essence of the phenomenon and some of the remaining problems have been spelled out.

Due to the inherently non-perturbative and many-particle nature of the process, it is not easy to get clear cut answers to these problems. In particular we emphasized in this talk that various energy barriers should be further examined. In our opinion, the role of the weak interaction is not yet clear and the effects of the higher partial wave sector need to be quantified. If these turn out to be small effects, one must still understand the most difficult part of the problem, namely, the strong interaction effects, in order to get a good estimate of the rate of the catalysis phenomenon. We certainly need ingenuity and a lot of luck to observe this fascinating effect. Let us hope for the best!

1. V.A. Rubakov, Pis'ma Zh Exsp. Teor. Fiz. 33 (1981) 658 [JETP Lett. 33 (1981) 644]; Nucl. Phys. B203 (1982) 311.
2. C.G. Callan, Phys. Rev. D25 (1982) 2141; *ibid.* D26 (1982) 2058; Nucl. Phys. B212 (1983) 391.
3. A. Sen, Phys. Rev. D28 (1983) 876.  
C.G. Callan and S.R. Das, Phys. Rev. Lett. 51 (1983) 1155.  
V.A. Rubakov and M.S. Serebryakov, Nucl. Phys. B218 (1983) 240.  
Y. Kazama, Prog. Theor. Phys. 70 (1983) 1166.  
Y. Kazama and A. Sen, preprint Fermilab-Pub-83/58-THY.  
Z.F. Ezawa and A. Iwazaki, Z. Phys. C20 (1983) 335.  
T. Yoneya, preprint UT-Komaba 83-3.  
T.M. Yan, preprint CLNS 83/563.  
J. Arafune and M. Fukugita, Phys. Rev. Lett. 50 (1983) 1901.  
K. Seo, Phys. Lett. 126B (1983) 201.  
B. Grossman, G. Lazarides and A.I. Sanda, preprint RU83/B/47.  
A.P. Balachandran and J. Schechter, preprint SU-4222-260 (1983).  
S. Dawson and A.N. Schellekens, preprint Fermilab-Pub-83/43-THY.  
T.F. Walsh, P. Weisz, and T.T. Wu, preprint DESY 83-022.  
F.A. Bais, J. Ellis, D.V. Nanopoulos, and K.A. Olive, preprint TH. 3383-CERN.  
W. Goldstein and M. Soldate, preprint SLAC-PUB-3240 (1983).
4. G. 't Hooft, Nucl. Phys. B79 (1974) 276, A.M. Polyakov, JETP Lett. 20 (1974) 194.
5. Y. Kazama, C.N. Yang, and A.S. Goldhaber, Phys. Rev. D15 (1977) 2287, A.S. Goldhaber, Phys. Rev. D16 (1977) 2287.
6. H. Yamagishi, Phys. Rev. D27 (1983) 2383; *ibid.* D28 (1983) 977.
7. B. Jancovici, Phys. Rev. 187 (1969) 2275, W. Tsai and A. Yildiz, Phys. Rev. D8 (1973) 3446, V.N. Baier, V.M. Katkov and V.M. Strakhovenko, Sov. Phys. JETP 40 (1975) 1415. M. Kobayashi and M. Sakamoto, Prog. Theor. Phys. 70 (1983) 1375.
8. K. Itoh and Y. Kazama, to be reported.

	$Q_{EM}$	$\frac{2}{\sqrt{3}}A_8^C$	$2A_3^C$	$T_3^W$	$Y^W$	B	$T_3$	
$d_{3L}$	-1/3	-2/3	0	-1/2	1/3	1/3	1/2	out
$e_L^+$	1	0	0	1/2	1	0	-1/2	in
$u_{2L}^C$	-2/3	-1/3	1	-1/2	-1/3	-1/3	1/2	out
$u_{1L}$	2/3	1/3	1	1/2	1/3	1/3	-1/2	in
$d_{3R}$	-1/3	-2/3	0	0	-2/3	1/3	1/2	in
$e_R^+$	1	0	0	0	2	0	-1/2	out
$u_{2R}^C$	-2/3	-1/3	1	0	-4/3	-1/3	1/2	in
$u_{1R}$	2/3	1/3	1	0	4/3	1/3	-1/2	out

(For anti-particle, change all the signs and replace "in" ("out") with "out" ("in").)

TABLE CAPTION

Table 1. Gauge and other quantum numbers of the SU(2) doublets in the SU(5) theory are tabulated, together with the incoming or outgoing nature of the fields.  $Q_{EM}$  is the electric charge,  $A_8$  and  $A_3$  are the color hypercharge and the color isospin, respectively,  $T_3^W$  and  $Y^W$  are the weak isospin and the weak hypercharge, B is the baryon number, and  $T_3$  is the SU(2)<sub>M</sub> charge.

FIGURE CAPTION

- Fig. 1. Profile of the function  $K(r)$  which describes the non-abelian nature of the monopole.
- Fig. 2. Charge polarization phenomenon around a monopole.
- Fig. 3. Structure of the monopole "core". A and B are, respectively, the abelian and the non-abelian regions, while C is the intermediate domain.
- Fig. 4. Lowest order bremsstrahlung process in the field of a monopole. The double line indicates that the wave functions are computed in the monopole background.

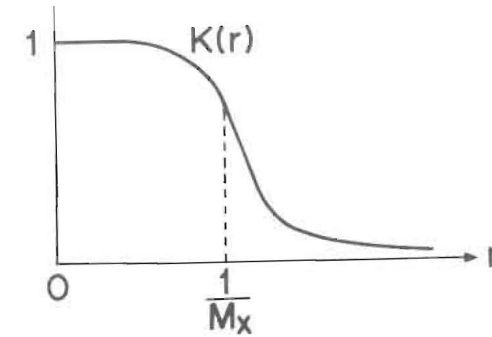


Fig. 1

ERRATA

for

"Issues in Monopole-Catalyzed Proton Decay"

Yoichi Kazama

National Laboratory for High Energy Physics (KEK)  
Tsukuba, Ibaraki, 305 Japan

In the preprint (KEK-TH 81) with the title above, there was a serious error concerning the conservation of the weak charge. First, in Table 1, the entries for the column under  $T_3^W$  should be changed to

$$-1/2, 0, 0, 1/2, 0, 1/2, -1/2, 0 \ .$$

Also the entries under  $Y^W$  should read

$$1/3, 2, -4/3, 1/3, -2/3, 1, -1/3, 4/3 \ .$$

With these corrections, the weak charge is conserved for all the processes listed in Eq. (23) of the text, and the discussions concerning its non-conservation should be disregarded. I am grateful to Ashoke Sen for pointing out these errors.

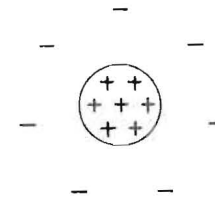


Fig. 2

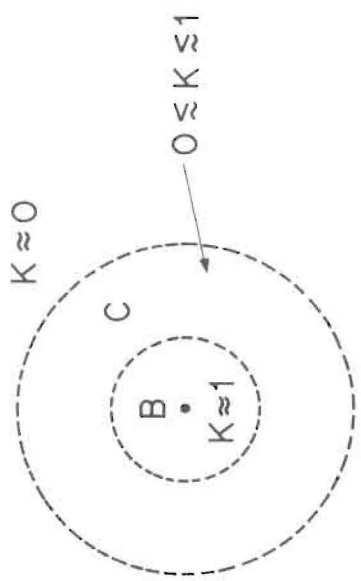


Fig. 3

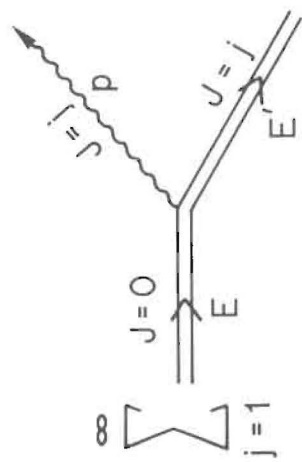


Fig. 4

## Present Status of Monopole Search

Fumiyoshi Kajino<sup>\*</sup>)

Institute for Cosmic Ray Research, University of Tokyo  
Tanashi, Tokyo, 188 Japan

### ABSTRACT

Searches for slowly moving magnetic monopoles have been performed with a combined detector of proportional chambers and scintillation counters. In the first stage, the scintillation counters were mainly used. In the second stage, the proportional chambers filled with helium gas using the Drell et al. mechanism were mainly used. We have no candidate for the monopole in either measurement.<sup>\*\*)</sup>

<sup>\*</sup>) Present address: Department of Physics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061, U.S.A.

<sup>\*\*)</sup> This work has been done with the collaboration of S. Matsuno, T. Kitamura, T. Aoki, Y.K. Yuan, K. Mitsui, Y. Ohashi and A. Okada.

### 1. Introduction

I reviewed recent experiments of monopole search in this conference. But there are some good reviews<sup>1-5</sup> about them. So in this paper I will describe only our monopole search experiment.

### 2. Experimental Apparatus

Figure 1 shows the experimental apparatus for the monopole search. This apparatus consists of nine layers of proportional chambers (PRC), six layers of scintillation counters and fourteen iron layers. The size of the apparatus is about 3.9 m (width) by 3.2 m (length) by 2.4 m (height). The number of SC's in each layer is nineteen. Each SC has area  $240 \times 20 \text{ cm}^2$  and thickness 1 cm. The number of PRC's in each layer is four. Each PRC has an effective area  $246 \times 92 \text{ cm}^2$  and thickness 2 cm. The thickness of the iron layers is 12 cm for the inner layers and 1 cm for the two outer layers. The time of flight was measured with time-to-digital converters (TDC) for the PRC's and the SC's. Ionization losses were measured by analog-to-digital converters (ADC) for the PRC's and SC's. The detector is situated on the ground near sea level at the Institute for Cosmic Ray Research in Tokyo. Details of the detector and the data acquisition system have already been reported elsewhere.<sup>6,7</sup>

Our experiment was divided into two stages. The experiment was performed mainly by using the SC's to set the energy threshold at a low level in the first stage, and mainly by the PRC's utilizing the Drell et al. mechanism and the Penning effect in the second stage.

### 3. The First Stage Experiment

In the first stage of our experiment, a trigger signal was generated by using signals from the SC's. The velocity region of the trigger was between  $10^{-4} c$  and  $0.1 c$ . The area-solid-angle product for this trigger was  $11.0 \text{ m}^2 \text{ sr}$ .

The ionization loss calculated by Ritson is shown in Fig. 2. Threshold levels for the SC's and the PRC's were set at  $1/20$  minimum ionization ( $I_{\text{min}}$ ) as shown in the figure. PR gas ( $\text{Ar} + 10\% \text{CH}_4$ ) was circulated through the PRC's. Monopoles with velocities from  $2.5 \times 10^{-4} c$  to  $0.1 c$  can be measured if Ritson's calculation<sup>11</sup> for the scintillator is valid. However, the lower limit on velocity to detect the monopoles with a scintillator may be about  $6 \times 10^{-4} c$

if the ionization loss mechanism suggested by Ahlen et al.<sup>12</sup> is valid. No candidate for the monopole was obtained during the live time of  $3.3 \times 10^3$  h. The upper flux limit of  $1.8 \times 10^{-12} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$  for the magnetic monopole is obtained at a 90% confidence level. Details of the first stage of the experiment have already been reported else where.<sup>7</sup>

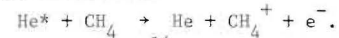
#### 4. The Second Stage Experiment

In the second stage, the gas in the PRC was changed to He + 10% CH<sub>4</sub>. A research work of the PRC with this gas has been reported.<sup>10</sup>

Drell et al.<sup>13</sup> have shown that large energy losses occur for low-velocity monopoles in helium gas. When the monopole goes through the helium gas, helium atoms are excited as



where He\* is a metastable state of the helium atom. This He\* collides with a methane molecule, which leads to an ionization of the methane through the Penning effect as follows:



Bortner and Hurst<sup>14</sup> have clearly demonstrated this effect in any arbitrarily mixed combination of helium and methane.

The calculated curve for ionization loss of the monopole in He + 10% CH<sub>4</sub> is shown in Fig. 3. The efficiency of ionization for the Penning effect is estimated as 83%,<sup>15</sup> and a new calculation of the Drell et al. mechanism,<sup>16</sup> which is smaller than the old one by a factor of 2, is used. The ionization loss of the monopoles with large velocity is calculated by using Ahlen's formula.<sup>17</sup>

The trigger signal for the second stage were generated by the successively delayed sixfold coincidence of respective layers of PRC's. The threshold level for the ionization loss in the PRC's was set at  $3 I_{\text{min}}$ . The trigger efficiency for sixfold coincidence is estimated to be 1.0 for an energy losses larger than about  $7 I_{\text{min}}$ . The area-solid-angle product for this trigger was  $24.7 \text{ m}^2 \text{ sr}$  which was more than twice that used in the first stage.

No candidate for the monopole was obtained during a live time of  $3.6 \times 10^3$  h in the second stage. The upper flux limit of  $7.2 \times 10^{-13} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$  for the magnetic monopole is obtained at a 90% confidence level over a wide velocity range from  $\sim 3 \times 10^{-4}$  c to 1 c.<sup>8,9</sup>

The present upper limit of magnetic monopole flux based on the first and the second-stage measurement is shown in Fig. 4 together with other experimental results.

#### 5. Conclusion

The upper limit of the monopole flux in the first stage measurement is  $1.8 \times 10^{-12} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$  at a 90% confidence level over a velocity range from  $2.5 \times 10^{-4}$  to 0.1 c if Ritson's conjecture is valid. The velocity threshold may be about  $6 \times 10^{-4}$  c if the energy loss calculated by Ahlen et al. is valid.

The upper limit of the monopole flux in the second stage measurement utilizing the Drell et al. mechanism and the Penning effect is  $7.2 \times 10^{-13} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$  at a 90% confidence level over a wide velocity range from  $\sim 3 \times 10^{-4}$  c to 1 c. These limits exceed the theoretical upper bound of  $\sim 3 \times 10^{-12} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$  presented by Arons and Blandford<sup>18</sup> for a monopole mass of  $10^{18}$  GeV. It is difficult to reconcile the first candidate measured by Cabrera<sup>19</sup> with the present null results.

#### References

1. R.A. Garrigan, Jr. and W.P. Trower, "Magnetic Monopoles: A Status report", FERMILAB-Pub-83/31, 2000.000.
2. G. Giacomelli, "Conference Highlights and Summation - Experimental", CERN/EP 83-203.
3. D.E. Groom, "Experimental Searches for GUT Monopoles", UU HEP 83/19.
4. J. Arafune, F. Kajino et al., Monthly Physics 4, No.8(1983), (in Japanese).
5. F. Kajino and T. Watanabe, Butsuri 39, 298(1984), (in Japanese).
6. F. Kajino, ICR-Report-108-83-2. And Proc. Workshop on Monopoles and Proton Decay, Kamioka, KEK 83-12, 74(1983).
7. F. Kajino, S. Matsuno, T. Kitamura, T. Aoki, Y.K. Yuan, K. Mitsui, Y. Ohashi and A. Okada, J. Phys. G: Nucl. Phys. 10, 447(1984). And Proc. 18th Int. Cosmic Ray Conf., Bangalore, India, 5, 56(1983).
8. F. Kajino, S. Matsuno, Y.K. Yuan and T. Kitamura, Proc. Monopole 83 Conf., University of Michigan, Ann Arbor, Michigan, 1983 (to be published).
9. F. Kajino, S. Matsuno, Y.K. Yuan and T. Kitamura, Phys. Rev., Lett., 52, 1373(1984).

10. F. Kajino and Y.K. Yuan, ICR-Report-114-84-3 and submitted to Nucl. Instr. and Meth.
11. D.M. Ritson, Stanford Linear Accelerator Center Report No. SLAC-PUB-2950, 1982(unpublished).
12. S.P. Ahlen and G. Tarlé, Phys. Rev. D 27, 688(1983).  
S.P. Ahlen, T.M. Liss and G. Tarlé, Phys. Rev. Lett. 51, 940(1983).
13. S.D. Drell, N. M. Croll, M.T. Mueller, S.J. Parke and M.A. Ruderman, Phys. Rev. Lett. 50, 644(1983).
14. T.E. Bortner and G.S. Hurst, Phys. Rev. 93, 1236(1954).
15. W.P. Jesse, J. Chemical Phys. 41, 2060(1964).
16. N.M. Croll, Proc. Monopole 83 Conf., University of Michigan, Ann Arbor, Michigan, 1983 (to be published).
17. S.P. Ahlen, Phys. Rev. D 17, 229(1978).
18. J. Arons and R.D. Blandford, Phys. Rev. Lett. 50, 544(1983).
19. B. Cabrera, Phys. Rev. Lett. 48, 1378(1982).

#### Figure Captions

Fig. 1. Schematic view of monopole detector.

Fig. 2. The curve shows the calculated ionization loss of magnetic monopoles having the Dirac charge, as a function of velocity in a scintillator (Carbon) and in argon gas. These curves are calculated by Ritson. Threshold levels for our detector are also shown. P, proportional chamber; S, scintillation counter.

Fig. 3. The curve shows the calculated ionization loss of monopoles having the Dirac charge, as a function of velocity in a mixed gas of helium and 10% methane.

Fig. 4. Compilation of upper limits on the flux of magnetic monopoles as a function of velocity  $\beta$  at a 90% confidence level for ionization/excitation experiments. The broken lines mean that it is impossible to measure if calculations for the energy losses taking into account the binary encounter approximation for scintillators and argon gas are valid.



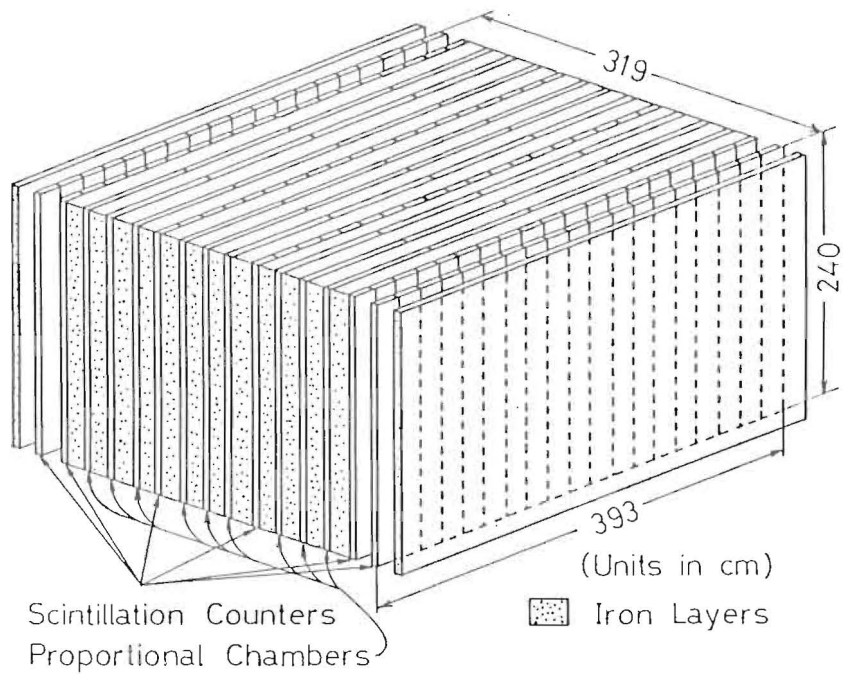


FIG. 1

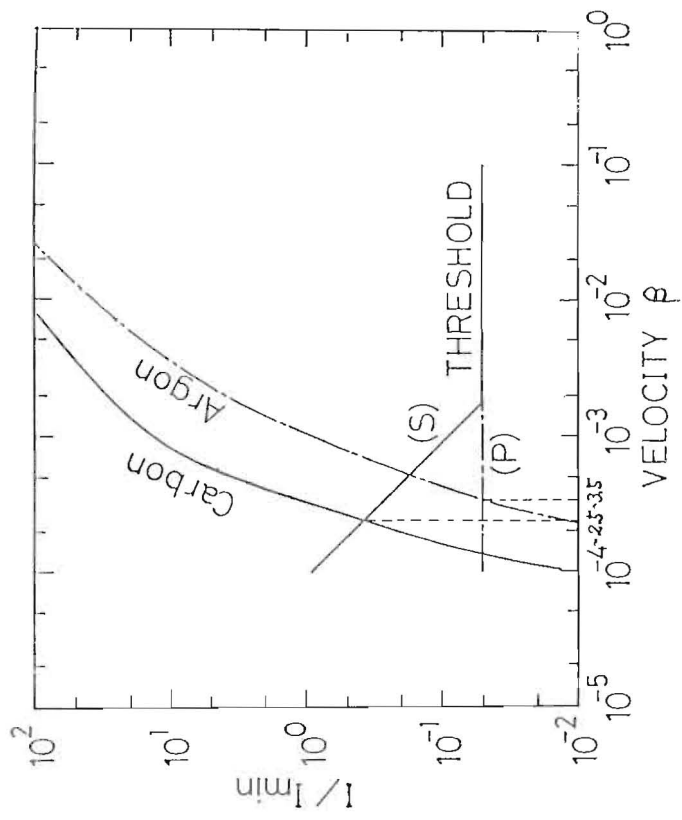


FIG. 2

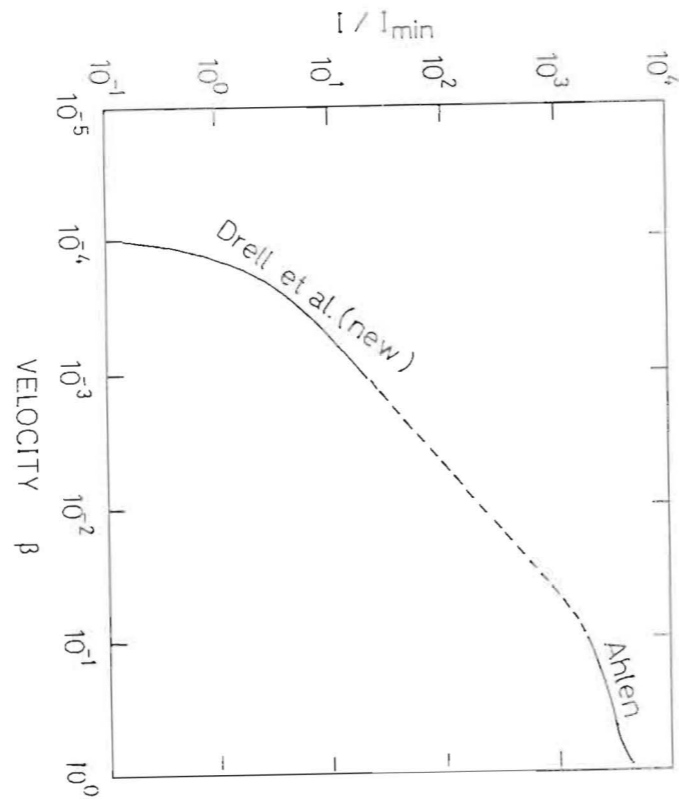


FIG. 3

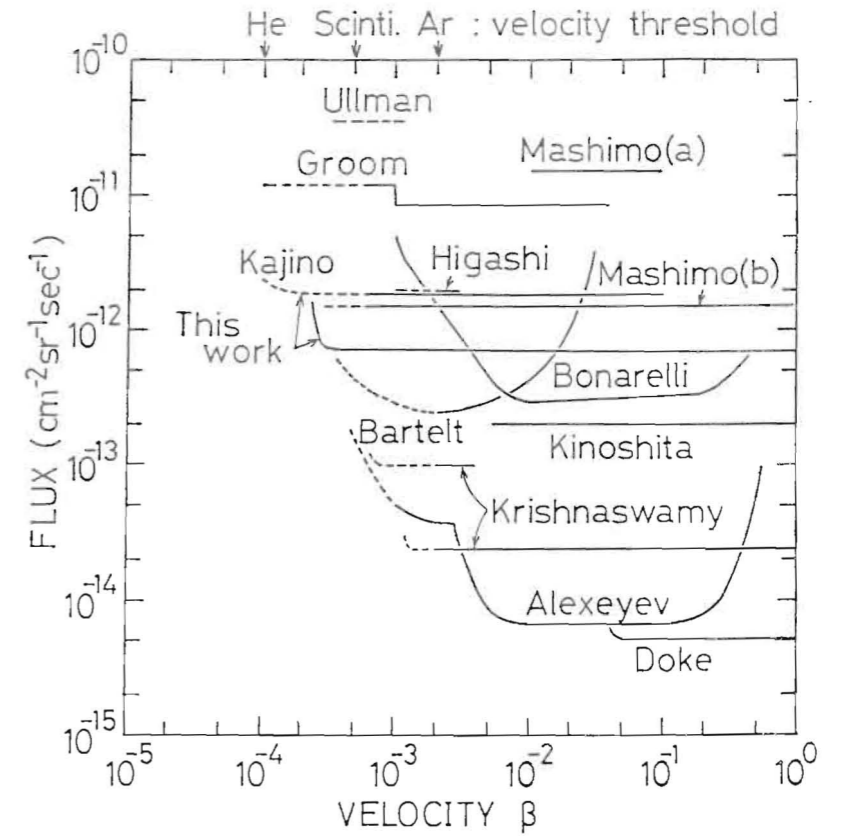


FIG. 4

Underground Searches for Anomalous Penetrating Particles

S. Orito

Department of Physics, Faculty of Science, University of Tokyo,  
Bunkyo-ku, Tokyo 113, Japan

ABSTRACT

A series of experiments are being performed at Kamioka Mine in search of anomalous cosmic ray particles such as GUT magnetic monopoles, heavy slow charged particles and relativistic fractionally charged leptons. We report on the results from or the status of: 1) Scintillator telescopes of  $22 \text{ m}^2 \text{Sr}^+$ , 2) Scintillator telescope of  $110 \text{ m}^2 \text{Sr}^+$ , 3) Plastic track detector CR39 of  $1500 \text{ m}^2 \text{Sr}^{++}$ , 4) Prototype study of a large area He drift chamber.

---

Physicists involved are:

- \* ) K. Kawagoe, S. Nakamura, M. Nozaki, T. Mashimo and S. Orito
- + ) K. Nagano, K. Anraku, T. Tsukamoto, K. Kawagoe, S. Nakamura, M. Nozaki and S. Orito
- ++) S. Nakamura, K. Kawagoe, K. Nagano, M. Nozaki, S. Orito, T. Doke, T. Hayashi, H. Tawara and K. Ogura

The main motivation of our experiments is the direct search for super-heavy GUT magnetic monopoles to a smallest possible fluxes given the limited available resources<sup>\*)</sup>. The experiments are however sensitive also to super-heavy slow charged (fractional or integer) particles as well. It is conceivable that such particles exist<sup>1), 2)</sup> and might be produced in very early Universe. Such particles could be stable either due to their fractional charges or the new conserved quantum number they might possess.

GUT magnetic monopoles or such anomalous charged particles, being super-heavy, would have had negligibly small energy dissipation relative to their kinetic energy in the process of galaxy formation, and are unlikely to have been trapped in the terrestrial material. Most promising way seems therefore to be the search in the primary cosmic rays.

If the superheavy particles are trapped in Galaxy or in the cluster of galaxies, they would have a typical virial velocity of  $3 \times 10^{-3}$  or  $3 \times 10^{-3}$  respectively. They might be a part of the invisible masses. The magnetic monopoles might have been accelerated to  $3 > 10^{-3}$ , escaped the galaxies, composing an isotropic extra galactic flux. Recent calculations<sup>3)</sup> of the  $dE/dx$  of the slow magnetic monopoles and charged particles indicate that the passage of such particles can be detected in the velocity range  $3 > 5 \times 10^{-4}$  by  $dE/dx$  detectors such as the scintillators and the drift chamber.

1) Experiment with scintillator telescope of  $22 \text{ m}^2 \text{Sr}$ :

This is the first experiment we have undertaken, the result of which were published<sup>4)</sup>. The detector shown in Fig. 1 consists of total 60 plastic

---

\*) The first two experiments described here are operated and analyzed mainly by graduate students utilizing second-hand detector components used in accelerator-experiments. We are indebted to Prof. M. Koshiya for letting us use the components.

scintillation counters composing a six-layer hodoscope, located 250 m underground. Each counter, viewed by two fast photomultipliers from two ends, has a  $dE/dx$  and the time of flight resolution of  $\pm 10\%$  and  $\pm 0.2$  nsec respectively. The trigger essentially requires the coincidence pulses from at least 4 out of 6 layers within 20  $\mu$ sec. The on-and off-line analyses then require a hit-hodoscope pattern and the timing information consistent with a passage of upgoing or downgoing single particle. The effective  $dE/dx$  threshold and the velocity range was 0.25 times the minimum ionization and  $3 \times 10^{-4} \leq \beta \leq 1$ .

Shown in Fig. 2 is the  $dE/dx$  versus  $\beta$  plot for  $2.7 \times 10^6$  events collected in 5200 hours of live time. Also shown are the expected  $dE/dx$  curves for Dirac magnetic monopole and for various charged particles. We observe in this plot no event outside the dominant "muon" peak at  $dE/dx = 1.0$  and  $\beta = 1$ ; thus no candidate for magnetic monopoles in  $5 \times 10^{-4} \leq \beta \leq 1$  nor for charged particles in  $5 \times 10^{-4} \leq \beta \leq 0.4$  with flux upper limit (90 % confidence level) of  $6.2 \times 10^{-13} \text{ m}^{-2}\text{sec}^{-1}\text{Sr}^{-1}$ .

Our  $dE/dx$  and the time of flight information were good enough also to search for any anomalously charged particles (leptons) among the relativistic "muons". Fig. 3 shows the  $dE/dx$  distribution for all events. We see no evidence for such particles giving flux limit  $9.8 \times 10^{-13}$  and  $7.5 \times 10^{-13} \text{ cm}^{-2} \text{sec}^{-1}\text{Sr}^{-1}$  for charge  $2/3$  and  $1/2$  particles respectively.

## 2) Experiment with plastic scintillator hodoscope of $110 \text{ m}^2\text{Sr}$ :

We are at present rearranging the scintillators into two-layer hodoscope of  $110 \text{ m}^2\text{Sr}$ . The trigger will require either wide pulses for slow particles or high pulse heights. Fast flash ADC's then record the counter pulse shapes every 5 nsec over 24 sec. The on-and off-line analyses by a personal computer based on the flash ADC informations should be sufficient to reject all accidental backgrounds. Consistency among the pulse widths and the time of flight

should be a strong enough constraint to positively identify a passage of slow particle, if any. The experiment should then be sensitive to magnetic monopoles in the range  $5 \times 10^{-4} \leq \beta \leq 1$  and for charged particles  $5 \times 10^{-4} \leq \beta \leq 0.1$ .

This experiment could be expanded to a larger scale with estimated cost of  $10^8$  Yen or  $10^6$  SF/ $1000 \text{ m}^2\text{Sr}$ .

## 3) Track detector CR39 with $2500 \text{ m}^2\text{Sr}$ :

Some of the recent inflational models incrementing the super symmetry<sup>4)</sup> predict a measurable monopole flux of order  $10^{-15} \text{ cm}^{-2}\text{sec}^{-1}\text{Sr}^{-1}$ , just below the Parker's limit. We are undertaking an exploratory experiment sensitive to such a minute flux. The solid track detectors seem to be most suitable to this purpose because of its low cost and the passive nature (no maintenance necessary). Among various track detectors, the plastic CR39 has the best sensitivity. The CR39 is at the same time totally insensitive to minimum ionizing particles such as muons, thus minimizing the background problem.

The optimization with the curing cycles and trials with various additives resulted to a product which is sensitive to  $Z/\beta \sim 6$ , while keeping clear surfaces after a heavy etching. A  $\pm 3\%$  uniformity in thickness have been obtained over an area of  $50 \times 50 \text{ cm}^2$ .

We have just completed the installation of the total  $450 \text{ m}^2$  of CR39 to  $650 \text{ m}$  underground. The detector consists of 7200 stacks, each consisting of four sheets of CR39 with dimension  $25 \text{ cm} \times 25 \text{ cm} \times 1.6 \text{ mm}$ .

After year's of exposure the top sheets will be collected and be heavily etched to the thickness of  $200 \mu\text{m}$ . By the etching, a track of particle with enough ionization will develop a hole, which then will be scanned. Only when a candidate hole is found in the first layer the corresponding subsequent layers will be etched. The four fold coincidence will eliminate such background as ratio-actives and pin-holes simulating a track. The experiment is expected to

be sensitive to magnetic monopoles in the region  $10^{-3} \leq \beta \leq 1$ .

4) Proto-type He drift chamber:

A proto-type large area (4 m x 80 cm) He drift chamber is constructed. The preliminary test results show that the chamber works with a proper He-Ar-CO<sub>2</sub> mixture with a low drift field of down to 200 V/cm up to a drift space of 40 cm. A self triggering scheme is being tested.

References

1. H. Terazawa, Phys. Rev. D22 (1980) 184 and the references therein.
2. S. M. Barr, D. B. Reiss and A. Zee, Phys. Rev. Lett. 50 (1983) 317.
3. S. P. Ahlen and K. Kinoshita, Phys. Rev. D26 (1982) 2347.  
D. M. Ritson, SLAC-Pub 2950 (82), S. P. Ahlen and G. Tarle, Phys. Rev. D27 (1983) 688.
4. T. Mashimo, S. Orito, K. Kawagoe, S. Nakamura and M. Nozaki, Phys. Letters 128B (1983) 327 and papers in preparation.

Figure Caption

- Fig. 1. The scintillator hodoscope of 6 layers. The distance between the top and the bottom layers is 80 cm.
- Fig. 2. The expected energy loss of a magnetic monopole and various charged particles as function of  $\beta$ . The energy loss was normalized to the minimum ionizing particle. The dashed line is calculated by taking into account the expected saturation effect of the scintillator. Also plotted are  $dE/dx$  and  $\beta$  of  $2.7 \times 10^6$  events.
- Fig. 3. The  $dE/dx$  distribution of the relativistic ( $\beta \approx 1$ ) particles. The dashed and the dash-dotted lines show the expected  $dE/dx$  distributions of the charge  $2/3$  and  $1/2$  particles with  $\beta = 1$  respectively.

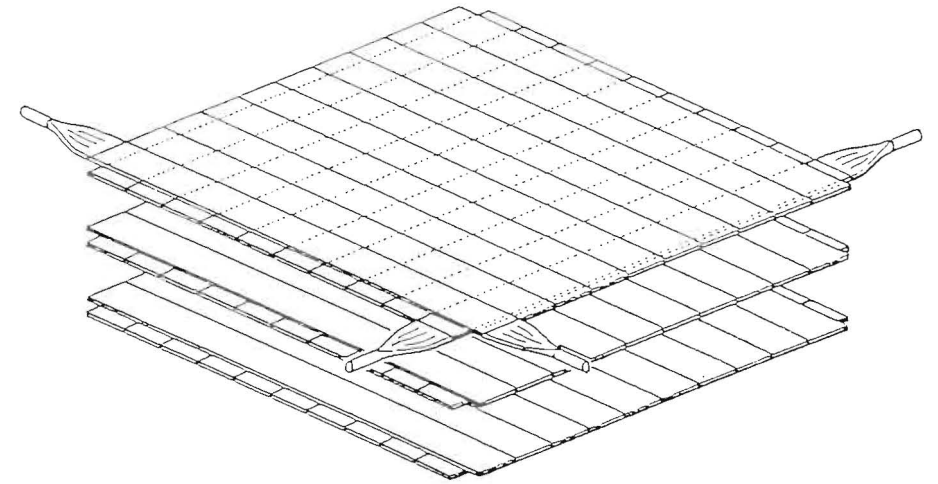


Fig. 1

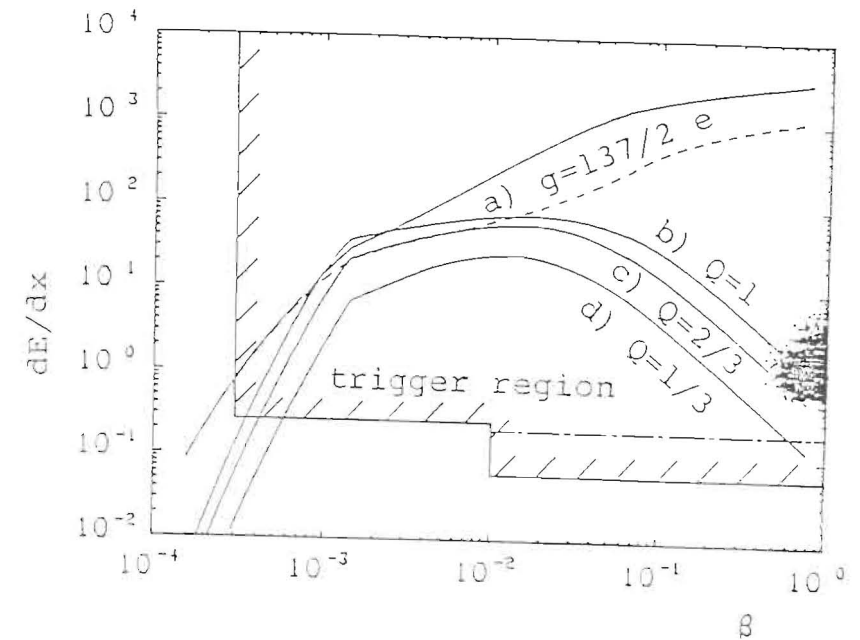


Fig. 2

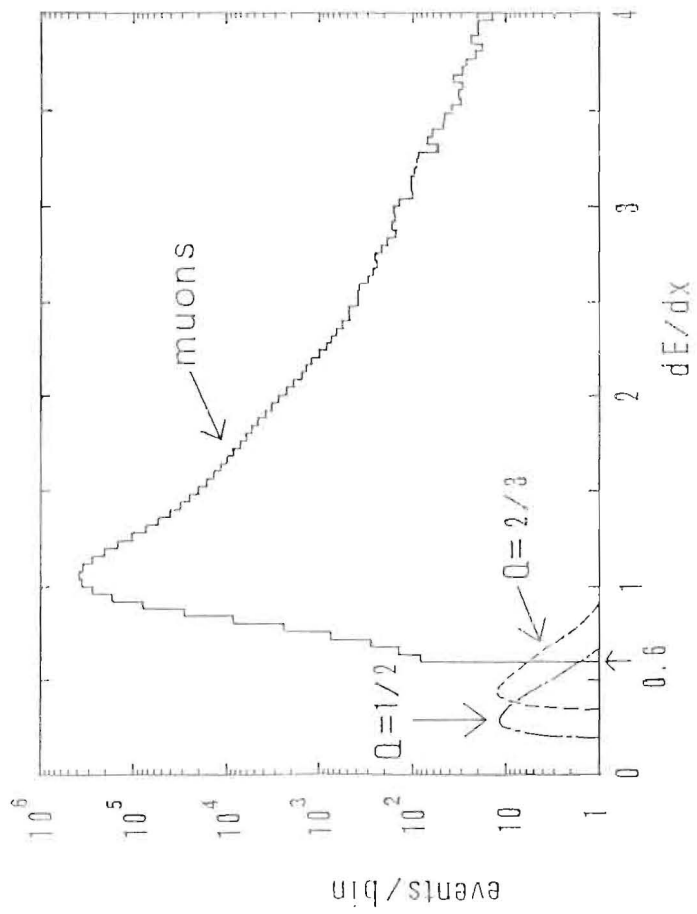


Fig. 5

A plan of monopole search experiment using 100m<sup>2</sup> calorimeter  
of Akeno air shower array

T.Hara, M.Honda and Y.Ohno  
Institute for Cosmic Ray Research, University of Tokyo,  
Tanashi, Tokyo, 188 Japan  
Y.Totsuka  
LICEPP, Faculty of Science, University of Tokyo,  
Bunkyo-ku, Tokyo, 113 Japan  
M.Kobayashi and T.Kondo  
National Laboratory for High Energy Physics(KEK),  
Oho-machi, Tsukuba-gun, Ibaraki, 305 Japan

ABSTRACT

In Akeno air shower observatory a large area calorimeter of an area of 100m<sup>2</sup> is operating for the study of air showers. We planed to use this calorimeter for the search of magnetic monopoles, and started the preparation for this experiment by examining the characteristics of proportional counters filled with He gas. It is found that proportional counters filled with He mixture gas(He85% + CH<sub>4</sub>15%) works satisfactorily for our purpose.

Adding 2 more layers of proportional counters to the calorimeter, the expected upper bound to the flux of monopoles with  $\beta > 5 \times 10^{-4}$  will be  $2 \times 10^{-14}$  cm<sup>-2</sup>sec<sup>-1</sup>str<sup>-1</sup> (90% confidence level) for one year observation.

---

\* This study is being performed in collaboration with members of Akeno group.

1. Introduction

Gut monopoles<sup>1)</sup> are expected to be accelerated to the velocity about  $10^{-2}c \sim 10^{-3}c$  by the galactic magnetic field. The lowest possible velocity of monopoles on the earth is about  $10^{-4}c$  corresponding to the velocity of the earth in the solar system. Therefore, it is important to search magnetic monopoles with velocity between  $10^{-2}c$  and  $10^{-4}c$ .

Up to now, scintillation detectors and proportional counters with a large area have been used to search monopoles. However, the availability of these detectors is not clear for such a slowly moving monopole. Recently, Drell et al.<sup>2)</sup> have calculated the exact energy loss of slowly moving monopoles( $10^{-4}c < v < 10^{-3}c$ ) in He gas and pointed out the energy loss is much larger than minimum ionization energy of a singly charged particle.

At the center of Akeno air shower array<sup>3)</sup>, a large area(100m<sup>2</sup>) calorimeter is operating for the study of air showers. It is composed of 4 layers of proportional counters being sandwiched between concrete blocks. The structure of the calorimeter is very suitable for the monopole search experiment. However, the gas put into the counters is P10(Ar90% + CH<sub>4</sub>10%), in which the energy loss of slowly moving monopole is unknown.

The characteristic of the proportional counters filled with He gas have been investigated to examine the possibility of using the calorimeter for monopole research. As a result, no substantial difference has been found in the characteristics of the proportional counter at high density of particles between He and Ar mixture gas. We report the characteristics of proportional counters filled with He mixture gas and a plan of monopole search experiment using the calorimeter with these proportional counters

2. The apparatus of calorimeter

In Akeno air shower array, the calorimeter was constructed for the study of the high energy particles near the core of large air showers. It is composed of 4 layers of proportional counters, 18 and 20 bunches of 10 proportional counters<sup>4)</sup> in upper 3 layers and the bottom, respectively, and the concrete layers inbetween as shown in Fig. 1. Each proportional counter(10cmx10cmx500cm) is placed parallel to each other. Each counter is provided with an independent amplifier so that the particle density can be recorded for each counter, giving the spacial resolution of 10cm on the plane perpendicular to the long axis of counters.

### 3. Characteristics of the proportional counter filled with He mixture gas

Though the proportional counters of the calorimeter have been stable with P10 gas for 5 years at Akeno, it was not clear that counters with He gas could be stable too. Therefore, counters with different kinds of He mixture gas were prepared, and their characteristics and stability were studied.

#### 1). Characteristics of particle response

Several different kinds of He mixture gas were put into the proportional counters in the calorimeter and the response to cosmic rays was investigated. Fig. 2 shows the typical pulse height distribution (PHD) for the P10 gas and He mixture gas (He95% + CH<sub>4</sub>15%) for the incidence of single cosmic muons. High voltage supplied to each proportional counter was adjusted so as to give the same peak value for each counter. In Fig. 2 an excellent agreement in PHD can be seen between the P10 gas and He mixture gas except the characteristic x-ray peak of Zn(8.6KeV) for the P10 gas.

Next, the dependence of the signal on supply voltage was studied for a few kinds of He mixture gas, and the result is shown in Fig. 3. Among He mixture gases tested here, the gas of He85% + CH<sub>4</sub>15% (hereafter 15%CH<sub>4</sub>) is the best for the proportional counter because the voltage sensitivity is weakest (good stability for the voltage change), and the higher voltage can be applied to get high drift velocity.

The response of the proportional counter to high particle density was studied by observing air showers. Fig. 4 shows the response of proportional counters filled with 15%CH<sub>4</sub> and P10. These two types of counter were set at the same place in the calorimeter and were operated for the same period by air shower trigger. As seen in Fig. 4, two curves are in good agreement with each other, showing that the response of proportional counters filled with 15%CH<sub>4</sub> gas is reasonably good at least up to 10<sup>3</sup> particles per counter.

#### 2). Timing resolution

For slowly moving monopoles, it is very important to observe not only their energy loss in the detectors but also their velocity. Monopole's velocity can be determined using the timing information from more than two separated proportional counters, with an accuracy limited by the drift time of electrons in the counters. To know the timing resolution, a small scintillation detector was placed just above the proportional counter and the time intervals between their pulses produced by single cosmic ray muons passed through them were measured. In Fig. 5 is shown the time intervals between pulses of the

scintillation detector and the signals at the level of 90% of the average pulse height in the proportional counter filled with He mixture gas. In this case, as single cosmic muons pass through at various distances from the anode wire, the time intervals between the singles from the scintillation detector and the proportional counter distribute in wide time range. In Fig. 5, we can see that the timing resolution is extremely poor for the He mixture gases when compared to P10 gas counters, and 15%CH<sub>4</sub> gas has the highest resolution of timing among the group of He mixture gas tested here.

So next, the counter filled with the gas of 15%CH<sub>4</sub> was exposed to the accelerator beam (external beam of pions with the energy of 2 GeV) in KEK to know the drift time of electrons in the gas accurately. The discrimination level to the signals was fixed to about 20% of the average pulse height of signals induced by pion beams passing at a right angle to long axis of the counter. In Fig. 6 is shown the drift time measured at the various distances between the beam and the anode wire for two extreme cases of different injection angles of the beam and different setting angles of the counter. The vertical bars in Fig. 6 indicate the full width at a half maximum of the frequency distribution. Fig. 6 shows that the drift velocity does not vary with the beam injection angle and the counter's cross section to the beam in the distance range of 0 - 5cm from the anode wire. The full width at a half maximum of each measurement point corresponds to the drift time of about 5mm, which is just the size of the beam collimator in this experiment. It means that the spread of the drift time distribution is expected to be smaller than that of this experiment if the beam can be collimated in smaller space. The proportional counter filled with He mixture gas (15%CH<sub>4</sub>) can also be used as a drift chamber to determine the position of charged particle trajectory.

### 4. Plan of monopole search experiment

As mentioned above, He mixture gas (He85% + CH<sub>4</sub>15%) can be used for the proportional counter in the calorimeter with just the same quality as P10 gas. For monopole search experiment, at least one layer of the proportional counters should be added to the present structure of the calorimeter because there are four unknown parameters, zenith angle ( $\theta$ ), velocity ( $\beta$ ), the intersection point ( $x_0$ ) and the absolute time ( $t_0$ ), respectively when the monopole passed across the top layer of the proportional counter arrays.



### 1). Effective area

In this plan, it is desirable to add two more layers of the proportional counters on the top of the present calorimeter in order to determine the unknown parameters. Fig. 7 shows the effective area for the arrangement of this plan which depends on the incident zenith angle. The maximum effective area of the equipment is about  $160 \text{ m}^2 \text{ str}$ .

### 2). Trigger system

The air shower registration system<sup>5)</sup>, which records the particle densities of all proportional counters together with those of other detectors of the air shower array, can be used for the monopole search experiment. Fig. 8 shows the block diagram of the trigger circuit for the monopole search experiment. For the trigger system, 10 proportional counters are combined to one group.

A discriminator is commonly set for 10 proportional counters provided by an analogue-OR circuit. Moreover two neighboring groups of the counters are combined by an OR-circuit to make 9 blocks(10 blocks) of proportional counters for the upper 5 layers(bottom layer).

The trigger pulse is generated when the large energy loss more than the discrimination level occurs in all layers of the proportional counters. In order to exclude local showers, the trigger pulse is killed whenever the signals of more than 2 blocks of proportional counters in one layer exceed the discrimination level. To decrease the chance coincidence, a successive delayed coincidence method is adopted. A signal from first layer generates the gate pulse for second layer. If the signal of second layer is generated within the gate width, the gate pulse for third layer is generated. Thus successive gate pulses are generated by the signals of successive layers. For monopole search, we have to consider two direction of the successive delay coincidence for both upward and downward monopoles.

The detectable minimum velocity of monopoles depends on the discrimination level of the trigger requirement. We can estimate the signals induced by slowly moving monopoles in  $15\% \text{CH}_4$  gas using the results calculated by Drell et al.. Fig. 9 shows the relation between the signal size(monopole's energy loss divided by minimum ionization energy) and monopole's velocity. In this case, the results calculated by Drell et al. is decreased by a factor of two according to the new calculation<sup>6)</sup>. At the first stage of this experiment, the registration system of Akeno air shower array will be commonly used for the monopole search experiment and air shower experiment as well. Therefore it is necessary that the trigger rate of the monopole search experiment should

be low. Therefore, the discrimination level of the trigger requirement is set at the level corresponding to 10 times of minimum ionization with the successive gate width of  $16 \mu\text{sec}$ (maximum gate time for 6 layers is  $96 \mu\text{s}$ ). In this case, the monopoles with the velocity of more than  $5 \times 10^{-4} c$  can be detected by the apparatus as seen in Fig. 9. Expected upper limit to the monopole flux in this condition is  $2 \times 10^{-14} \text{ cm}^{-2} \text{ sec}^{-1} \text{ str}^{-1}$ (90% C.L.) for the observation time of one year, which is shown in Fig. 10 with the results of other groups.

To know the velocity of the monopoles passed through the apparatus, relative time differences between the signals of all layers are measured. In this case, the time differences are measured with an unit of every 10 proportional counters which are connected with an analogue-OR circuit and the discrimination level for this purpose is set at about 20% of the trigger threshold to exclude the effects of the corner region of the proportional counter.

For the monopoles with the velocity  $\beta < 5 \times 10^{-4} c$ , it is planned in the second stage of monopole research to introduce a new registration system because the trigger rate is expected to be so high that it is impossible to use the present registration system of Akeno air shower array commonly with the monopole search.

### References

- 1) H.Georgi and S.L.Glashow, Phys. Rev. Lett. 32, 438 (1974).
- 2) S.D.Drell et al., Phys. Rev. Lett. 50, 644 (1983).
- 3) T.Hara et al., Proc. 16th Int. Conf. on Cosmic Ray, Kyoto 8, 135 (1979).
- 4) N.Hayashida and T.Kifune, Nucl. Inst. and Meth. 173, 431 (1980).
- 5) T.Hara et al., Proc. 16th Int. Conf. on Cosmic Ray, Kyoto 11, 166 (1979).
- 6) N.M.Croll, Monopole 83 Conf. at Univ. of Michigan (1983).
- 7) P.C.Bosetti, 4th Workshop on Grand Unification, Philadelphia, (1983).

### Figure captions

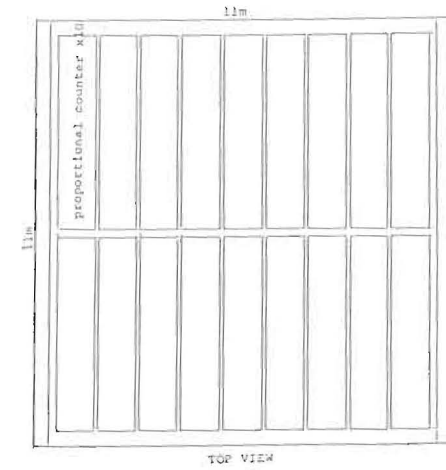
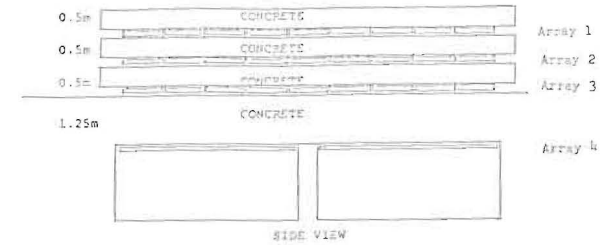
Fig. 1 Structure of the calorimeter.

a) whole view.

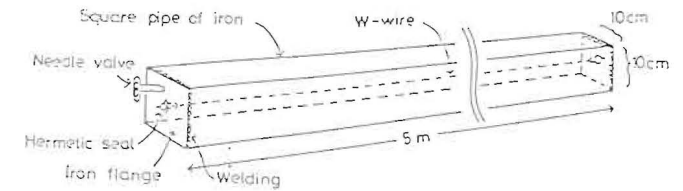
b) proportional counter.

Fig. 2 Pulse height distribution of proportional counter.

- Fig. 3 Correlation between signals and supply voltage for various gases.
- Fig. 4 Density spectrum of air shower particles observed by the proportional counter.
- Fig. 5 Arrival time distributions of 90% level of average signal voltage in various gases.
- Fig. 6 Drift time in the proportional counter.  
The discrimination level is set at about 20% of the peak voltage induced by beam passed at a right angle to long axis of the counter.
- Fig. 7 Effective area of the calorimeter for monopole search experiment.
- Fig. 8 Block diagram of the trigger system.
- Fig. 9 Ratio of the energy loss of magnetic monopole to the minimum ionization energy as a function of monopole's velocity.
- Fig. 10 Upper limits on the flux of magnetic monopoles<sup>?)</sup>.



(a)



(b)

FIG. 1

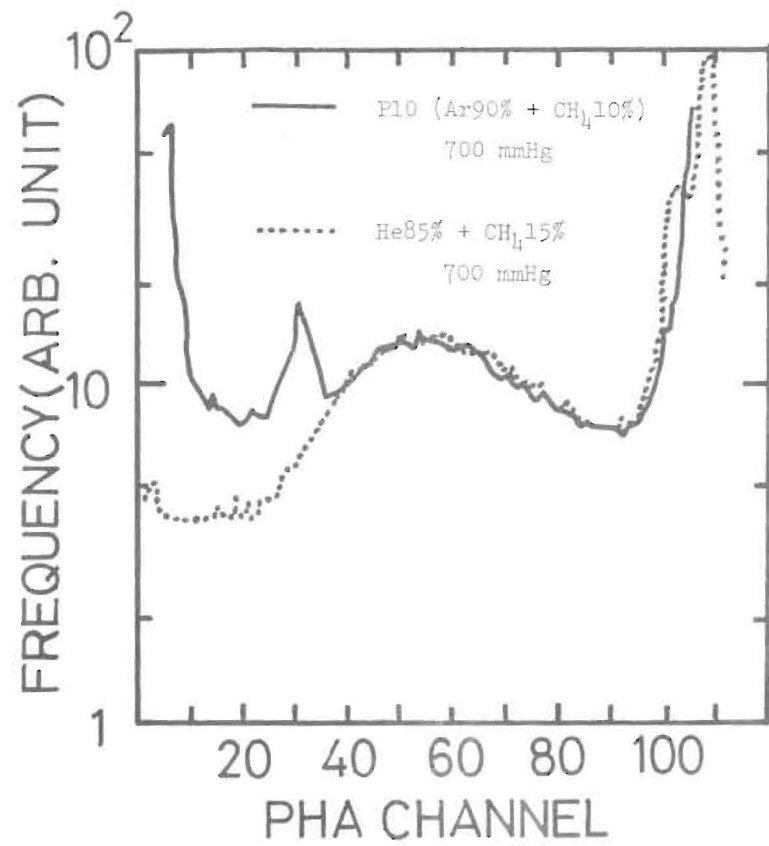


FIG. 2

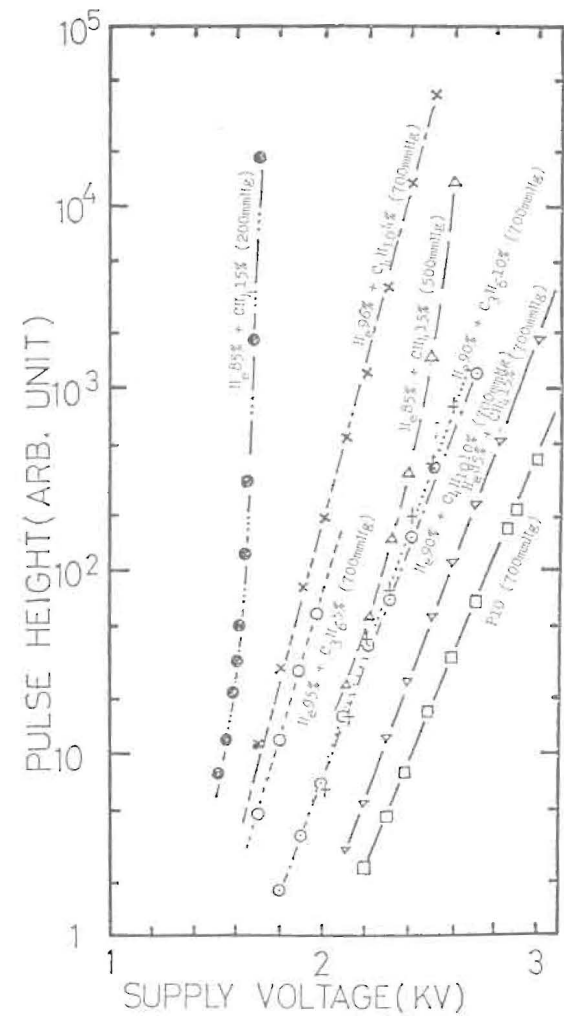


FIG. 3

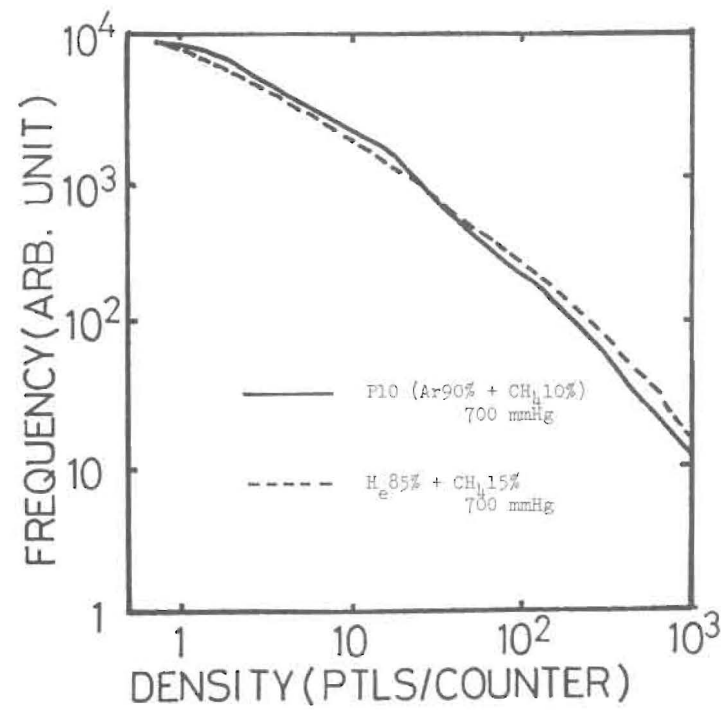


FIG. 4

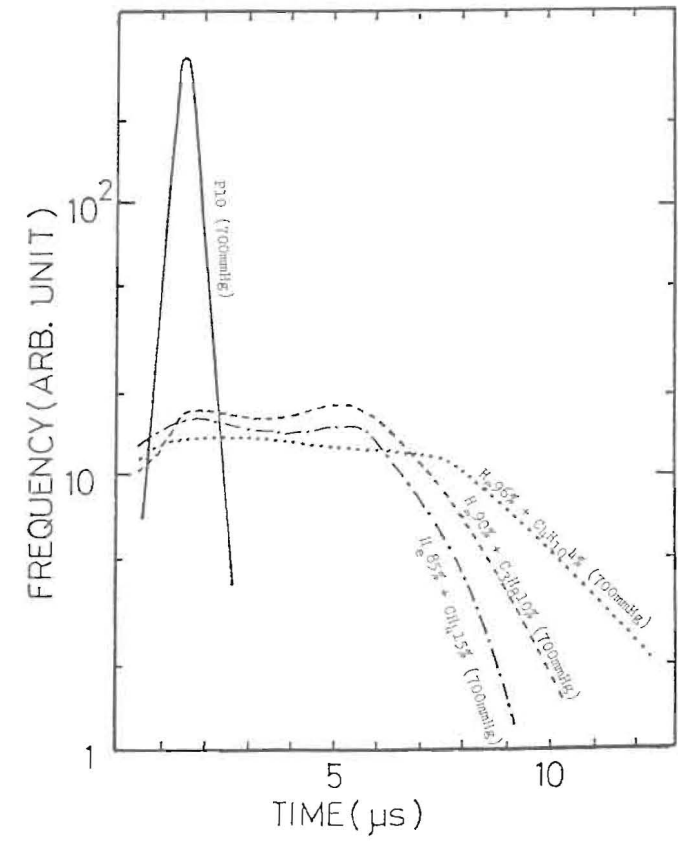


FIG. 5

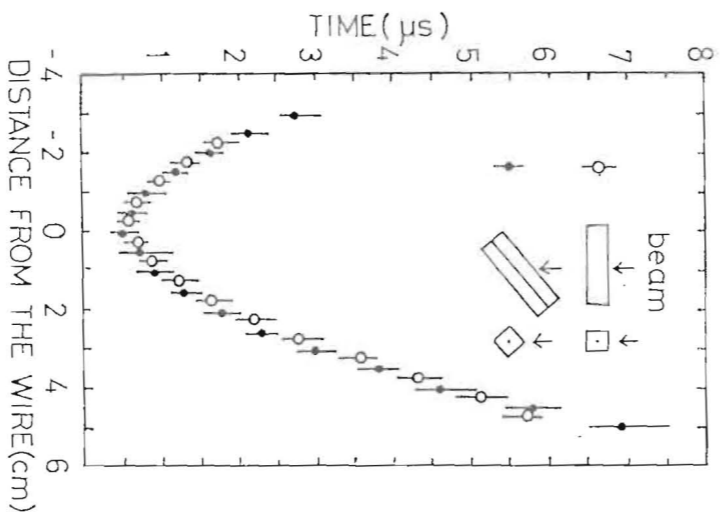


FIG. 6

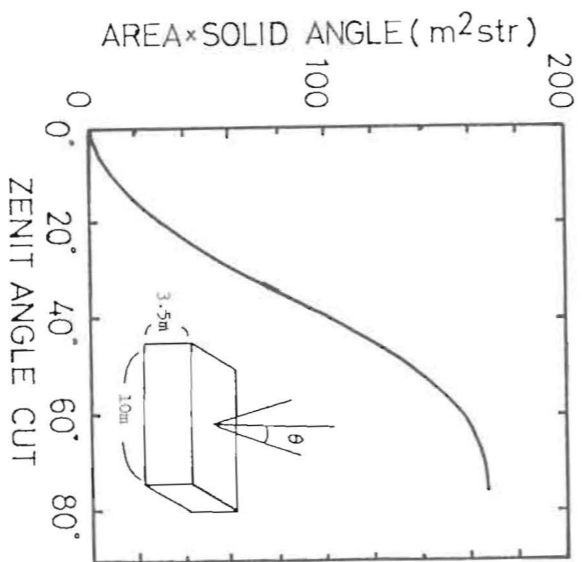


FIG. 7

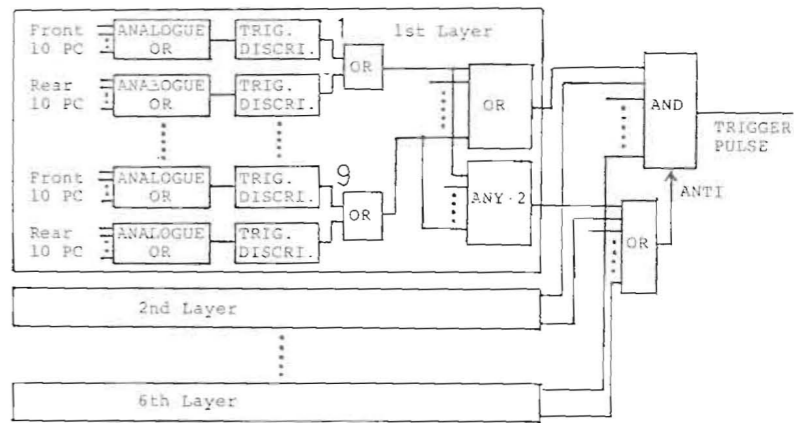


FIG. 8

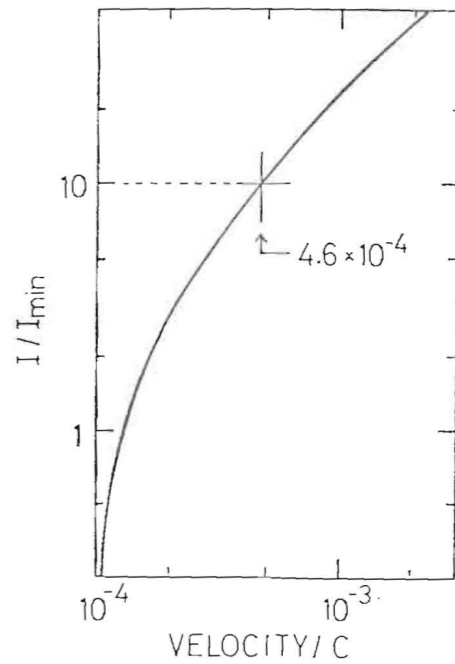


FIG. 9

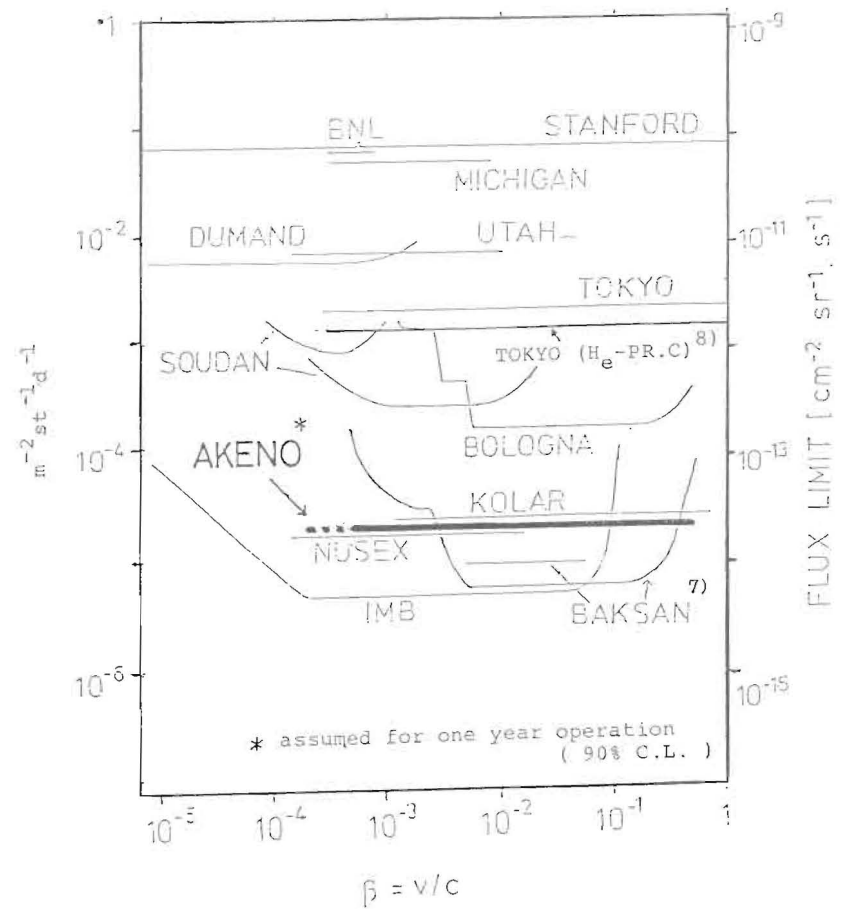


FIG. 10

Present Status of Monopole-Search  
with Superconducting Induction Coils

Takeo EBISU and Tadashi WATANABE

Department of Physics, Faculty of Science, Kobe University  
Nada-ku, Kobe 657, Japan

March 1984

ABSTRACT

An overview is given on the status of search for magnetic monopoles using superconducting induction coils. First, the elements of the technique are recapitulated. Second, some features of operating fluxmeters, of Stanford Univ., IBM, Chicago-Fermilab.-Michigan and Kobe Univ., are reviewed and the upper limit of monopole flux obtained thereof is reported.

§1. Introduction

The race to find traces of magnetic monopoles was accelerated by Cabrera's candidate event and has involved many scientists from various fields; not only cosmic-ray, high-energy and low-temperature experimentalists but also particle theorists, cosmologists and astrophysicists. Thus so many research groups are pursuing them with various kinds of detecting devices: Emulsion, Scintillation counter, Proportional chamber, Solid state track detector, Induction coils, Detection of successive nucleon decay and so forth<sup>1)</sup>.

This short report intends to survey the status of search for magnetic monopoles with superconducting induction coils; first, to recapitulate the elements of the technique and second, to review some features of operating fluxmeters and the upper limit of monopole flux obtained thereof.

§2. Elements of Superconducting Induction Coil Technique

Magnetic monopole search with superconducting ring is now a well-known method. It is based on two principles of physics, extended electromagnetism and superconductivity.

According to Faraday's law of induction the change of magnetic flux through a conducting loop causes a current in it given by

$$\Delta I(t) = - \Delta \phi_L(t) / L \quad (1)$$

with  $L$  the loop self-inductance. The simplest way to obtain the flux  $\phi_L$  is perhaps to combine fluxes  $\phi_M$  and  $\phi_T$  due to magnetic charge  $g$  and magnetic current  $\vec{j}_m$ , respectively.

$$\phi_T(t) = \phi_M(t) + \phi_I(t) \quad (2)$$

When a monopole is bound for the ring (Fig. 1a) with velocity  $v$  along the axis, the charge  $g$  makes a flux

$$\phi_M(t) = g\Omega(t) - 4\pi g\theta(t) \quad (3)$$

where  $\Omega = 2\pi(1 + \gamma vt / \sqrt{(\gamma vt)^2 + r^2})$  is the solid angle subtended by the ring at the monopole ( $\gamma = v/c$ ). The total flux in the ring area is derived by using Faraday's law including magnetic current,

$$\text{rot } \vec{E} + \frac{1}{c} \frac{\partial \vec{B}}{\partial t} = -\frac{4\pi}{c} \vec{j}_m(t), \quad \vec{j}_m = g\vec{v}\delta(x)\delta(y)\delta(z-vt) \quad (4)$$

Integrating Eq. (4) over the ring area and neglecting safely the electromotive force along the superconducting loop, we get

$$\dot{\phi}_T(t) = -4\pi g \dot{\theta}(t) \quad (5)$$

Thus induced flux is obtained as

$$\Delta\phi_I = \phi_I(\infty) - \phi_I(-\infty) = 4\pi g = hc/e \quad (6)$$

Dirac's quantization condition<sup>2)</sup> gives  $g = n \hbar c / 2e = 3.29 \times 10^{-8}$  CGS Gaussian units ( $n=1$ ) and  $\Delta\phi_I = 4.13 \times 10^{-7} \text{G}\cdot\text{cm}^2$ . The flux change  $\Delta\phi_I$  is exactly twice the magnetic flux quantum  $\phi_0$  in superconductivity. We note here some typical values of a fluxmeter. When a three-turn, 8-cm-diam search coil is made of  $5 \times 10^{-3}$ -inch-diam Nb wire, the loop self-inductance turns out to be around 3  $\mu\text{H}$ . Correspondingly we have induced current  $\Delta I$  of around  $1.3 \times 10^{-9} \text{A}$ .

How can we detect this minor current? It will be transient and decay instantaneously in the characteristic time  $L/R$  in the normal state coil ( $R$ : resistance of the coil), while in the superconducting one it will be persistent. With a persistent current, even if it were smaller by several orders of magnitude than one caused by a monopole, the signal can be caught by superconducting quantum interference device, SQUID<sup>3)</sup>, coupled to the search coil. On the bases of two phenomena in superconductivity, London's fluxoid quantization and tunneling of the Cooper pairs, the device is composed of a superconducting ring with one/two Josephson weak links and has a flux sensi-

tivity of  $\Delta\phi \approx 10^{-3}\phi_0$ , while the induced flux is given by  $\Delta\phi_I = 2\phi_0$ . A schematic diagram (Fig. 2) shows the arrangement for a superconductive induction detector.

On considering that the detector is steeped in the geomagnetic field ( $\approx 0.3\text{G}$ ) and in the fluctuation ( $\approx 10^{-5}\text{G}$ ), it is imperative to shield magnetically the search coil. Usage of the mumetal cylinder can reduce the ambient field by around 2 orders of magnitude. By shielding the coil by something like superconducting lead foil, the trapped field could be frozen in place and the field inside the shield case will be stabilized on account of the Meissner effect. Furthermore the Stanford Univ. group has succeeded to obtain the ultra-low magnetic field through dilution by expanding several bags of lead foil from the outside in ( $\approx 10^{-8}\text{G}$ ).

Magnetic shield case shares the flux to modify the magnitude of induced current in the ring. As a result of the effect the current of penetrating monopole will be diminished and that of near-miss one will be given rise to (Fig. 1b).

Representative methods for detecting monopoles are summarized in Table 1 for comparison<sup>1)</sup>. It will be confirmed that an unambiguous result is expected only when induction coil technique, since SQUID detector can distinguish the passage of monopole and/or nucleus-monopole compound from that of magnetic dipole and is insensitive to other properties such as the electric charge, mass, velocity and above all the energy-loss rate of monopoles in media. The only drawback will be difficulty in increasing detection area. However, some trials and proposals to accommodate larger loops or search for monopoles accumulated in bulk matter during long term will be found in the next section.

### §5. Current Working Superconducting Detectors

Some features of current working detectors and the results thereof are reviewed. Arguments presented here are mainly based on talks given at



Monopole '83 conference held at the University of Michigan, 6-9 October, 1983.

Stanford Univ.

Starting with the 5-cm-diam ring<sup>4)</sup>, which succeeded to detect the candidate event, they are now searching for cosmic ray monopoles with triaxial 10-cm-diam<sup>5)</sup> rings, with 70-cm<sup>2</sup> isotropic sensing area (Fig. 3) in the ambient field  $5 \times 10^{-8}G$  stated in §2. The flux is obtained as less than  $2.1 \times 10^{-11} \text{cm}^{-2} \text{sr}^{-1} \text{sec}^{-1}$  (90% C.L.) as of Oct. 4 '83.

In order to enlarge the area it is proposed to scan the shield bag itself with a sensitive magnetometer and to detect the twin magnetic vortices left by the passage of a monopole.

IBM

They have developed the planar gradiometer<sup>6)</sup> as the pick-up coil which consists of coplanar superconducting loops, wound in opposite directions and connected in series (Fig. 4). The important feature of the coil is to allow large detection area. While providing low sensitivity to external magnetic field changes and low self-inductance, it remains sensitive to local flux changes such as that caused by a monopole passing through any one of the cells.

The upper limit of monopole flux is set as  $1.1 \times 10^{-10} \text{cm}^{-2} \text{sr}^{-1} \text{sec}^{-1}$  in 165 days (Mar.-Sept. '83) with the prototype gradiometer ( $\approx 50\text{-cm}^2$  sensing area). They will set it as  $1.3 \times 10^{-12}$  and  $5 \times 10^{-14} \text{cm}^{-2} \text{sr}^{-1} \text{sec}^{-1}$ , by Oct. '84 with the 2000-cm<sup>2</sup>-area detector and by one year later from '84 with 5-m<sup>2</sup>-area detector, respectively, if no candidates observed.

Chicago-Fermilab.-Michigan

This group has also developed a kind of planar gradiometer, "Macrame"<sup>7)</sup>. They set up 2 macrames of around 60 cm diam. into double layers (Fig. 5) and have obtained the upper bound as  $2.2 \times 10^{-10} \text{cm}^{-2} \text{sr}^{-1} \text{sec}^{-1}$  (90% C.L.) in 13 days 17 hours (Aug. 29 - Oct. 4 '83). It is planned to construct more than 1-m-diam

macrame.

Kobe Univ.

The detector is a very simple one; an 8-cm-diam, 3-turn ring<sup>8,9)</sup>. They are searching for not only incident monopoles but also ones trapped magnetically in old iron ore, which would have accumulated them though the flux in cosmic ray is very small.

By heating old iron ore, magnetic sand ( $6 \sim 25 \times 10^6$  years old) and maghemite ( $\sim 100 \times 10^6$  y) above their Curie point, they have tried to detect superheavy monopoles passing through the search coil, pulled downward by the gravitational force. They have not observed any signal consistent with the passage of Dirac-charge monopole and set the limit as  $2.3 \times 10^{-6}$  monopoles/gram. In addition from the running time of the detector, more than 1000 hours (Jan. 13 - Aug. 12 '83) the upper limit of monopole flux is set as  $4.6 \times 10^{-10} \text{cm}^{-2} \text{sec}^{-1} \text{sr}^{-1}$ .

As for the next experiment is planned a search for trapped monopoles in iron ore aged  $10^9$  years by the heat-treatment method with a larger-scale coil (Fig. 6).

#### References

1. A complete discussion of the material will be found in *Magnetic Monopoles*, edited by R. A. Carrigan, Jr. and W. P. Trower (Plenum, New York, 1983) and in *Proceedings of Monopole '83* (Plenum New York, 1984 to be published).
2. P. A. M. Dirac, Proc. Roy. Soc. London, Ser. A, 133 (1931) 60.
3. See, for instance, O. V. Lounasmaa: *Experimental Principles and Methods Below 1K* (Academic Press 1974) Chap. 7.
4. B. Cabrera, Phys. Rev. Lett. 48 (1982) 1378.
5. B. Cabrera, et al., Report at *Monopole '83* and Phys. Rev. Lett. 51 (1983) 1933.
6. J. Chi, Report at *Monopole '83*.
7. J. Incandela, Report at *Monopole '83*.
8. T. Ebisu and T. Watanabe, J. Phys. Soc. Jpn. 52 (1983) 2617.
9. T. Ebisu and T. Watanabe, Preprint "SQUID Fluxmeter Search for Monopoles in 400 kg of Old Iron Ore" (KOB-83-11), presented at *Monopole '83*.

#### Figure and Table Captions

- Fig. 1a A one-turn coil shielded with a cylindrical superconducting case co-axially and typical seven trajectories of monopoles.
- Fig. 1b Flux responses  $A \sim B$  and  $a \sim c$  when non-shielded, and  $A' \sim B'$  and  $a' \sim c'$  when shielded.
- Fig. 2 Experimental arrangement for a superconductive induction detector. The flux transformer is shielded superconductively as a whole and the temperature is kept at 4.2K.
- Fig. 3 Current detector at Stanford Univ..
- Fig. 4 Current detector of IBM and their gradiometer.
- Fig. 5 Current detector of Chicago-Fermilab-Michigan group.
- Fig. 6 Detectors of Kobe Univ..
- Table 1. Comparison of several methods detecting monopoles.

Table 1.

Detecting Methods	Properties	Physics
Emulsion, Scintillation c. Proportional c. Plastic sheet	$\frac{g^2}{\hbar c} = 34 \gg \frac{1}{137}$	Ionization
Induction coils	Charge quantum $g$ Mag. current $j_m$	Faraday's law
Proportional c.	Angular mt. of e-g system	Excitation (Drell's mech.) 1983
Successive nucleon decay	Non-Abelian structure	Rubakov's effect 1982

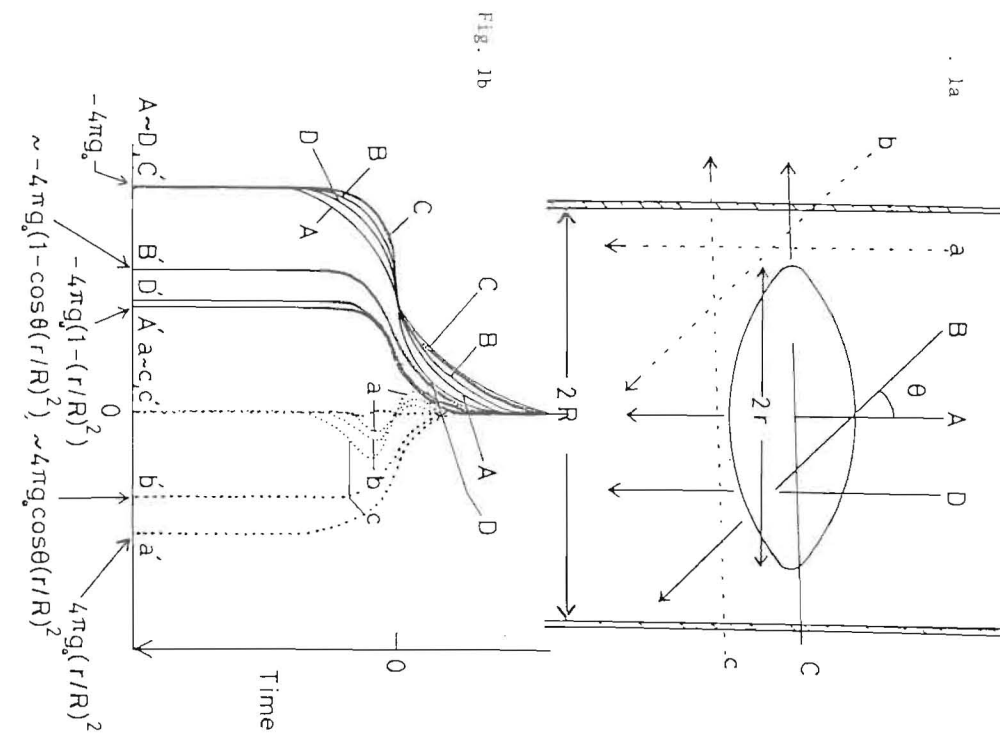


Fig. 1b

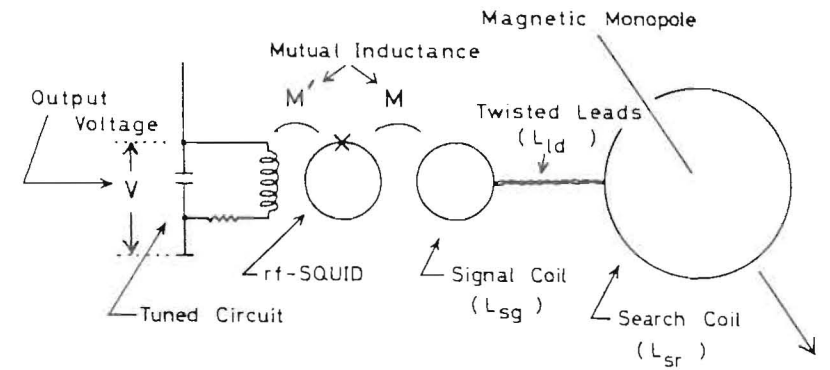


Fig. 2 Flux Transformer

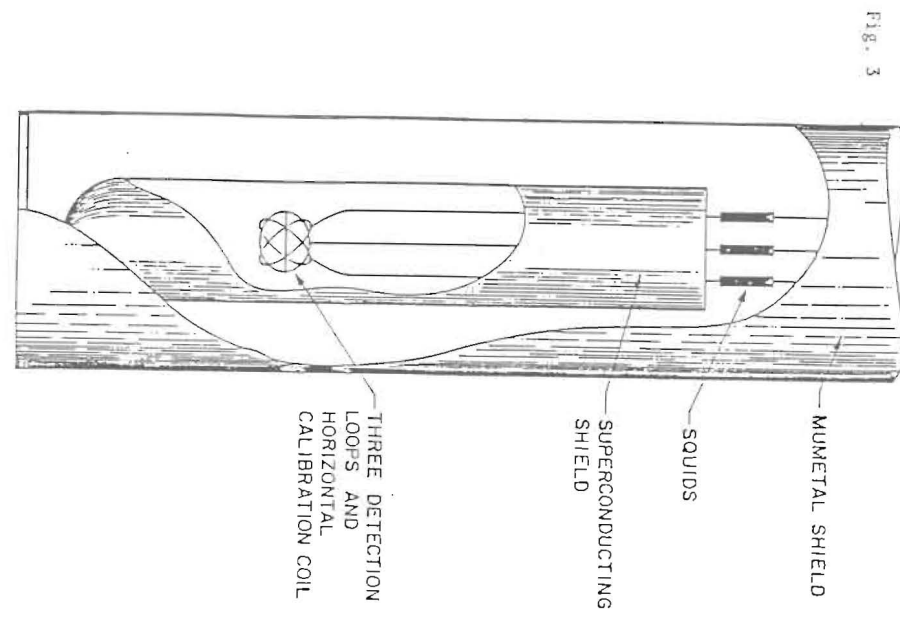


Fig. 3

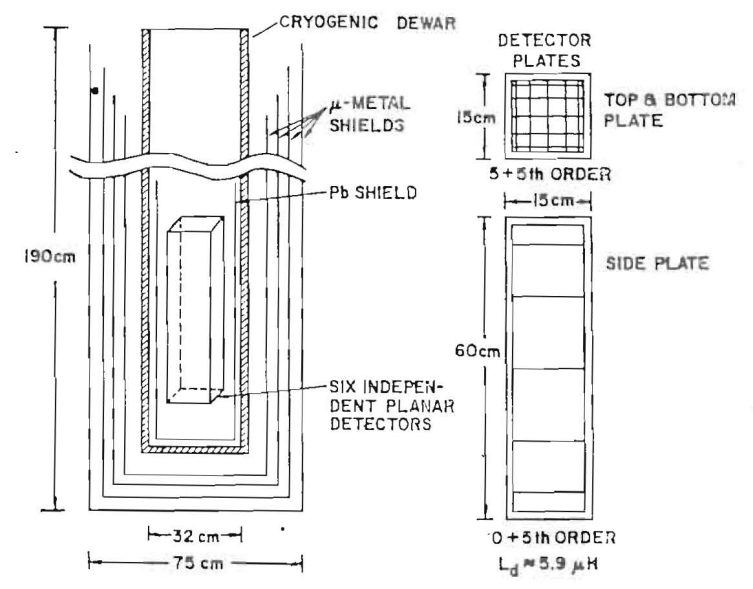


Fig. 4

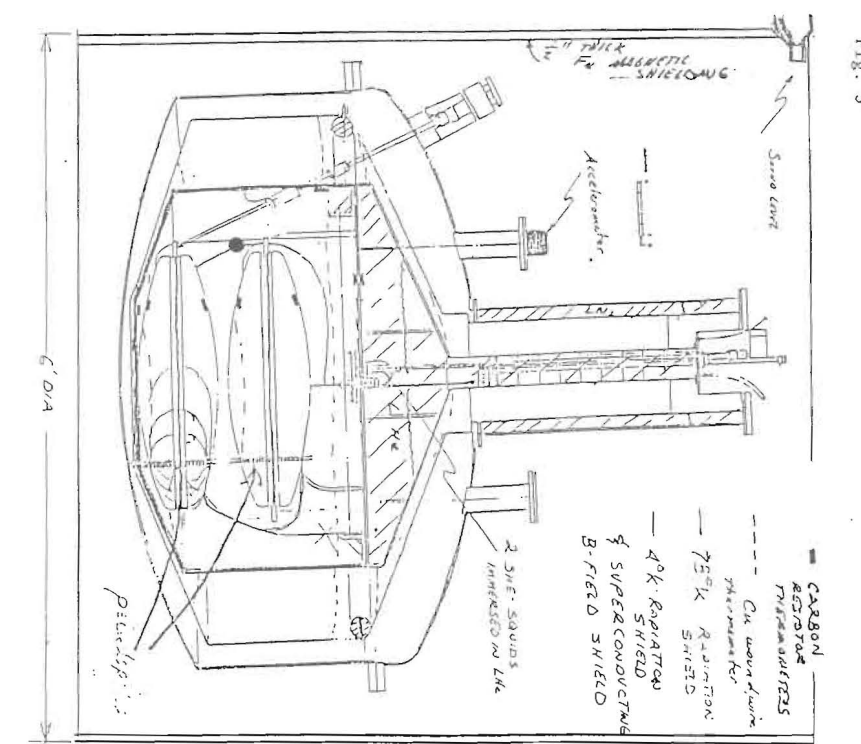
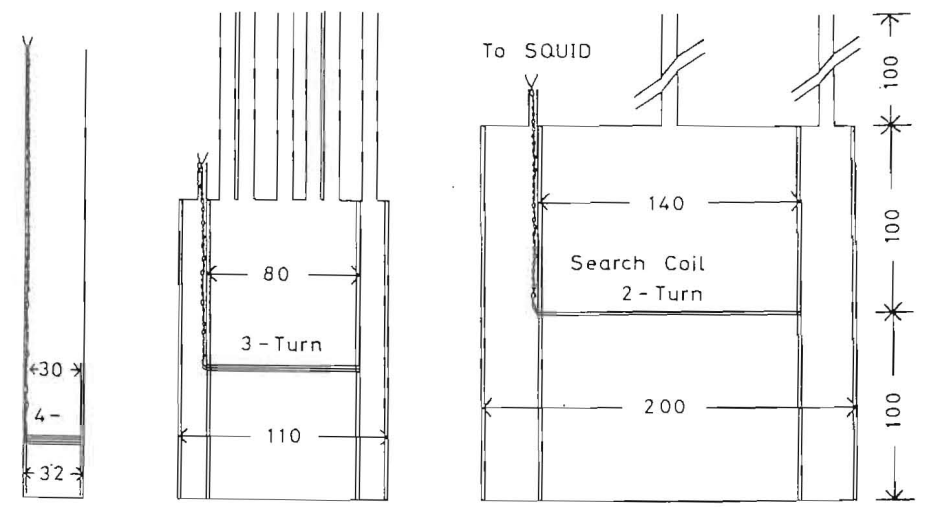


Fig. 5



700 gr	428.4 kg	2000 kg
50 H	1010 H	(2000 H)
July, 1982	Aug., 1983	?, 1984

Fig. 6

## Recent Observations of the Galactic Center

Junji Inatani

Nobeyama Radio Observatory,  
Tokyo Astronomical Observatory, University of Tokyo  
Nobeyama, Minamisaku, Nagano 384-13, Japan

(to appear in the proceedings of KEK Symposium on  
"GUT and Astrophysics" held on 7-10 Dec. 1983)

### Abstract

Recent observational topics on the galactic center are briefly reviewed. Main interest is gas conditions in the vicinity of the nucleus. A rapidly rotating molecular cloud and a spiral-like ionized gas are proposed to be in an intimate physical correlation.

## 1. Introduction

The galactic center was recognized for the first time with its strong radio continuum emission. The dominant emission at long wavelengths (roughly longer than 10 cm) is synchrotron radiation due to high energy electrons, whose distribution shows a smooth concentration to the galactic center<sup>1)</sup>. As we go to shorter wavelengths, many discrete sources appear, which are either supernova remnants (synchrotron radiation) or thermally ionized regions (HII regions)<sup>2)</sup>. The brightest source among them is called Sgr A, which is located at the center of the galactic rotation<sup>3)</sup>. Sgr A is divided into two components "East" and "West" (Fig. 1)<sup>4)</sup>. Sgr A East is a supernova remnant, and Sgr A West is a thermal source which includes a very compact nonthermal source within it. This compact core is regarded as the central object (nucleus) of our galaxy.

Another prominent property of the galactic center is that it is not only the center of the galactic rotation but also the origin of radial motions observed in a large scale gas distribution<sup>5)</sup>. For example, a molecular ring with a radius of about 250 pc (1 pc =  $3 \times 10^{18}$  cm) is regarded as a remnant of an explosion which occurred at the nucleus a million years ago<sup>6)</sup>.

The third property of the galactic center is a strong mass concentration. Mass density is derived either from kinematics of neutral hydrogen gas (emission at 21 cm) or from IR photometry of the stellar luminosity<sup>5)</sup>. Those results agree with each other and shown in Table 1. It is clear from this table that the effect of tidal disruption is important in the vicinity of the nucleus.

## 2. Nucleus

The nucleus has been investigated with VLBI observations<sup>7)</sup>. Its observed size varies according to  $\lambda^2$  ( $\lambda$  = observed wavelength), which is interpreted as the effect of scattering or optical depth in the source. The actual diameter is estimated to be  $5 \times 10^{14}$  cm. The brightness temperature is then calculated to be  $4 \times 10^8$  K, so this radiation is regarded to be nonthermal. Time variation of the radio flux is also reported, which has a time scale of several months.

One of the recent important observations on the nucleus is 511 keV line emission, which is attributable to positron annihilation<sup>8)</sup>. It is reported that the line flux was as high as  $2 \times 10^{37}$  ergs<sup>-1</sup> and that it decreased by a factor of three within half a year.

Another important result is a very broad emission of He 4857  $\text{cm}^{-1}$  line. The line width corresponds to a velocity dispersion of 1500  $\text{kms}^{-1}$ , if it is due to the Doppler effect. If this is a typical circular velocity in the nucleus, the central mass (probably a black hole) should be  $10^5 M_{\odot}$ .

### 3. Spiral-like ionized gas

Gas distribution and its kinematics in the vicinity of the nucleus will be reviewed in the following. There are two important observations on the ionized gas within a few pc of the nucleus. One of them is Ne II 12.8  $\mu\text{m}$  observation by Lacy et al.<sup>10)</sup>. They found 14 clouds of ionized gas which are moving back and forth around the nucleus at high velocities up to 260  $\text{kms}^{-1}$ . They concluded from this fact that it is most probable to assume a point mass of  $3 \times 10^6 M_{\odot}$  at the nucleus.

Another interesting result is obtained with the Very Large Array in USA. Brown et al. and Ekers et al.<sup>11)</sup> revealed a spiral-like feature of ionized gas around the nucleus (Fig. 2). Velocity of this ionized gas is measured with a radio recombination line of hydrogen<sup>12)</sup>. The north arm of this spiral is moving back and the south arm is moving forth at velocities of 50 - 100  $\text{kms}^{-1}$ .

### 4. Rotating Molecular Cloud

A new information on neutral gas distribution is recently obtained with the 45 meter telescope of Nobeyama Radio Observatory<sup>13)</sup>. We have observed a HCN emission at 3.4  $\mu\text{m}$  with a spatial resolution of 18 seconds of arc. This has revealed the existence of a rapidly rotating molecular cloud in approximately the same region as the spiral-like ionized gas (Fig. 3). According to a CO observation by Liszt et al.<sup>14)</sup>, we can further recognize an outer part of this rotating cloud. Several parameters of this cloud are summarized in Table 2.

We can compare this molecular cloud with IR distributions. Becklin et al. have shown the existence of a warm dust cloud (60 - 100 K) with a doughnut-like shape within 3 pc of the nucleus<sup>15)</sup>. Molecules and dusts are therefore considered to coexist in the same rotating cloud.

A possible model of gas distributions and kinematics is presented in Fig. 4<sup>16)</sup>. It is assumed that the gravitational field is slightly non-axisymmetric. A supersonic gas flow in this field forms a pair of shocked layers, which will ionize the gas to make a spiral-like feature as observed.

### References

1. For example,  
H.V. Cane, Aust. J. Phys. 31 (1978) 561.  
J.C. Novaco and L.W. Brown, Astrophys. J. 221 (1978) 114.
2. W.J. Altenhoff, D. Downes, L. Goad, A. Maxwell and R. Reinhart, Astron. Astrophys. Suppl. 1 (1970) 319.  
W.J. Altenhoff, D. Downes, T. Pauls and J. Schraml, Astron. Astrophys. Suppl. 35 (1978) 23.
3. D. Downes, W.M. Goss, U.J. Schwarz and J.G.A. Wouterloot, Astron. Astrophys. Suppl. 35 (1978) 1.
4. R.D. Ekers, W.M. Goss, U.J. Schwarz, D. Downes and D.H. Rogstad, Astron. Astrophys. 43 (1975) 159.
5. For a review,  
J.H. Oort, Ann. Rev. Astron. Astrophys. 15 (1977) 295.  
C.H. Townes, J.H. Lacy, T.R. Geballe and D.J. Hollenbach, Nature 301 (1983) 661.
6. N. Kaifu, T. Iguchi and T. Kato, Publ. Astron. Soc. Japan 26 (1974) 117.
7. K.Y. Lo, M.H. Cohen, A.S. Readhead and D.C. Backer, Astrophys. J. 249 (1981) 504.  
K.Y. Lo, in AIP Conf. Proc. No. 83 "The Galactic Center", ed. by G.R. Riegler and R.D. Blandford (1982).  
K.Y. Lo, in IAU Symp. No. 110 "VLBI and Compact Radio Sources" (1983).
8. M. Leventhal, C.J. MacCallum and P.D. Stang, Astrophys. J. (Letters) 225 (1978) L11.  
G.R. Riegler, J.C. Ling, W.A. Mahoney, W.A. Wheaton, J.B. Willett, A.S. Jacobson and T.A. Prince, Astrophys. J. (Letters) 248 (1981) L13.
9. D.N.B. Hall, S.G. Kleinmann and N.Z. Scoville, Astrophys. J. (Letters) 262 (1982) L53.
10. J.H. Lacy, C.H. Townes, T.R. Geballe and D.J. Hollenbach, Astrophys. J. 241 (1980) 132.  
J.H. Lacy, C.H. Townes and D.J. Hollenbach, Astrophys. J. 262 (1982) 120.
11. R.L. Brown, K.J. Johnston and K.Y. Lo, Astrophys. J. 250 (1981) 155.  
R.L. Brown, Astrophys. J. 262 (1982) 110.  
R.D. Ekers, J.H. van Gorkom, U.J. Schwarz and W.M. Goss, Astron. Astrophys. 122 (1983) 143.

12. J.D. Bregman and U.J. Schwarz, *Astron. Astrophys.* **112** (1982) L6.
13. N. Kaifu, J. Inatani, T. Hasegawa and M. Morimoto, in *IAU Symp. No. 106 "The Milky Way"* (1983).
14. H.S. Liszt, J.M. van der Hulst, W.B. Burton and M.P. Ondrechen, *Astron. Astrophys.* **126** (1983) 341.
15. E.E. Becklin, Ian Gatley and M.W. Werner, *Astrophys. J.* **258** (1982) 135.
16. T. Matsuda, K. Sawada and I. Hachisu, preprint (1983).

Captions

- Fig. 1 5 GHz map of Sgr A with an angular resolution of  $6.3'' \times 34''$  (R.A. x Dec.). Structures of both Sgr A East and Sgr A West are resolved. This figure is taken from Ekers et al. (1975)<sup>4)</sup>.
- Fig. 2 Spiral-like ionized gas observed at 2 cm. The angular resolution is  $2'' \times 3''$  (R.A. x Dec.). Small circles represent the size of the Ne II clouds, the numbers are their radial velocities ( $\text{kms}^{-1}$ )<sup>10)</sup>. This figure is taken from Ekers et al. (1983)<sup>11)</sup>.
- Fig. 3 Rotating molecular cloud observed with a HCN emission. These maps show the spatial distribution of molecules which have the certain velocities just in the range given in each map. Central cross of each map indicates the position of the nucleus.  $\Delta l$  and  $\Delta b$  are spatial offsets from the nucleus parallel and perpendicular to the galactic plane. The upper-left (northern) part of the molecular cloud is moving away from us, and the lower-right (southern) part is approaching us.
- Fig. 4 A model of gas conditions in the vicinity of the nucleus. Dotted area shows a spiral-like ionized gas, which is regarded to be made by a supersonic gas flow in a non-axisymmetric gravitational field. This picture is taken from Matsuda et al. (1983)<sup>16)</sup>.
- Table 1 Gravitational field around the nucleus. R: radius,  $M_x$ : mass within R,  $V_x$ : circular velocity,  $T_x$ : time of revolution,  $n_x$ : critical density for tidal disruption,  $\tau_{tid}$ : time scale for tidal disruption.
- Table 2 Parameters of the rotating molecular cloud.  $\tau_{rot}$  and  $\tau_{exp}$  are time scales for: the rotation and the probable expansion of the cloud.

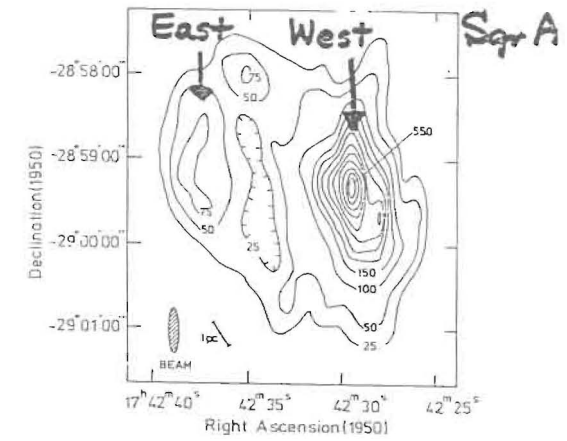


Fig. 1

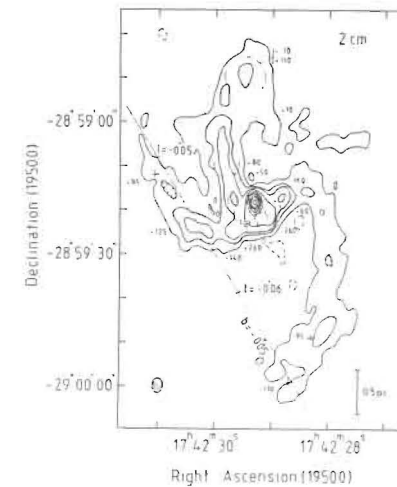


Fig. 2

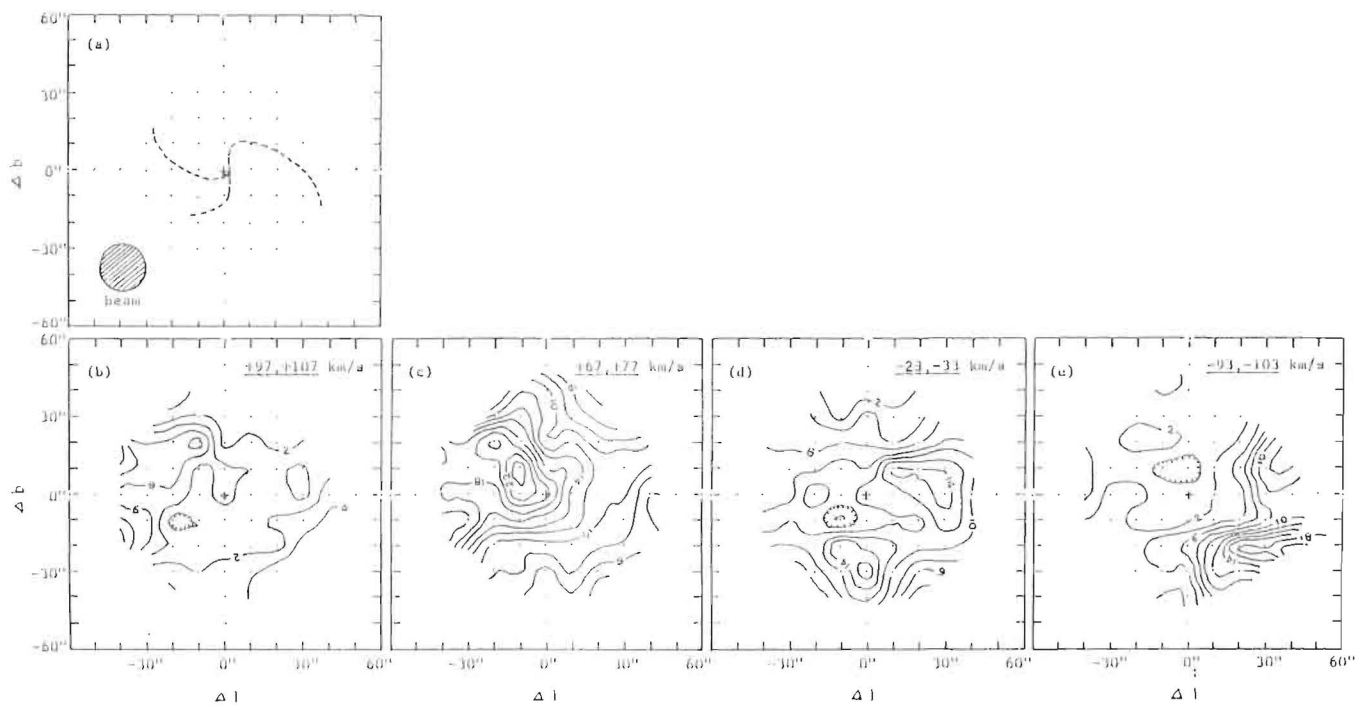


Fig. 3

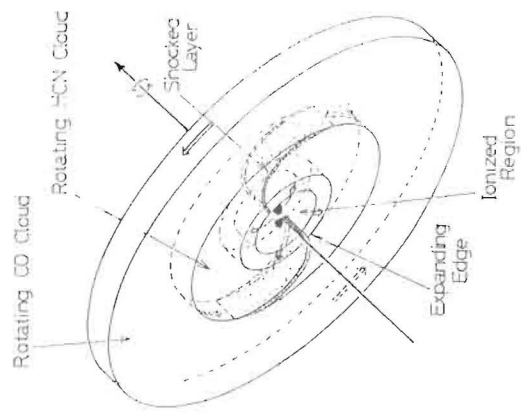


Fig. 4



R	$M_*$	$V_*$	$T_*$	$n_*$	$\tau_{tid}$
pc	$10^6 M_\odot$	$\text{km s}^{-1}$	$10^4 \text{ yr}$	$10^6 \text{ cm}^{-3}$	$10^3 \text{ yr}$
0.1	0.3	114	0.6	1500	0.9
1	4.4	138	4.5	22	7.6
5	30	161	20	1.2	33
10	70	174	36	0.4	60

Table 1

	" HCN cloud "	" CO cloud "
diameter	4.5 pc	18 pc
thickness	3 pc	6 pc
mass	$2 \times 10^3 \left( \frac{n_{\text{H}_2}}{10^3 \text{ cm}^{-3}} \right) M_\odot$	$7 \times 10^4 \left( \frac{n_{\text{H}_2}}{10^3 \text{ cm}^{-3}} \right) M_\odot$
$V_{\text{rot}}$	60 $\text{km s}^{-1}$ at 2.2 pc	100 $\text{km s}^{-1}$ at 9 pc
$\tau_{\text{rot}}$	$2 \times 10^5 \text{ yr}$	$6 \times 10^5 \text{ yr}$
$V_{\text{exp}}$	70 $\text{km s}^{-1}$	70 $\text{km s}^{-1}$ (?)
$\tau_{\text{exp}}$	$3 \times 10^4 \text{ yr}$	$1 \times 10^5 \text{ yr}$ (?)

Table 2

## Infrared Observation of The Early Universe

T. Matsumoto

Department of Astrophysics, Nagoya University  
Nagoya , 464 Japan

### ABSTRACT

The rocket observation of the near-infrared extragalactic background radiation and its influence on the cosmology are described. The future plans to observe the near-infrared and far-infrared backgrounds are also presented.

### I. INTRODUCTION

Early history of the universe has been one of the most important problems in cosmology. Since the discovery of 3K cosmic background radiation (Penzias and Wilson 1965) the big bang origin of the universe has been extensively investigated. However, physical processes taking place at epoch  $5 < z < 1000$  are not well known due to the lack of observational materials.

Partridge and Peebles (1967) have proposed evolutionary models that galaxies may have experienced a very bright phase at its formation. Having been stimulated by their result, many people have tried to observe individual young galaxies and/or integrated background light, but only upper limits have been so far obtained. A rather low upper limit in the optical region (Dube, Wickes, and Wilkinson 1977) implies that either galaxies might have been not so bright, or that the redshift at the galaxy formation might be larger than expected (Davis 1980).

Since the previous observations were performed in the optical wavelength band, near-infrared observation is regarded to be more advantageous on searching for young galaxies for the following reasons. First, in the optical band the extragalactic component is much weaker than other diffuse components such as zodiacal light, star light, and airglow. On the other hand, the extragalactic component in the near-infrared band can be observed at a lower background level, because the spectra of zodiacal light and star light decrease rapidly towards longer wavelength and extragalactic component is expected to have a rather flat spectrum. Second, the near-infrared extragalactic component is originated in the radiation at high redshift, while only the nearby galaxies contribute to the optical part. Thus infrared observation is profitable to study the universe at early epochs.

At far infrared region (100  $\mu\text{m}$  - 1 mm), 2 kinds of the background radiations are expected. One is the Wien's end of the 2.7K cosmic microwave background which appears at  $\lambda > 500 \mu\text{m}$ . The distortion of its spectrum and spatial fluctuation at  $\lambda < 1\text{mm}$  will provide the valuable informations on the early history of the universe. The other is the integrated radiation of the distant galaxies which is originated in the thermal emission of the dust. The violent activities at the early epoch of the universe, such as pre-galactic pop III era, will be able to be observed in this wavelength regions as a counterpart of the visible and near-infrared background.

In this paper, we present the recent result of the rocket observation of the near-infrared night sky and the forthcoming rocket observations at near and far infrared regions.

## II. Rocket Observation of The Near Infrared Night Sky

### 1. Observation

A conceptual design of the instrument is shown in Figure 1. Optical system was composed of 5 sets of telescope each of which consisted of 26 mm silicone lens and InSb detector forming 4° beam parallel to the rocket axis. Each telescope corresponded to the specific wavelength (band width), that is, 1.6  $\mu\text{m}$  (0.3  $\mu\text{m}$ ), 2.2  $\mu\text{m}$  (0.4  $\mu\text{m}$ ), 3.8  $\mu\text{m}$  (0.7  $\mu\text{m}$ ), 4.2  $\mu\text{m}$  (0.7  $\mu\text{m}$ ) and 4.7  $\mu\text{m}$  (0.6  $\mu\text{m}$ ). Whole optical system was cooled by a solid nitrogen which realized no background radiation from the instrument itself. The cold shutter in front of the detector was used to check the zero-levels every 15 seconds.

The infrared photometer on board the sounding rocket, K-9M-75, was launched on 13 Sept. 1983 at 21:30 JST (12:30 UT) from Kagoshima Space Center, Institute of Space and Astronautical Science. At 288 sec after launch, the rocket reached the apogee of 322 km altitude. After the lid open at 80 sec after launch, the sky was surveyed by means of the precession of the rocket axis. Yo-despin was executed at the apogee and the precession half cone angles of 5° and 21° were obtained before and after the despin, respectively. Figure 2 shows the obtained trajectory of the optical axis.

### 2. Results

An absolute calibration of the photometer was attained in laboratory using the standard blackbody source and was confirmed well during the flight by observing the bright stars and the galactic plane. The errors are estimated to be  $\pm 10\%$  in all wavelength bands.

In order to obtain the extragalactic component, other diffuse components are subtracted carefully as follows.

It is assumed that no residual atmospheric emission exists above a certain altitude, since the signals did not depend on the altitude above the specific height for each wavelength band. The time dependent component probably due to the debris of the fuel of the rocket engine was observed, but it dissipated well before the apogee. After all, atmospheric effects can

be neglected above the altitude of 250 km in the descending phase.

As the optical axis approached to the earth surface, the stray light through the baffles caused the large contaminations, especially for the longer wavelength bands. This restricted the available range of the elevation angle,  $\theta$ , between the optical axis and the earth surface to be  $\theta > 80^\circ$ , since signals became flat at  $\theta > 80^\circ$  and  $\theta$  went down to  $100^\circ$ .

It must be noted that above two effects cause severe contamination at longer wavelength bands, but only a little effect for 1.6 and 2.2  $\mu\text{m}$  band.

In the interplanetary space, there are two kinds of diffuse sources, the zodiacal light (ZL) and the thermal emission from the interplanetary dust (IPD). Although the correlation of the signals with the ecliptic coordinate was not found, their contributions are inferred as follows. ZL is estimated adopting the optical data at  $\lambda - \lambda_{\odot} \sim 130^\circ$ ,  $\beta \sim 25^\circ$  (Levasseur-Regourd and Dumont 1980) and the solar spectrum (Hayakawa et al. 1970, Hoffmann et al. 1973). There are so many unambiguities for IPD due to the lack of reliable observations that we assumed two cases referring to the different observations (Soifer et al. 1971, Price et al. 1980).

Finally, the contribution of the integrated star light (SL) should be taken into account. During the second phase of the precession, the telescope scanned the galactic plane at  $|b| < 30^\circ$ . The profiles of the galactic plane are modelled based on the infrared luminosity function at the solar neighborhood (Ishida and Mikami 1982) and the model of the Galaxy. As a result of the fitting which is shown in Figure 3, two parameters (the surface brightness at the galactic pole and the constant non-galactic component) are obtained. The non-galactic component thus estimated is significantly larger than the interplanetary components.

Figure 4 shows the deconvolution of signals to stray light, star light (SL) and non-galactic component (CL + ZL + IPD). Here, CL means the extragalactic background light (or cosmic light).

Figure 5 shows the observed spectrum of the darkest sky at  $l = 52^\circ$ ,  $b = -23^\circ$  where no IRC star ( $m_k < 3.0$  mag) was in the beam. Other known diffuse sources described above are also indicated in this figure. The spectrum of SL is plotted so as to be consistent with optical observations (Leinert and Richter 1981) adopting the same color derived for the galactic pole. Figure 5 shows clearly that there remain excess fluxes in all

wavelength bands which cannot be explained by the known sources.

### 3. Discussions

Regarding that the excess flux observed is the extragalactic origin, we discuss its influence on the cosmology.

Figure 6 shows the residual diffuse radiation after subtracting the known diffuse sources, in which an upper limit of the extragalactic radiation in the optical band (Dube et al. 1977) is included. Two theoretical estimation by Partridge and Peebles (1967) are also shown. Model 1 assumes the constant luminosity of galaxies without evolution and corresponds to the lowest estimation. Model 4, the brightest case, assumes that all helium observed at present were synthesized in the stars during the first bright phase of galaxies. Our result shows that infrared sky is much brighter than the brightest case. This disagreement can be ascribed to the assumption that Partridge and Peebles (1967) took only the luminous mass into account. In other word, the observed near-infrared background necessitates the new energy sources which had activities at the early epoch of the universe.

One possible candidate is pregalactic objects (Thorstensen and Partridge 1975, Carr, Bond and Arnett 1983). In this picture, the very massive pop III stars were first formed after the decoupling of the matter and radiation. These stars emitted the radiation very efficiently in UV and optical band, which forms now the near-infrared background due to the large redshift. After burn-out, massive stars collapsed to the blackholes which are now composing the missing mass in the universe. The redshift,  $z$ , of pop III era is estimated to be 50 - 100, assuming the temperature of pop. III stars of  $10^5$ K, and energetics requires the density parameter,  $\Omega$ , close to 1.

The gravitational energy can liberate more energy than the nuclear energy. Carr, McDowell and Sato (1983) proposed another origin, that is, radiation from super massive blackholes. In this case, the redshift and density parameter are supposed to be  $\sim 10$  and  $\sim 0.1$  respectively.

There may be other possibilities to explain the observed excess flux, however, the future detailed observations will make the physical processes at the early universe clear.

## III. Forthcoming Observations

### 1. The second rocket observation of the near infrared background radiation

In our first rocket flight, an unexpectedly bright surface brightness at  $1 \sim 5 \mu\text{m}$  was observed, which is supposed to be the extragalactic origin. However, the observed sky was so limited that the isotropy, which is an evidence of its extragalactic origin, was not well confirmed. Therefore, we planned a new rocket observation with an improved instrument. The optical system is composed of the following 2 parts.

#### a. Wide band photometry

This system consists of 4  $14\text{mm}\phi$  Si lenses each of which correspond to the specific wavelength (J.K.L.M) with  $4^\circ$  beam. The main objective is the confirmation of the previous result.

#### b. Narrow band photometry

The optical system consists of 2  $26\text{mm}\phi$  lenses with  $4^\circ$  beam. 2 sets of 12 filters on the filter wheel are changed every 3 seconds in front of the 2 InSb detectors to obtain the course spectrum of the diffuse light with a spectral resolution of 0.1 at  $0.7 - 5.5 \mu\text{m}$ . This system is designed to search for the redshifted Lyman limit, Lyman  $\alpha$  and other feature.

Cryogenics and other parts are almost same as that of the previous one. The instrument was installed on the sounding rocket k-9M-77 which was launched on Jan. 14, 1984 towards the galactic north pole. The instrument worked well during the flight and the wide sky range was surveyed. The data are in analysis and will be open soon.

### 2. The rocket observation of the far infrared background

The spectrum of the 2.7 K cosmic background radiation has been extensively observed, however, the wavelengths observed from the ground are restricted to the radio wavelength region due to the atmospheric contamination. Woody and Richards (1979) attained the balloon observation and provided a reliable data above 1 mm. Gush (1981) carried out the rocket observation to obtain the spectrum below 1 mm, however, his data was not so reliable due to the contaminated radiation from the ejected nose cone. On the other hand, the recent infrared astronomical satellite (IRAS) has provided some data of the diffuse radiation but these are restricted at  $120 \mu\text{m}$  and

shorter wave length bands. At present stage no reliable data are available between 100  $\mu\text{m}$  and 1 mm, therefore, we planned to make a rocket observation under the collaboration with Prof. Richards, U.C. Berkeley.

The design of the instrument is shown in Figure 7. The light concentrator (Winstone cone) and photometer are cooled by superfluid He down to 1 K, while the HeI in the annular tank is responsible for the heat load from the warm part. The photometer (Figure 8) consists of 6 detectors co-operated with 45° incident dichroic filters. Central frequency, band-width, detectors are as follows.

Band	Central Frequency	$\Delta\nu/\nu$	Detector
1	10 $\text{cm}^{-1}$	25 %	bolometer
2	14	"	"
3	20	"	"
4	30	30	"
5	65	50	Ge:Ga stressed
6	95	50	Ge:Ga

The band 1 and 2 are dedicated to measure the 2.7 K cosmic background, while band 4, 5, 6 are used to estimate the contribution from the zodiacal and galactic emission.

The instrument will be installed on K-9M-78 rocket and launched on August or September, 1985.

#### References

- B.J. Carr, J.R. Bond, and W.D. Arnett, (1983) preprint  
 B.J. Carr, J. McDowell, H. Sato, (1983) preprint  
 M. Davis, 1980, in objects of High Redshifts, ed. G.O. Abell and P.J.E. Peebles (Dordrecht: Reidel), p. 57.  
 R.R. Dube, W.C. Wickes, and D.T. Wilkinson, Ap. J. (Letters) 215 (1977) L51.  
 H.P. Guch, Phys. Rev. Lett. 47 (1981) 745.  
 S. Hayakawa, T. Matsumoto, and T. Nishimura, Space Research X (1970) 248.  
 W. Hoffmann, D. Lemke, C. Thum, and U. Fahrback, Nature Phys. Sci. 243 (1973) 140.  
 K. Ishida, and T. Mikami, Publ. Astron. Soc. Japan 34, (1982) 89.

- C. Leinert, and I. Richter, Astr. Ap. Supple. 46 (1981) 115.  
 A.C. Levasseur-Regourd, and R. Dumont, Astr. Ap. 84 (1980) 277.  
 R.B. Partridge, and P.J.E. Peebles, Ap. J. 148 (1967) 377.  
 S.D. Price, T.L. Murdock, and L.P. Marcotte, A.J. 85 (1980) 765.  
 B.T. Soifer, J.R. Houck, and M. Harwit, Ap. J. (Letters) 168 (1971) L73.  
 J.R. Thorstensen, and R.B. Partridge, Ap. J. 200 (1975) 527.  
 D.P. Woody, and P.L. Richards, Ap. J. 248 (1981) 18.

Figure Captions

- Fig.1. Conceptual view of the instrument.
- Fig.2. Trajectory of the optical axis on  $l$ - $b$  plane. The small and large circles correspond to different precession phases before and after a despin at 280 sec after launch, respectively. Dashed line represents the zenith angle of  $90^\circ$
- Fig.3. Observed signals from 350 sec to 395 sec after launch. The angle,  $\theta$ , between the optical axis and the earth limb, and the galactic latitude are indicated at the bottom of the figure. The events occurred in this period are shown at the top of the figure.  
Thin lines and dot-dashed lines show integrated starlights (SL) and non-galactic component (CL+LPD+ZL) derived by the model fitting. The dotted lines are drawn by subtracting the above two components from the observed signals and are regarded as the stray light of the earthshine.
- Fig.4. Dependence of signals on  $\sec(90^\circ - |b|)$ . Solid lines represent the best-fit model. Nongalactic components derived from the model are indicated at the left end of the figure by the solid circles. Zodiacal light (ZL) and thermal emission of the interplanetary dust (IPD) for two different estimations are also shown.
- Fig.5. Observed spectrum of the surface brightness at  $l = 52^\circ$ ,  $b = -23^\circ$ , where the signals recorded the lowest levels at 372 sec after launch. Other background components, SL, ZL, and IPD are also indicated.
- Fig.6. The residual background components after subtracting the contribution of SL, ZL, and IPD in Figure 12. Filled and open circles correspond to cases 1 and 2 for IPD, respectively. Dotted line shows the spectrum of 1500 K blackbody. Solid lines represent the two extreme cases in the models by Partridge and Peebles (1967). Upper limit at optical band (Dube et al. 1977) is also indicated.
- Fig.7. Cross sectional view of the rocket-borne instrument to observe the far infrared background.
- Fig.8. Photometer for the far infrared observation. The light concentrater (Winstone cone) is placed perpendicular above the paper.

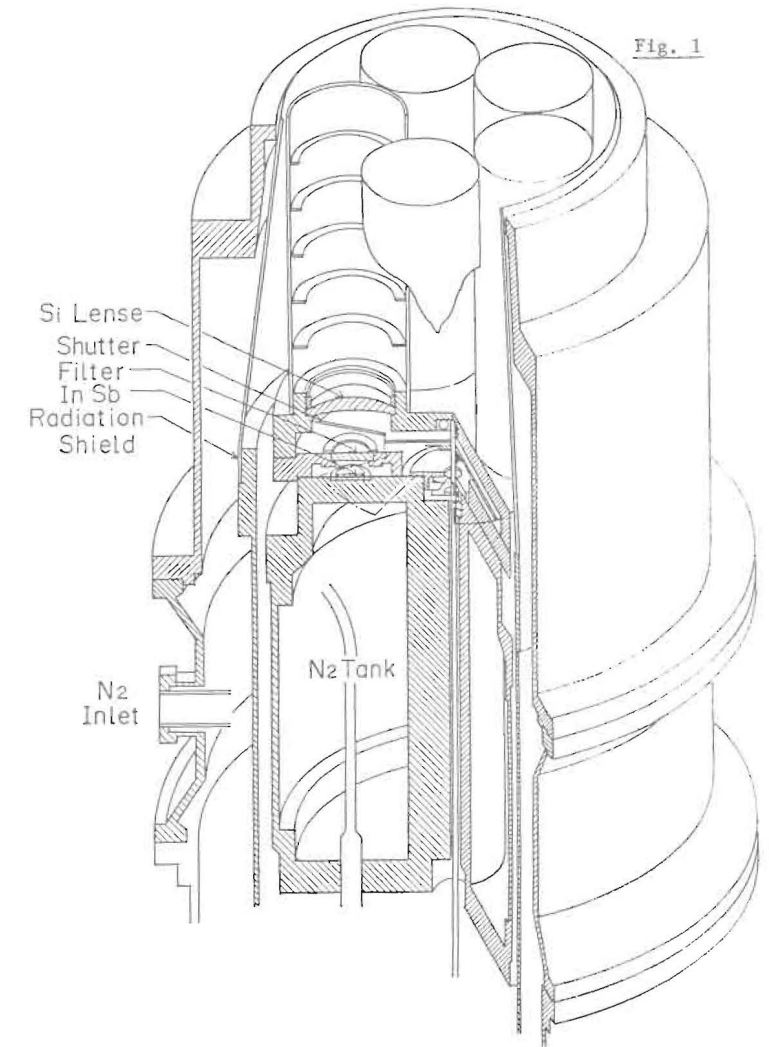


Fig. 1

Fig. 3

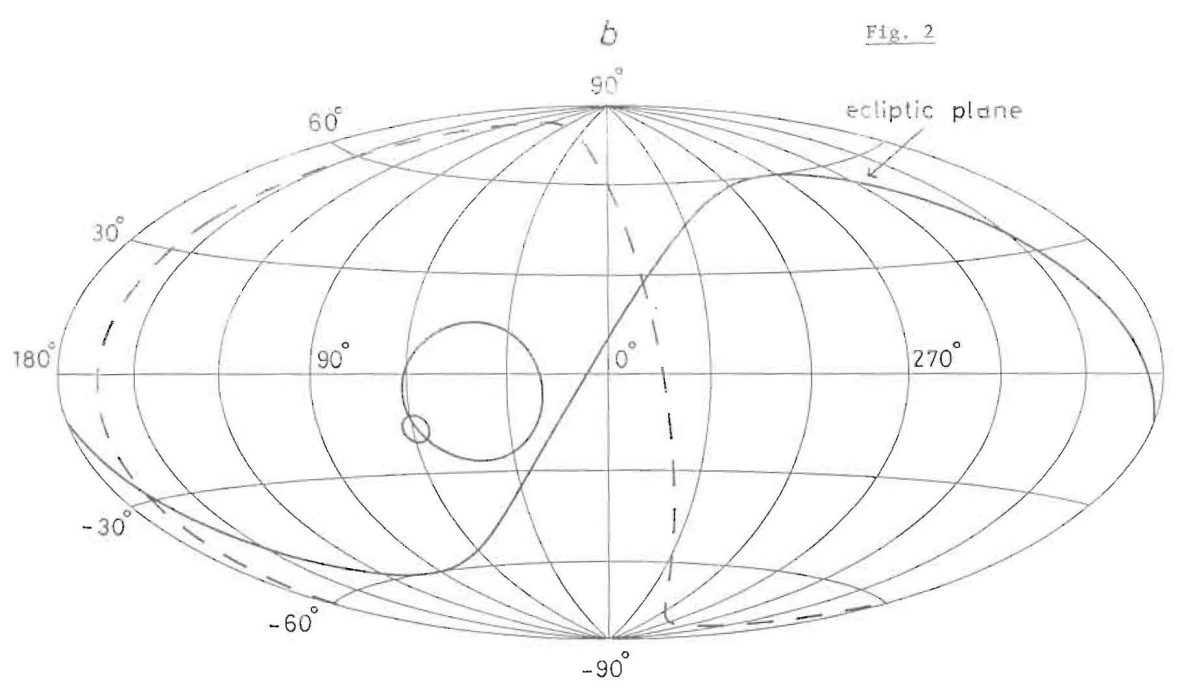
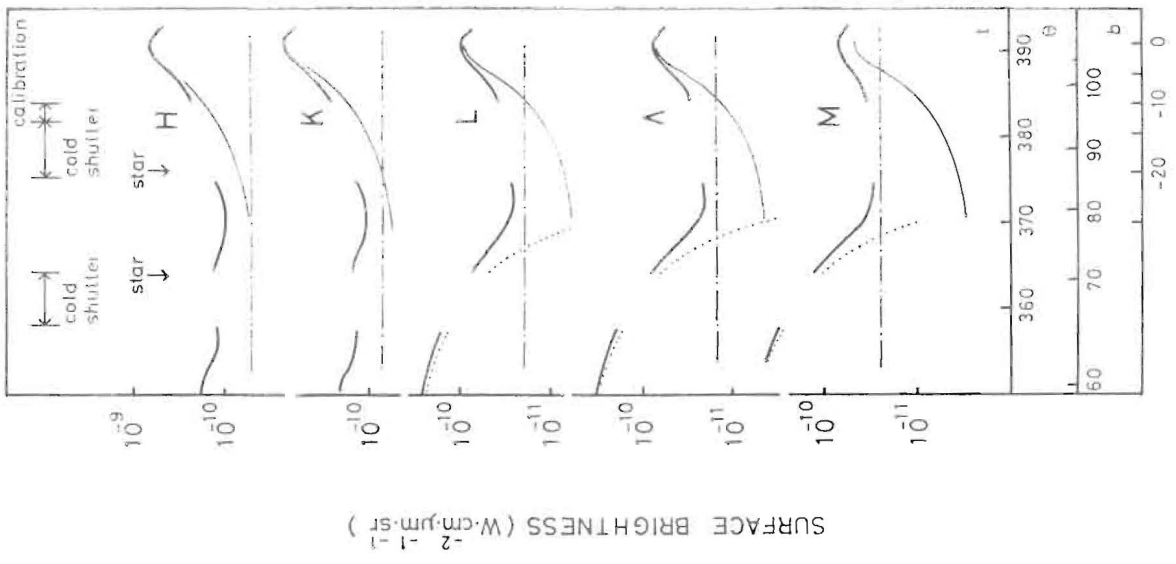


Fig. 4

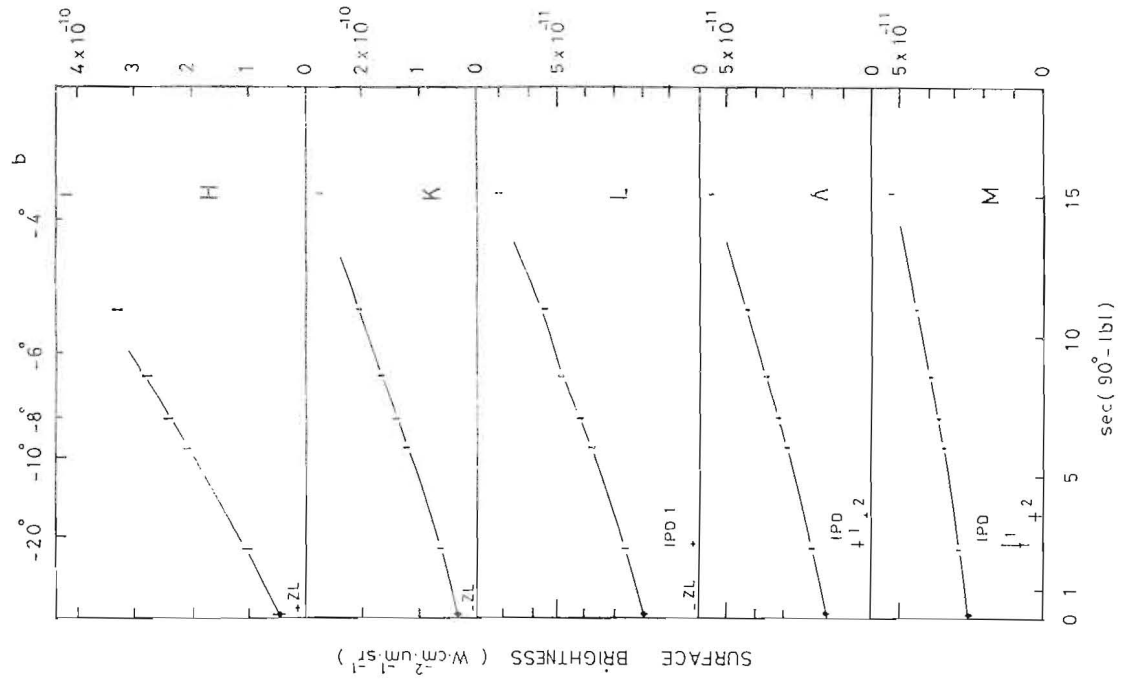


Fig. 5

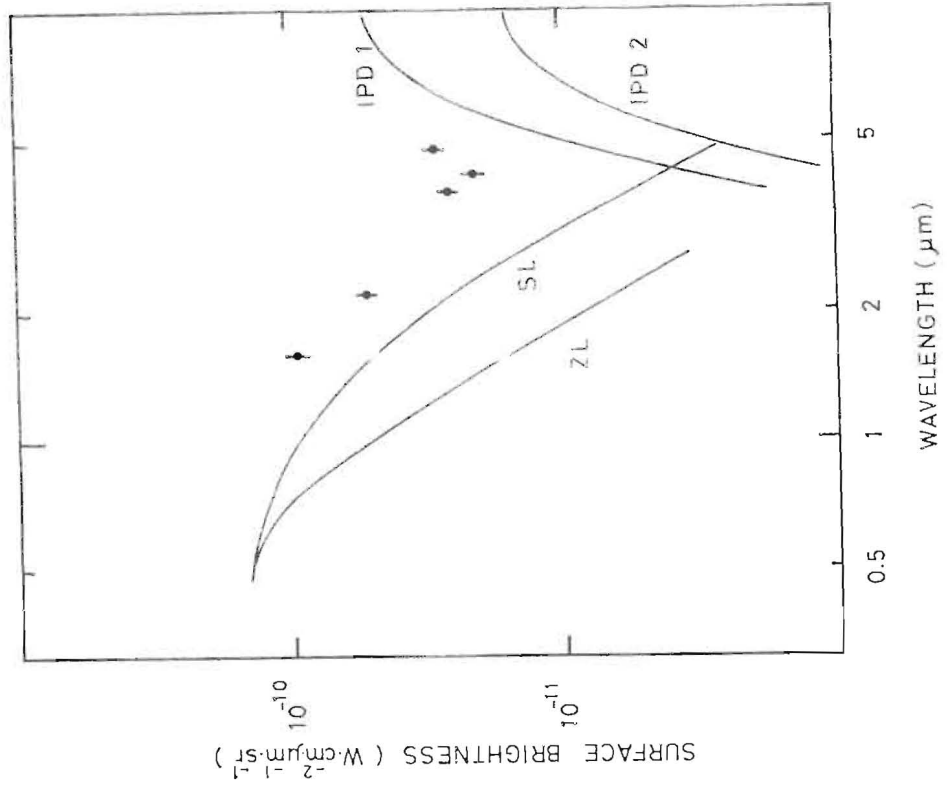




Fig. 6

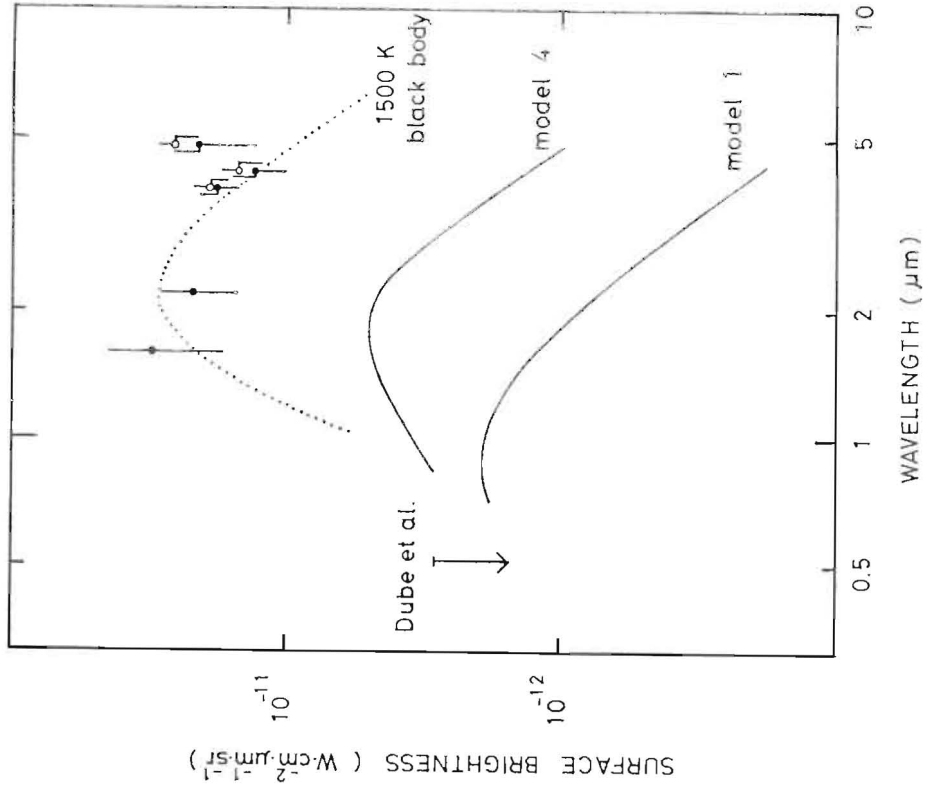


Fig. 7

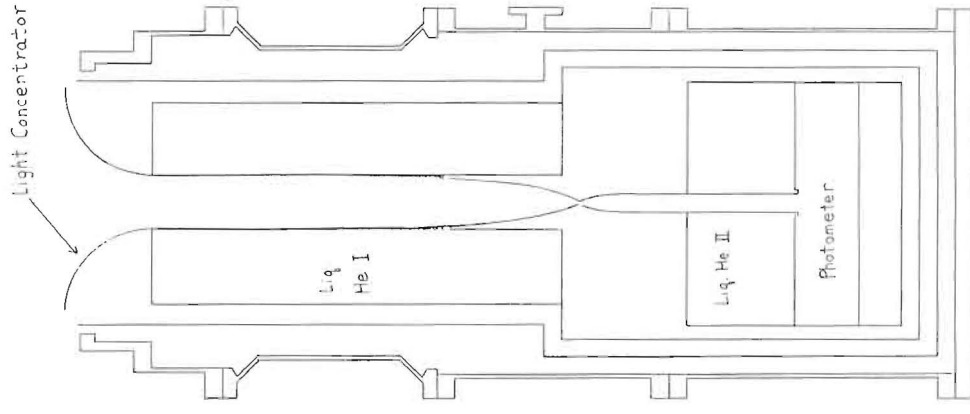
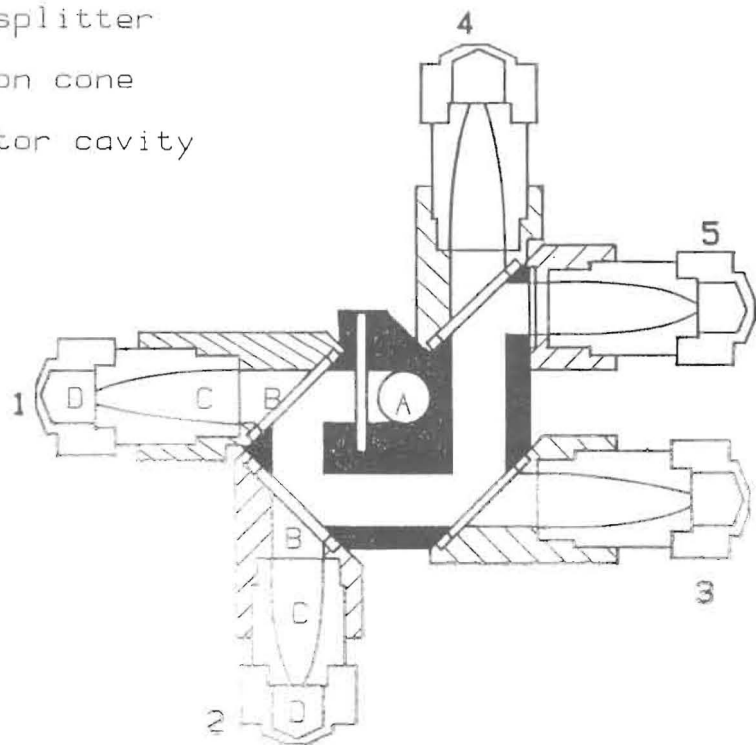


Fig. 8

- A : diagonal mirror
- B : beam splitter
- C : winston cone
- D : detector cavity



Current Status of Missing Mass Problem

Fumio Takahara

Nobeyama Radio Observatory  
Tokyo Astronomical Observatory, University of Tokyo  
Nobeyama, Minamisaku, Nagano 384-13, Japan

Abstract

Current status of missing mass problem is reviewed with emphasis on recent progress in observational cosmology. Topics include mass to light ratios of astronomical objects of various scales, infall of the

local group of galaxies towards the Virgo cluster and the redshift survey of galaxies. Problems with nucleonic and non-nucleonic matter as candidates for missing mass are discussed. It is concluded that non-nucleonic matter dominates at least on scales larger than rich clusters of galaxies.

§ 1. Introduction

It is well known that the dynamically inferred mass of rich clusters of galaxies far exceeds the mass inferred from the mass to light ratio of galaxies. This discrepancy which also exists in galactic haloes, binary galaxies and small groups of galaxies is called the missing mass problem although really missing is not mass but light. Missing mass problem is related not only to the structure and evolution of various astronomical objects but also to cosmology and elementary particle physics. The determination of the parameters of Friedmann universe is critically affected by the mean mass density of the universe which is directly related to the missing mass problem. Views before 1974 was beautifully summarised by Gott et al.<sup>1)</sup> who favored open universe without non-nucleonic matter. However since 1980 experimental suggestions on finite neutrino mass and theoretical prediction of the existence of many species of weakly interacting elementary particles based on unified theories have stimulated the idea that the universe is dominated by non-nucleonic matter.

In this article I will review the missing mass problem

emphasizing the recent progress in observational cosmology. In § 2, mass to light ratios of various astronomical objects are reviewed including recent topics on dwarf spheroidal galaxies. In § 3, the infall of local group of galaxies towards the Virgo cluster is used to estimate the cosmological density parameter. In § 4, the estimation of mean mass density is made based on the large scale dynamics of galaxies using the recent redshift survey of galaxies. Finally in § 5, I discuss problems with nucleonic and non-nucleonic matter as candidates for missing mass in connection with the primordial nucleosynthesis and galaxy formation.

## § 2. Mass to Light Ratios of Astronomical Objects

We can determine the mass of an astronomical system by the Newton mechanics. If a test particle rotates around the central mass of  $M$  in a circular orbit of radius  $r$  and the velocity  $v$ , we get by the force balance

$$GM/r^2 = v^2/r, \quad (1)$$

where  $G$  is the gravitational constant. From Eq. (1) we can determine the dynamical mass  $M_{dyn}$  as

$$M_{dyn} = rv^2/G. \quad (2)$$

For an isolated many particle system in dynamical equilibrium, total gravitational energy  $W$  and the total kinetic energy  $T$  are related through the virial theorem as

$$W + 2T = 0. \quad (3)$$

From Eq. (3) we can get a similar expression to Eq. (2) with suitable definitions of mean radius  $r$  and velocity dispersion  $v$ .

Although Eq. (2) is simple, mass determination is inevitably uncertain since we can measure only the angular distance projected on the celestial sphere and the velocity component along the line of sight. We need certain statistical assumptions and treatments for the proper estimation of the dynamical mass. It is also to be noted that thus inferred mass is inversely proportional to the Hubble constant  $H_0$ . On the other hand the luminosity of an object is deduced from the apparent luminosity and inversely proportional to the square of  $H_0$ . Thus mass to light ratio  $M/L$  is proportional to  $H_0$ . Hereafter  $M/L$  ratio is represented in units of  $M_\odot/L_\odot$  and  $H_0$  is measured in units of  $100\text{km s}^{-1}\text{Mpc}^{-1}$  and represented by  $h \equiv H_0/100\text{km s}^{-1}\text{Mpc}^{-1}$ . The luminous mass  $M_{lum}$  is defined by

$$M_{lum} = L \times (M/L)_c, \quad (4)$$

where  $(M/L)_c$  denotes the mass to light ratio of constituents.

### (2-1) galactic haloes

Mass to light ratios of individual galaxies within the Holmberg radius, i.e., in the part where stellar light dominates are summarized by Faber and Gallagher<sup>2)</sup>. Spiral galaxies have  $M/L$  of about  $10h$ , while SO and elliptical galaxies have  $M/L$  of  $10h \sim 20h$  in the blue band. These values are compared to the value  $2.3 \sim 3.3$  in the solar neighborhood, and these differences may well be ascribed to the differences of stellar populations.

Mass distribution of spiral galaxies outside the Holmberg radius

can be evaluated by observing the 21cm line emission of neutral hydrogen and optical emission line of ionized gas. If mass distribution is the same as the light distribution, the rotation curve in the outer region would decrease as  $r^{-1/2}$ . Observations have not shown such a decrease but rotational velocity has been shown to be constant as far as observations are made. In Fig.1 rotation curves obtained by 21cm observations by Bosma<sup>3)</sup> are shown. In Fig.2 those obtained by optical emission line observations by Rubin et al.<sup>4)</sup> are shown. Flat rotation curves thus shown imply that the mass within the radius  $r$  increases in proportion to  $r$ . This mass which distributes far extending from the optical image is called massive haloes. The density profile of massive halo is proportional to  $r^{-2}$ , which is different from the profile of halo stars of  $r^{-3}$ . Although we cannot yet reach the end point of rotation curve, mass to light ratio of spiral galaxies should increase at least to  $20h \sim 50h$ .

For elliptical galaxies which contain little gas, we have not evidence for dark mass for a large sample of galaxies. Only one example is M87 which resides at the center of the Virgo cluster. M87 has a hot gas halo which reveals the extended X-ray emission. Imaging observation by the Einstein satellite<sup>5)</sup> has shown that the hot gas extends to  $100'$  from the center far exceeding the optically determined radius. Since this hot gas is confined by the gravitational potential of M87, we can determine the mass distribution. Although the details depend on the temperature distribution of hot gas, the existence of missing mass is clearly shown. In Fig.3 are shown the X-ray brightness distribution and derived mass distribution. The inferred mass to light

ratio is about 180 at  $20'$  and may become larger at larger distances.

#### (2-2) systems of galaxies

There are various systems of galaxies from binaries to superclusters of galaxies. The situation is essentially the same as that described by Faber and Gallagher<sup>2)</sup> except the results by redshift survey. For binary galaxies at separation  $25h^{-1} \sim 50h^{-1}$  kpc,  $M/L$  of  $35h \sim 70h$  has been reported by several authors. There are still problems such as the statistics of orbits and the existence of spurious pairs. For small groups of galaxies  $M/L$  of  $60h \sim 80h$  has been reported by several authors. Those groups which contain several galaxies may not be in a dynamical equilibrium state since the crossing time is comparable to the Hubble time. Also there is the membership problem.

While binaries and small groups of galaxies contain largely spiral galaxies, rich clusters of galaxies contain mainly elliptical galaxies. As a typical example of rich clusters of galaxies,  $M/L$  of the Coma cluster is estimated about  $650h$  by the virial theorem. It is to be noted that  $M/L$  of rich clusters of galaxies is much larger than that of galactic haloes, binary galaxies and small groups. Rich clusters are strong X-ray emitter through thermal bremsstrahlung of hot gas. The mass of hot gas is about 10% of the dynamical mass but the confinement of hot gas requires the dynamical mass comparable to the mass inferred from virial theorem.

Recent progress in  $M/L$  ratio determination for systems of galaxies has been made using the Center for Astrophysics (CfA)

redshift survey<sup>6)</sup>. Press and Davis<sup>7)</sup> selected virialized clusters and found that the total mass of a cluster is roughly proportional to the size of the cluster. They concluded that the contribution to the cosmological density parameter of virialized clusters is 0.07. This corresponds to the  $M/L$  of  $180h$ , using the luminosity density of  $1.1 \times 10^8 h L_{\odot} \text{Mpc}^{-3}$ . On the other hand Huchra and Geller<sup>8)</sup> selected groups of galaxies according to the criterion of number density enhancement. The resultant  $M/L$  ratio is  $170h$  similar to that of Press and Davis. But they found that  $M/L$  ratio does not depend on the scale of groups. See Fig.4 for the situation.

Superclusters which are largest scale structure known have not yet collapsed and the method described here cannot be used. It is to be noted that there is the suspect that many of rich clusters may not be in a relaxed states but collections of groups, i.e., superclusters<sup>9)</sup>.

#### (2-3) dwarf galaxies

The existence of missing mass from galactic haloes to rich clusters of galaxies has been established and it seems that  $M/L$  ratio increases as the scale length increases although the result of Huchra and Geller reveals no such trend. Recently Aaronson<sup>10)</sup> and Faber and Lin<sup>11)</sup> asserted that the missing mass problem also exists in dwarf spheroidal galaxies around our galaxy. Based on the resemblance of color to globular clusters  $M/L$  ratio of  $\sim 2$  has been adopted for a long time. Faber and Lin asserted that  $M/L$  ratio should be  $\sim 30$  if the radius of these galaxies is determined by tidal cutoff. Aaronson

measured radial velocities of several carbon stars in Draco and found the velocity dispersion of  $\sim 6\text{km/s}$ , which also implies the  $M/L$  ratio of  $\sim 30$ . It seems that there exists the discrepancy of an order of magnitude if stellar population in dwarf spheroidals is similar to that in globular clusters. Since detailed observations of dwarf spheroidals have just begun, we should reserve conclusions until more definite observations, e.g., by space telescope will be done.

Another topic is the discovery of large intergalactic HI cloud in M96 group of galaxies<sup>12)</sup>. This cloud comprises of HI mass of  $\sim 10^9 M_{\odot}$  and maximum rotation velocity of  $\sim 80\text{km/s}$  on a scale of  $\sim 100\text{kpc}$ , which results in a gravitational mass of  $\sim 10^{11} M_{\odot}$ . This cloud does not show any evidence of stellar light so that star formation efficiency should be very low because of low gas density. Invisible dark mass is shown to exist in this cloud and we may be looking at truly primordial cloud. However the contribution of primordial clouds to the mean mass density seems to be negligibly small.

#### § 3. Infall to the Virgo cluster

As was stated in § 2,  $M/L$  ratio increases as scale size increases. Recent observations have shown the existence of larger scale structure such as superclusters and voids. What value of  $M/L$  does it take for such structure? Superclusters are just beginning to collapse or to deviate from the general cosmic expansion and are not in a dynamically relaxed state. So we cannot use the methods described

in § 2 and we must take account of cosmic expansion. Nearest such structure is the local supercluster, the center of which is the Virgo cluster and at the periphery of which locates the local group of galaxies. The motion of local group can be studied by examining the motion of our galaxy with respect to nearby galaxies and the background radiation. Recent status of this peculiar motion is summarized in the article by Davis and Peebles<sup>13)</sup>.

Here I describe three observations detecting the peculiar motion. First dipole anisotropy of 2.7K microwave background radiation has been firmly established by independent groups<sup>14)</sup>. With a small value of upper limit of quadrupole anisotropy, the dipole anisotropy is considered to be due to the peculiar motion of our galaxy. After subtracting the motion of the sun around the galactic center, the motion of the local group is estimated to be  $\sim 600\text{km/s}$  towards the direction  $45^\circ$  apart from the center of the Virgo cluster. The component towards the Virgo cluster is about  $410\text{km/s}$ .

Second estimation is based on the dipole anisotropy of galaxy distribution. Rubin et al.<sup>15)</sup> found that our galaxy is moving towards the completely different direction from that microwave background suggests, using a sample of Sc galaxies. On the contrary, Hart and Davies<sup>16)</sup> found that the direction is coincident with the microwave result using the HI data. The speed of peculiar motion is similar among these results, i.e.,  $400\text{km/s} \sim 600\text{km/s}$ .

The cause of the differences among these results is not well known and there seem to exist still some selection effects and systematic errors in galaxy data. Or else there exist peculiar motions

of very large scales. In Fig.5 various results of peculiar motion of local group are shown.

Hereafter we discuss the dynamics of the local supercluster assuming that local group is infalling to the Virgo cluster with  $400\text{km/s}$  with respect to the cosmic expansion. According to the result of CfA redshift survey, mean overdensity of galaxies within the local supercluster  $\bar{\delta}$  is about 2<sup>17),18)</sup>. Assuming that peculiar motion is generated by this overdensity, we can infer the density parameter by investigating the dynamics of the overdensity. If we assume further the spherical symmetry, peculiar acceleration  $g$  is given by

$$g = (4\pi/3) \times G\bar{\delta}\rho_b R, \quad (5)$$

where  $R$  and  $\rho_b$  are the distance between the Virgo cluster and the local group and mean mass density of the universe, respectively. We can follow the nonlinear dynamics for given  $\bar{\delta}$  and  $\rho_b$ . Comparing the resultant peculiar velocity with observation Davis et al.<sup>17)</sup> has given  $\Omega_0$  of 0.4 as a preferable value, where the cosmic density parameter  $\Omega_0$  is the ratio of  $\rho_b$  to the critical density  $\rho_{cr} = 3H_0^2/8\pi G$ .

If we abandon the assumption of spherical symmetry, we can use linear approximation and connect the peculiar velocity  $v_p$  and peculiar acceleration as

$$v_p = 2g\Omega_0^2 / (3H_0\Omega_0), \quad (6)$$

Inserting the observed values, similar value to the above result is obtained<sup>18)</sup>. The corresponding  $M/L$  ratio amounts to about 1000. This result implies that the linear trend between  $M/L$  ratio and scale length continues to supercluster size, i.e., about 20Mpc. It should be noted that the adopted overdensity represents the overdensity in

galaxy distribution, i.e., in light distribution so that overdensity in mass distribution may be noticeably different from that in light distribution.

#### § 4. Statistical Methods to Evaluate Mean Mass Density

Instead of treating individual structures, we may deduce mean mass density by extracting the information about peculiar velocities from three dimensional galaxy distribution. Before several systematic redshift surveys have been completed recently, we could only use two dimensional galaxy distribution projected onto the celestial sphere. Using the CfA redshift survey which covers  $2.66\text{sr}$  and includes 2400 galaxies within  $100h^{-1}\text{Mpc}$ , Davis and Peebles<sup>19)</sup> found several important results. They analyzed the apparent elongation of galaxy correlation function along the line of sight due to the existence of peculiar velocities.

Their analysis shows that the one dimensional peculiar velocity of a pair  $\sigma$  is given by

$$\sigma = 340 \pm 40 \times (hr/1\text{Mpc})^{0.13 \pm 0.04} \text{ km/s}, \quad (7)$$

for  $10\text{kpc} < hr < 1\text{Mpc}$ . This velocity may be used to estimate mean mass density by the cosmic virial theorem or cosmic energy equation. Cosmic energy equation is written as

$$v_p^2 = (2/7) \times H_0^2 J_2 \Omega_0, \quad (8)$$

where  $J_2 = \int \xi(r) r dr$  and  $\xi$  denotes the pair correlation function. It is to be noted that  $v_p$  is the root mean square of peculiar velocity field

of galaxies which is different from  $\sigma$ . If we tentatively identify these two and use the CfA result of  $h^2 J_2 = 150\text{Mpc}^2$ , then we obtain

$$\Omega_0 = (v_p/660\text{kms}^{-1})^2. \quad (9)$$

As stated above, if we set  $v_p = \sigma = 330\text{km/s}$ , then we get  $\Omega_0 = 0.25$ .

Davis and Peebles also examined the statistical stability condition on correlations  $hr < 1\text{Mpc}$  and obtained the similar value for  $\Omega_0$ . Other redshift surveys which are deeper for smaller angular coverage show the similar values for  $\Omega_0$  (20), (21).

In concluding this section, I note that Bahcall and Soneira<sup>22)</sup> suggested that correlation in matter distribution may exist on much larger scale. They analyzed three dimensional distribution of rich clusters of galaxies and found that spatial correlation between rich clusters is detected up to a separation of  $150h^{-1}\text{Mpc}$  and that velocity dispersion between a pair amounts to  $\sim 2000\text{km/s}$ . This probably corresponds to the large scale structure such as superclusters and voids. The peculiar velocity may imply that there are large scale streaming motions and affect the estimation of peculiar velocities of galaxies. The estimated cosmic density parameter based on cluster-cluster correlation is similar to that of Davis and Peebles.

#### § 5. Concluding Remarks

It seems now clear that dark invisible matter exists in a manner roughly in proportion to scale length from galactic haloes to



superclusters. There are several candidates for missing mass. They are classified into two kinds: one is nucleonic matter and the other is non-nucleonic matter.

(5-1) nucleonic matter

The possibility of dark nucleonic matter has been widely discussed in relation to population III star problem. They may be stellar remnants such as black holes, neutron stars and white dwarfs, or they may be very low mass stars or uncondensed primordial gases. However, severe constraints are set on the mean mass density of nucleonic component by the primordial nucleosynthesis argument, which is one of the strong supports for the big bang cosmology.

As is well known,  ${}^4\text{He}$ , D,  ${}^3\text{He}$  and Li are produced in the primordial nucleosynthesis. Their abundances are functions of the ratio of nucleon density to photon number density when the number of species of neutrinos are fixed. In Fig. 6<sup>23)</sup> the resultant abundances are shown assuming 3 species of neutrinos. As is shown the abundance of  ${}^4\text{He}$  increases while that of D decreases as nucleon density increases.

The observed abundances of He and D suffer from the effects of galactic evolution and various chemical and physical processes. For example galactic evolution is considered to make  ${}^4\text{He}$  abundance increase and D abundance decrease. Taking account of all these effects and uncertainties, the allowed ranges for the deduced primordial abundances are shown in Fig. 6 as boxes. All data seem to be consistent with the nucleon density of  $2 \times 10^{-31} \text{ g/cm}^3$  and allowed range is very

small. This density corresponds to the density parameter of nucleons  $\Omega_N$  of  $0.01h^{-2}$  or  $M/L$  ratio of  $25h^{-1}$ , much less than the values inferred in previous sections. This nucleon density is marginally consistent with the missing mass of galactic haloes, binary galaxies and small groups of galaxies, but inconsistent with that in rich clusters and superclusters.

(5-2) non-nucleonic matter

The most extensively discussed candidate is massive relic neutrinos. Recently various types of hypothetical particles introduced by variants of GUT such as axions, photinos and gravitinos are discussed. These candidates should be assessed by the confrontation with theories of formation of galaxies and large scale structure. If the rest mass of neutrinos is an order of 20eV, it naturally predicts the scale of superclusters and rich clusters. Perturbations of smaller scale structures had been washed out due to the free streaming of neutrinos and large amplitude isothermal perturbations in nucleonic component are required for the efficient galaxy formation<sup>24)</sup>. This may be a fatal difficulty of massive neutrino theory. Also neutrinos cannot explain the missing mass problem in dwarf spheroidals if it is really a problem.

References

- 1) J.R.Gott III, J.E.Gunn, D.N.Schramm and B.M.Tinsley, *Astrophys.J.*,

- 194 (1974), 543.
- 2) S.M.Faber and J.S.Gallagher, *Ann.Rev.Astron.Astrophys.* **17** (1979), 135.
  - 3) A.Bosma, *Astron.J.*, **86** (1981), 1825.
  - 4) V.C.Rubin, W.K.Ford and N.Thonnard., *Astrophys.J.*, **238** (1980), 471.
  - 5) D.Fabricant and P.Gorenstein, *Astrophys.J.*, **267** (1983), 535.
  - 6) J.Huchra, M.Davis, D.Latham and J.Tonry, *Astrophys.J.Suppl.*, **52** (1983), 89.
  - 7) W.H.Press and M.Davis, *Astrophys.J.*, **259** (1982), 449.
  - 8) J.P.Huchra and M.J.Geller, *Astrophys.J.*, **257** (1982), 423.
  - 9) M.J.Geller, *Comm.Astrophys.*, **10** (1984), 47.
  - 10) M.Aaronson, *Astrophys.J.Lett.*, **266** (1983), L11.
  - 11) S.M.Faber and D.N.C.Lin, *Astrophys.J.Lett.*, **266** (1983), L17.
  - 12) S.E.Schneider, G.Helou, E.E.Salpeter and Y.Terzian, *Astrophys.J.Lett.*, **273** (1983), L1.
  - 13) M.Davis and P.J.E.Peebles, *Ann.Rev.Astron.Astrophys.*, **21** (1983), 109.
  - 14) P.M.Lubin, G.L.Epstein and G.F.Smoot, *Phys.Rev.Lett.*, **50** (1983), 616.  
D.J.Fixsen, E.S.Cheng and D.T.Wilkinson, *Phys.Rev.Lett.*, **50** (1983), 620.
  - 15) V.C.Rubin, N.Thonnard, W.K.Ford and M.S.Roberts, *Astron.J.*, **81** (1976), 719.
  - 16) L.Hart and R.D.Davies, *Nature*, **297** (1982), 437.
  - 17) M.Davis, J.Tonry, J.Huchra and D.W.Latham, *Astrophys.J.Lett.*,

**238** (1980), L113.

- 18) M.Davis and J.Huchra, *Astrophys.J.*, **254** (1982), 437.
- 19) M.Davis and P.J.E.Peebles, *Astrophys.J.*, **267** (1983), 465.
- 20) R.P.Kirshner, A.Oemler, Jr., P.L.Schechter and S.A.Schectman, *Astron.J.*, **88** (1983), 1285.
- 21) A.J.Bean, G.Efstathiou, R.S.Ellis, B.A.Peterson and T.Shanks, *Month.Not.Roy.Astron.Soc.*, **205** (1983), 605.
- 22) N.A.Bahcall and R.M.Soneira, *Astrophys.J.*, **270** (1983), 20.
- 23) B.E.J.Pagel, *Phil.Trans.Roy.Soc.London*, **A307** (1982), 19.
- 24) H.Sato and F.Takahara, *Prog.Theor.Phys.*, **66** (1981), 508.

#### Figure Captions

Figure 1. Rotation curves of spiral galaxies obtained by 21cm line observations by Bosma<sup>3)</sup>. In this figure the Hubble constant is taken as  $75 \text{ km s}^{-1} \text{ Mpc}^{-1}$ .

Figure 2. Rotation curves of spiral galaxies obtained by optical line observations by Rubin et al.<sup>4)</sup>. In this figure the Hubble constant is taken as  $50 \text{ km s}^{-1} \text{ Mpc}^{-1}$ .

Figure 3. X-ray surface brightness distribution (a) and the inferred dynamical mass (b) of M87 obtained by the Einstein satellite by Fabricant and Gorenstein<sup>5)</sup>. For the distance to M87 of 15Mpc,  $1'$  corresponds to 4.4kpc.

Figure 4. The relation between  $M/L$  and size for groups of galaxies. Figs. (a) and (b) are taken from Press and Davis<sup>7)</sup> and Huchra and Geller<sup>8)</sup>, respectively.

Figure 5. A representation of various observations of the peculiar motion of the local group in supergalactic coordinates by Hart and Davies<sup>16)</sup>.

Figure 6. The comparison of theoretical and observed abundances of light elements by Pagel<sup>23)</sup>.

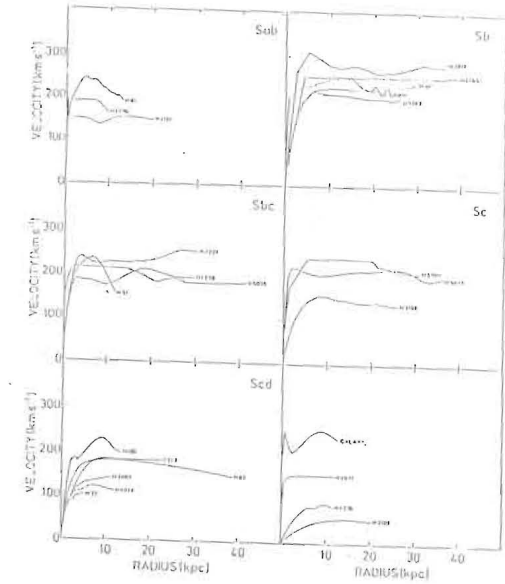


Fig. 1

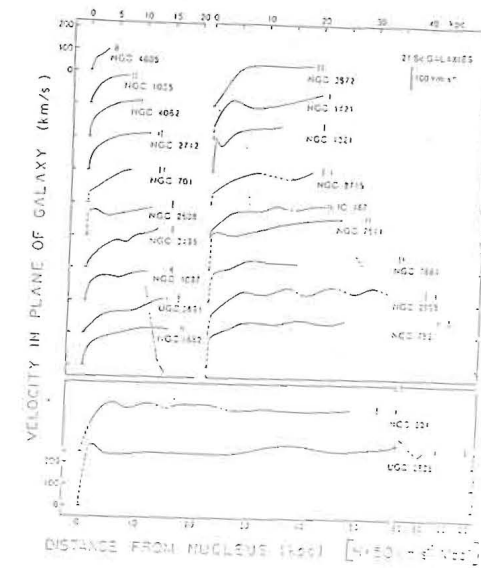


Fig. 2

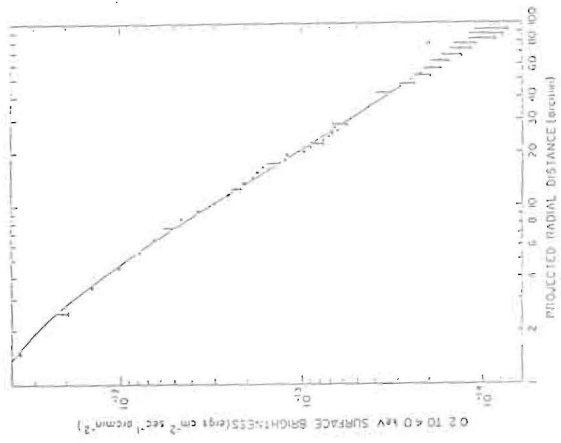


Fig. 3(a)

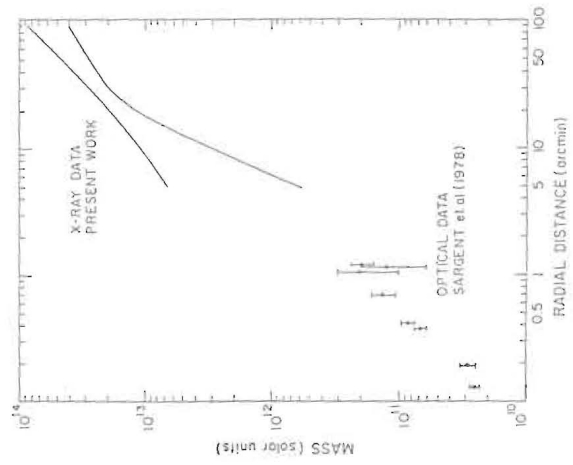


Fig. 3(b)

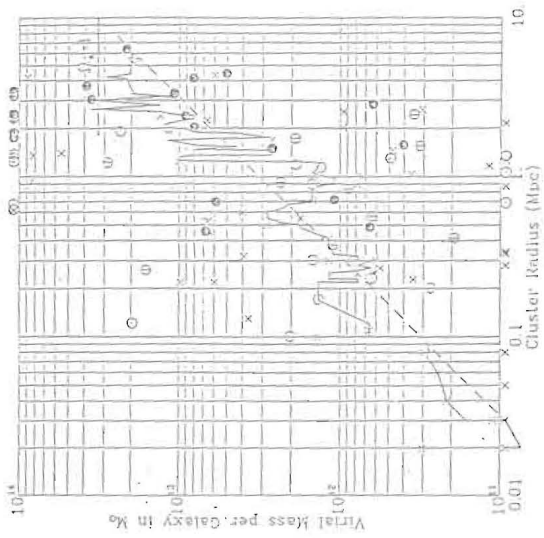


Fig. 4(a)

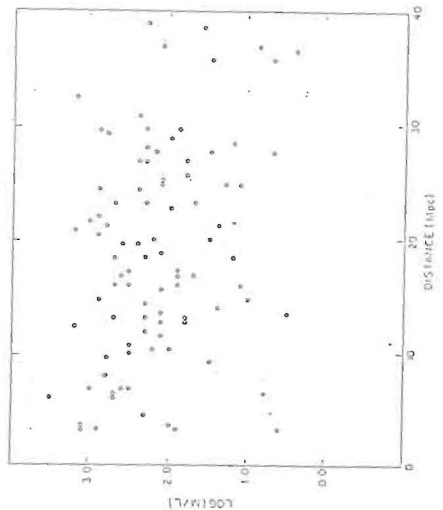


Fig. 4(b)

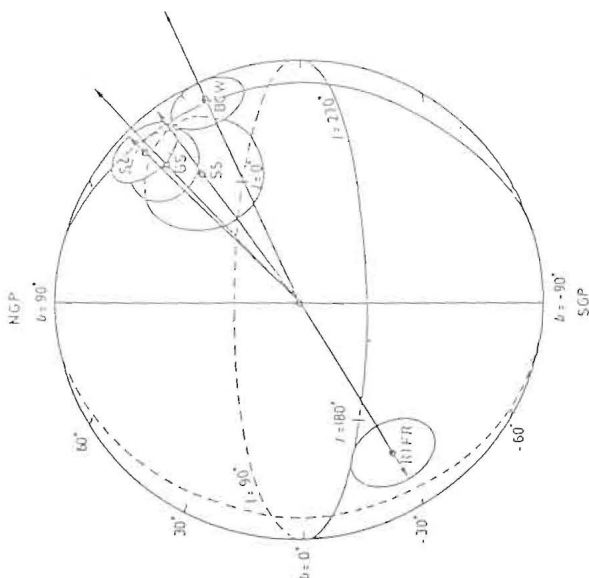


Fig. 5

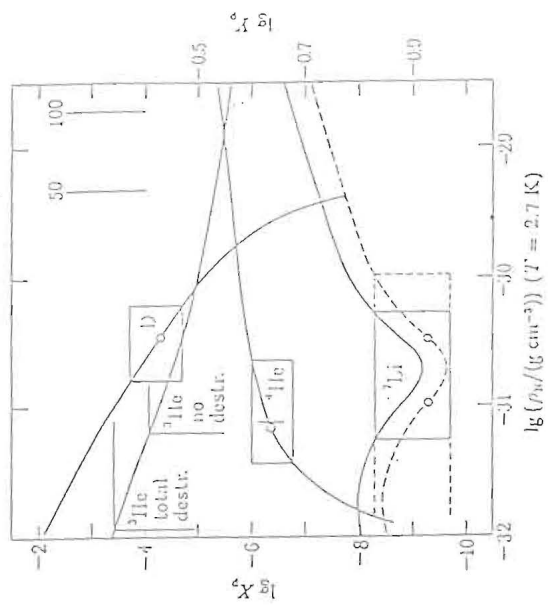


Fig. 6

## Some Topics in Cosmic Rays

Jiro Arafune

Institute for Cosmic Ray Research, University of Tokyo

Midori-cho, Tanashi-shi, Tokyo 188

### Abstract

Some topics of cosmic ray physics are introduced, namely, antiproton flux, cosmic ray flux at knee, and  $10^{20}$  eV cosmic rays. Brief introduction of elementary knowledges of cosmic rays are also given, which are closely related to the above topics.

### Introduction

Cosmic rays are initially born as ionized atoms, and may be pre-accelerated and injected like an accelerator beam, and are again accelerated. They propagate in the galactic magnetic field of  $\sim 2 \mu\text{G}$  for about  $10^7$  years and finally arrive at the earth and collide with air nuclei. The details of these processes are not well established, but let us give a rough idea of cosmic rays before introducing specific topics.

Cosmic rays are mainly composed of proton and nuclei. At low energies proton dominates and the elemental composition is measured<sup>[3]</sup>. At higher energies composition of some heavy nuclei are measured up to  $\sim 100$  GeV<sup>[4]</sup>. Above  $10^3$  TeV the composition is quite uncertain. We do not know if proton dominates or iron dominates or whatever. At low energies there are about 2% of electrons and positrons. Electrons are dominant by an order of magnitude<sup>[5]</sup>.

Cosmic ray flux for  $E_{\text{tot}} \geq 10^{16}$  eV is about  $1/\text{m}^2\text{st}\cdot\text{year}$ . The integrated cosmic ray flux decreases with  $\propto E^{-2}$  above this energy. Below this energy the flux decreases more slowly

with  $\propto E^{-1.6 \sim 1.7}$ . This bent of the spectrum is called "knee"<sup>[6]</sup>.

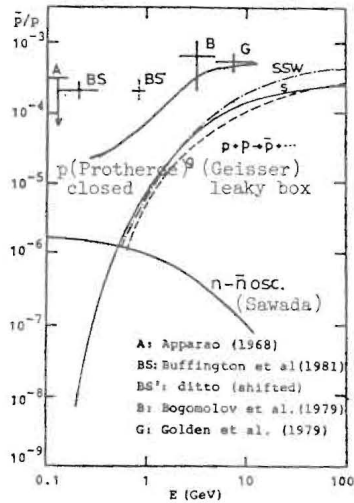
The mechanism of acceleration of cosmic rays are not well understood yet. Historically the second order Fermi acceleration was proposed by Fermi first<sup>[7]</sup>. The cosmic ray collide with randomly moving magnetic clouds  $n$  times, then the particle is accelerated by a factor  $(1 + v^2/c^2)^n$  with  $v$  the velocity of the cloud, and  $n = ct/\lambda$ . (The collision mean free path is  $\lambda$ .) If we assume the survival rate of the cosmic ray particle  $P(t) = e^{-t/\tau}$ , we obtain the energy spectrum  $E^{-\gamma}$  with  $\gamma = (\lambda/c\tau)c^2/v^2 + 1$ . This  $\gamma$  seems to be too large. Another mechanism called first order Fermi acceleration at the shock front is attractive and will be briefly explained in the appendix. There are other proposals related with pulsars<sup>[9]</sup>, and it is common that these are based on the energy released by supernova explosion.

The cosmic rays propagate in the galaxy, and let us here introduce three typical models of propagation, which will be discussed later in relation to the antiproton flux. (a) Leaky Box Model: The cosmic ray, after accelerated, propagates in our galaxy and is lost by leaking out of it. The mean path length traversed by the cosmic ray is  $X_{\text{esc}} \sim (8\text{gr}/\text{cm}^2) \cdot E^{-0.4}$ , and the distribution of the  $X$  is  $P(X) \sim e^{-X/X_{\text{esc}}}$ . This model explains the elemental composition ratio B/C, Sc-Cr/Fe and  $^{10}\text{Be}/^9\text{Be}$  etc. (b) Nested Leaky Box Model: This modifies the leaky box model by assuming that the cosmic rays, just after accelerated, traverses a certain thickness of matter ( $X_s$  in the average) surrounding the source. After this they propagate and escape from the galaxy with the mean path length  $X_{\text{esc}}$ . The total path length distribution becomes  $P(X) \sim e^{-X/X_{\text{esc}}} - e^{-X/X_s}$ . With  $X_s \ll X_{\text{esc}}$  this model also explains the elemental composition.<sup>[10]</sup>

(c) Closed Galaxy Model: The cosmic rays never leave the galaxy, but they lose the energy by collision with interstellar gases.

Antiproton flux

Cosmic ray antiproton is less than cosmic ray proton by a factor  $10^4$ . ( See Fig.1 ). This is one of the reasons why we believe our galaxy is made of matter rather than antimatter.



If we could find cosmic anti-nucleus, it would extremely be interesting, for it may strongly suggest the existence of anti-matter stellar objects. Cosmic ray antiproton flux has so far been measured with use of ballon. To our surprise the experimental results show there are much more anti-proton than theoretically expected. [13] Especially at low energies below .4 GeV, the disagreement is by factor  $10 \sim 10^2$ .

At low energies the antiproton flux should be small by the following reason: If there is no anti-stellar objects, the only source of such antiproton is from collisions of cosmic rays with interstellar gases. The energy threshold of antiproton production process is about 6 GeV. If the cosmic ray energy is  $6\text{GeV} \sim 20\text{GeV}$ , the antiproton produced in p-p reaction should have a kinetic energy larger than .7GeV. If the cosmic ray energy is larger than 20GeV, the probability of such high energy is small. Thus we expect a small probability for low energy antiproton.

In the leaky box model or nested leaky box model explained in

the introduction the theoretical predictions are below the low energy data of Buffington et al., by a factor of  $10^2$  [11] [12] [13]. The closed galaxy model gives a larger antiproton flux, for in this model the proton traverses the interstellar space ten times more than the above models. The prediction is, however, smaller by a factor 10 still. Apparently we need more experimental and theoretical efforts in this problem. One way is to use the nuclear emulsion, for there is no technical difficulties in this method.

Composition at Kee

The differential or integral spectrum of cosmic rays has a kink at  $E \sim 10^{16}$  eV, as explained in the introduction, and it is called "knee". The reason of this knee may be either from the acceleration mechanism or from the propagation or confinement mechanism, but we do not know which is more important. [14] There are also exotic explanations, like monopole annihilation [15] or extragalactic anti-protons [16]. It is important to examine the elemental composition at this knee region to have better understanding of this knee. Some crucial experiment should be desirable.

$10^{20}$  eV Cosmic Rays

The cosmic rays with energy larger than  $10^{20}$  eV can have an inelastic reaction with 2.7K back ground photon, either producing pion or  $e^+e^-$  or disintegrating the nucleus itself. It might be a good place of test of Lorentz invariance or quantum mechanics. The presently available data seem not consistent with each other. [17] It is also expected that if there are cosmic rays above  $10^{20}$  eV or  $10^{21}$  eV, they should exhibit the arrival direction asymmetry, for they will not be disturbed by cosmic magnetic field very much.

The flie's eye system in Utah is one of those detection facilities for such high energy cosmic ray air shower. This has another attractive ability to detect a possible  $\nu_e$ -induced air

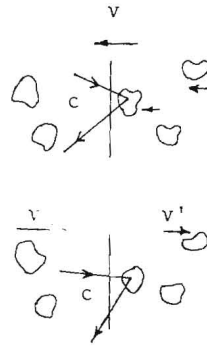
shower which may develop from the earth upward to the sky. This<sup>[18]</sup> idea is based on the naive extrapolation of the Landau-Pomeranchuk effect which suppresses the Brehmsstrahlung cross section of electron in rock through density effect at very high energy by a factor more than  $10^3$ . It may be that we need more careful theoretical treatment of this problem.

#### Appendix

This is a brief and simplified explanation of the derivation of power spectrum of 1st order Fermi acceleration at shock front, and more details are given in the references [8, 19], and also a text book [20].

Let us consider a shock front moving with velocity  $v$  to the left. Let us assume the magnetic field behind the front is vertical to the front surface for simplicity. Cosmic ray on the left-hand side enters the front and is scattered back to the left by magnetic clouds on the right hand side. Some time later this cosmic ray will be caught up by the shock front during its being randomly scattered by magnetic clouds of the left-hand side space. Then it may repeat such processes, and each time of reflection at the front it acquires an energy increase.

To be more quantitative let us consider the Lorentz frame in which the front is at rest. Then the plasma space of the left-hand side now moves with velocity  $v$  to the right. The plasma on the right hand side moves to the right with velocity  $v'$  ( $<v$ ) to the right.



In this frame we can easily calculate the probability of the cosmic ray not being reflected but escaping to the right  $p_{esc}$ , as  $p_{esc} = \frac{4v'}{c}$ ; for the cosmic ray flux coming from the left is  $\langle \rho v_z \rangle |_{v_z > 0} = \frac{v_c}{4}$ , and the escaping flux at the right end should be asymptotically  $\rho v'$ , giving the ratio  $\frac{4v'}{c}$ . After  $n$  times reflection the survival rate of the cosmic ray particle should be  $dP_s = e^{-P_{esc}n} dn$ .

Let us calculate the energy increase at one reflection. Net effect of the cosmic ray in the original galaxy frame should be due to the reflection by a magnetic cloud with velocity  $\Delta v = v - v'$  to the left. For the relative velocity of the two plasma spaces is  $v - v'$  in the front rest frame, and the relative velocity is nearly the same in the both frames. Let the cosmic ray energy before entering the magnetic cloud be  $E$  (neglecting the mass). The energy after entering the cloud in the cloud rest frame is approximately  $E' = E(1 + \frac{\Delta v}{c} \cos \theta_{in})$ . Here  $\theta_{in}$  is the angle between the cosmic ray and the  $z$ -axis. The energy of the cosmic ray after it is emitted out of the cloud should be  $E'' = E'(1 + \frac{\Delta v}{c} \cos \theta_{out})$ . Averaging over the angles with the flux weight  $\cos \theta d\Omega$ , we obtain  $\Delta E = \langle E'' - E \rangle = \frac{4\Delta v}{3c} E$ . After  $n$  times reflection the energy becomes  $E = E_0 e^{\frac{4\Delta v}{3c} n}$ . Combining this equation with the equation for  $dP_s$ , we obtain

$$dP_s \sim E^{-\gamma} dE \quad \text{with } \gamma = p_{esc} / (\frac{4\Delta v}{3c}) + 1.$$

If we remember Rankin-Hugoniot relations for the two plasma systems,

$$\begin{aligned} \rho v &= \rho' v', & r &= \frac{\rho'}{\rho} = \frac{v}{v'}, \\ \rho v^2 + p &= \rho' v'^2 = p', \\ \frac{1}{2} v^2 + \frac{dp}{\rho} &= \frac{1}{2} v'^2 + \frac{dp'}{\rho'} \end{aligned}$$

and the relations within each plasma,  $\rho p^{-\gamma} = \text{const.}$ ,  $\frac{\gamma p}{\rho} = c_s^2$ ,  $\gamma = \frac{5}{3}$ ,



we obtain  $r = \frac{4M^2}{M^2+3}$  and  $\gamma = \frac{r+2}{r-1}$ . Here  $M = v/c_s$  is the Mach number of the plasma. In the strong shock we may take the compression rate  $r = 4$ , the maximum value, we get  $\gamma = 2$ .

We should also take account of the propagation effect, or the life time effect  $t_{esc} \sim E^{-\delta}$ , and finally we obtain

$$dP \sim E^{-\gamma-\delta} dE$$

If we take  $\gamma = 2.3$ , and  $\delta = .4$  for nuclei, and the  $\delta = 1$  for electrons due to synchrotron energy loss, we have a consistent spectra.

#### References

1. For example, see R.N. Manchester Ap.J 188, 673 (1974).
2. M.E. Wiedenbeck, et al., Ap.J Letters 239, L139, 91980), for ex.
3. J.A. Simpson, in "Composition and Origin of Cosmic Rays", edited by M.M. Shapiro.
4. See also J.A. Simpson, Ann. Rev. Of Nucl. Part. Phys. 33, (1983).
5. W.R. Webber, in the same book as [2].
6. See for ex. A.M. Hillas, in the same book as [3].
7. E. Fermi, Phys. Rev. 75, 1169 (1943).
8. W.I. Axford, E. Leer, and G. Skadron, 15th ICRC 11, 132 (1977)  
R.D. Blandford, and J.P. Ostriker, Ap.J Letters 221, 129 (1978).
9. J.P. Ostriker, Nature 217, 1127 (1968); Goldreich and Julian, Ap. J. 157, 869 (1969).
10. Garcia-Munoz, M and J.A. Simpson, 16th ICRC 1, 270 (1979).
11. See e.g. T.K. Geisser and R.H. Mauer Phys. Rev. Letters 30, 1296 (1973).
12. S.A. Stephens Nature 289, 267 (1981); R.J. Protheroe, Ap.J 251, 387 (1981). See also C.J. Cesarsky and T.M. Montmerle, 17th ICRC 9, 207 (1981), and R. Coswik and T.K. Geisser, ibid., 2, 218 .
13. R.L. Golden et al., Phys. Rev. Letters, 43, 1196 (1979).  
E.A. Bogomolov et al., 16th ICRC 1, 330 (1979).  
M.V.K. Apparao, Canadian J. Phys. 46, S652 (1968).  
A. Buffington et al, Ap.J 248, 1179 (1981).
14. See the review by W.R. Webber, in the same book as [3], p.25.
15. P. Rottelli, Proc. of Workshop on Monopoles and Proton Decay, ed. by J. Arafune and H. Sugawara, KEK83-12, 1983
16. private communication, A.W. Wolfendale
17. see the talk by Nagano in this workshop.
18. private communication, J. Elbert.
19. A.R. Bell, M.N.R.A.S. 182, 147 (1978).
20. C.J. Cesarsky, in the same book as [3], p.161 and references therein.

## Observation of Ultra High Energy Cosmic Rays ( $\geq 10^{10}$ GeV)

M. Nagano

Institute for Cosmic Ray Research, University of Tokyo,  
3-2-1 Midoricho, Tanashi, 188 Tokyo

### ABSTRACT

Recent measurements of energy spectrum, arrival directions and mass compositions of ultra high energy cosmic rays ( $\geq 10^{10}$  GeV) are summarized. The surface air shower array of area over  $20 \text{ km}^2$  under construction in Akeno and a future plan for a huge array of over  $250 \text{ km}^2$  are described.

### 1. Introduction

It is of special interest to know the origin of the highest energy cosmic rays; are they Galactic or extragalactic?

The origin of cosmic rays of energy lower than  $10^{10}$  GeV is still far from conclusion. Nevertheless, the reason for studying the higher energy region is that the problem of propagation of cosmic rays becomes simpler, the higher the primary energy is. This is because cosmic rays of energy higher than  $10^{10}$  GeV cannot be confined in the Galaxy by the Galactic magnetic field of  $3 \mu\text{G}$ , if they are protons. Furthermore, the possible astronomical objects which can be considered to be able to accelerate the cosmic rays up to  $10^{10}$  GeV are limited and hence anisotropy of their arrival direction may be expected.

In this report, some important aspects of the propagation of cosmic rays are briefly discussed and the data from recent measurements on the primary energy spectrum, arrival direction and mass composition are summarized. Finally a surface array of area over  $20 \text{ km}^2$  under construction in Akeno and a plan of huge array of over  $250 \text{ km}^2$ , currently under discussion are described.

### 2. A few important remarks on ultra high energy cosmic rays

There are some excellent reports on the origin and propagation of ultra high energy cosmic rays <sup>(1)-(3)</sup>. A recent reviews by Hillas <sup>(4)</sup> will be helpful in understanding the problem of acceleration of these cosmic rays. Here a few important remarks are given.

(i) Size of accelerating regions <sup>(3)(4)</sup>

In order to accelerate a particle in a statistical acceleration process, the size  $L$  of the accelerating region containing the magnetic field  $B$  (in microgauss) must be much larger than twice the gyroradius ( $r_L$ ) of the particle.

$$L \gg 2 r_L = 2 E_{18}/ZB$$

$$B \times L \gg 2 E_{18}/Z \text{ (in } \mu\text{G} \times \text{kpc),}$$

where  $E_{18}$  is energy of the particle of charge  $Ze$  in unit of  $10^{18}$  eV. A similar limit is also obtained in the case of one-shot acceleration. In Fig.1 is shown a plot of  $L$  vs  $B$  for several astronomical objects. It is seen that all objects below the line are excluded from being the sources.

(ii) Propagation in the Galactic space

The gyroradii of protons and iron nuclei in a uniform magnetic field of  $3 \mu\text{G}$  are tabulated in Table 1 for various energies. It is known that there is a strong interaction between the charged particle and the magnetic field only when the inhomogeneities in the galactic magnetic field are on the same scale as the gyroradius of the charged particles. Since  $r_g$  for protons of  $10^{10}$  GeV is of the order of the scale of the halo and much greater than the scale of the magnetic irregularities, they can not be confined in the Galaxy. Actually Karakula et al<sup>(5)</sup> and Osbone et al<sup>(6)</sup> calculated the trajectories of particles in the galactic magnetic field by taking into account the magnetic irregularities in the disc and showed that protons above  $10^9$  GeV are anisotropic. Even if we take into account the magnetic field in the halo of about  $1 \mu\text{G}$  on the scale of  $2\text{--}3\text{kpc}$ <sup>(40)</sup>, the situation may not change for  $10^{10}$  GeV protons.

(iii) Propagation in the intergalactic space

As is well known, the particles with energy of  $\gamma mc^2$  ( $\gamma$  is the Lorentz factor) produce pions by interaction with the  $2.7^\circ\text{K}$  primordial radiation if the energy of photon  $\gamma \epsilon(1+\cos\theta)$  exceeds the threshold energy about 140 MeV. The density of the photon is  $400/\text{cm}^3$  and photon energy  $\epsilon$  at the upper part of the Planck distribution is about  $10^{-3}$  eV. Attenuation length for protons in the  $2.7^\circ\text{K}$  radiation field calculated by Giler et al<sup>(7)</sup> is shown in Fig.2. It is understood that a sudden decrease of the flux is expected above  $5 \times 10^{10}$  GeV, if the cosmic rays are extragalactic, and the spectral shape above  $10^{11}$  GeV may be modified from the production spectrum even when they are produced in the Virgo cluster of galaxies. In Fig.3 is shown the predicted spectrum<sup>(7)</sup> in which the Virgo cluster is the sole source of particles and the particles diffuse outwards under the influence of randomly directed intergalactic magnetic field.

### 3. Summary of recent measurements on ultra high energy cosmic rays

#### 3.1. Primary energy spectrum

In Fig.4 are compared the detector arrangements of large air shower experiments in the world: Volcano Ranch<sup>(8)</sup>, SUGAR<sup>(9)</sup>, Haverah Park<sup>(10)</sup>, Yakutsk<sup>(11)</sup> and Akeno<sup>(12)</sup>. In Table 2, important parameters of these

experiments such as locations, kinds of detectors, exposure time in  $\text{km}^2$  year, number of events detected up to 1983, etc. are listed. Though the SUGAR experiment, located in the southern hemisphere, covered the largest area, it was already shut down in 1979. Also the detector spacing is so large that the threshold energy for the whole array is a few times  $10^{10}$  GeV.

It should be noted that the largest sites under operation, Yakutsk and Haverah Park, are situated at the  $62^\circ\text{N}$  and  $54^\circ\text{N}$  latitude and the Galactic center can not be observed by them. In Fig.5 is shown the differential energy spectrum of primary cosmic ray measured by the groups listed in Table 2. The vertical axis is multiplied by  $E_0^{2.5}$  in order to see the spectral shape in detail. The spectrum by the Akeno group is converted from the electron and the muon size spectrum based on the methods which are considered to be insensitive to either mass composition or hadronic interaction model<sup>(13)</sup>. This spectrum smoothly joins the spectra obtained by the Proton satellite<sup>(16)</sup>, and the Haverah Park<sup>(14)</sup> and Yakutsk<sup>(15)</sup> groups. The results from the SUGAR<sup>(17)</sup> experiment are also plotted in the same figure. Here the SUGAR spectrum is converted from their muon size spectrum after normalizing their muon energy to 1 GeV and using the conversion factor  $E_0 = 1.2 \times 10^{17} (N_\mu / 10^6)^{1.15}$  in eV. The differences among the groups are less than 30% in energy. By admitting the ambiguity of a factor of 1.5 in energy estimation, we can draw the following conclusions :

- (i) The slope of the spectrum changes around  $(1 \sim 2) \times 10^{10}$  GeV from 3.0 to 2.5.
- (ii) This change of spectral shape is observed both in northern and southern hemispheres.
- (iii) The existence of the spectrum cutoff above  $5 \times 10^{10}$  GeV does not seem to be established.
- (iv) There are certainly several showers whose energies exceed  $10^{11}$  GeV and hence the spectrum should be explored to the higher energies.

#### 3.2 Arrival direction

In Fig.6 is shown the ratio of the observed to the expected number of showers as a function of Galactic latitude for ten energy bins reported by the Haverah Park group.<sup>(18)</sup> The expected one is calculated under the assumption of isotropic distribution. The gradient of the least square fitted straight line in the figure is plotted in Fig.7 as a function of primary energy. It is

remarkable that the sign of gradient changes above  $10^{10}$  GeV, suggesting that those cosmic rays come from higher Galactic latitude. The Haverah Park group interpreted these results, together with the change of the slope in the energy spectrum, as evidence that the origin of the highest energy cosmic rays is extragalactic. However, a similar analysis by the Yakutsk group<sup>(19)</sup>, plotted in the same figure, does not show such tendency above  $10^{10}$  GeV, though they agree with the Haverah Park results below  $10^{10}$  GeV. It is also noted that the SUGAR experiment which can not observe the center of the Virgo Cluster shows a similar flattening of the energy spectrum.

In order to see the arrival direction of the largest showers above  $10^{10}$  GeV, their arrival direction are plotted in Fig.8(a), (b), and (c) from the three groups separately. The distribution from the Haverah Park group<sup>(18)</sup> seems to show an enhanced flux in the direction of the Galactic pole, while that from Yakutsk<sup>(20)</sup> in the Galactic plane. The results from SUGAR<sup>(21)</sup>, which is located in the southern hemisphere, are not inconsistent with isotropic distribution.

The arrival directions of 82 showers of energies larger than  $4 \times 10^{10}$  GeV are plotted all together in equatorial coordinates (Fig.9). Here the points from the SUGAR experiment are for showers of energy larger than  $2 \times 10^{10}$  GeV according to their energy assignment. It should be noted that the exposure time is uniform only in the same declination band. Nevertheless, it is clear from the figure that the directions are not concentrated in the Galactic plane, which is expected if these are protons and produced near supernova or new born pulsars in the Galaxy.

Conclusions are summarized as follows :

- (i) The change of arrival direction distribution and hence origin at around  $10^{10}$  GeV is suggested by the Haverah Park group, but not yet conclusive.
- (ii) The possible origin of  $10^{10}$  GeV showers from the center of Virgo Cluster suggested by the Haverah Park group is not yet confirmed by other groups. The arrival direction from both the Yakutsk group and the SUGAR experiment are not inconsistent with isotropic distribution and the energy spectrum from SUGAR shows a similar change of slope at  $10^{10}$  GeV as observed in the northern hemisphere.

### 3.3 Mass composition

Information on mass composition is inevitable to determine the origin of the cosmic rays around the energy concerned. Determination of mass composition by indirect method is most difficult.

Mass composition of primary cosmic rays is measured up to  $10^5$  GeV directly by emulsion chambers flown in balloons<sup>(22)</sup>. Above  $10^5$  GeV, the method is indirect and hence the results are of much dispute among various groups. One of the reasons is that most conclusions are drawn from the observation of one or two observables, which are related not only to the mass composition but also to the characteristics of hadronic interactions.

One of the best methods to estimate the composition in the EAS energy region is to measure the starting point distribution of the EAS. This distribution reflects the collision mean free path of various species of primaries. The Fly's Eye<sup>(23)</sup> is primarily intended to measure this distribution by observing the fluorescent light from EAS. However, since the inelastic cross-section of p-Air collision increases with energy as  $290 E_0^{0.06}$  mb ( $E_0$  in TeV)<sup>(24)</sup>, the resolution required for the determination of mass composition is less than  $10 \text{ g/cm}^2$  which can not be achieved by the present Fly's Eye (more than  $50 \text{ g/cm}^2$ ). Improvement of the resolution of the Fly's Eye is under planning<sup>(25)</sup> and this kind of experiment at high altitude by a large  $\checkmark$  telescope of fine resolution is proposed by Tanahashi.<sup>(26)</sup>

The next approach is to investigate the relation between the depth of the shower maximum ( $t_{\text{max}}$ ) and its fluctuation. As illustrated schematically in Fig.10, the average  $t_{\text{max}}$  is at higher level for heavier primaries than for proton primaries and the fluctuation in  $t_{\text{max}}$  ( $\sigma(t_{\text{max}})$ ) for heavy primaries is smaller than for proton primaries.

In Fig.11, the values of  $\sigma(t_{\text{max}})$  are plotted against the associated values of  $t_{\text{max}}$ . The curves represent the predictions by Chantler et al.<sup>(27)</sup> for different proportions of iron and protons in the primary composition on the basis of a two-component approximation. Each curve corresponds to different interaction models. Open and closed circles correspond to the values when the primaries are 100% iron and 100% protons, respectively. The large cross is from the Cerenkov experiments by Chantler et al.<sup>(27)</sup>. The large open square and the shaded square are the measurements from Akeno and the Fly's Eye<sup>(28)</sup>, respectively. At Akeno, the  $t_{\text{max}}$  is obtained by Cerenkov telescope<sup>(29)</sup> and  $\sigma(t_{\text{max}})$  from both the  $\checkmark$  measurements and the indirect

method using the age parameter distribution for a fixed muon size<sup>(30)</sup>, developed by Coy et al<sup>(31)</sup>. The triangle is the summary of the world survey of measurements of  $t_{\max}$  and  $\sigma(t_{\max})$  at the Cosmic Ray Conference at Bangalore in 1983<sup>(32)</sup>. The results suggest a large fraction of proton primaries at around  $10^8$  GeV, irrespective of interaction models. For  $10^{10}$  GeV, the data is limited to the estimates of  $\sigma(t_{\max})$  determined from the fluctuation study of various observables. The results are summarized in Fig.12 together with those at lower energies<sup>(32)</sup>. The solid and dotted lines are calculated by Walker and Watson<sup>(33)</sup> for pure proton and pure iron primary, respectively. There seems no change of composition up to  $10^{10}$  GeV.

Another method to estimate the primary composition from the ground based parameters is to use the fluctuation of number of muons in a shower. If we assume the superposition model in nucleus-Air collision, the number of muons induced by the proton and the nucleus A are

$$N_{\mu}(p) = c E_0^b$$

$$N_{\mu}(A) = c E_0^b A^{1-b}.$$

$N_{\mu}(A)$  is larger than  $N_{\mu}(p)$  by a factor of  $A^{1-b}$ , where b is related to the multiplicity in p-Air collision and less than 1.0. Therefore

$$\log N_{\mu}(A) = \log c + (1-b) \log A + b \log E_0.$$

If  $\langle A \rangle$  changes with  $E_0$ , the value  $N_{\mu}(A)$  changes resulting in change of the slope in  $\log N_{\mu}$  vs.  $\log E_0$  relation.

Fluctuations in  $N_{\mu}$  for a given energy  $E_0$  is also sensitive to mass composition, since not only the  $N_{\mu}$  depends on A, but also its dispersion decreases with A. The observed dispersion for a mixed composition is expressed by<sup>(34)</sup>

$$\sigma^2 = \sum w_i \sigma_i^2 + \sum w_i \langle \log N_{\mu} \rangle_i^2 - \langle \log N_{\mu} \rangle^2,$$

where  $w_i$  and  $\sigma_i$  represent the intensity ( $\sum w_i = 1$ ) and the dispersion in  $\log N_{\mu}$  of ith component, respectively. The variance consists of two parts, one due to fluctuations (the first term), the other due to the width of  $\langle \log N_{\mu} \rangle$  of each component. (See the general formulation and the detailed discussion by Linsley<sup>(34)</sup>).

Actually we can not analyse the  $N_{\mu}$  distribution for a given energy, but for a given electron size ( $N_e$ ). Furthermore, the experimental error in determination of  $N_{\mu}$  is not far smaller than  $\sigma_p$  even in the case of Akeno experiment which has the largest muon detectors in the world.

In Fig.14(a) and (b) are shown the results of the Akeno experiment<sup>(35)</sup>,  $\log N_{\mu}$  vs  $\log N_e$  and the  $\log N_{\mu}$  distribution for a fixed  $N_e$ . Curves are the accepted combinations of primary compositions shown in Fig.13 and hadronic interactions among many other combinations which can be excluded. In cases A and B, scaling of production spectrum of secondary particles is assumed at  $x \geq 0.05$  and multiplicity at central region increases as  $(\ln s)^2$ . In case of C, fireball model similar to CKP is assumed. In all cases, the cross-section of p-Air collision increases with energy as  $290E^{0.06}$  mb(E in TeV).

There are many other observables such as the pulse shape of arrival time of muons<sup>(36)</sup>, the longitudinal development curves of electrons and muons<sup>(35)(36)</sup>, frequency attenuation length of the showers for a fixed  $N_e$ <sup>(35)</sup>, etc. All these quantities must be consistently used to distinguish between the effects of mass composition and hadronic interactions. This analysis is now in progress.

#### 4. Plan of the giant air shower observation in Japan

In order to clarify the ambiguities stated in the preceding sections, a plan for a huge array of over  $250 \text{ km}^2$  is currently under discussion. The main goals of this experiment are as follows.

- (i) To increase the total observed number of showers above  $10^9$  GeV world wide by an order of magnitude and to establish their arrival direction distributions.
- (ii) To extend the primary energy spectrum above  $10^{11}$  GeV.
- (iii) To investigate the existence of cutoff in the primary spectrum above  $5 \times 10^{10}$  GeV in a direction where candidates of possible sources can not be found within 30 Mpc.
- (iv) To investigate whether there is a change of characteristics of EAS at  $10^{10}$  GeV, where the shape of the primary spectrum changes. If there is any change, it may be related to the change of composition around this energy.
- (v) To calibrate the new detection method of giant air showers, in order to extend the primary spectrum above  $10^{11}$  GeV.
- (vi) To calibrate the response of detectors used in other experiments in order to analyse the data with the same energy scale and hence to increase the total number of events in the world.

The detector arrangement under consideration is shown in Fig.15. The area of each electron detector is  $2.25 \text{ m}^2$  and the detectors will be located at about 1 km separation. Muon detectors of area about  $20\text{-}30 \text{ m}^2$  will also be arranged at about 2 km separation. The solid large and small circles indicate the areas within which the detectors would record by more than one or ten particles, respectively, when an EAS of larger than  $10^{10}$  GeV fell. The dotted circle shows the case of muons for 1 particle larger than 1 GeV. The accuracy of the size estimation and arrival direction would be 20-40% and  $3^\circ$ , respectively, for  $10^{10}$  GeV and less for  $10^{11}$  GeV<sup>(38)</sup>.

The whole area is divided into four sections, each of which is called a "Branch". The details of "Akeno Branch", part of which is now under construction, are illustrated in Fig.16. Each detector is connected to the next one with two optical fiber cables successively on a string as shown in Fig.17. Each string is controlled by the central CPU through the master CPU(MPC). One cable is used for the control of each slave unit(DCU) and the other for transmission of data from the detector to the center. The details will be described in the reports by Ohoka<sup>(37)</sup> and Teshima<sup>(38)</sup>.

The main reasons for using optical fiber are (i) to avoid the radio noise from lightning which often causes serious damage to electronics in Akeno in summer and (ii) to get a timing signal of better quality than can be obtained by coaxial cable. The optical fiber cable is hung on the electricity poles or the utility poles of the Telegraph and Telephone Corporation.

Besides electron and muon density, measurements of arrival time profile of electrons and muons are also planned. Recently, Linsley<sup>(39)</sup> proposed that the time dispersion of the shower front is a function of core distance and hence by making use of pulse width information as well as pulse height information of arrival time of electrons, the size of large showers can be estimated. If both dispersions of arrival time and density are small enough, this method is applicable to extend the effective area as shown by solid circles in Fig.16. The pulse width information from muon arrival time distribution is expected to measure the stage of the longitudinal development of the shower.

In Fig.4 is shown our plan of expansion of array. Area surrounded by the solid line will be in operation this year. The dotted area is the plan of the whole "Akeno Branch". The chain area is the planned array of the first stage covering  $75 \text{ km}^2$ . The  $250 \text{ km}^2$  array covers the whole area of Fig.4.

#### Acknowledgement

The author is grateful to all members of the Workshop on Observation of Ultra High Energy Cosmic Rays for the detailed discussions of this project. He thanks to Prof. M.V.S. Rao for the reading of the manuscript.

#### References:

- (1) V.S. Berezinsky ; Invited Talk, Proc 15th Int. Cosmic Ray Conference (ICRC), Plovdiv 10 (1977) 84.
- (2) W.I. Axford ; Invited Talk, Proc 17th ICRC, Paris 12 (1981) 155.
- (3) A.M. Hillas ; Composition and Origin of Cosmic Rays (ed. by M.M.Shapiro, Reitel Publishing Company) (1983) p.125. ; Proceedings of Cosmic Ray Workshop at University of Utah(ed. by T.K.Gaisser, Bartol Research Foundation) (1983) 1 and 16.
- (4) A.M. Hillas ; To appear in Annual Review of Astronomy and Astrophysics vol.22 (1984).
- (5) S. Karakula et al. ; J. Phys. A 5 (1972) 904.
- (6) J.L. Osborne et al. ; J. Phys. A 6 (1983) 421.
- (7) M. Giler et al. ; J. Phys. G : Nucl. Phys. 6 (1980) 1561.
- (8) J. Linsley ; Proc. 8th ICRC, Jaipur 4 (1963) 295.
- (9) C.J. Bell et al. ; J. Phys. A 7 (1974) 990.
- (10) D.M. Edge et al. ; J. Phys. A 6 (1973) 1612.
- (11) O.S. Diminstein et al. ; Proc. 15th ICRC 8 (1977) 154.
- (12) T. Hara et al. ; Proc. 18th ICRC, Bangalore (1983) EA1.2-2.
- (13) M. Nagano et al. ; to be published in J. Phys. G : Nucl. Phys. (1984).
- (14) A.J. Bower et al. ; Proc. 17th ICRC, Paris 9 (1981) 166.
- (15) N.N. Efimov et al. ; cited from Hillas(ref. 3) (1981).
- (16) N.F. Grigorov et al. ; Proc. 12th ICRC, Hobart 5 (1971) 1746.
- (17) L. Horton et al. ; Proc. 18th ICRC, Bangalore 6 (1983) 124.
- (18) S.M. Astley et al. ; Proc. 17th ICRC, Paris 2 (1981) 156.
- (19) N.N. Efimov et al. ; Proc. 18th ICRC, Bangalore 2 (1983) 149.
- (20) D.D. Krasilnikov et al. ; Proc. 18th ICRC, Bangalore 2 (1983) 145.
- (21) L. Horton et al. ; Proc. 18th ICRC, Bangalore 2 (1983) 153.
- (22) T.H. Burnett et al. ; Workshop on Very High Energy Cosmic Ray Interactions (ed. by M.L.Cherry, K.Lande and R.I.Steinberg, University of Pennsylvania) (1982) 221.

Table 1. Gyroradii of protons and irons in Galactic magnetic field of 3μG

log E(eV)	Protons	Iron(Z=26)
15	0.3 pc	0.01 pc
17	30 pc	1 pc
19	3 kpc	100 pc
21	300 kpc	10 kpc

Table 2. Important parameters of large air shower experiment

Place	Latitude	Longitude	Elevation	Area	Exposure	Main detector	Number of events
Volcano Ranch	35.16°N	106.79°W	1770 m	8 km <sup>2</sup>	110 km <sup>2</sup> yr	9cm Scinti.	44
Haverah Park	53.96°N	1.63°W	220	12	350	120cm Water C.	144
Pilliga Forest	30.52°S	149.63°E	250	55	540 for 10 <sup>11</sup> GeV 70 for 10 <sup>10</sup> GeV	15g/cm <sup>2</sup> Liq.Sci.	32
Yakutsk	62°N	130°E	100	17		8g/cm <sup>2</sup> Scinti.	115
Akeno	35.78°N	138.5°E	920	4		5cm Scinti. Proportional D.	11

- (23) G.L. Cassiday et al. ; Workshop on Very High Energy Cosmic Ray Interactions (ed. by M.L.Cherry, K.Lande and R.I.Steinberg, University of Pennsylvania) (1982) 72.
- (24) T. Hara et al. ; Phys. Rev. Lett. 50 (1983) 2061.
- (25) G.L. Cassiday ; Talk at Cosmic Ray Workshop at University of Utah, January (1983).
- (26) G. Tanahashi ; to appear in Proc. of GAS Workshop at ICRR, February (1984).
- (27) M.P. Chantler et al. ; J. Phys. G : Nucl. Phys. 9 (1983) L27.
- (28) R. Cady et al. ; Proc. 18th ICRC, Bangalore (1983) EA4-22.
- (29) M. Daigo et al. ; Proc. 18th ICRC, Bangalore (1983) EA4-12.
- (30) M. Nagano et al. ; to be published in J. Phys. Soc. Japan No.5 (1984).
- (31) R.N. Coy et al. ; Proc. Workshop on Very High Energy Cosmic Ray Interactions (University of Pennsylvania) (1982) p.120.
- (32) M. Nagano ; Rapporteur paper, Proc. 18th ICRC, Bangalore (1983).
- (33) R. Walker and A.A. Watson ; Proc. 18th ICRC, Bangalore 6 (1983) 114.
- (34) J. Linsley ; Rapporteur paper, Proc. 18th ICRC, Bangalore (1983).
- (35) M. Honda et al. ; Proc. 18th ICRC, Bangalore (1983) EA3-13.
- (36) F. Kakimoto ; Proc. 17th ICRC, Paris 11 (1981) 254.
- (37) H. Ohoka ; to appear in Proc. of GAS Workshop at ICRR, February (1984).
- (38) M. Teshima ; to appear in Proc. of GAS Workshop at ICRR, February (1984).
- (39) J. Linsley ; preprint (1983).
- (40) Sofue ; Talk at GAS Workshop at ICRR, February (1983).

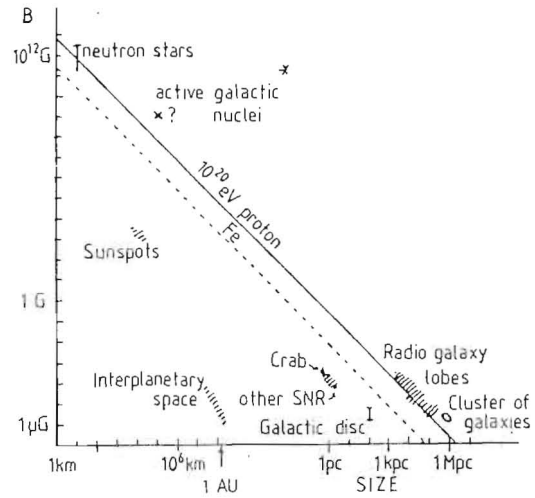


Fig.1. Size, and estimated magnetic field strength, of several sites of relativistic particle acceleration. Objects below the diagonal line cannot accelerate particles to  $10^{20}$  eV. (Hillas(3))

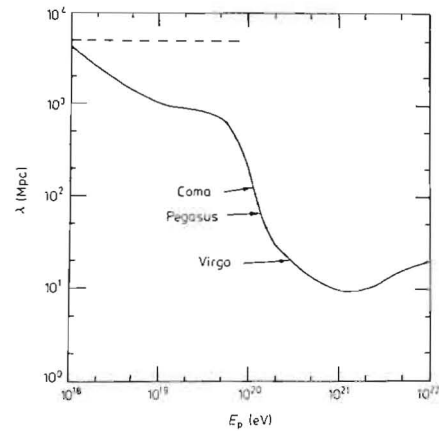


Fig.2. Attenuation length for protons in the 2.7 K radiation field. The distances to the three most prominent galactic clusters are indicated (based on  $H_0 = 60 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ). The broken line gives the Hubble loss. (Giler et al(7))

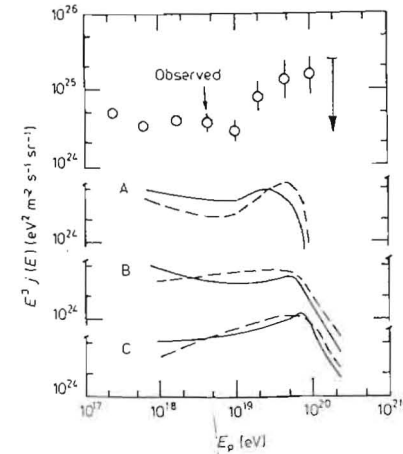


Fig.3. Observed primary spectrum (average of the data from the Volcano Ranch, Yakutsk and Haverah Park experiments; see I). Predictions for the model in which the Virgo cluster is the sole source of particles, with differential production spectrum  $E_p^{-7}$ , which diffuse with diffusion coefficient  $D$  and undergo attenuation in the 2.7 K radiation (figure 1). A,  $\gamma = 2.44$ ,  $D = 5 \times 10^{23} E_{19}^{1/2} \text{ cm}^2 \text{ s}^{-1}$ ; B,  $\gamma = 2.44$ ,  $D = 1.58 \times 10^{24} E_{19} \text{ cm}^2 \text{ s}^{-1}$ ; C,  $\gamma = 2.17$ ,  $D = 1.58 \times 10^{24} E_{19} \text{ cm}^2 \text{ s}^{-1}$ . The broken curves represent approximation A and the full curves the more accurate treatment of approximation B. (Giler et al.(7))

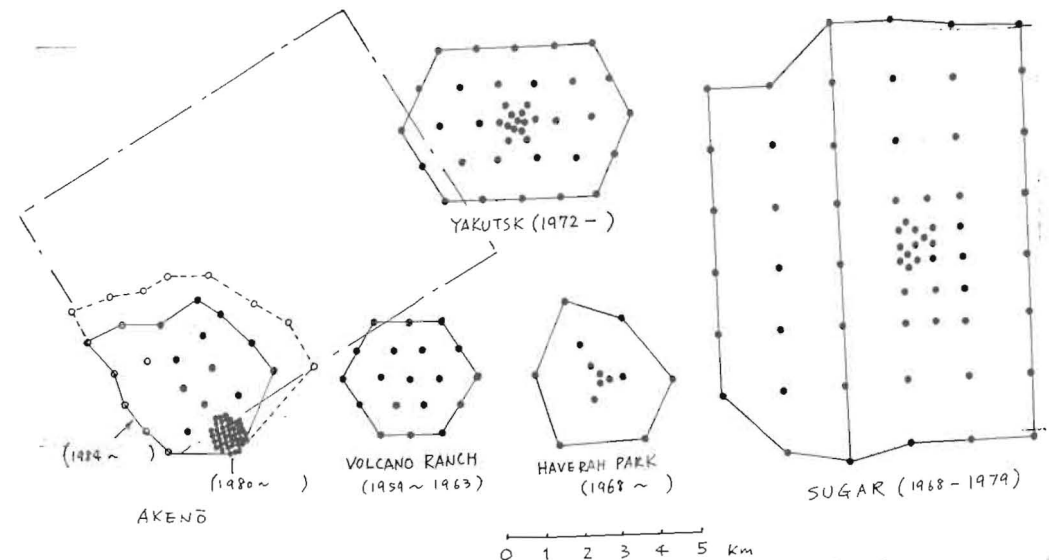


Fig.4. Detector arrangement of large air shower array. Details of Akeno array are described in Section 4.



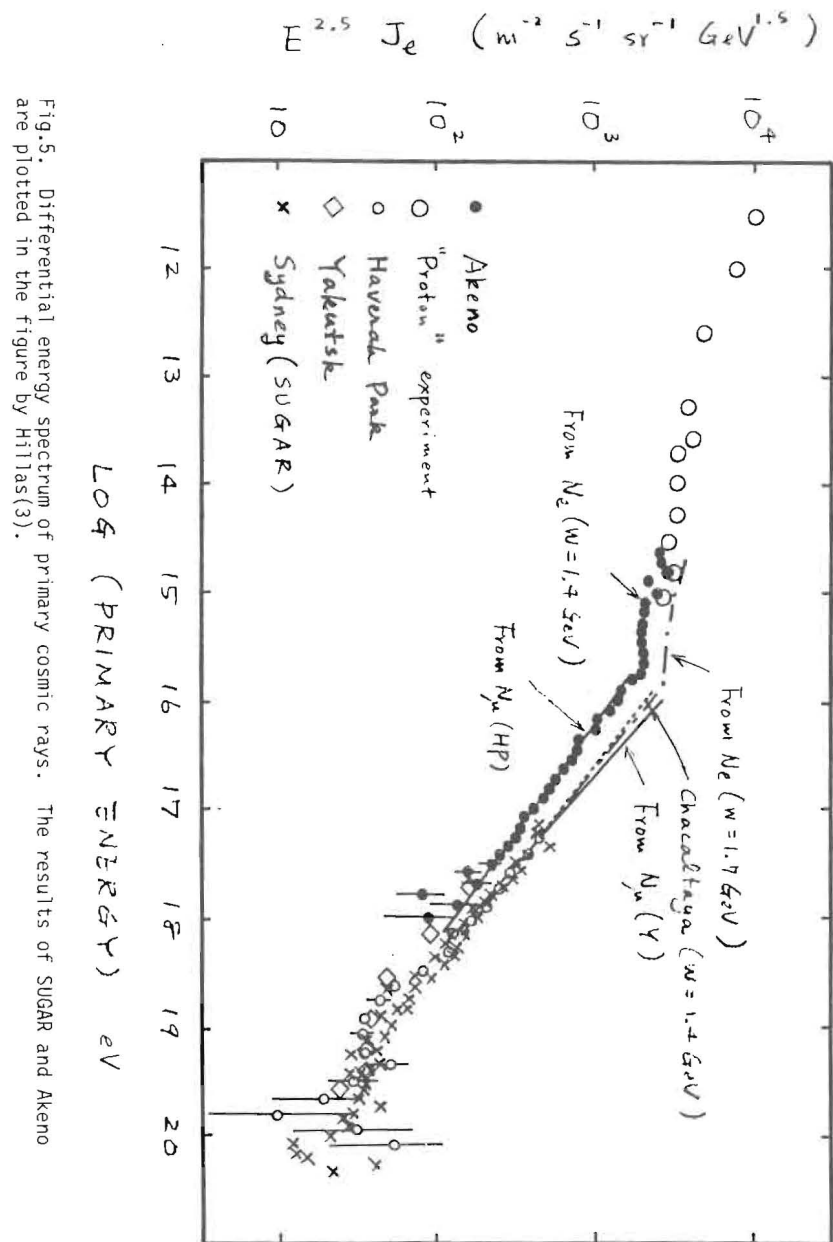


Fig.5. Differential energy spectrum of primary cosmic rays. The results of SUGAR and Akeno are plotted in the figure by Hillas(3).

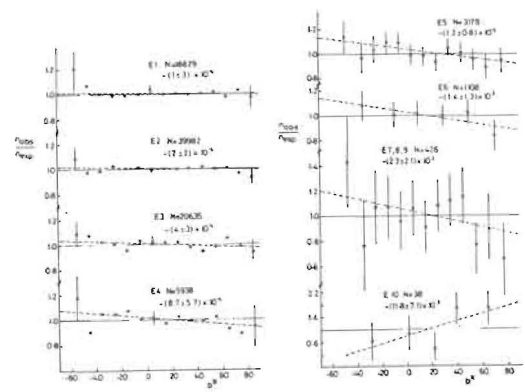


Fig.6

The ratio of the observed to the expected number of showers as a function of Galactic latitude. Showers are grouped into ten energy intervals (see text), with the number of showers and the gradient of a least squares fit indicated. Note that E7, E8, E9 are treated as a single group and that there is a change in vertical scale for group E10. (Astley et al.(18))

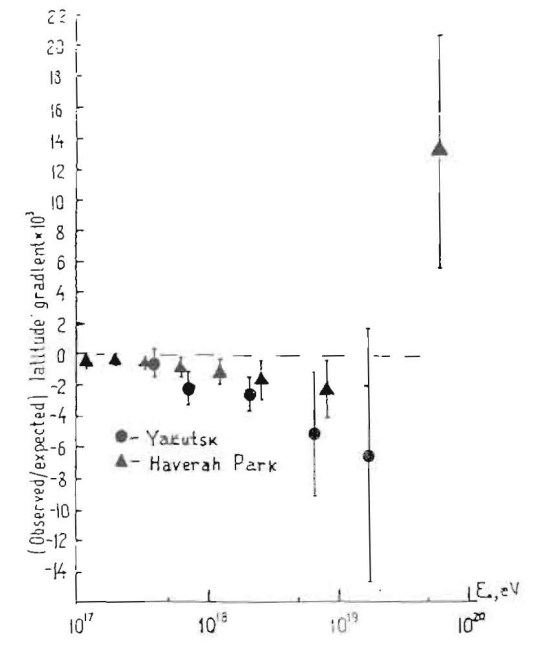


Fig.7.

Value of gradient  $N_{obs.}/N_{expect.}$  on the galactic latitude  $b''$   
 (Efimov et al.(19))

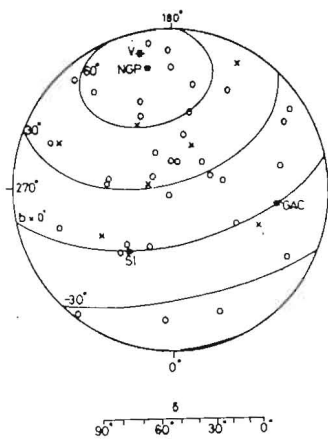
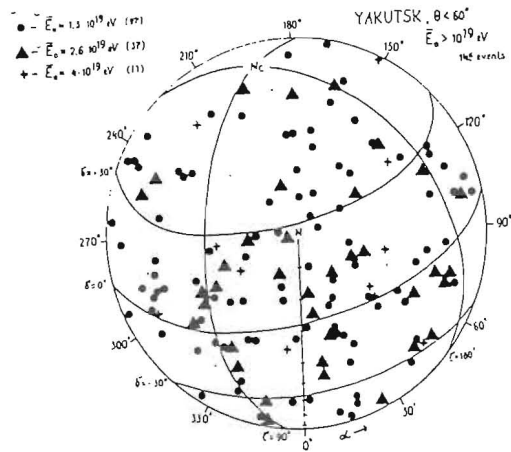
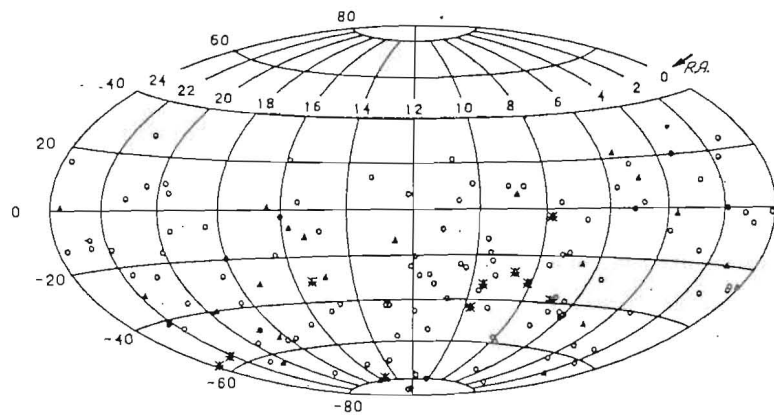


Fig. 8(a) An equal area plot of 41 events ( $E \geq 4 \times 10^{19}$  eV,  $\delta \geq 0^\circ$ ) recorded at Volcano Ranch (x) and Haverah Park (o). (Astley et al(18))

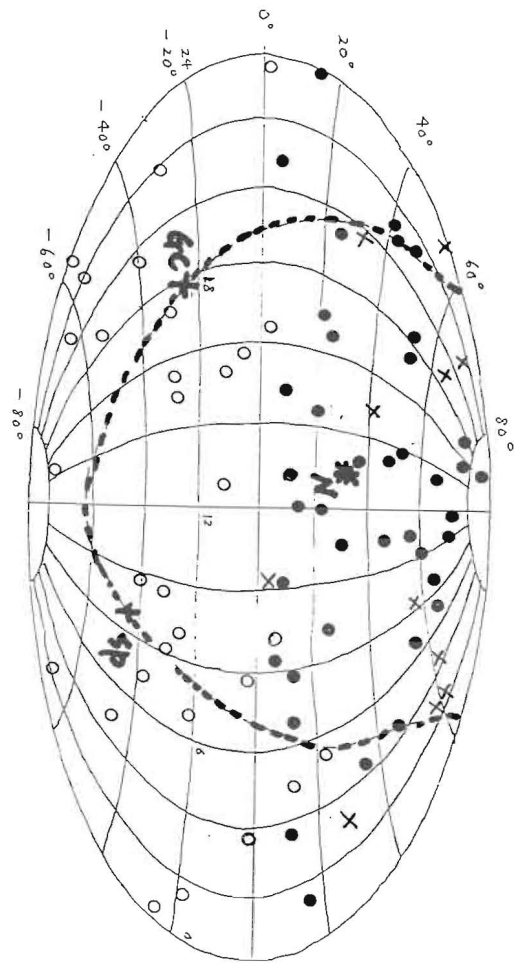


(b) An equal exposure plot of 145 events ( $E \geq 10^{19}$  eV) recorded at Yakutsk. (Krasilnikov et al.(20))



(c) The arrival directions of air showers with energy  $E > 10^{19}$  eV and zenith angle  $< 60^\circ$  as observed by S.U.G.A.R. (Equatorial coordinates; Aitoff equal area projection) (\*  $E > 4 \times 10^{19}$  eV;  $\Delta$   $4 \times 10^{19}$  eV  $> E > 2 \times 10^{19}$  eV,  $\circ$   $2 \times 10^{19}$  eV  $> E > 1 \times 10^{19}$  eV). (Horton et al(21)).

Fig. 9. The arrival directions of 82 showers with energy  $E \geq 4 \times 10^{19}$  eV. (Haverah Park (●), Yakutsk (x) and SUGAR (○)) in equatorial coordinates. Dotted line indicates the galactic plane.



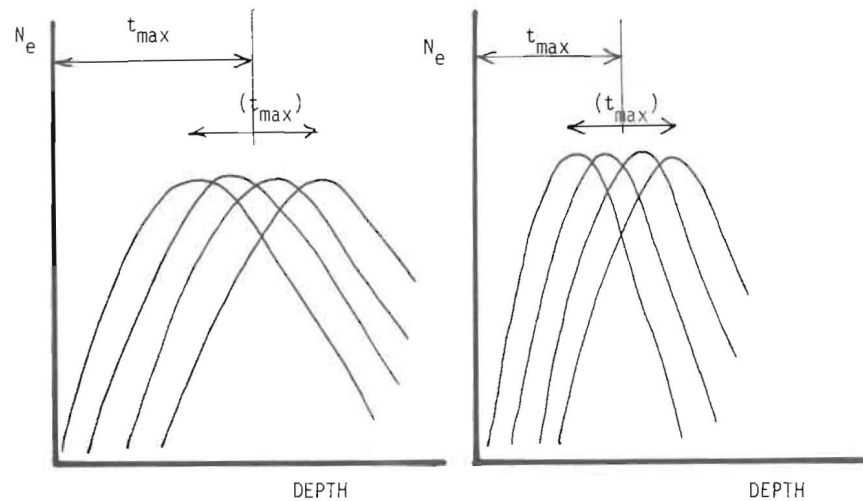


Fig.10. Schematic illustration of longitudinal development curves of protons(left) and heavy particles(right).

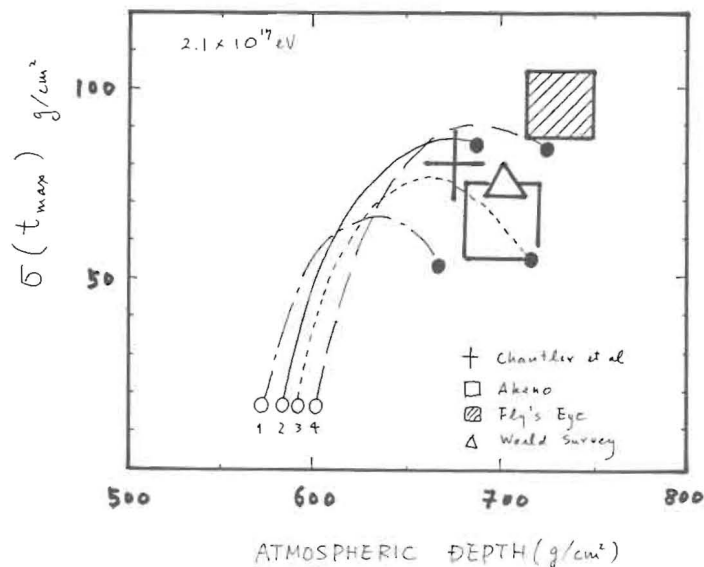


Fig.11. The relation between  $t_{max}$  and  $\sigma(t_{max})$ . The curves represent the predictions for different proportions of iron (○100% iron) and protons (●100% proton) in the primary composition for four kinds of interaction models. (see Chantler et al(27) for details of models).

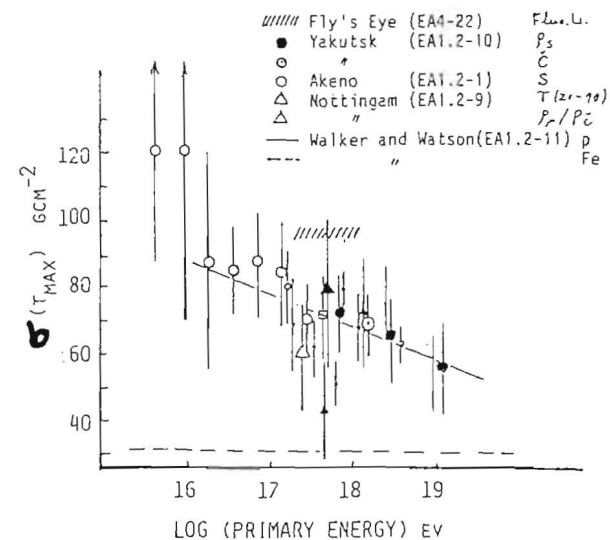


Fig.12.

Fluctuation in the depth of shower maximum vs. primary energy from the various experimental methods. References are in the Proceedings of 18th ICRC(1983). (Nagano (3z))

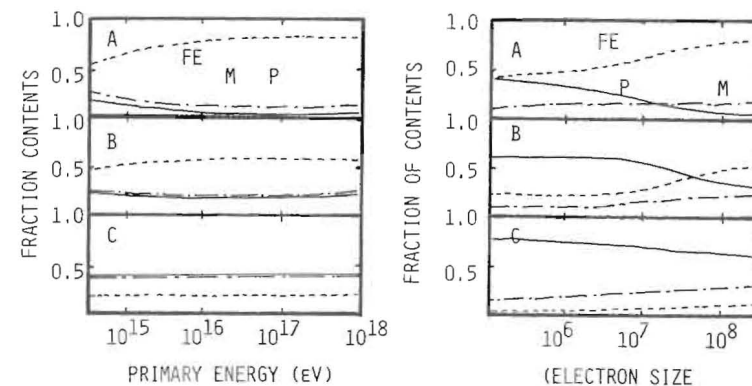


Fig.13 Three models on primary composition A, B and C. Each curve represents the fractional contents of iron(Fe), proton(p) and others (M), respectively. Left figure is drawn as a function of total energy of primary particles and right one as electron size of EAS at Akeno altitude.

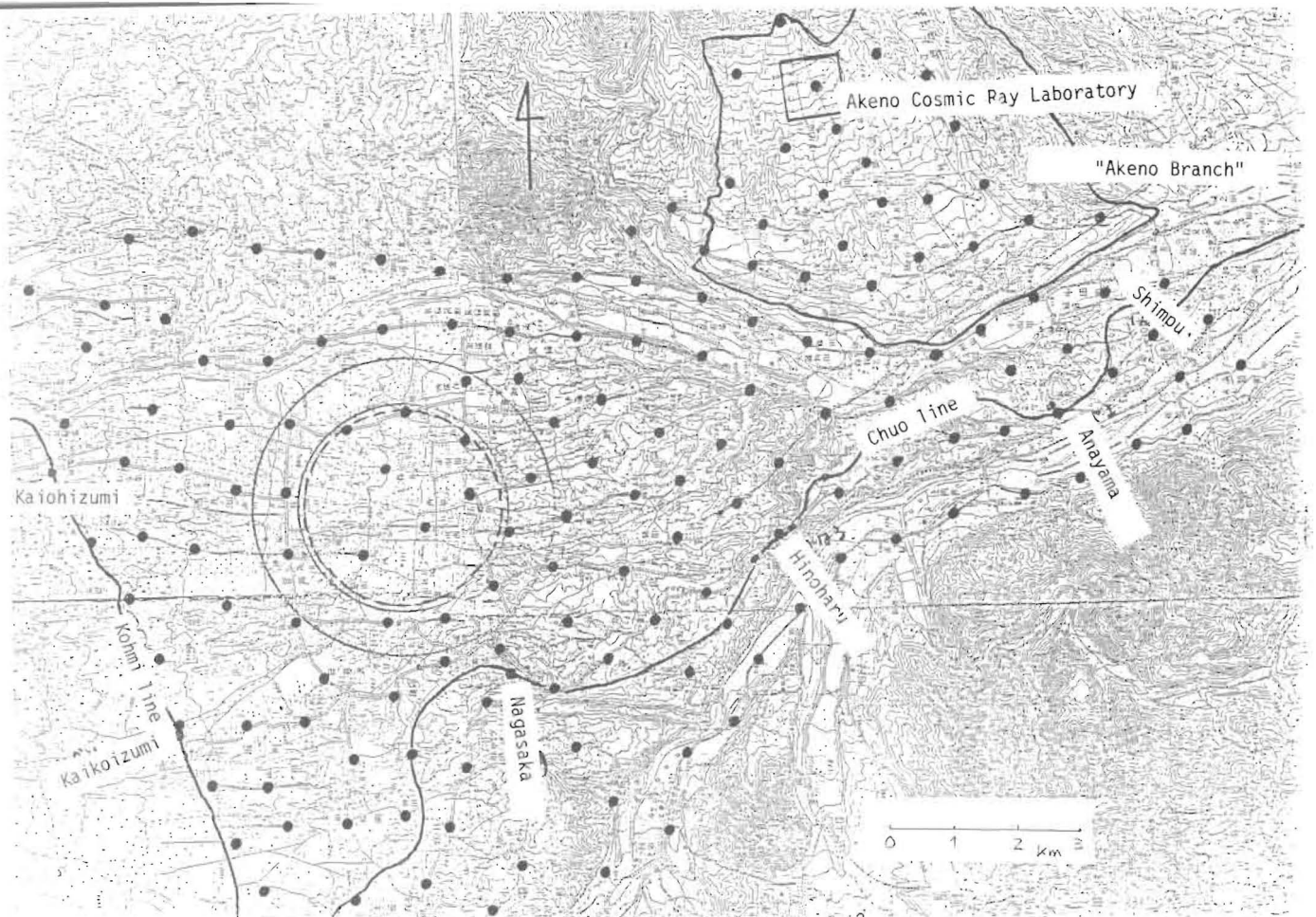
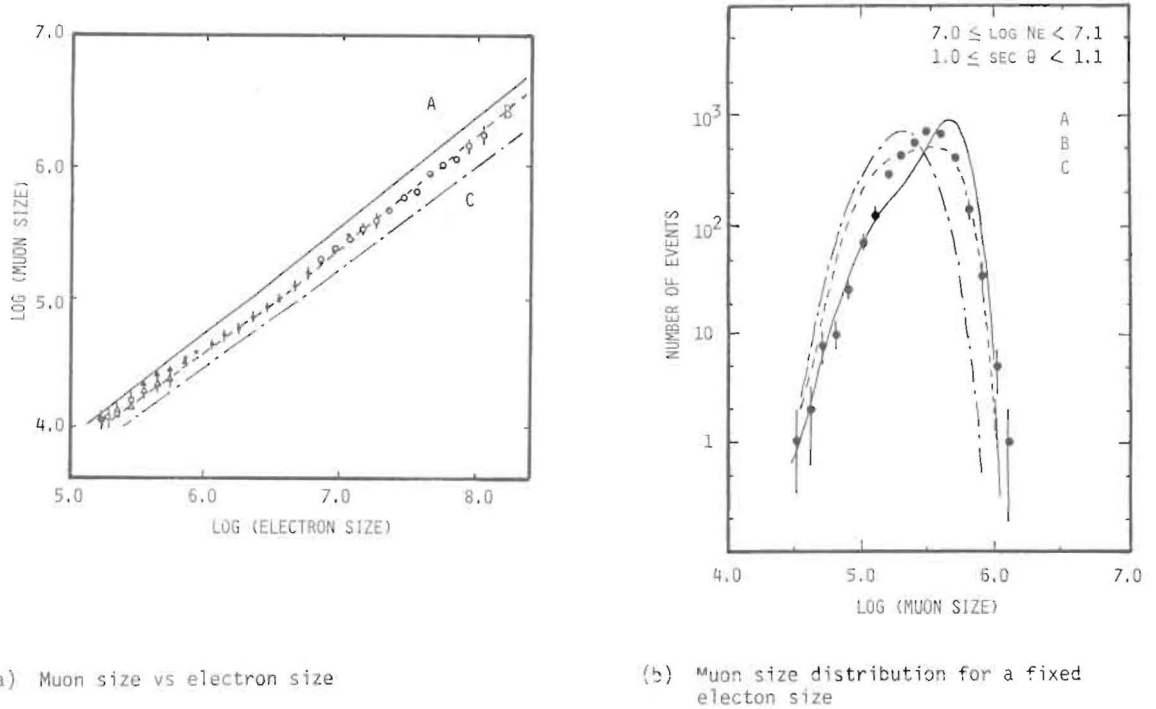


Fig.15. Plan view of the detector arrangement of the first stage of  $250 \text{ km}^2$  array. Positions of detectors are tentative and are placed just in about 1 km separation along the road. The large solid and dotted circles show the area hit by more than one electron and muon, respectively, by an EAS of  $10^{10} \text{ GeV}$ .



(a) Muon size vs electron size

(b) Muon size distribution for a fixed electron size

Fig.14 Results of Akeno experiment on electron and muon size relation are compared with the expectations from three combinations of primary composition and hadronic interaction model.

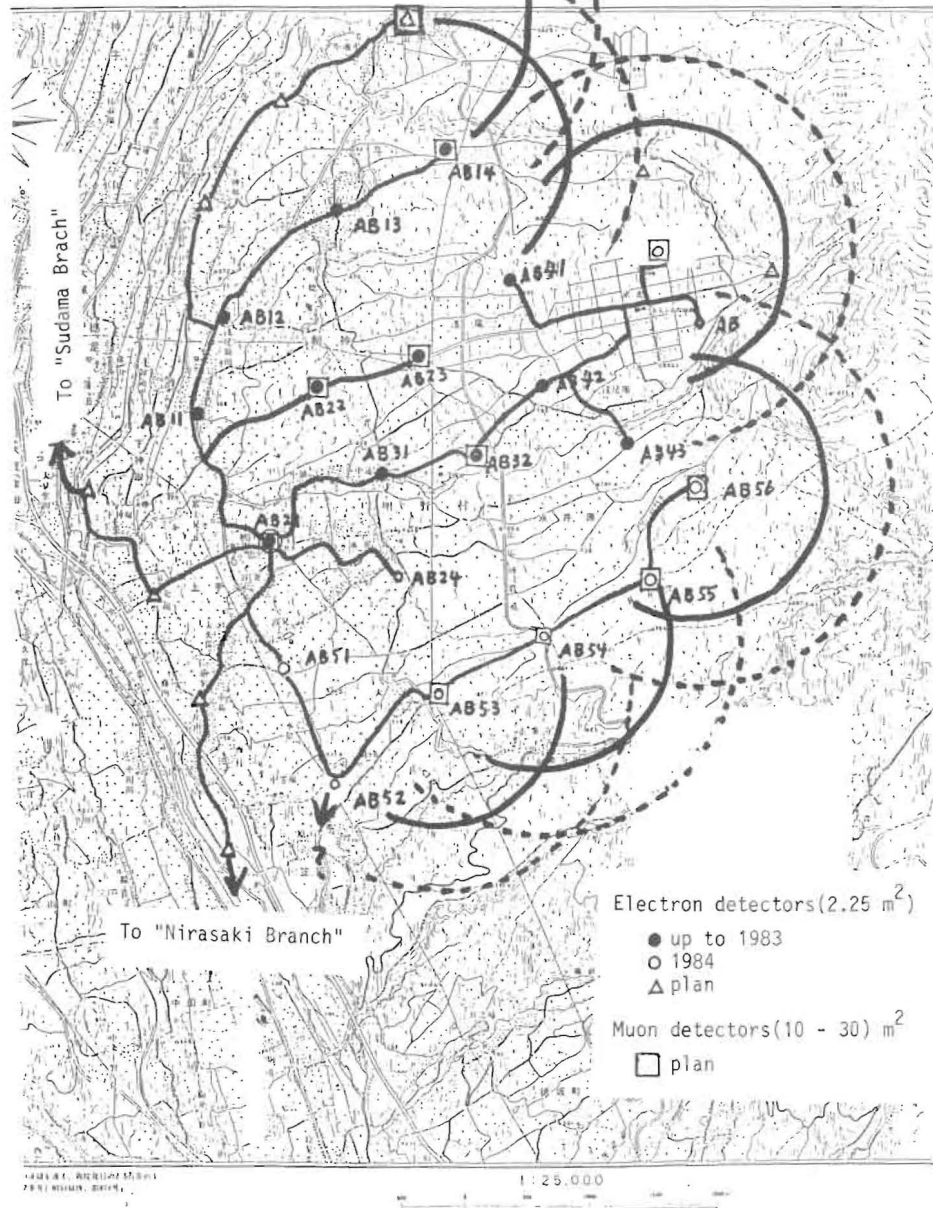


Fig.16 The detailed arrangement of "Akeno Branch". Detectors are connected with two optical fiber cables shown by solid lines. Large solid and dotted semicircles represent the area which may be covered by the measurements of arrival time distribution of electrons.

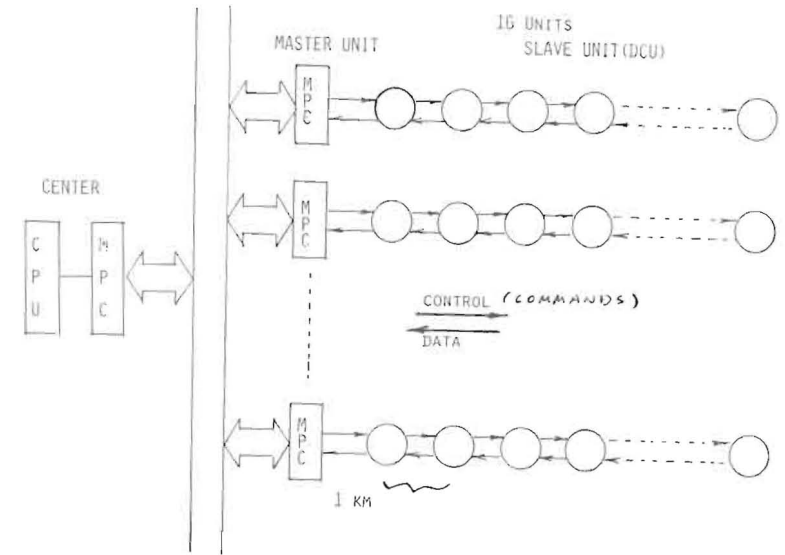


Fig.17. Data collection system in each "Branch"

POINT SOURCE OF VERY HIGH ENERGY GAMMA RAY

T. KIFUNE

Institute for Cosmic Ray Research  
University of Tokyo  
Tanashi, Tokyo, 188 Japan

ABSTRACT

Recent experiments show the positive evidences of point sources of very high energy gamma rays above  $10^{12}$  eV. In the energy region of  $10^{15-16}$  eV, the most reliable evidences are found on Cyg X-3 among the reported sources. A brief review of the experimental works on the source is introduced. The emission mechanism of very high energy gamma ray in Cyg X-3 is also discussed. Intensive fluxes of cosmic ray proton at the source can explain the observed value of the gamma ray flux. This indicates the possibility that the high energy cosmic ray particles are originated at the point source of very high energy gamma ray.

1. Introduction

Several recent observations have shown that the gamma ray spectrum from the point sources of hard gamma ray extends up to the very high energy region of  $10^{12}$  eV. In the even higher energy region larger than  $10^{15}$  eV, some positive evidences from Cyg X-3 has been reported by observing extensive air showers (EAS).

Presented in this report are a brief review of the experimental studies to detect gamma ray above  $10^{12}$  eV from Cyg X-3. The clarification of production mechanism of the gamma ray will be very important in understanding the dynamics of pulsar and its close binary system and in identifying the origin of cosmic ray particles.

2. Detection of energetic gamma ray

The gamma ray from a point source is identified essentially by detecting an excess flux from the narrow area of the sky including the source. The backgrounds from charged particles of cosmic ray are serious in the ultra high energy region, which is quite a different situation from what is in the lower energy region like X-ray. The signal to noise ratio is, therefore, determined by the angular resolution of the experimental apparatus. Another difference is in that the detection must be ground-based because of the very small value of the flux.

In the energy region of TeV, the gamma ray (together with the backgrounds of charged particle cosmic ray) is detected on the ground by observing cherenkov light emitted high in the atmosphere through the cascade showers initiated by the primary particles. The other components of the cascade except the cherenkov light are attenuated in the atmosphere and do not reach the observation level. The direct detection of the electron-photon cascade at the observation level becomes available above the primary energy as large as  $10^{15}$  eV, which is nothing but the detection of EAS.

For an example, the characteristics of the apparatus for TeV gamma ray at Dugway in Utah by Durham University group are shown as follows : The field of view is  $1.7^\circ$  FWHM and parabolic mirrors of 1.5 m diameter are used together with 5" phototubes. Twelve mirrors are set within the range of about 100 m. The mirror telescopes automatically trace the source under the control of a computer. The effective area of detection is about  $3 \times 10^4 \text{ m}^2$  and the threshold energy is about 2 TeV.

There have been several positive reports for Cyg X-3 from different groups <sup>1),2),3)</sup>, which will not be shown in this article. In their works the identification of the point source of gamma ray has been confirmed by detecting the synchronized signal with the periodical intensity variation in X-ray signal from the source. This technics of discriminating the signal is utilized in the search in the higher energy region.

### 3. Observation in the region of $10^{15}$ eV

There have been two reports, Samorsky and Stamm(1983) of Kiel University <sup>4)</sup> and Lloyd-Evans et al.(1983) of Haverah Park group <sup>5)</sup>, claiming the affirmative detection of gamma ray signals in the energy range of  $10^{15}$  eV. The detection is by observing EAS from the source direction. The signals are overnumbered by the isotropic incidences of charged particle backgrounds even in the small field of view around the source. The observed events in the field of view are plotted versus the phase of X-ray variation. A considerable portion of the total data appears concentrated in one phase bin and the period which gives the strongest concentration coincides with the one predicted by the X-ray variation.

The number of observed events is plotted as a function of the phase of X-ray variation in Fig.1 and Fig.2 in the case of Kiel and Haverah Park data, respectively. The period is about 4.8 hours and phase = 0 corresponds to the time of the minimum intensity of X-ray. The angular resolution of the apparatus is about 2 and 5 degree in each group, respectively. Fig.2a, 2c and 2d correspond to the calculation of phase using slightly different values of X-ray period. The X-ray period obtained at the most recent observation <sup>6)</sup> produces a peak around about 0.3 phase (Fig.2a) consistent with the result in Fig.1. The both data of two groups in EAS are accumulated for a long time more than three years. A slight change in 4.8 hours period results in a considerable change of phase during the course of time.

The observed flux is about  $10^{-13}$  cm<sup>-2</sup>sec<sup>-1</sup> around  $10^{15}$  eV. this value of flux gives a very flat spectrum when combined with the data in lower energies. The index of power of the integral spectrum would be even flatter than -1. The flux is given as an average over all the phase bins. The peak value of the periodical variation can be much larger and the spectrum might be much flatter.

The spectrum at the source position may be even flatter and of more intense if the distance to Cyg X-3 is as far as 10 Kpc as is taken usually.

Because the gamma ray can interact with the universal microwave background of 2.7<sup>0</sup>K to be converted into positron and electron pair. The threshold of the process is about  $10^{14}$  eV and the cross section has a maximum of about  $10^{-25}$  cm<sup>2</sup> a little above the threshold. The attenuation length is about 10 Kpc at  $10^{15}$  eV.

The intensity of the gamma ray flux from Cyg X-3 is shown in Fig.3. The logarithm of the integral spectrum is in the vertical axis in the unit of cm<sup>-2</sup>sec<sup>-1</sup> and the logarithm of energy (in eV) in the horizontal axis. The diamond mark and closed circle represent EAS data from Kiel and Haverah Park, respectively.

The data obtained at EAS Array at Akeno is now under way of analysis and are not shown in the present report. The preliminary results do not show a strong positive evidence at most, which may indicate a transient nature of gamma ray in the ultra high energy region. The data at Akeno cover the most recent time of observation. The Akeno Array has an advantage in identifying gamma ray from the backgrounds of the ordinary showers by observing muon number in EAS with a better accuracy. If this information of muon number is taken into consideration, the Akeno data give somewhat lower upper limits than the other data in EAS.

### 4. Emission mechanism

The gamma ray emission below MeV region together with the periodical features is well understood with the so called cocoon model by Milgrom and Pines <sup>7)</sup>. The model explains the sharper dependences of signal upon the orbital phase of the binary system for the higher energy of gamma ray. 4.8 hours is taken as the period of orbital motion of the close binary system of a rapid rotating pulsar and a nondegenerate companion. The system is assumed surrounded by a shell of gas of radius  $10^{12}$  cm. The hydrogen column density is taken as  $2 \times 10^{24}$  cm<sup>-2</sup> through the shell.

The high energy gamma ray in TeV region can be explained by Stepanion <sup>8)</sup> and Apparao <sup>9)</sup> with synchrotron radiation and/or inverse Compton processes by energetic electrons. Stepanion has assumed the existence of energetic electrons as high as  $10^{15}$  eV. Apparao has obtained the estimated value of number of energetic electrons larger than 150 MeV as  $6 \times 10^{46}$  at the source region within  $10^{12}$  cm. These models must assume an existence of electrons of higher energies to explain the gamma rays of  $10^{15}$  eV.

An alternative model is the production of gamma ray through nuclear interaction of proton. Vestrand and Eichler <sup>10)</sup> have calculated the gamma ray

flux through the decay of neutral pi-mesons, which are produced in the interactions between cosmic ray proton and the surrounding matter of the binary system. This model, however, leads to the spectrum of gamma ray as steep as the spectrum of the parent proton.

A flatter energy spectrum of gamma ray can be obtained with a model in which the energetic gamma rays are produced by the decay of neutral pions created in the photo-nuclear interactions of proton with the ambient number of photons in the energy region of X-ray to soft gamma ray.

The number of photons,  $n(k)$ , is described by a power law as a function of photon energy  $k$  as  $n(k)dk = a k^{-\alpha} dk$ , where  $\alpha$  is about -2 for  $k$  greater than 100 eV and tends to 0 below 100 eV. The flux of proton is also approximated by a power law  $N(E)dE = b E^{-\beta} dE$ , where  $E$  is the energy of proton and  $\beta$  is about 2.9 when observed at the earth. The probability that the nuclear interaction takes place is given by the product of cross section of pion photo-production and the 'luminosity' of photon and proton beams. Although the photon energy depends upon the pion energy and the kinematics of each decay, let us take a simple view that the decayed photons carry a half of pion energy and are emitted into  $90^\circ$  relative to the initial beam direction in the center of mass system. With this approximation and near the threshold, we can estimate the energy of final gamma ray  $K \sim E/10$  almost independent of  $k$ . The spectrum of gamma ray is given by

$$f(K)dK = \int dk dE \sigma abk^{-\alpha} E^{-\beta}$$

If we approximate the cross section as a constant, the above equation is easily integrated in  $k$  above the threshold to give

$$f(K)dK \sim dK K^{\alpha-\beta-1}$$

The index of power of gamma ray spectrum is thus given as  $\alpha-\beta-1$ , which is flatter than than proton spectrum and consistent with the observation. It is due to the fact that the number of target photons above the threshold increases with the increasing proton energy, which corresponds to the increasing energy of the final gamma ray. More careful treatments give essentially the same results.

#### 4. Discussions

By comparing the result of photo-production process with the observed

value, we can estimate the absolute intensity of cosmic ray proton at the source. The normalization of photon number,  $a$ , is determined by multiplying the flux of X-ray observed at the earth by  $(D/R_0)^2$ , where  $D$  is the distance to the source Cyg X-3 and  $R_0$  is the characteristic dimension of X-ray emission, about  $10^{12}$  cm. The calculated value of  $n(k > 1\text{KeV})$  is about  $10^{46}$ , which is as large as the one used by Apparao<sup>9)</sup> who has calculated the flux of energetic electron to explain TeV gamma ray by inverse Compton process.

With the above number of target photons the cosmic ray proton is calculated at the source as intense as  $10^{22}$  times of the flux observed at the earth, which is about  $10^{-10} \text{ cm}^{-2} \text{ str}^{-1} \text{ sec}^{-1}$ . This flux at the source is comparable with the flux of electron extrapolated from 1 GeV region to the considered high energy region, when the electron flux estimated for Crab Nebula or for Cyg X-3 is applied.

The phase of maximum flux in the lower energy region than TeV is about 0.6 of 4.8 hours period<sup>2) 3) 11)</sup>. On the other hand, the phase in  $10^{15}$  eV is around 0.3. This may suggest that the production mechanism changes from  $10^{12}$  to  $10^{15}$  eV energy region. It is quite necessary to accumulate more reliable and detailed data in EAS energy region.

It is also very interesting to know if the flat spectrum of gamma ray is affected as a dip by the interaction with  $2.7^\circ\text{K}$  radiation around  $10^{15}$  eV, as a function of distance to the various sources of very high energy gamma ray. The universality of the microwave radiation can be checked in the Galactic scale.



References

1. Neshpor et al., *Astrophysics and Space Science* 61 (1979) 349.
2. Douthwaite et al., *Astron. Astrophys.* 126 (1983) 1.
3. Weekes et al., *Astron. Astrophys.* 104 (1981) L4.
4. Samorsky and Stamm, *Astrophys. Journ.* 268 (1983) L17.
5. Lloyd-Evans et al., to be published in *Nature*.
6. Van der Klis and Bonnet-Bidault, *Astron. Astrophys.* 95 (1981) L5.
7. Milgrom and Pines, *Astrophys. Journ.* 220 (1978) 272.
8. Stepanian, *Proceeding 17th Int. Cosmic Ray Conf.* 1 (1981) 50.
9. Apparao, *Month. Not. Roy. Astron. Soc.* 179 (1977) 763.
10. Vestrand and Eichler, *Astrophys. Journ.* 261 (1982) 251.
11. Lamb et al., *Astrophys. Journ.* 212 (1977) L63.

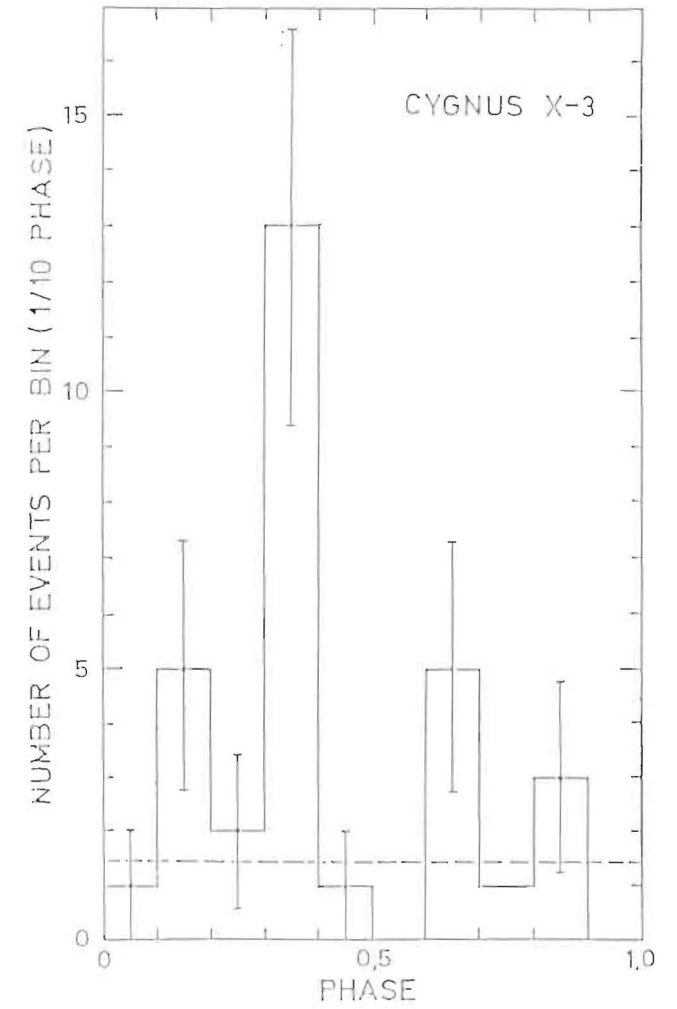


Fig. 1

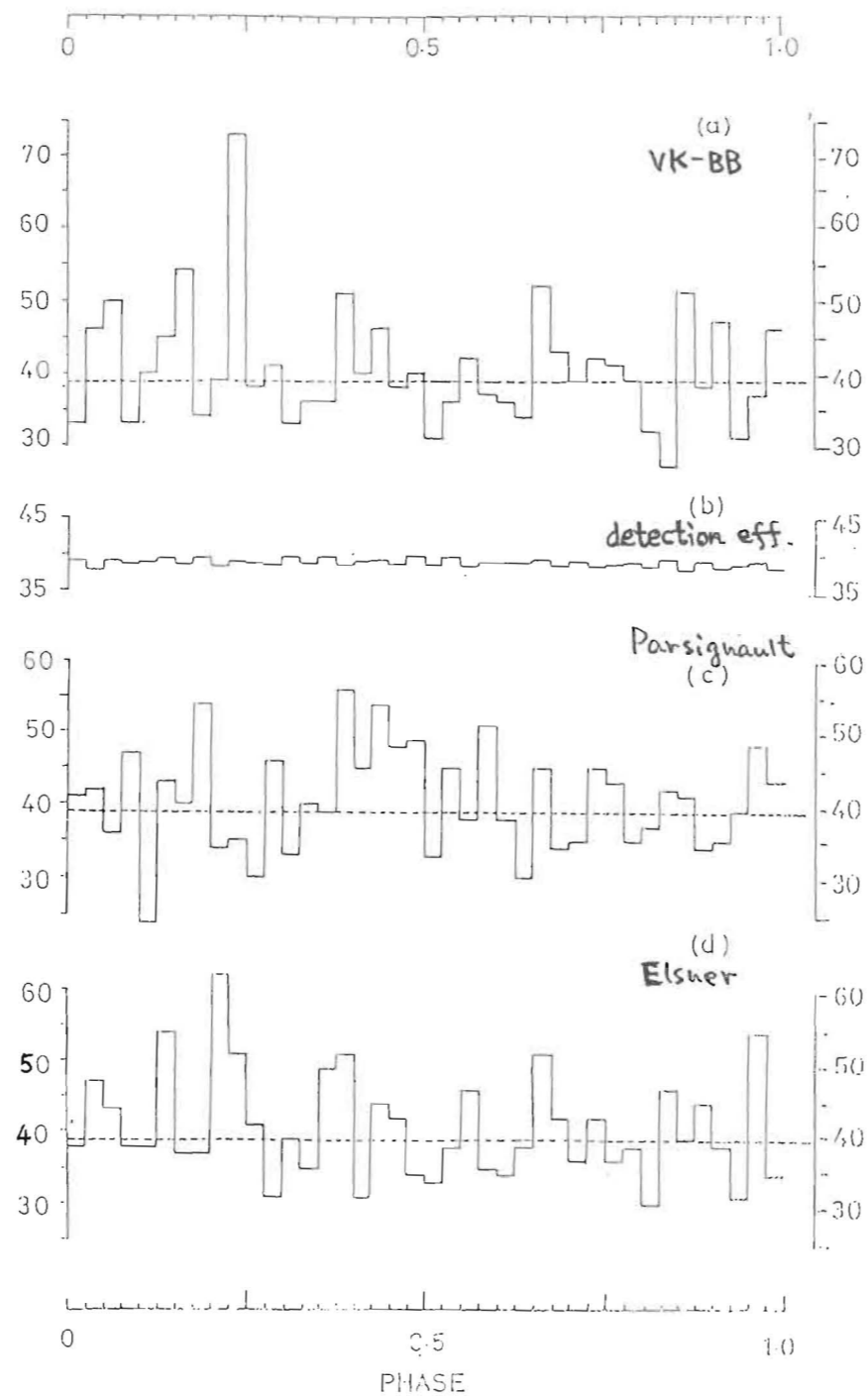


Fig. 2 - 209 -

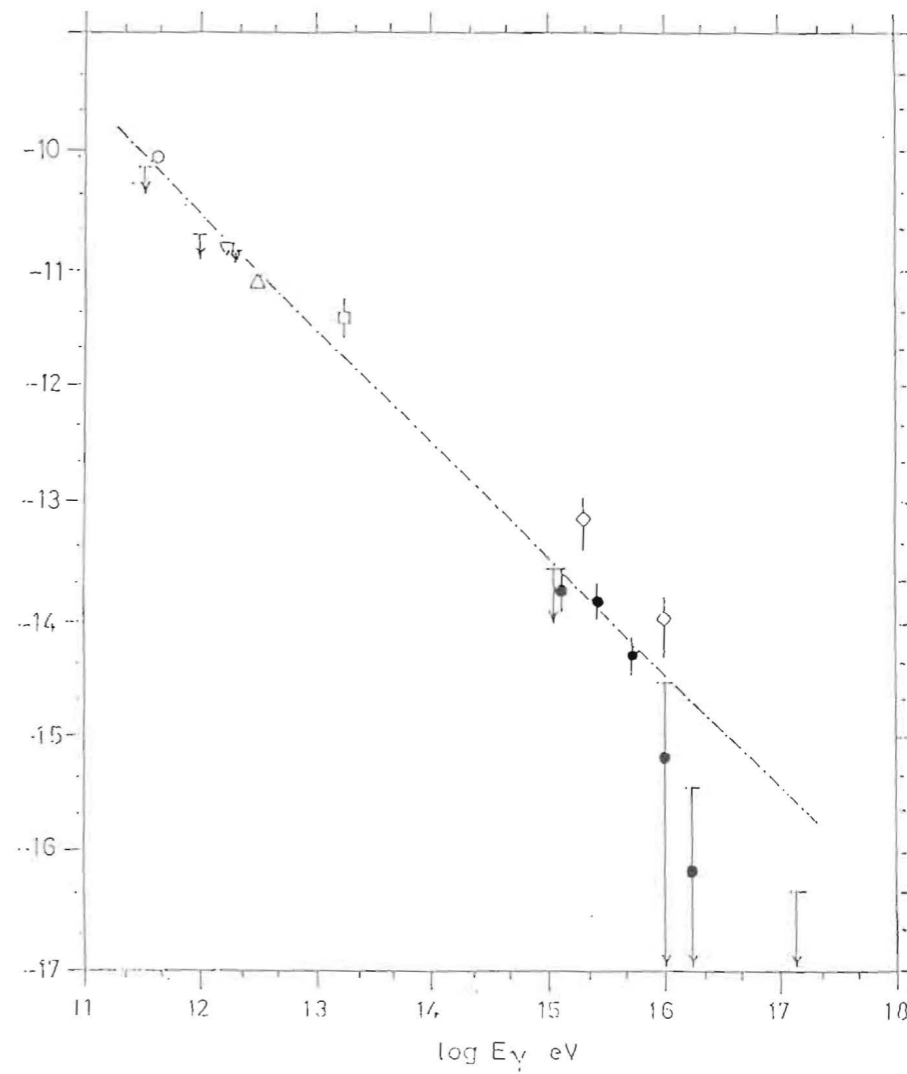


Fig. 3

A Brief Review on Neutrino Mass Measurements

T. Yasuda and T. Ohshima

Institute for Nuclear Study

University of Tokyo

Tanashi, Tokyo 188

I. Introduction

In recent years, the question of whether the neutrino has the finite mass or not has been of growing interest. The precise measurement of it may give a clue to construct a new understanding of the elementary particle physics beyond the standard model of the electroweak and strong interactions. In the schemes such as the grand unification in which the left-right symmetry is fundamentally assumed, the neutrino necessarily acquires the mass. Also the missing mass problem of the universe in the astrophysics may be resolved satisfactorily, if the neutrino has a mass of a few tens of eV's.

Since the success of the  $\beta$ -decay theory originated from E. Fermi based on the neutrino hypothesis proposed by W. Pauli, many attempts have been made to measure the neutrino mass experimentally. By far the most accurate and unambiguous information on the electron anti-neutrino mass has been obtained from the measurement of the phase space modification due to the non-zero mass near the end-point of the energy spectrum of  $^3\text{H}$   $\beta$ -decay as shown in Fig. 1. The reason tritium is used as the  $\beta$ -decay object is due to its following properties:

- 1) The small Q-value of  $\sim 18.6$  keV
- 2) The moderate life time of  $\sim 12$  years
- 3) The simple structure of nuclei.

Thus as listed in Table 1, tritium has been used in most of the experiments, except for that of Beck<sup>4)</sup> in which  $^{22}\text{Na}$  was used as the  $\beta^+$ -decay source and which therefore measured the electron neutrino mass, not the electron anti-neutrino mass. ITEP group presented their results in 1980<sup>11)</sup> and 1983<sup>12)</sup> (hereafter we refer to them as ITEP-80 and ITEP-83, respectively.) reporting that the electron anti-neutrino has a non-zero mass of  $> 20$  eV. This announcement became the focus of the world's attention. Since then many experiments were started

at various laboratories in the world trying either to verify or disprove the results given by ITEP-80 and ITEP-83.

We presented at various opportunities the brief experimental reviews on the former neutrino mass measurements including those on the muon- and tau-neutrinos, and also our proposal on the electron anti-neutrino mass measurement at INS. Today, we will try to review the currently proceeding experiments that are aiming to pin down the electron anti-neutrino mass (which is hereafter referred to simply as the neutrino mass  $m_\nu$ ) with the accuracy of  $< 20$  eV using the traditional way of  $\beta$ -spectrum measurement for  ${}^3\text{H}$  decay, and will also report our present preparatory status for the mass measurement.

## II. Background and Mass Sensitivity

There are three experimental key-points essential for achieving high mass-sensitivity. They are:

- 1) High statistics
- 2) High energy resolution
- 3) Small background contribution

All of the experiments listed in Table 1 were made in an effort to improve the above former two items which can be easily understood by looking at Fig. 1. In the following, we would like to show how the background suppression is as important as the former two.

Firstly, let us see the mass sensitivity ( $\Delta m_\nu$ ) from the viewpoint of the integrated S/N ratio. Let us define  $\Delta R$  as:

$$R = 1 - \frac{\int_{\Delta E} n(E) |_{m_\nu \neq 0} dE}{\int_{\Delta E} n(E) |_{m_\nu = 0} dE} = \frac{3}{2} \left( \frac{m_\nu}{\Delta E} \right)^2 \quad (1)$$

Where  $n(E)$  is the  $\beta$ -spectrum and the integration is carried out over the energy interval of  $\Delta E$  up to the end-point

energy ( $E^{\text{max}}$ ). As  $\Delta R$  is a difference of the integrated  $\beta$ -spectra in the cases of  $m_\nu \neq 0$  and  $m_\nu = 0$ , the upper mass limit we could set is obtained by estimating how accurately we can measure  $\Delta R$ . Here,  $\alpha$  (an inverse of the integrated S/N ratio) is defined as:

$$\alpha = \frac{N_{BG}}{N_\beta}$$

$$N_{BG} = \int_{\Delta E} n_{BG}(E) dE$$

$$N_\beta = \int_{\Delta E} n_\beta(E) dE$$

where  $n_{BG}(E)$  and  $n_\beta(E)$  are the background and  $\beta$ -spectra, respectively. By solving Eq. (1),  $\Delta m_\nu$  can be expressed as:

$$\Delta m_\nu = \sqrt{\frac{2n}{3}} (1+2\alpha)^{1/4} \Delta E / N_\beta^{1/4} \quad (2)$$

where  $n$  is the number of standard deviations of  $\Delta R$  from the zero. Fig. 2 shows the relations between  $N_\beta$  (total number of events) and  $\alpha$  to achieve the upper limit value of  $\Delta m_\nu$  with  $n=3$  standard deviation in the case of  $\Delta E = 100$  eV. It has to be kept in mind that in the above calculation the energy resolution, which relates proportionally to  $\Delta m_\nu$ , is not taken into account.

Secondly, let us see the mass sensitivity  $\Delta m_\nu$  from the viewpoint of the differential S/N ratio. This will show the reason why the error bars of Kurie spectrum in ITEP-83 is  $\sim 10$  times smaller than those in ITEP-80 while the former has  $\sim 2$  times higher accumulated events compared to the latter. Near the end-point region, as the momentum ( $p$ ) of the  $\beta$ -particle and Fermi function ( $F$ ) have almost constant numbers, the behaviour of error bars of Kurie spectrum ( $\sqrt{n(E)} / p^2 F$ ) can be estimated by looking at those of  $\sqrt{n(E)}$ . It can be written as:

$$\begin{aligned} \sqrt{\frac{1}{n(E)}} &= \frac{1}{2} \left[ 1 + 2 \left( \frac{n_{BG}(E)}{n_{\beta}(E)} \right) \right]^{1/2} \\ &= \frac{1}{2} \left[ 1 + \left( \frac{2}{3} \right) \left( \frac{\alpha}{\xi^2} \right) \right]^{1/2} \end{aligned} \quad (3)$$

where  $\xi = (E^{\max} - E) / \Delta E$ . Fig 3(a) shows the behaviours of  $\sqrt{1/n(E)}$  in the case of  $\Delta E = 100$  eV. With smaller  $\alpha$  value, the error reaches to a certain value, of 0.5 in this case, and the error rapidly increases near the end-point. When the mass sensitivity is defined to be the energy of  $E^{\max} - E^*$  at  $n_{BG}(E^*) = n_{\beta}(E^*)$ , the sensitivity varies with the  $\alpha$  as shown in Fig. 3(b).

From the above considerations, we can set up a standard for the three key-points, for instance, to achieve  $\Delta m_{\nu} = 5$  eV:

$$\begin{aligned} N_{\beta} &= 10^4 - 10^5 \\ \alpha &\leq 0.01 \end{aligned} \quad (4)$$

and the energy resolution of 5 eV.

### III. Present and Future Experiments

Let us briefly review the experiments, which will have the results regarding the neutrino mass in the near future, including ITEP-83. Table 2 is a list of the experiments originally presented by Shaevitz<sup>13)</sup> at 1983 Cornell Conference and was changed and added a few items by us.

Experiments 1) - 6) are the traditional types of measurement that employ the focusing type of magnetic spectrometer as the momentum analyzer for the  $\beta$ -particle, while 8) - 11) are the new types of experiments using the atomic  $^3\text{H}$  or the solid molecular  $^3\text{H}$  sources.

#### ITEP-83 (Boris et al.)

Details of ITEP-80 and ITEP-83 can be seen in ref. 12, 13). A multi-loop iron-free toroidal magnetic spectrometer with a rotation angle of  $4\pi$  was used as seen in Fig 4. This  $4\pi$  rotation of the electron trajectory might reduce substantially the background contribution that originates from the  $^3\text{H}$  labelled

molecules evaporated from the source. The source was the  $^3\text{H}$  labelled valine ( $\text{C}_5\text{H}_{11}\text{NO}_2$ ) compound with the specific activity of 58 Ci/mmol and the thickness of  $\sim 2\mu\text{g}/\text{cm}^2$ . The  $\beta$ -detector was the proportional counter with several anode wires from which the pulse height was also measured to discriminate the signal from the background. One of their terrific works is to study the spectrometer line shape by using the standard  $^{169}\text{Yb}$  source mounted between the working  $^3\text{H}$  sources. Unfortunately, in spite of their careful study ITEP-80 did not take the intrinsic line width of  $^{169}\text{Yb}$  into account at the deduction of the spectrometer line shape.

ITEP-83 reported a new result for the neutrino mass to be  $> 20$  eV. They had improved the following aspects in this new measurement compared with those in ITEP-80:

- 1) Energy resolution was improved to 20 eV compared to 45 eV at ITEP-80.
- 2) Background was reduced by a factor of  $\sim 12$  compared with ITEP-80, which resulted in the mass sensitivity of  $\sim 20$  eV from the viewpoint of the differential S/N ratio discussed above.
- 3) Line shape was studied considering the intrinsic line width of  $^{169}\text{Yb}$ .
- 4) Spectrum was measured over wider energy range of  $\sim 1.8$  keV compared with 0.7 keV at ITEP-80.
- 5) Higher statistics.

ITEP-83 employed the electrostatic scanning method keeping the magnetic field at a constant value. With varying applied electric potential,  $\beta$ -particle was accelerated or decelerated so that the  $\beta$ -particle detected could always take the central orbit of the spectrometer and the same energy. By this method, the detection efficiency of the  $\beta$ -detector could not cause any bias on the spectrum. Also by setting the above accelerated  $\beta$ -particle energy at 22 keV and using pulse height discrimination by the detector, their efficient background suppression could be attained.

The data analysis in ITEP-83 was made for several final states: Valine molecule,  $^3\text{H}$  molecule,  $^3\text{H}$  atom, and  $^3\text{H}$  nucleus. They obtained the neutrino mass  $m_\nu = 33.0 \pm 1.1$  eV as the weighted mean value for the results of  $\chi^2$  fit and the end-point energy  $E^{\text{max}} = 18575$  eV at a 95% confidence level. They concluded that it was possible to set a model independent lower limit for the neutrino mass:

$$m_\nu \geq 20 \text{ eV}$$

Stockholm (Bergkvist)<sup>7)</sup>

Fig. 5 shows the  $\pi\sqrt{2}$  spectrometer used in 1972 by Bergkvist's experiment which was the pioneering work in this field. Many ideas for most of other proceeding experiments are based on his magnificent works. It is heard that he will use  $^3\text{H}$  labelled valine compound as same as that used by ITEP. Chalk River (Graham et al.)

Their  $\pi\sqrt{2}$  air-core magnetic spectrometer with the central orbital radius of 100 cm is the largest  $\beta$ -spectrometer in the world. Thus the best instrumental energy resolution is expected among experiments 1) - 6). Tritiated titanium is being manufactured as the working source.

Zürich (Kündig et al.)<sup>15)</sup>

They are in the process of constructing a  $\beta$ -spectrometer using a magnetic spectrometer similar to that of ITEP and a retarding electrostatic field surrounding the source. In principle, by retarding the  $\beta$ -particle energy with the electric potential opposite to ITEP method, a large improvement on the energy resolution could be attained. For instance, if  $\sim +18$  kV is applied on a source,  $\beta$ -particle emitted from the source has the energy of  $E - 18$  keV. While the intrinsic energy resolution  $\Delta E/E$  of the magnetic spectrometer is not affected by applying the electric potential. Then, absolute energy resolution  $\Delta E$  is improved by a factor of  $\sim E/(E-18 \text{ keV})$ . They accelerate  $\beta$ -particle once again up  $\sim 20$  keV near the focal plane of the spectrometer so that it has enough energy to penetrate the detector window.

Peking (Sun Hun-Cheg et al.)<sup>16)</sup>

$\pi\sqrt{2}$  iron  $\beta$ -spectrometer with the central radius of 40 cm is used in combination with a single wire proportional counter on the focal point. Their source is a  $^3\text{H}$  labelled Oestriadol compound. They measured the intrinsic spectrometer resolution of 20 eV with a standard source, and then began the first measurement of  $^3\text{H}$   $\beta$ -spectrum with 1 mCi/cm<sup>2</sup> source.

Institute for Nuclear Study (INS)

See next section.

U.C. Berkeley (Heller et al.)

We guess that the principle of this measurement may be the same with that of the experiment by J.J. Simpson.<sup>10)</sup> While this method is free from the atomic interplay problem by detecting sum of  $\beta$ -particle energy and X-ray energy released at the transition to the ground state from the excited one of  $^3\text{He}$ , the achievable energy resolution can not be high enough.

Rockefeller, Fermilab & Lll (Fackler et al.)

We know only a scheme of their experiment shown in Fig. 6.<sup>13)</sup> It seems that they use the pure frozen  $^3\text{H}_2$  with very high activity so that the complex atomic interplay problem arose from the  $^3\text{H}$  labelled compound can be simplified. This scheme uses the electric field for selecting and focusing the  $\beta$ -particle with very high accuracy.

O.S.U. (Boyd et al.)

We have no information.

LAMPF (Bowles et al.)

As seen in Fig. 7,<sup>17)</sup>  $^3\text{H}_2$  will be dissociated to  $^3\text{H}$  atom by shaking with microwave so that the atomic interplay problem becomes the simplest. The  $\beta$ -particle may be trapped by the magnetic field of a superconducting magnet so as to enhance the transmission to the  $\beta$ -spectrometer. Also,  $\beta$ -particle produced at the decay region inside of the magnet may be accelerated to transport to the spectrometer so that the

background will be suppressed by the same way with ITEP-83.  
IBM (Clark et al.)

We have no information.

IV. INS Experiment (Collaboration of INS/Univ. of Tokyo/  
Tokyo Metro. Univ./Tokyo Inst. of Tech./Tohoku Univ.)

The goal of our experiment is to measure the neutrino  
mass with the sensitivity of 5 - 7 eV.

INS double focusing  $\pi\sqrt{2}$  air-core  $\beta$ -ray spectrometer with  
the central radius of 75 cm is one of our big weapon. This  
spectrometer has the following properties: i) dispersion = 4,  
ii) field stability =  $5 \times 10^{-5}$ /day, iii) measurable energy  
range = 0.2 keV - 4 MeV and iv) maximum momentum resolution  
so far obtained = 0.013% for  $^{137}\text{Cs}$  (662 keV). Also, it has  
a wide momentum acceptance of 4% so that by setting a position  
sensitive  $\beta$ -detector on the focal plane,  $\beta$ -spectrum in the  
vicinity of the end-point can be measured at once over the  
energy interval of  $\sim 1.5$  keV. This method is called as the  
spectrographic detection whose scheme is shown in Fig. 8(a).  
In comparison with the magnetic or electrostatic scanning  
method, this results in shorter experimental period or higher  
statistics by a factor of  $\sim 100$  or more. Realization of the  
above approach depends only on the successful development  
of a position sensitive  $\beta$ -detector with high position  
resolution. In addition to the spectrographic detection,  
a widely extended source is used to have high luminosity  
with an application of the electric potential on it to  
compensate the divergence of the  $\beta$ -particle on the focal  
plane. Acceleration or deceleration of the  $\beta$ -particles with  
the electric potential depending on the source position can  
focus them on the focal plane as shown in Fig. 8(b). This  
idea was introduced by Bergkvist<sup>7)</sup> and is called as a non-  
equipotential method. This approach makes the luminosity  
higher by a factor of a few hundreds compared to the ordinary  
measurement with a single trip-source of a few tenth of mm  
in width. Wide resistive membranes are prepared as the source  
backing material. As mentioned earlier, the background

suppression is substantially important to achieve the high  
mass-sensitivity. This background arises mainly from  $^3\text{H}$   
contamination on the spectrometer wall, according to ITEP  
experience. The way of suppression in many experiments is  
based on the different properties between the signal and the  
background as shown in the following:

- 1) The background from the wall takes different  
trajectory from that of the signal so that some of them can be  
rejected by putting masks along its long flight path.
- 2) The discrimination on the pulse height distribution  
in the  $\beta$ -detector also rejects the background. Especially,  
the acceleration scheme as used in ITEP-83 makes the  
difference between the distributions of the signal and of the  
background larger and results in a large suppression  
together with the effect of 1).

Our approach for the suppression is different from the  
above. Our principle is to suppress the evaporation rate of  
 $^3\text{H}$  labelled compound from the source in vacuum by cooling  
down to  $\sim -100$  °C. At this temperature, the evaporation rate  
could be reduced by an order of  $10^{-6}$  -  $10^{-7}$  compared to that  
at the room temperature.

Today, we would like to report our preparatory works  
on two items: 1) measurement of the evaporation rate of  $^3\text{H}$   
source and ii) test of the non-equipotential method.

#### Evaporation rate of $^3\text{H}$ source

$^3\text{H}$  labelled compound with the specific activity of  
 $\sim 200$  Ci/mmol and with the thickness of a few molecular layers  
is mounted on the resistive plate. These several sources  
were set in a vacuum chamber shown in Fig. 9 which was  
evacuated to the pressure of  $\sim 10^{-6}$  torr.  $^3\text{H}$  sources and  
a calibration source of  $^{14}\text{C}$  were mounted on a rotatory table  
and the yield of the  $\beta$ -particle from the source was continuously  
monitored by a scintillation counter. Inner wall of the chamber  
was surrounded by Al foil on which some part of the evaporated

$^3\text{H}$  compound was attached. Also, the activity of the evacuated gas was monitored. With this set-up, the following observations were obtained:

1) At the room temperature, the accumulated activity on the Al foil for  $\sim 1$  day was  $\sim 1\%$  of the source activity.

2) At the temperature of  $\sim 100^\circ\text{C}$ , the  $\beta$ -yield monitored over 10 days is as shown in Fig. 10. Data does not show any decreasing trend related with time within the detection accuracy of  $\sim \pm 2\%$ .

3) After 10 days of measurement, the accumulated activity on the Al foil was measured. Activity integrated over the surface area of the chamber was  $\sim 10^{-4}$  of the source activity (3 mCi).

4) Same measurements with 2) and 3), but done with various time durations, and same results with 2) and 3) were obtained as shown in Fig. 11.

5) Exhausted gas from pumping system was accumulated, and their activities were measured. For  $\sim 10$  days accumulation,  $\sim 0.2 \times 10^{-4}$  of the source activity was measured.

Measurement of 2), 3), 4) and Fig. 11 indicates that the evaporation rate and the accumulated activity on the Al foil do not show any dependence on time. In our test system, it takes 10 - 20 minutes for cooling down the source to 10 - 20  $^\circ\text{C}$  below room temperature after evacuating the chamber. The  $^3\text{H}$  compound continuously evaporated during this period. This rate was estimated to be  $\sim 10^{-4}$  from 1). Then, it can be reasonably concluded that the measured activity accumulated for 10 days is dominated by the evaporation in the above-mentioned cooling process. Therefore, the accumulated activity at  $-100^\circ\text{C}$  might be much smaller than  $10^{-4}$  and the ratio of the evaporation rates between  $-100^\circ\text{C}$  and room temperature might be  $\ll 10^{-3}$ .

The  $\alpha$ -value discussed earlier, an indicator of the background contribution, will be, with a rough and not very reliable estimate, smaller than 0.002 in the case that the

evaporation rate is less than  $10^{-5}$  of the source activity.

#### Test of the non-equipotential method

For this test, 62.5 keV  $\beta$ -particle (88.0 keV - K line) from  $^{109}\text{Cd}$  was used as a mono-energetic source. Five  $^{109}\text{Cd}$  sources with the area of  $0.5 \times 20 \text{ mm}^2$  were made by electrically depositing on thin Ag foils. The momentum resolution of each source was measured by the magnetic field scanning method and was obtained to be 0.03 - 0.04%. These five sources were placed on a plane, as illustrated in Fig. 8(b), with distance of 46 mm between the sources and were surrounded by the electrodes for applying non-equipotential. A position sensitive single-wire proportional counter<sup>18)</sup> was used as a  $\beta$ -detector which had the position resolution of  $\sim 1.3 \text{ mm}$  in this energy region. Thus, at this time, the spectrographic detection made the overall momentum resolution of 0.06% for 62.5 keV  $\beta$ -particle.

Firstly,  $\beta$ -spectrum was measured with no electric potential applied. Fig. 12(a) shows the measured spectrum, and dotted histogram in Fig. 13 is a numerical sum over the five measured individual spectra by lining up the peak positions. Secondly, the optimum electric potential was applied on the sources to unify five peaks. Fig. 12(b) and a solid histogram in Fig. 13 show the resultant spectra.

By comparing spectra with and without the electric potential in Fig. 13, we were convinced that the non-equipotential method worked very well without any deterioration of the resolution over the source spread of 180 mm.

Further test with higher resolution is in progress.

#### V. Summary

Since ITEP-80, many experiments with new interesting ideas for the measurement of the electron (anti-) neutrino mass have been and are currently proceeding in the world. And their results are eagerly awaited in the field of the elementary particle physics, astrophysics and cosmology to confirm that the neutrino has the mass of a few tens of eV as ITEP reported.



Fig. 14 is an expected Kurie spectrum obtained by a stimulation with the condition of:

$$N_{\beta} = 10^4 \\ = 0.01$$

$\Delta E$  (energy resolution) = 5 eV  
which our experiment is aiming to achieve.

#### References

- (1) L.M.Langer and R.J.D.Moffat, Phys. Rev. 88(1952)689.
- (2) D.R.Hamilton, W.P.Alford and L.Gross, Phys. Rev. 92(1953)1521.
- (3) L.Friedman and L.Smith, Phys. Rev. 115(1959)2214.
- (4) E.Beck and H.Daniel, Zeit. Phys. 216(1968)229.
- (5) R.C.Salgo and H.H.Staub, Nucl. Phys. A138(1969)417.
- (6) R.Daris and C.St-Pierre, Nucl. Phys. A138(1969)545.
- (7) K.E.Bergkvist, Nucl. Phys. B39(1972)317.
- (8) B.Rode and H.Daniel, Lett. Nuov. Cim. 5(1972)139.
- (9) E.T.Tretyakov et al., Proceedings of the International Neutrino Conference, Aachen 1976.
- (10) J.J.Simpson, Proceedings of the International Neutrino Conference 1980 : Phys. Rev. D23(1981)649.
- (11) V.A.Lyubimov et al., Phys. Lett. 94B(1980)266.
- (12) S.Boris et al., Talk presented at the International Europhysics Conference on High Energy Physics, Brighton 1983.
- (13) Shaevitz, Talk presented at the International Symposium on Lepton and Photon Interactions at High Energies, Cornell University 1983.
- (14) PR-P-134, -136 and -137, Progress Report Physics Division, Chalk River Nuclear Laboratories.
- (15) W.Kündig et al., SIN Annual Report.
- (16) M.Fujioka, private communication.
- (17) Proceedings of the Neutrino Mass Mini-conference and Workshop, Telemark, Wisconsin, Oct. 2-4, 1980.
- (18) Y.Fujita, H.Kawakami and M.Hosoda, Nucl. Inst. & Meth. 196(1982)271.

Table Captions

Table 1 : A list of the electron anti-neutrino mass measurements so far performed.

Table 2 : A list of the currently proceeding experiments. This table was originally presented by Shaevitz<sup>13</sup>).

Figure Captions

Fig.1 : Kurie plot for the  ${}^3\text{H}$   $\beta$ -decay.  $E^{\text{max}}=18.6$  keV is applied.

Fig.2 : Relation between  $N_{\beta}$  and  $\alpha$  for setting the upper mass limit of  $\Delta m_{\nu}$  with 3 standard deviation in the case of  $m_{\nu} \approx 0$ .

Fig.3(a): Behaviour of error bars in the Kurie spectrum.

Fig.3(b): Mass sensitivity from view-point of differential S/N ratio.

Fig.4 : A sectional plan of ITEP iron-free toroidal spectrometer.

Fig.5 : A schema of the Bergkvist's spectrometer.

Fig.6 : A conceptual figure of Rockefeller, Fermilab & Lll experiment cited from ref.13.

Fig.7 : A conception of LAMPF experiment cited from ref.17.

Fig.8(a): A schema of the spectrographic detection.

Fig.8(b): A schema of the non-equipotential method.

Fig.9 : The detection system used for the measurement of the evaporation rate of  ${}^3\text{H}$  source.

Fig.10 : The source activities monitored continuously by a scintillation counter for 10 days. 4 sources whose thickness correspond to 2, 4, 6 and 8 molecular layers, were set under the condition of  $10^{-6}$  torr and  $-100$  °C.

Fig.11 : Accumulated activities on the Al foil for various time durations of the measurement.

Fig.12(a) : The spectrographically measured spectra for 5  ${}^{109}\text{Cd}$  sources without the electric potential. One horizontal channel corresponds to  $\Delta p/p = 0.0025\%$ .

Fig.12(b) : The spectra obtained by the optimum potential applied.

Fig.13 : Dotted histogram is the spectrum summed 5 individual spectra by lining up the peak positions. The spectrum in Fig.12(b) is also plotted with solid line, for comparison.

Fig.14 : Expected Kurie spectrum at INS experiment, obtained by a simulation under the condition of  $N_{\beta} = 10^4$ ,  $\alpha = 0.01$  and the energy resolution  $\approx 5$  eV. Horizontal axis shows  $\Delta E - E^{\text{max}} - (\beta\text{-particle energy})$ .

MEASUREMENTS OF  $\bar{\nu}_e$  MASS

UPPER LIMIT ON MASS ( $eV/c^2$ )	CONFIDENCE (%)	REFERENCE
< 250		LANGER 1952(1)
< 500		HAMILTON 1953(2)
< 550		FRIEDMAN 1958(3)
< 4100	67	BECK 1968(4)*
< 200	90	SALGO 1969(5)
< 75	90	DARIS 1969(6)
< 60	90	BERGKVIST 1972(7)
< 86	90	RÖDE 1972(8)
< 35	90	TRETYAKOV 1976(9)
< 70	95	SIMPSON 1979(10)
$14 \leq m_{\nu} \leq 46$	99	LYUBIMOV 1980(11)
> 20	99	BORIS 1983(12)

\*  $\nu_e$  MASS MEASUREMENT USING  $^{22}\text{Na}$ 

Table 1

Table 2

EXPERIMENT	SOURCE	RESOLUTION	SENSITIVITY
1) ITEP-83 (Boris et al.)	$^3\text{H}$ in valine	20 eV	20 eV
2) Stockholm (Bergkvist)	$^3\text{H}$ in valine	25 eV	> 19 eV
3) Chalk River (Graham et al.)	Tritiated titanium	10 eV	$\approx$ 20 eV
4) Zurich (Kundig et al.)		5 eV	> 10 eV
5) Peking (Sun Hun-Cheng et al.)	$^3\text{H}$ in Oestradiol	20 eV	
6) INS	$^3\text{H}$ labelled compound	5 eV	5 - 7 eV
7) U.C. Berkeley (Heller et al.)	$^3\text{H}$ in semiconductor	100 eV	> 30 eV
8) Rock-FNAL-LLL (Fackler et al.)	Solid molecular $^3\text{H}$	1-2 eV	> 4 eV
9) O.S.U. (Boyd et al.)		10 eV	> 10 eV
10) LAMPF (Bowles et al.)	Atomic $^3\text{H}$	40 eV	> 10 eV
11) IBM (Clark et al.)	Solid $^3\text{H}$	5 eV	

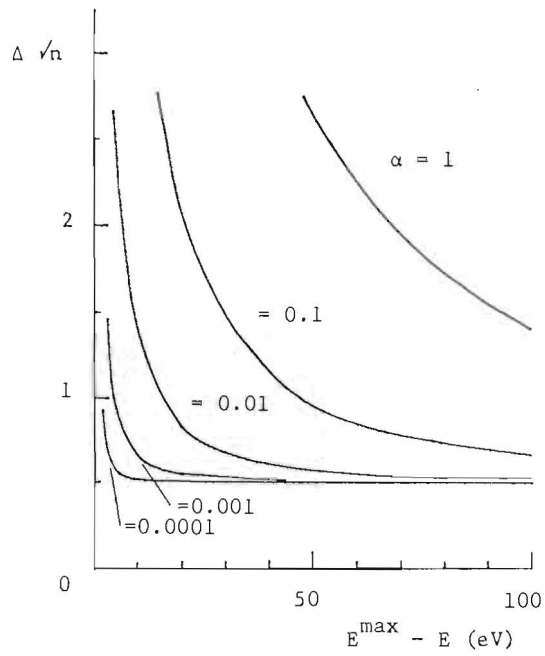


Fig. 3 (a)

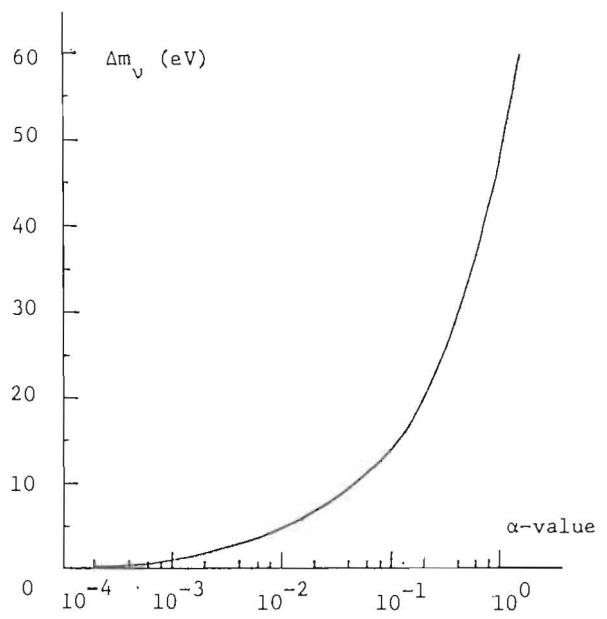


Fig. 3 (b)

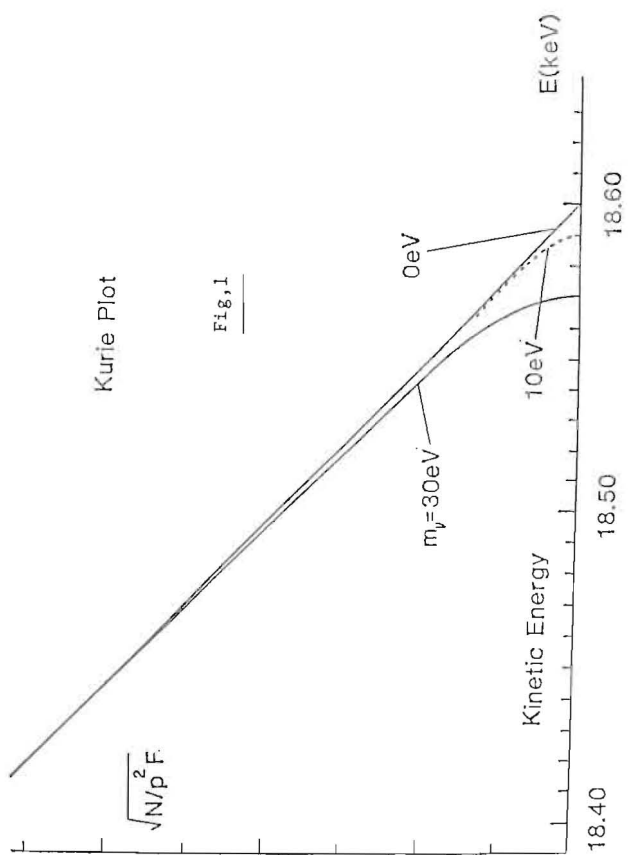


Fig. 1

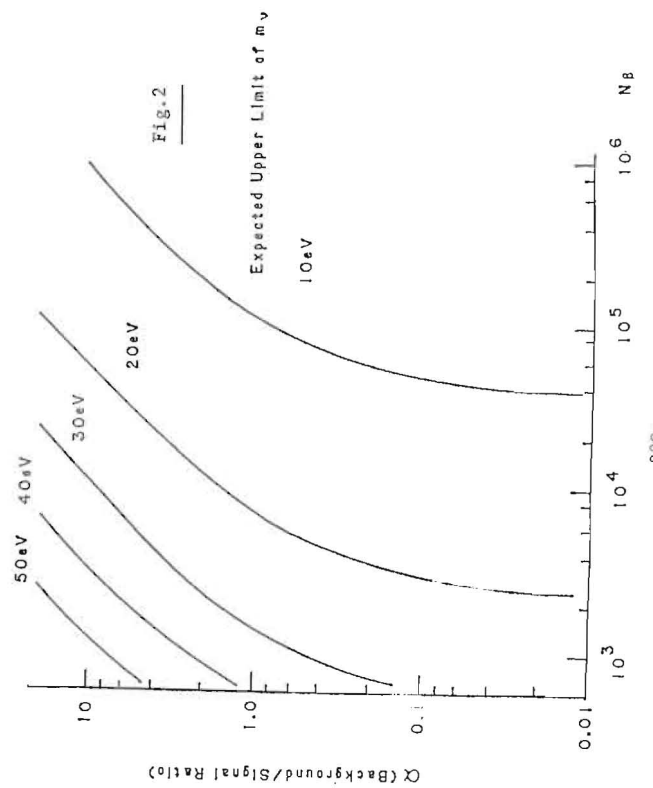
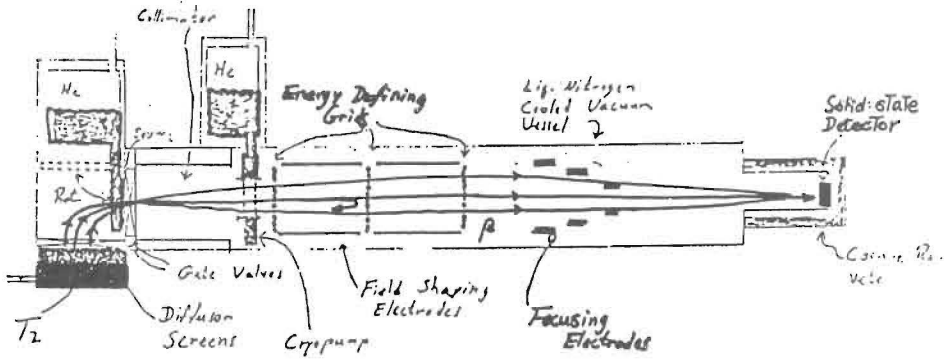


Fig. 2



Source: Pure Frozen Tritium (T<sub>2</sub>) (10-50 Ci)

Resolution:  $\sigma_E = 1.3 \text{ eV}$

Sensitivity:  $M_{D_2} \geq 4 \text{ eV}$

Fig. 6

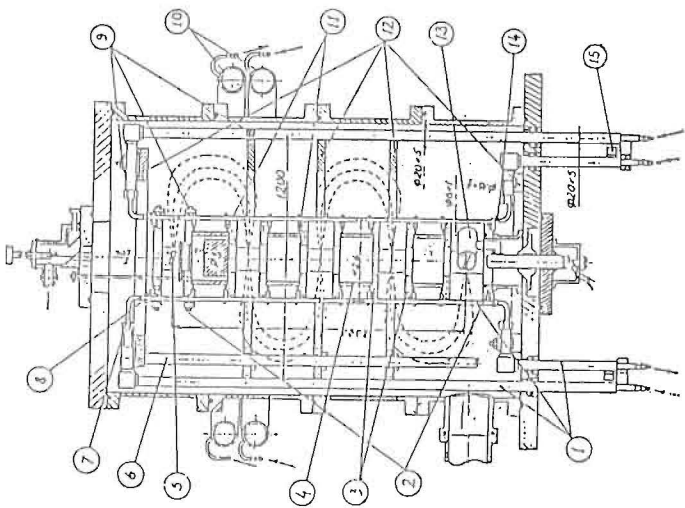


Fig. 4

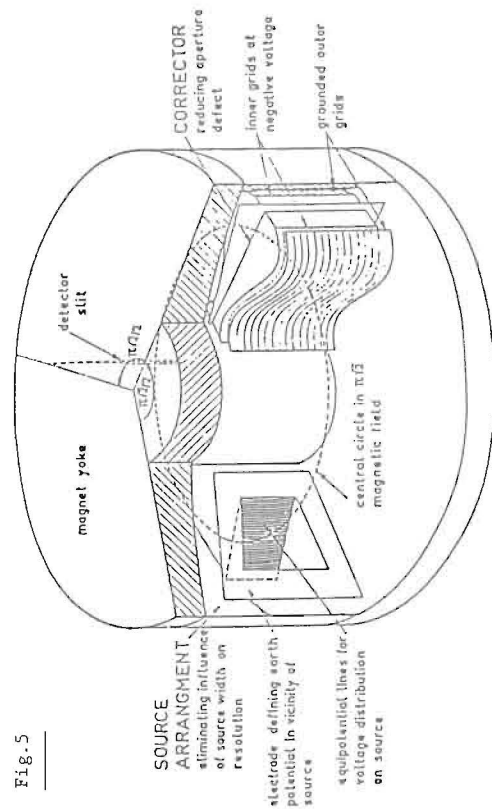


Fig. 5

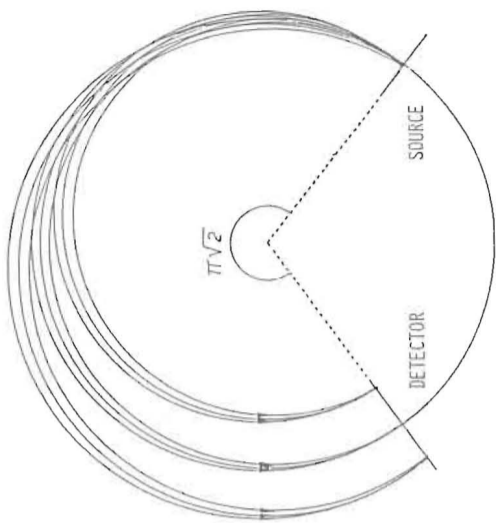


Fig. 8 (a)

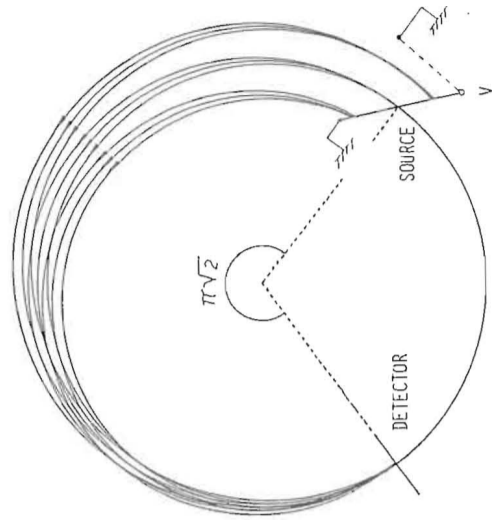


Fig. 8 (b)

LOS ALAMOS ( Bowles et. al. )

From Proc. of Mj Conf., Wisconsin (1980)

\* Point : overcome following concerns ;

- 1) Final state interactions
- 2) Scattering in solid source
- 3) Backgrounds
- 4) Step by step operation of spectrometer

} monoatomic Tritium atomic beam  
with a dispersive spectrometer

\* Apparatus

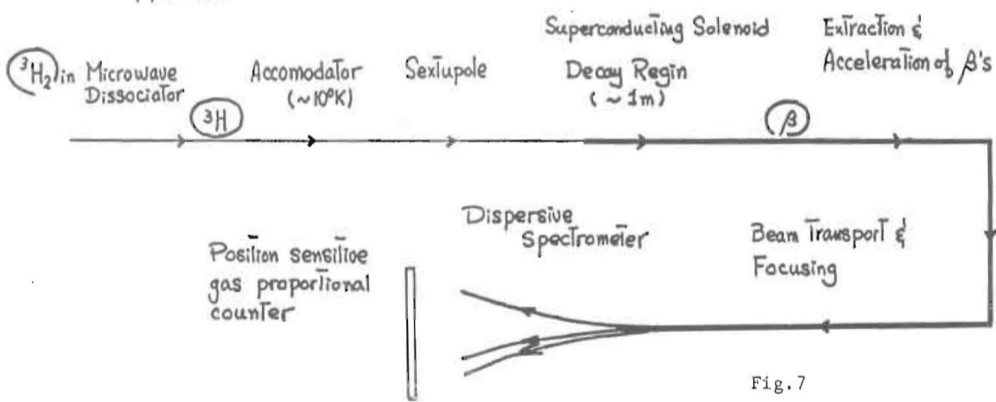


Fig. 7

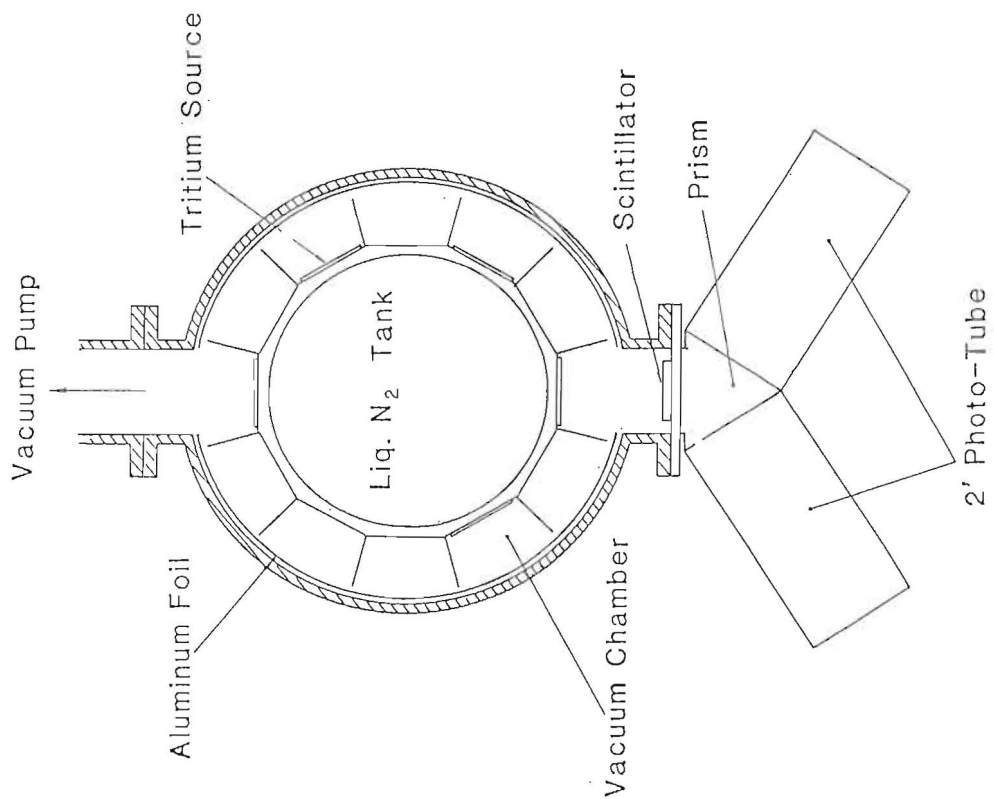
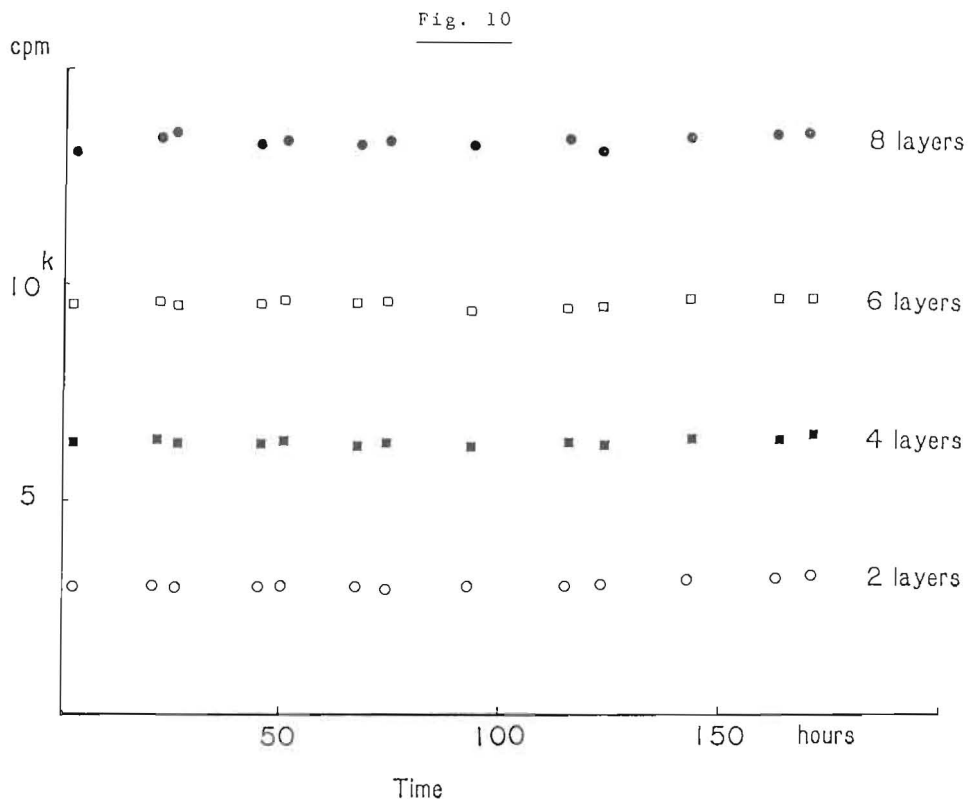
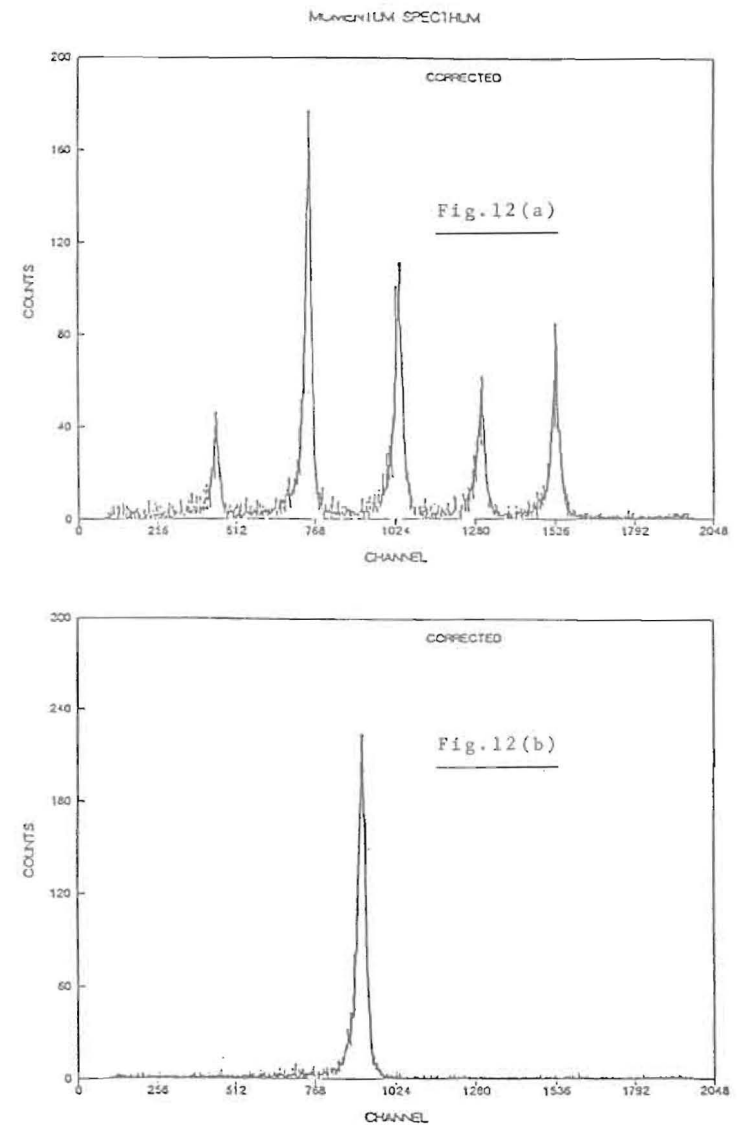
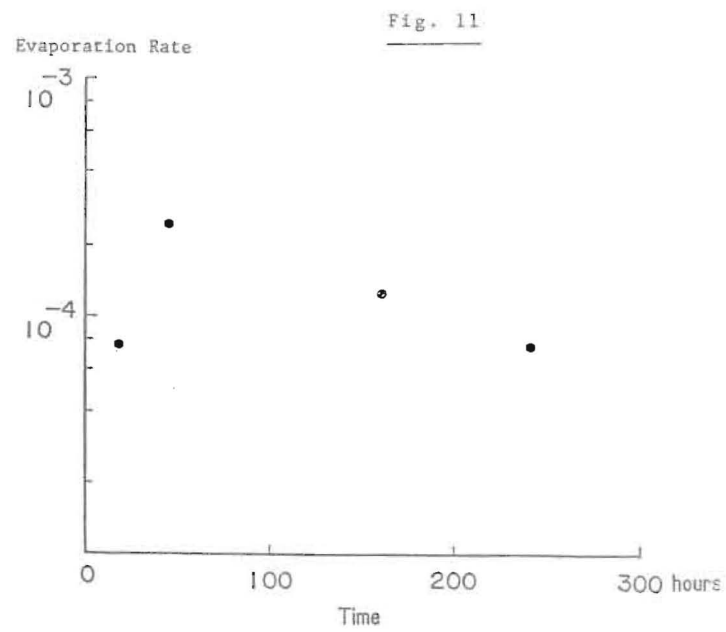
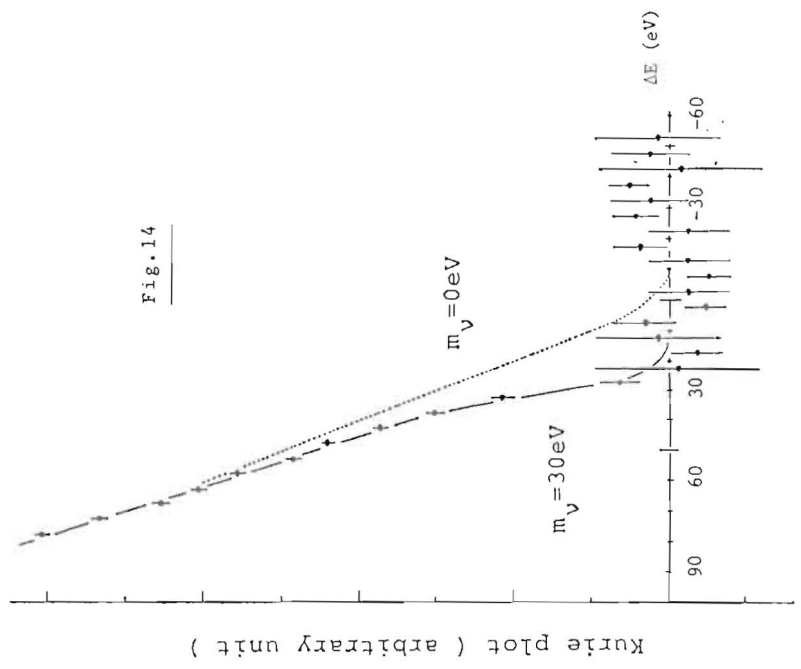
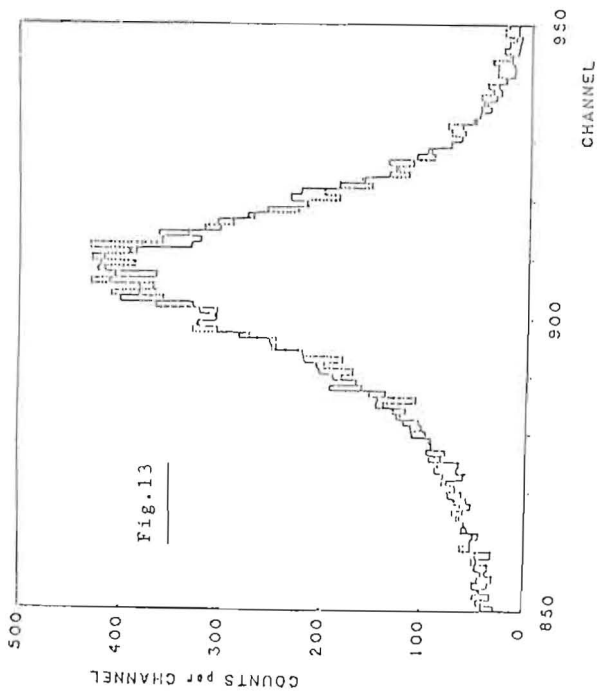


Fig. 9







Measurement of the Mass of the Electron Neutrino  
using Electron Capture in  $^{163}\text{Ho}$

Shinjiro Yasumi

National Laboratory for High Energy Physics (KEK)

Oho-machi, Tsukuba-gun, Ibaraki-ken, 305, Japan

This experiment is being done by the following collaboration;

KEK: S. Yasumi, F. Ochiai, M. Ando, H. Maezawa, H. Kitamura

Osaka University: M. Maruyama

Tohoku University: M. Fujioka, K. Ishii, T. Shinozuka, K. Sera, T.

Omori, G. Izawa, M. Yagi, K. Masumoto

University of Tsukuba: K. Shima

Kyoto University: T. Mukoyama, Y. Inagaki

T.I.T.: H. Taketani

INS: I. Sugai

University of Tokyo: A. Masuda.

Our  $m_{\nu_e}$ -studies are now well going on along the lines as described in my talk at the Brighton Conference<sup>1)</sup>. We already obtained the relationship between  $m_{\nu_e}$  and the Q-value of  $^{163}\text{Ho}$  using the value of  $T_{1/2}^M$  and the nuclear matrix element relevant to the transition  $^{163}\text{Ho} \rightarrow ^{163}\text{Dy}^{1,2)}$ . We are now trying to improve the precision on  $m_{\nu_e}$ , as determined from the Q-value by reducing experimental uncertainties both in the total number measurement for  $^{163}\text{Ho}$  atoms in a source and in the MX-ray intensity measurement. For that we are measuring the total number of  $^{163}\text{Ho}$  atoms in a source using the Isotope Dilution Mass Spectrometry<sup>3)</sup> as well as the PIXE method. Further, in order to remove the uncertainty in the thickness of a beryllium window of a Si(Li) detector used to measure MX-rays from  $^{163}\text{Ho}$ , a windowless Si(Li) detector has been purchased from HORIBA Company Ltd. With this device we have succeeded in obtaining a very beautiful MX-ray spectrum from a  $^{163}\text{Ho}$  source as shown in Fig. 1.

If the Q-value of  $^{163}\text{Ho}$  is determined independently,  $m_{\nu_e}$  can be obtained from the  $m_{\nu_e}$ -Q relation mentioned above.

On the other hand, in order to determine both  $m_{\nu_e}$  and the Q-value simultaneously, we are doing studies on the M shell of the dysprosium atom using monochromatic photons from the 2.5 GeV Electron Storage Ring in our Photon Factory.<sup>4)</sup>

The experimental setup is schematically shown in Fig. 2. Undulator radiations from BL-2 line of the storage ring are monochromated through a double reflection monochromator made of a beryl crystal and impinge upon a Dy target. Incident photon beams are monitored with a photon detector. We use beams with five different energies,  $E_a, E_b, E_c, E_d$  and  $E_e$  where  $E_a > E_{M1} > E_b > E_{M2} > E_c > E_{M3} > E_d > E_{M4} > E_e > E_{M5}$ , and  $E_{Mi}$  ( $i=1\sim 5$ ) stands for the binding energy of Mi subshell as indicated in Fig. 3. The energy widths of these photon beams are a few eV and sharp enough to remain distinct from each subshell level. MX-rays emitted by dysprosium atoms excited by the incident photon beams, are measured with two Si(Li) detectors, one of which is set in the direction of the polarization of photon beams (horizontal) and the other is set at an angle of  $90^\circ$  to the direction of the polarization of photon beams (vertical) as shown in Fig. 2. If  $S(E_\alpha)$  denotes MX-ray fluorescence spectrum from Dy atoms excited by monochromatic photons having an energy  $E_{\alpha_A}$ ,  $S(E_\alpha)$  is represented by the following equation;

$$S(E_\alpha) = N m \cdot \sum_{i=1}^5 \sigma_{i-1}(E_\alpha) \cdot f_i \quad (i = 1\sim 5) \quad (1)$$

where

$f_i$  : MX-ray spectrum in the case where there is one vacancy in  $M_i$  subshell only,

$\sigma_i(E_\alpha)$  : photo electric cross section of  $M_i$  subshell for a photon of an energy of  $E_\alpha$ ,

$N$ : total number of incident photons per second,

$m$ : number of dysprosium atoms in a target per  $\text{cm}^2$ .

Rewriting equation (1) in detail, we have

$$S(E_a) = Nm\{\sigma_1(E_a) \cdot f_1 + \sigma_2(E_a) \cdot f_2 + \sigma_3(E_a) \cdot f_3 + \sigma_4(E_a) \cdot f_4 + \sigma_5(E_a) \cdot f_5\},$$

$$S(E_b) = Nm\{\sigma_2(E_b) \cdot f_2 + \sigma_3(E_b) \cdot f_3 + \sigma_4(E_b) \cdot f_4 + \sigma_5(E_b) \cdot f_5\}$$

$$S(E_c) = Nm\{\sigma_3(E_c) \cdot f_3 + \sigma_4(E_c) \cdot f_4 + \sigma_5(E_c) \cdot f_5\}$$

$$S(E_d) = Nm\{\sigma_4(E_d) \cdot f_4 + \sigma_5(E_d) \cdot f_5\}$$

$$S(E_e) = Nm\{\sigma_5(E_e) \cdot f_5\}$$

A typical raw data of  $S(E_a)$  recently obtained by us is shown in Fig. 4, where  $E_a = 2.13$  KeV. We measure five spectra,  $S(E_a)$ ,  $S(E_b)$ ,  $S(E_c)$ ,  $S(E_d)$  and  $S(E_e)$ , changing the energy of an incident photon beam.

If  $\sigma_i(E_\alpha)$  is known, one can obtain  $f_i$  ( $i=1 \sim 5$ ) by turns using the above five equations.  $f_1$  and  $f_2$  in Fig. 3 are theoretical spectra which were calculated by T. Mukoyama.<sup>5)</sup>

If  $S_p^{163\text{Ho}}$  stands for a photon spectrum from  $^{163}\text{Ho}$  where number of photons per atom per second is plotted as a function of an energy of photons, as shown in Fig. 3,

then we have

$$\begin{aligned} S_p^{163\text{Ho}} &= \frac{1}{N_0} \frac{d}{dt} [f_1 \cdot N_{M1} + f_2 \cdot N_{M2}] \\ &= \frac{d}{dt} [f_1 \cdot n_{M1} + f_2 \cdot n_{M2}] \\ &= f_1 \frac{dn_{M1}}{dt} + f_2 \frac{dn_{M2}}{dt} \\ \therefore S_p^{163\text{Ho}} &= \lambda_{M1} \cdot f_1 + \lambda_{M2} \cdot f_2. \end{aligned} \quad (2)$$

where

$N_{M_i}$  ( $i=1,2$ ): number of vacancies produced in  $M_i$  subshell in the decay  $^{163}\text{Ho} \xrightarrow{\text{E.C.}} ^{163}\text{Dy}$ ,

$n_{M_i}$  ( $i=1,2$ ):  $\equiv N_{M_i} / N_0$ ,

$N_0$  : total number of  $^{163}\text{Ho}$  atoms in a source,

$\lambda_{M_i}$  ( $i=1,2$ ): partial  $M_i$ -capture decay constant.

Equation (2) tells us that when we reconstruct  $S_p^{163\text{Ho}}$  using spectra  $f_1$  and  $f_2$ , these coefficients of  $f_1$  and  $f_2$  correspond  $\lambda_{M1}$  and  $\lambda_{M2}$  respectively.

Alternative way to get  $\lambda_{M1}$  and  $\lambda_{M2}$  is in the following, which is based on the fact that peak 1 in Fig. 3 comes from  $M_1$  subshell only and peak 2 in Fig. 3 comes from  $M_1$  and  $M_2$  subshells:

If intensities of peak 1 and peak 2 in  $S_p^{163\text{Ho}}$  denote  $I_1$  and  $I_2$  respectively as shown in Fig. 3, we have

$$I_1 \div P_1^{M1} = \frac{dn_{M1}}{dt} = \lambda_{M1},$$

$$[I_2 - (P_2^{M1} \times \frac{dn_{M1}}{dt})] \div P_2^{M2} = \lambda_{M2} \quad (3)$$

where  $P_1^{M1}$  and  $P_2^{M1}$  stand for counterparts in  $f_1$  for peaks 1 and 2 in  $S_p^{163\text{Ho}}$  respectively, and  $P_2^{M2}$  stands for a counterpart in  $f_2$  for peak 2 in  $S_p^{163\text{Ho}}$  in Fig. 3. On the other hand  $\lambda_{M1}$  and  $\lambda_{M2}$  are expressed as

$$\lambda_{M1} = 2.7 \times (Q - 2.047) \cdot \sqrt{(Q - 2.047)^2 - m_{ve}^2} \times 10^{-12} \text{ per atom per sec.},$$

$$\lambda_{M2} = 0.13 \times (Q - 1.841) \cdot \sqrt{(Q - 1.841)^2 - m_{ve}^2} \times 10^{-12} \text{ per atom per sec.} \quad (4)$$

where  $m_{ve}$  and  $Q$  are given in KeV.

Using  $\lambda_{M1}$  and  $\lambda_{M2}$  as determined experimentally (equations (2) and (3)), both  $m_{ve}$  and the  $Q$ -value can be obtained from equations (4).

Finally we summarize several ways to determine  $m_{ve}$  using electron capture in  $^{163}\text{Ho}$  in Fig. 5. It seems similar ways to these are also applicable to other electron capturing nuclides.

References:

- 1.) S. Yasumi, PROC. Int. Europhys. Conf. High Energy Phys. Brighton July 1983 p.391.
- 2.) S. Yasumi, G. Rajasekaran, M. Ando, F. Ochiai, H. Ikeda, T. Ohta, P.M. Stefan, M. Maruyama, N. Hashimoto, M. Fujioka, K. Ishii, T. Shinozuka, K. Sera, T. Omori, G. Izawa, M. Yagi, K. Masumoto and K. Shima, Phys. Lett. 122B (1983) p.461.
- 3.) K. Kawakami, K. Yoshida and A. Masuda, KEK 83-24 March 1984 E p.37.
- 4.) S. Yasumi, F. Ochiai, M. Ando, H. Maezawa, H. Kitamura, K. Itoh, M. Maruyama, M. Fujioka, K. Ishii, T. Shinozuka, K. Sera, T. Omori, G. Izawa, M. Yagi, K. Masumoto, K. Shima, T. Mukoyama, Y. Inagaki and H. Taketani, KEK Preprint 83-17 August 1983 E.
- 5.) T. Mukoyama, PROC. of the 9th Workshop on the Mass of the Electron Neutrino KEK 83-24 March 1984 E p.28.

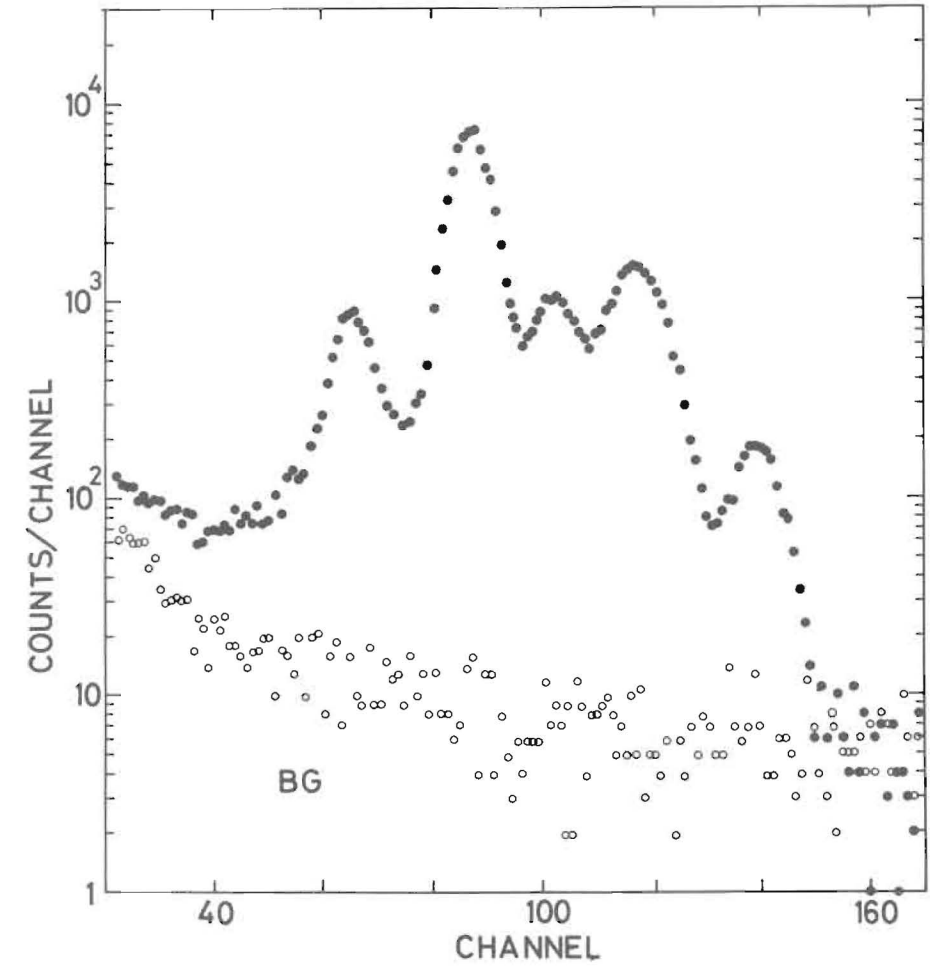


FIG. 1

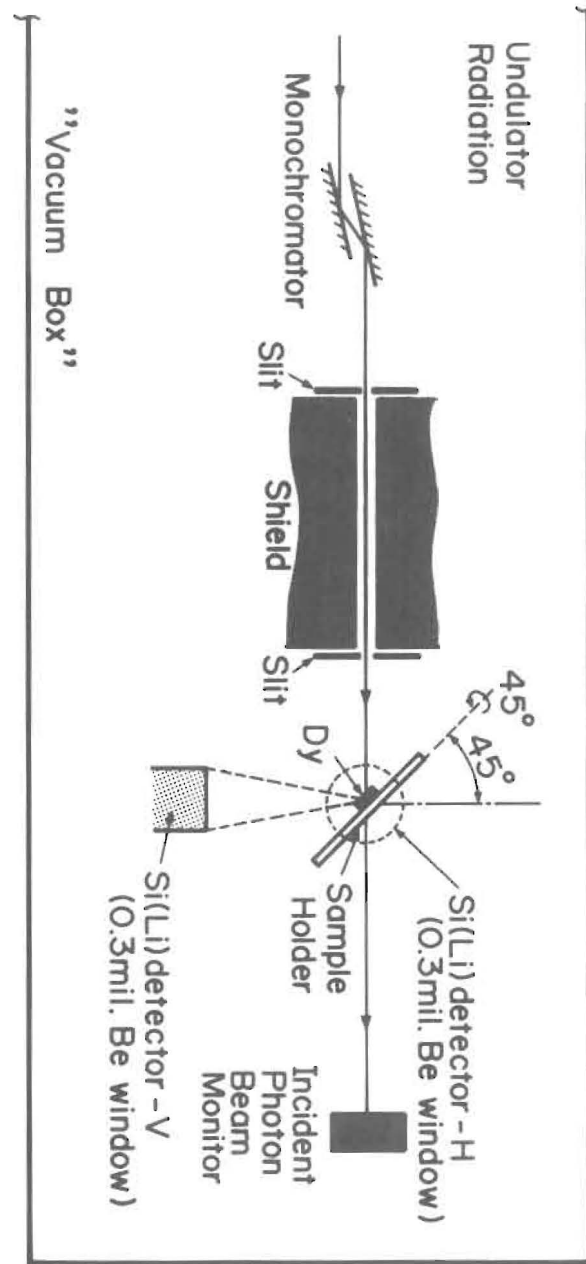


Fig. 2

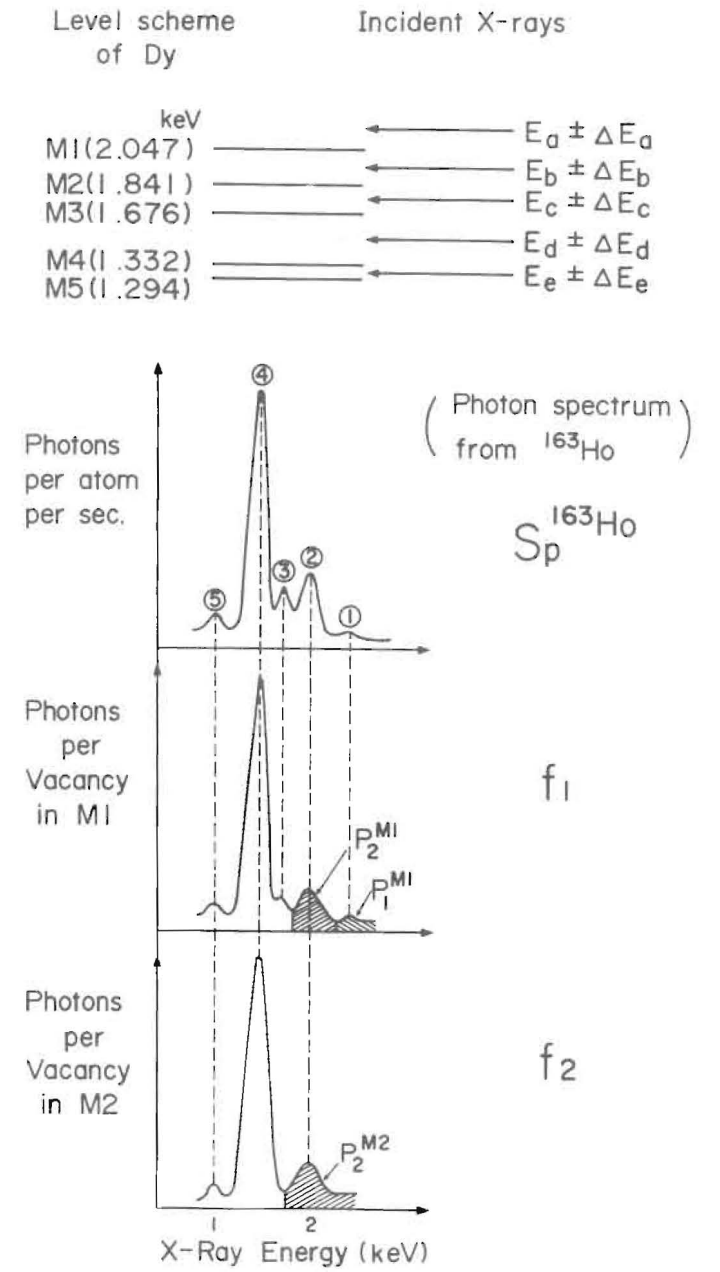


Fig. 3

Dy + monochromatic photons (2.13 keV)  
 from BL-2, PF-KEK  
 ( Run No. D111604-HA )

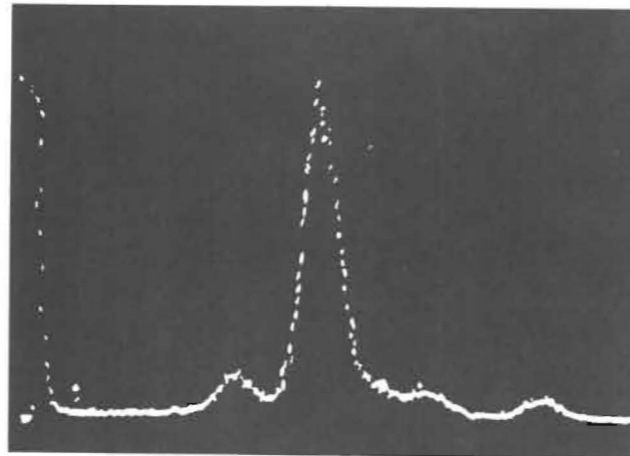


FIG. 4

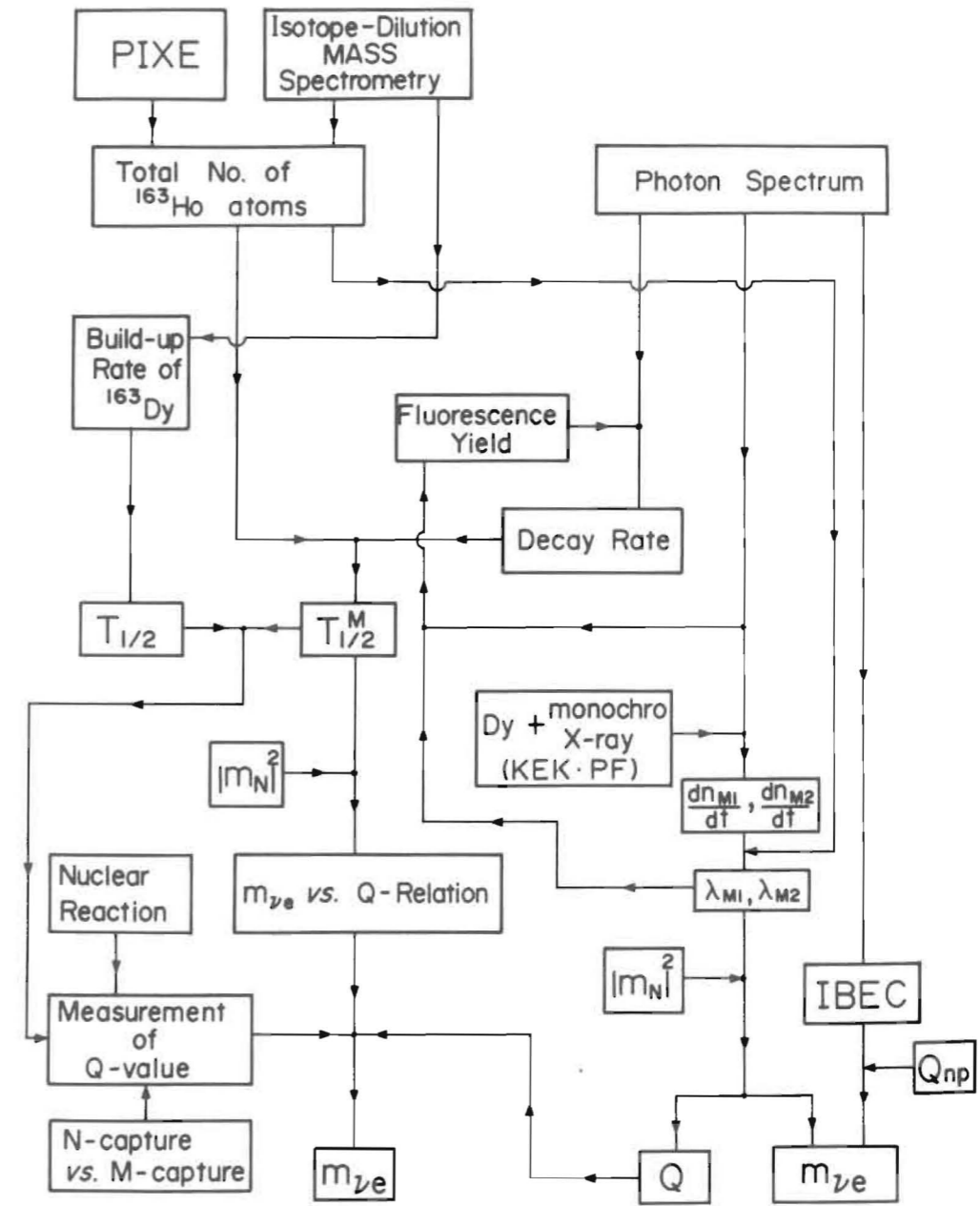


FIG. 5

Present Status of the Solar Neutrino Problem

Naoki Itoh

Department of Physics, Sophia University  
Kioi-cho, Chiyoda-ku, Tokyo, 102 Japan

ABSTRACT

The present status of the solar neutrino problem is critically reviewed. It is suggested that further measurements of the low-energy nuclear cross sections, the controversial  ${}^3\text{He}({}^4\text{He}, \gamma){}^7\text{Be}$  cross section in particular, should be made before taking up more exotic explanations.

Nuclear Physics in the Solar Neutrino Problem

Kunio Nagatani<sup>†</sup>

Institute for Nuclear Study, University of Tokyo

Tanashi, Tokyo 188, Japan

It is with great sorrow that we have to announce the death of Professor Kunio Nagatani, on 18th of March 1984, at the age of 48, in a tragic accident. He gave a very lucid talk, unfortunately his last, at this Workshop. There is no manuscript for this Proceedings, and therefore we have attempted to collect the transparencies that he presented. It is naturally impossible to reconstruct his actual talk that so fascinated the audience. Nevertheless, we hope that this material will help to keep his memory alive for the participants of the Workshop and others who know him.

M. Fukugita

---

†) Deceased.



## Solar Neutrino Problem

1. Why is it so important?  
It is even called a "puzzle".
  - i) Astronomical Aspects Fig 1.
  - ii) Nuclear Physics Aspects Fig 1.
2. Nuclear Physics in the Solar Neutrino Problem.
  - i) Historical events Fig 2
    - ${}^3\text{He}(\alpha, \gamma){}^7\text{Be}$ :  $\alpha$  is 1000 times larger!! '58
    - ${}^{37}\text{Cl}(\nu) \rightarrow {}^{37}\text{Ar}$ : analog state!! '63
    - S-factor measurements followed.
3. What is the present status? Fig 3
4. Brief glance at Davis' exp. Fig 4, 4a, b
  - ${}^{37}\text{Cl}$ -capture rate  
use of the analogs Figs 5a, 6, 6a, b
  - How sure is the Davis' result Fig 7, 8, 9

## 5. Nuclear Reaction Rates.

- i) p-p chain Fig 10, 11, 12
  - Expected capture rates Fig 13
6. • The present result again Fig 14, 14a
7. • So what is wrong (mistake...)?  
Is it in the nuclear physics?  
Several examples: how we have been trying.
  - i)  ${}^3\text{He} + {}^3\text{He} \rightarrow \alpha + 2p$  Fig 15  
Willy asks:  
"What is cooking the solar neutrinos?"  
Resonances at  $\sim$  threshold?  
Fig 16, 17
  - ii)  $\alpha({}^3\text{He}, \gamma){}^7\text{Be}$ ? Fig 18, 19, 20, 21, 22

iii)  ${}^7\text{Be}(p,\gamma){}^8\text{B}$

Fig 23

5. What is next?

i) Believe more checks shall be made

ii) New  $\gamma$  detectors

iii) Or some new ideas?

Fig 24, 25

Fig 26, 27.

The following figures are omitted:

Fig. 2

Fig. 3

Figs. 4a, 4b

Figs. 5, 5a, 6a

Figs. 7, 8, 9

Fig. 11, 12

Fig. 13

Figs. 14, 14a

Figs. 26, 27

For general references, see:

J. N. Bahcall and R. Davis, Jr., An Account of the Development of the Solar Neutrino Problem, Contribution to the Festschrift for Willy Fowler;  
J. N. Bahcall, W. F. Huebner, S. H. Lubow, P. D. Parker and R. K. Ulrich, Rev. Mod. Phys. 54, 767 (1982).

### The $^{37}\text{Cl}$ Experiment

Reaction:  $\nu + ^{37}\text{Cl} \rightarrow ^{37}\text{Ar} + e$

Size: 610 tons  $\text{C}_2\text{Cl}_4$

Location: Homestake Gold Mine  
depth 4400 m. water equivalent

Procedure: 1. Exposed for period 35-100 days  
with 0.1  $\text{cm}^3$  STP  $^{36}\text{Ar}$  or  $^{38}\text{Ar}$   
carrier

2.  $^{37}\text{Ar}$  removed by helium purge  
Tank  $\rightarrow$  Condenser  $\rightarrow$  Absorber  $\rightarrow$   
Charcoal  $\rightarrow$  Tank

3. Argon purified by gas  
chromatography and getter

4.  $^{37}\text{Ar}$  measured by gas propor-  
tional counter

5. Mass-spectrometer analysis of  
 $^{36}\text{Ar}$  or  $^{38}\text{Ar}$  to measure yield

Fig. 4

" within our store of observational information about stars, only the chlorine  $^{37}$  experiment is clearly inconsistent with the standard theory of stellar evolution. It is the only place where we don't see a way out of observational difficulties", "We know more about the sun than any other star. Because it is so close, we can measure its mass, radius, age, surface composition, and luminosity much better than for any other star. Also the sun is in the static main sequence stage where the calculations are the simplest and most reliable. So the sun is the easiest and the best determined case for stellar evolution. It is a serious embarrassment that the theoretical answer doesn't come out right."

We believe that the nuclear synthesis in star is quite well understood. In the sun, most of the energy comes from the very initial stage of the whole chain of the synthesis where four protons are converted into an alpha, so called proton-proton (p-p) chain. Therefore, the study of the solar neutrinos boils down to investigating the p-p chain reactions, in particular their cross sections at very low energies. Here again we notice that if we are not able to determine these simple reaction rates, we can not be credible for any other processes.

Fig. 1

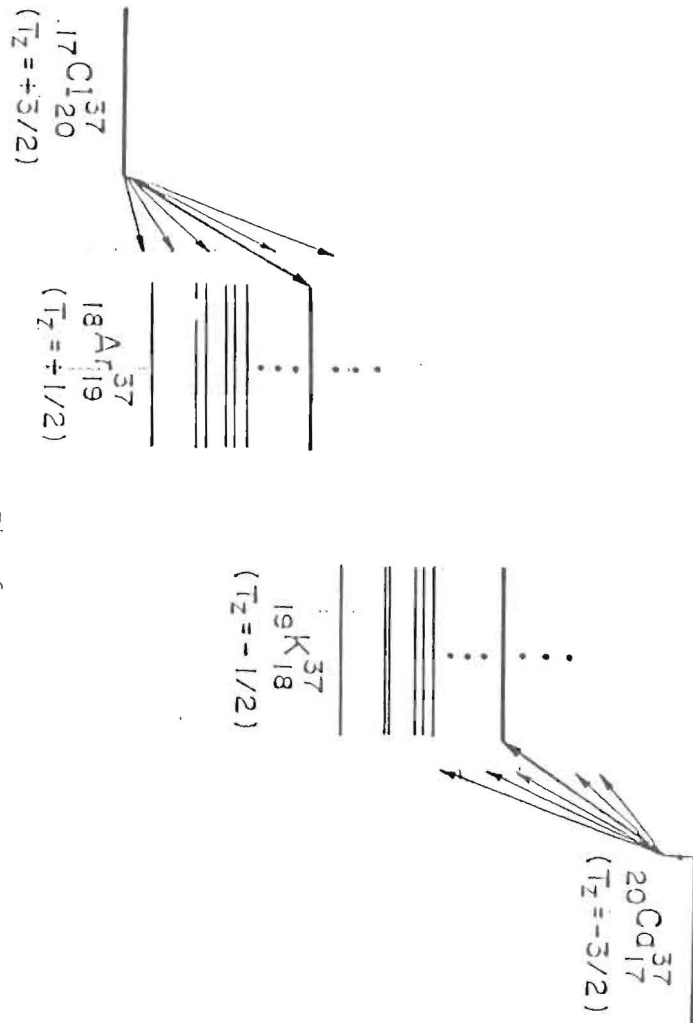


Fig. 6

TABLE V. Individual  $ft$  values and cross sections for  $^8\text{B}$  neutrinos incident on  $^{37}\text{Cl}$ . The  $ft$  values are from measurements by Sextro, Gough, and Cerny (1974) of  $^{37}\text{Ca}$  decay ratios.

Excitation energy (MeV)	$(10^5 \text{s}/ft)^a$	$\sigma_{2\nu}$ [Eq. (17)]	$\beta^b$	$\sigma$ ( $10^{-46} \text{cm}^2$ )
0.0	0.794	397	0.872	596
1.41	0.826	272	0.857	418
2.80	1.351	176.5	0.833	429
3.53	1.471	135	0.820	354
3.84	1.919	120	0.809	405
4.40	0.752	96	0.803	126
4.50	1.449	92	0.800	231
4.66	0.909	86	0.798	135
4.95	1.923	75.5	0.790	249
4.98	50.0	74.5	0.790	6,375
5.12	6.711	70	0.788	799
5.32	1.408	64	0.785	153
5.45	0.769	60	0.782	78
6.02	1.667	45	0.773	126

<sup>a</sup>Taken from Sextro, Gough, and Cerny (1974).

<sup>b</sup>See Eqs. (24) and (34).

TABLE IV. Various estimates of the  $^8\text{B}$  solar neutrino cross section on  $^{37}\text{Cl}$ .

Model	$^{37}\text{Cl}$ data	Assumption ...	$\sigma$ ( $10^{-42} \text{cm}^2$ )
A	Sextro <i>et al.</i> (1974)	$\log ft = 3.30$ analog	1.05
B	Sextro <i>et al.</i> (1974)	$\log ft = 3.256$ analog (a minimum)	1.11
C	Sextro <i>et al.</i> (1974)	2.1% decay to $E_x = 6.5$ MeV	1.10
D	Poskanzer <i>et al.</i> (1966)	18% decay to the 1.41 MeV state	1.05
E	Poskanzer <i>et al.</i> (1966)	18% decay to the 2.80 MeV state	1.09
F	Poskanzer <i>et al.</i> (1966)	$\lambda(1.4 \text{ MeV})/\lambda(2.8 \text{ MeV}) = 2.4^a$	1.06
G	Poskanzer <i>et al.</i> (1966)	$\lambda(1.4 \text{ MeV})/\lambda(2.8 \text{ MeV}) = 0.18^b$	1.08

<sup>a</sup>Haxton and Donnelly (1977).

<sup>b</sup>Lanford and Wildenthal (1972).

Fig. 6b

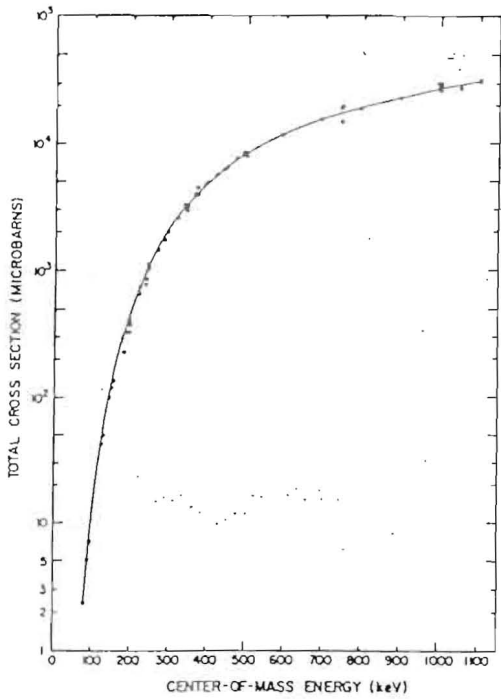


FIG. 7.  ${}^3\text{He}({}^3\text{He}, 2p){}^4\text{He}$  total cross sections as a function of the center-of-momentum energy. The solid line is a guide to the eye.

$$S(E) = \sigma(E) E e^{2\pi\eta}; \quad \eta = \frac{Z_1 Z_2 e^2}{\hbar v}$$

$$S(E_{c.m.}) = \sigma(E_{c.m.}) \times E_{c.m.} \times \exp\left(\frac{4.860}{E_{c.m.}^{1/2}}\right),$$

$E_{c.m.}$  in MeV.

Fig. 15

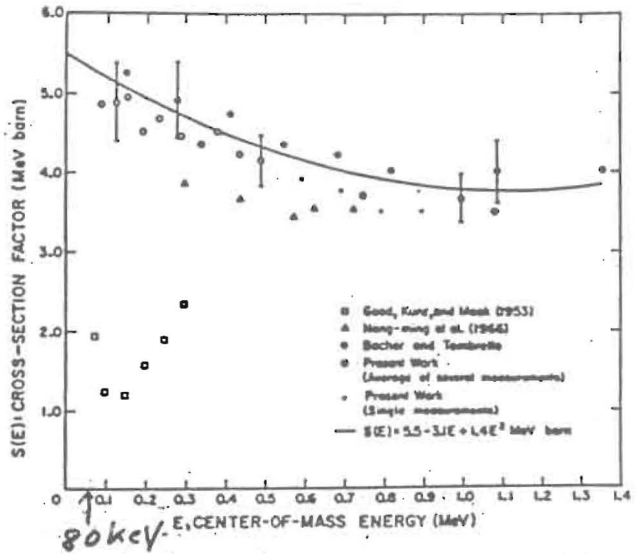
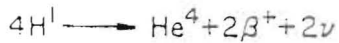
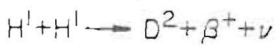


FIG. 8.  ${}^3\text{He}({}^3\text{He}, 2p){}^4\text{He}$  cross-section factors as a function of center-of-momentum energy. Error bars indicated are typical, include statistical and estimated systematic errors in both measured total cross sections and center-of-momentum energy. Solid line is the curve  $S(E_{c.m.}) = 5.5 - 3.1E_{c.m.} + 1.4E_{c.m.}^2$ , which gives the best quadratic representation of the data of the present work and the data of Bacher and Tombrello.



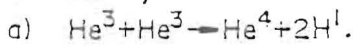
$$Q = +26.731 \text{ MeV}$$



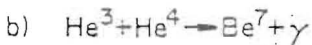
$$1.442 \text{ MeV} \quad \overline{E}_\nu = 0.25 \text{ MeV}$$



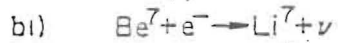
$$5.493 \text{ MeV}$$



$$12.859 \text{ MeV}$$



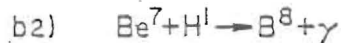
$$1.587 \text{ MeV}$$



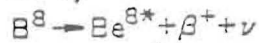
$$0.862 \text{ MeV} \quad E_\nu = 0.862 \text{ MeV}$$



$$17.347 \text{ MeV}$$



$$0.135 \text{ MeV}$$



$$15.079 \text{ MeV} \quad \overline{E}_\nu = 7.4 \text{ MeV}$$



$$2.995 \text{ MeV}$$

Fig. 10

$$N_r = \frac{N_{^3\text{He}} n_{^3\text{He}} S(E_{\text{c.m.}})}{4\pi} \times \int_{-\infty}^{+\infty} dx \frac{\Omega(x)}{E_{\text{c.m.}}(x)} \exp\{-4.860/[E_{\text{c.m.}}(x)]^{1/2}\}.$$

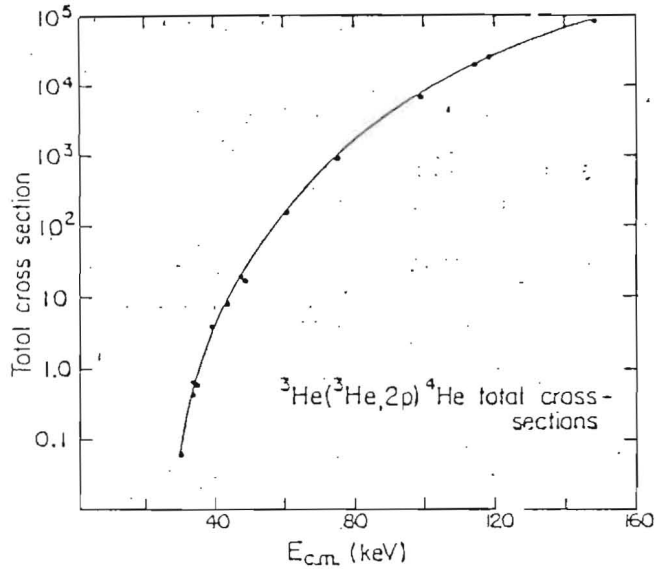


FIG. 3.  $^3\text{He}(^3\text{He}, 2p)^4\text{He}$  total cross sections as a function of c.m. energy. The solid curve is a guide to the eye.

Fig. 17

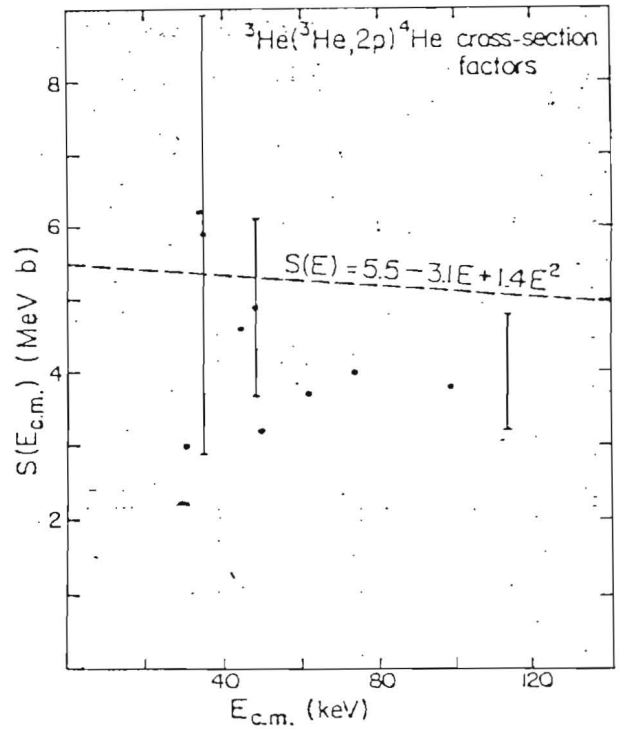


FIG. 4.  $^3\text{He}(^3\text{He}, 2p)^4\text{He}$  cross-section factors as a function of c.m. energy. The broken curve  $S(E_{\text{c.m.}}) = 5.5 - 3.1E_{\text{c.m.}} + 1.4E_{\text{c.m.}}^2$  gave the best quadratic representation of the data at higher energies (Ref. 5).

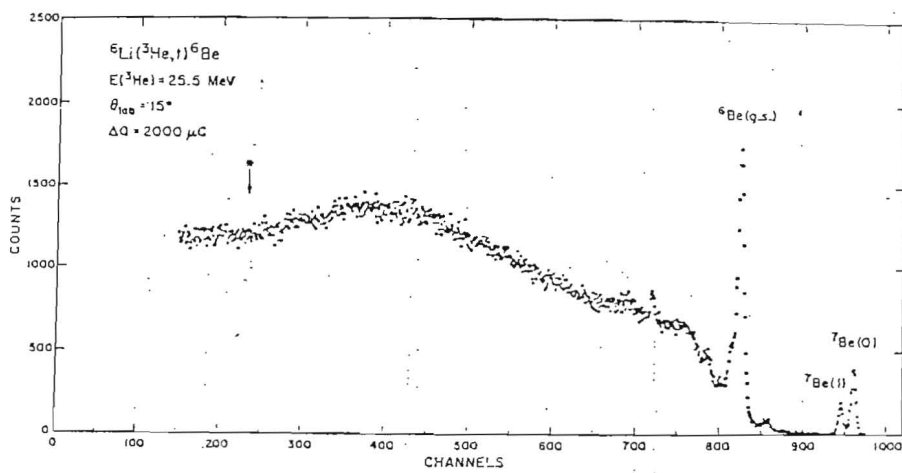


Fig. 16

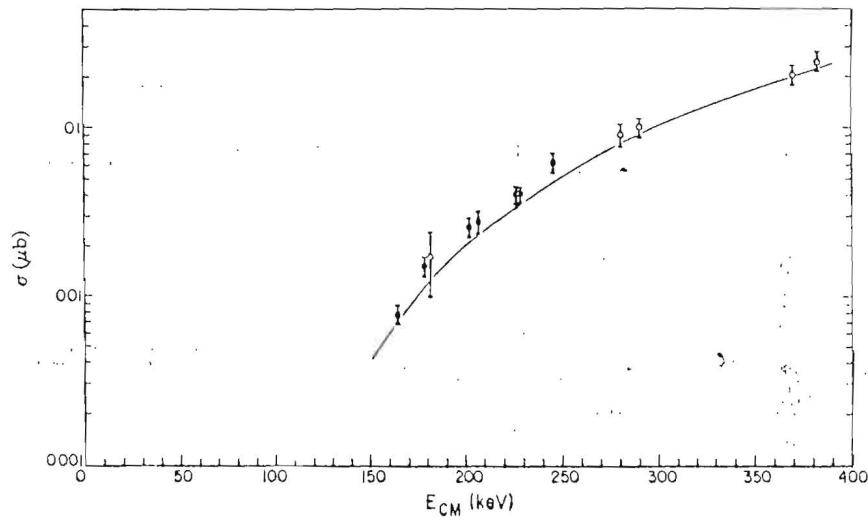


Fig. 2. The cross sections  $\sigma(E)$ : The solid points are the present results while the open circles are those of Parker's corrected values<sup>12)</sup>. The solid curve indicates the theoretical prediction of ref. 4).

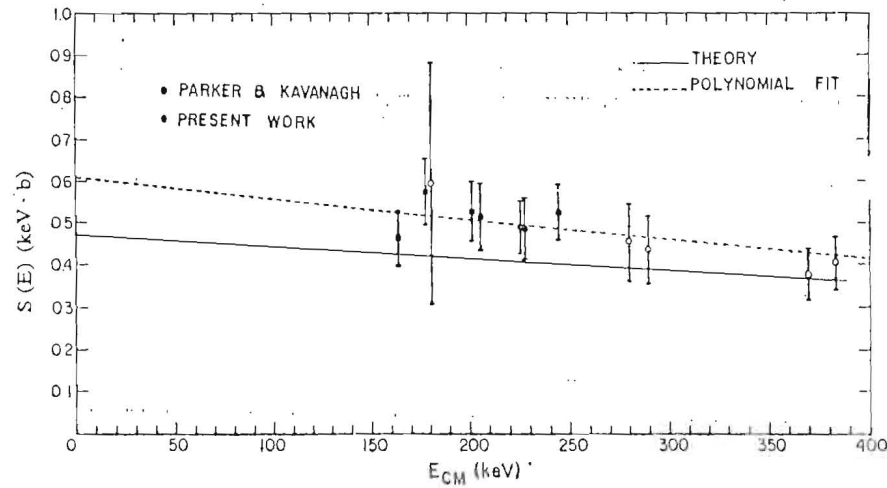


Fig. 3. The cross-section factor  $S(E)$ : The solid points are the present results while the open circles are those of Parker's corrected values<sup>12)</sup>. The solid curve indicates the theoretical prediction of ref. 4), while the dashed curve shows the result of the least squares polynomial fit.

Fig. 18

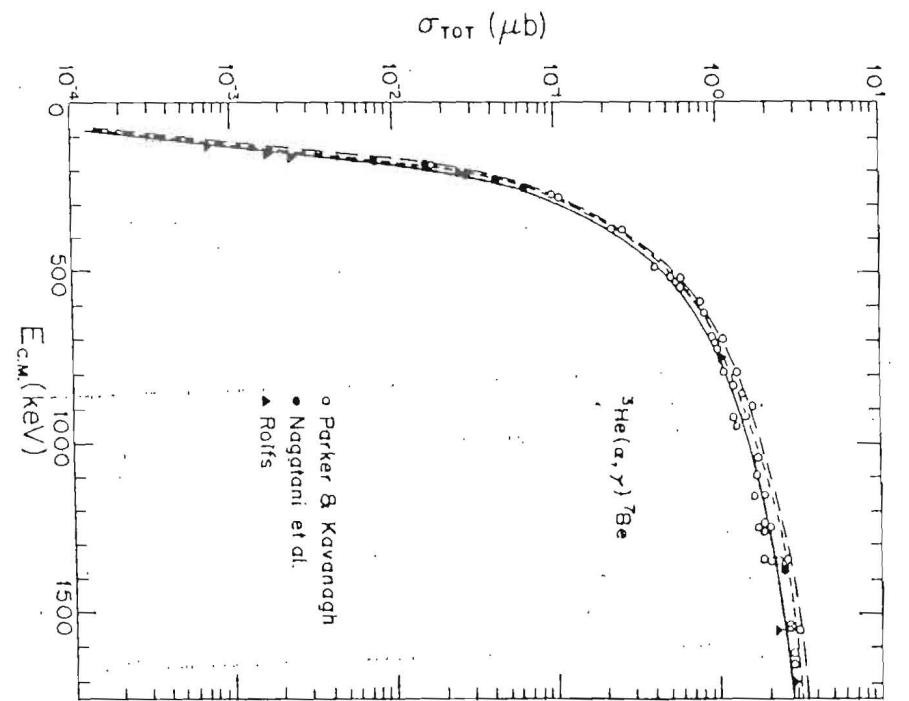
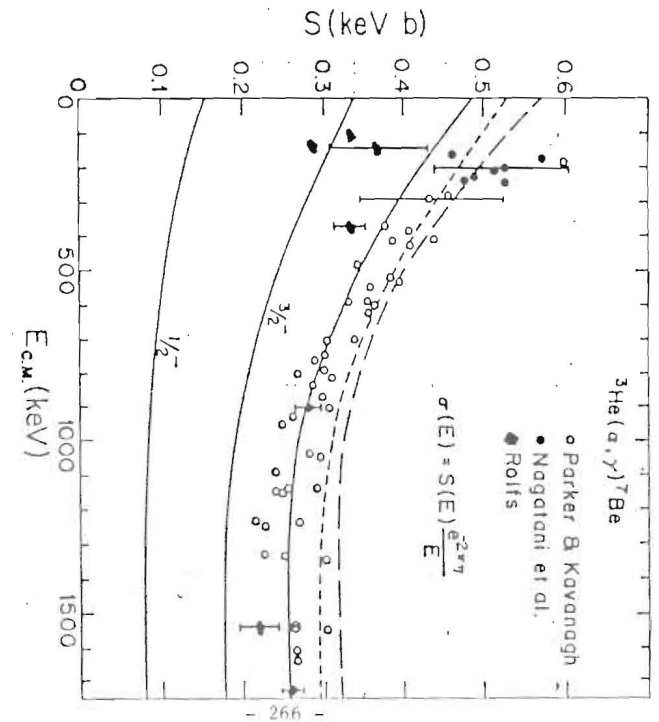
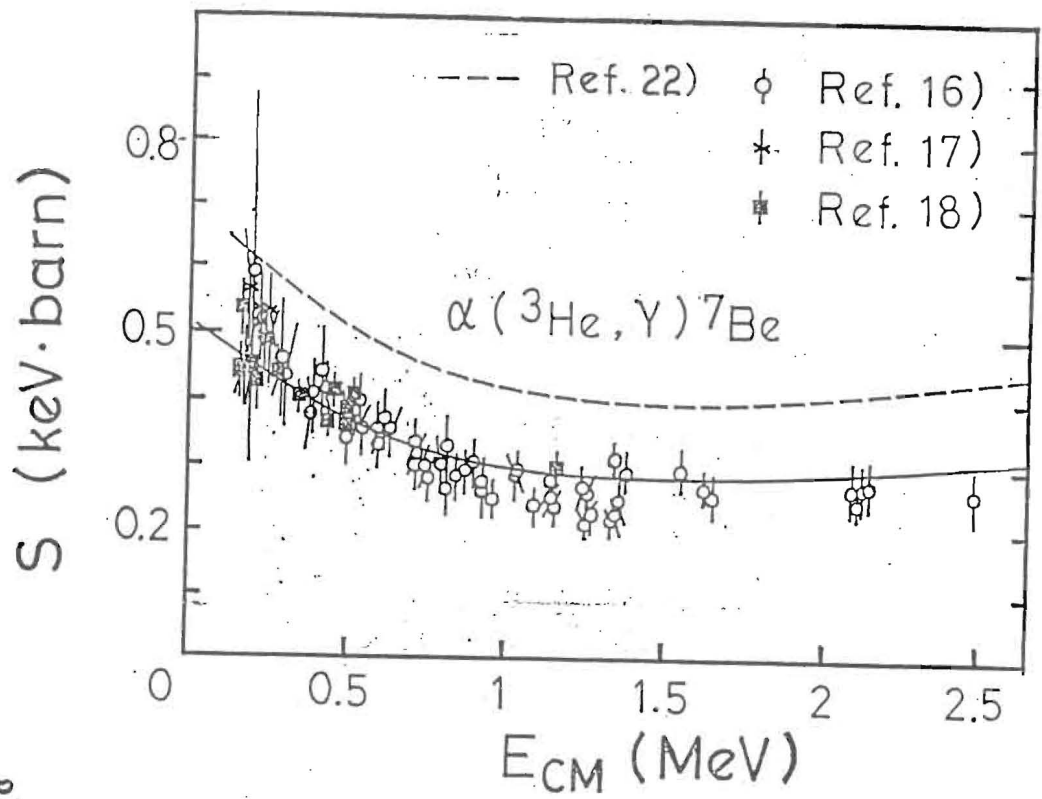


Fig. 19





Kajimo

Fig. 21

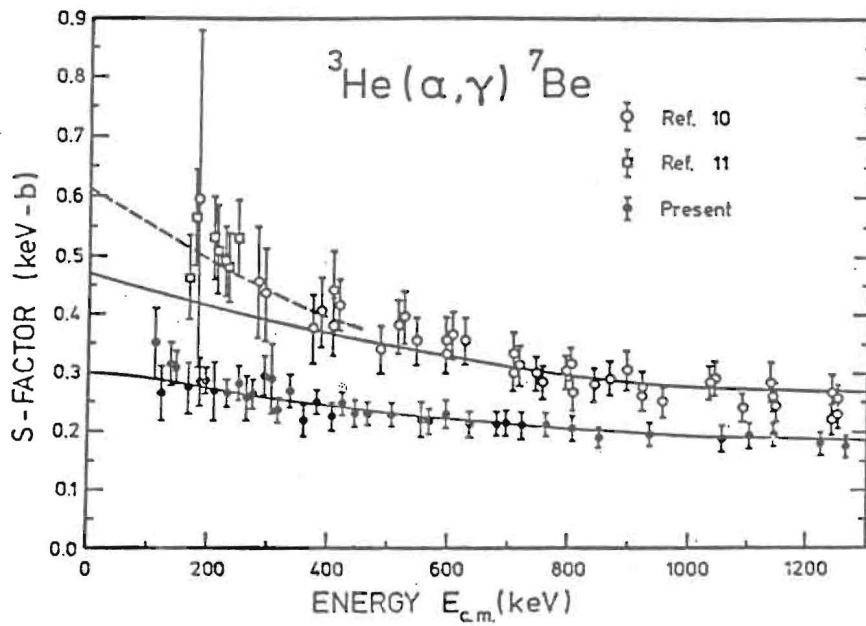


Fig. 20



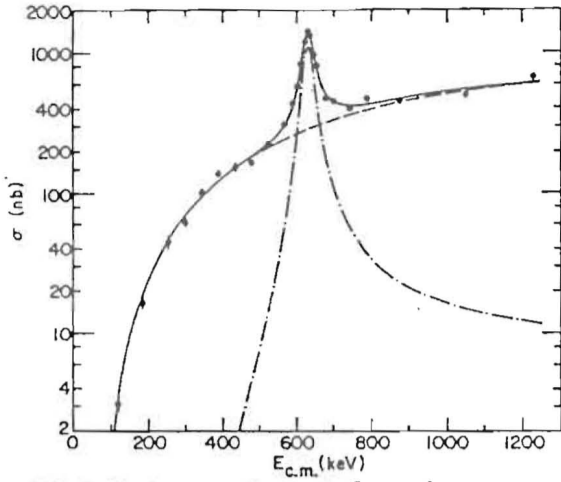


FIG. 7. Total cross section for the  ${}^7\text{Be}(p,\gamma){}^8\text{B}$  reaction as a function of the center of mass energy. The dashed curve is the nonresonant direct capture part of the cross section, the dashed-dot curve is the resonant cross section using the parameters discussed in the text, and the solid curve is the sum of the two. If not shown, the error bars (statistical only) are smaller than the data points.

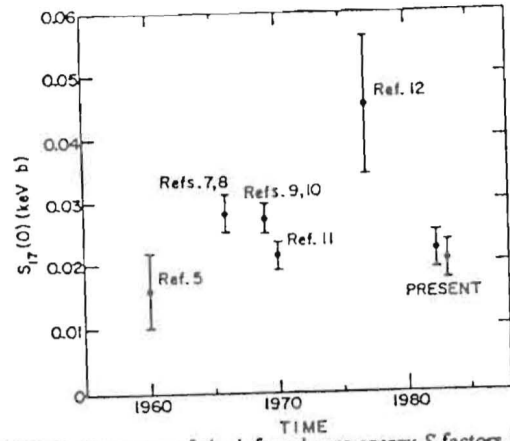


FIG. 9. Summary of the inferred zero-energy  $S$  factors from the measurements of the  ${}^7\text{Be}(p,\gamma){}^8\text{B}$  reaction. The two values for the present measurement are for the two independent methods used to determine the  ${}^7\text{Be}$  areal density.

Ref 5. Kavanagh (80)  $E_p = 0.8$  and  $1.4$  MeV  
 7,8 Parker (86-'68)  $0.483 - 1.952$  MeV, 8 points  
 9,10 Kavanagh (89-72)  $0.185 - 10.0$  MeV  
 11. Vaughn ('72)  $0.953 - 3.281$  MeV, 20 points  
 12. Wiezorek ('77)  $0.360$  MeV  
 prst, Fillipone ('83)

Fig. 23

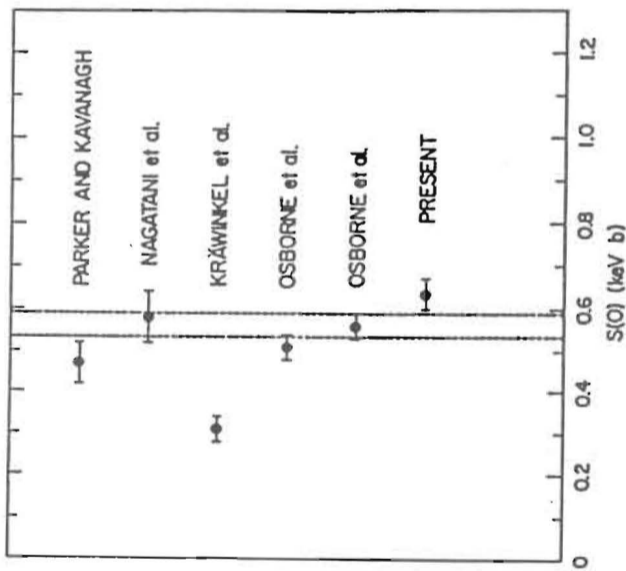


FIG. 4. Graphical representation of  $S(0)$  from various experiments. Reference numbers from top to bottom are 4, 5, 3, 12, and 12. The band represents  $1-\sigma$  limits from the recommended value. As discussed in the text, further uncertainty is contributed by the choice of theoretical extrapolation. The data in this figure have all been reduced to zero energy with the aid of the Tombrello-Parker theory (Ref. 11).

Fig. 22

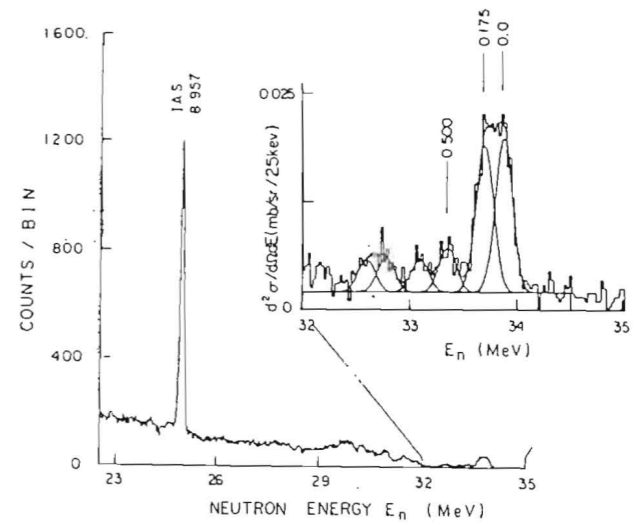


TABLE I. Results of reactions  $^{71}\text{Ga}(p, n)^{71}\text{Ge}$  and  $^{58}\text{Ni}(p, n)^{58}\text{Cu}$ (g.s.).

Final state $E_x$ (MeV), $J^\pi$	$(\frac{d\sigma}{d\Omega})_{\theta=0}^{expt}$ (mb/sr)	$(p, n)$	$\log ft^a$ $\beta^+$ decay
$^{71}\text{Ge}(0.0, 1/2^-)$	0.153	4.676 <sup>b</sup>	4.375 <sup>c</sup>
$(0.175, 5/2^-)$	0.145	4.69	...
$(0.500, 3/2^-)$	0.036	5.31	...
$^{58}\text{Cu}(0.0, 1^+)$	0.325	4.36	4.83

<sup>a</sup>A  $\log ft$  value is connected to the GT matrix element  $B(\text{GT})$  through  $ft = 6163.4 / (g_A/g_V)^2 B(\text{GT})$  (for this constant, see Ref. 14), the ratio of the axial to vector coupling constants being 1.250 (Ref. 15).

<sup>b</sup> $B(\text{GT}; p, n)$ , which is also the GT matrix element for neutrino absorption, is related to  $B(\text{GT}; \beta^+)$ , the GT matrix element for inverse  $\beta^+$  decay, through  $B(\text{GT}; p, n) = [(2J_f + 1)/(2J_i + 1)] B(\text{GT}; \beta^+)$ , where  $J_i$  and  $J_f$  are the spins of the initial and final states for the  $(p, n)$  reaction, respectively.

<sup>c</sup>Taken from Ref. 16.

Fig. 25

PREDICTIONS FROM STANDARD SOLAR MODEL

NEUTRINO SOURCE	ENERGY MEV	CAPTURE RATE, SNU
$^{37}\text{Cl}$		
$\text{H} + \text{H} \rightarrow \text{D} + e^+ + \nu$	0-0.42	0
$\text{H} + \text{H} \rightarrow \text{D} + \nu$	1.42	0.23
$^7\text{Be}$ DECAY	0.86	1.02
$^8\text{B}$ DECAY	0-14	6.05
$^{13}\text{N}$ DECAY	0-1.20	0.08
$^{15}\text{O}$ DECAY	0-1.74	0.26
		<u>7.64</u>
		365
$^{71}\text{Ga}$		
		0
		67.2
		2.4
		28.5
		1.7
		2.7
		<u>3.8</u>
		106
		32

TOKS ELEMENT FOR ONE  
NEUTRINO CAPTURE/DAY:

Fig. 24

Dear Willy:

Thank you very much for sending me a copy of Dr. Rolfe's paper. Indeed, I have been worried about it, and looked over my old note book with a hope to see something extra. However, I just do not see

2) However, no matter what he gets, the serious discrepancy is in its slope.

I think that Rolfe's points should be plotted together with all previous points to see exactly how they compare, especially his  $\frac{ds}{dE}$  is the crucial factor.

At this point, I am not able to check all the details, on the other hand it is quite hard to imagine that we missed by a factor of 8. Especially our result was in good harmony with Parker's (I remember that Dworakowith and I joyfully shook hands when the first point came out and it was in line with Parker's, and said "you can't fool mother nature..."). On the other hand, it is quite possible we did fool. At any rate it is very nice that Rolfe (though I have not yet met him) is doing a beautiful work on this. It is indeed a time to redo it with a modern technique. Under the circumstance (having a 88" cyclotron alone), I could not do this here, but it would be a great service if additional experiments are independently pursued (do not mean to distrust anybody, but if a machine is available, this will be a challenging problem, in this

sense, all these proposals to NUSAC — including ours — won't be helpful, perhaps Charlie Barns makes sense) I can not deny that I'm very much worried considering if our result has been fooling these small  $\beta$ 's. But I shall remain so for now.

Thank you again, and will be looking forward hearing the result.

Sorry for my poor handwriting.

Sincerely yours,

Fig. 28

- 273 -

Ken Taylor

THE INSTITUTE FOR ADVANCED STUDY

PRINCETON, NEW JERSEY 08540

Telephone 609-734-8000

SCHOOL OF NATURAL SCIENCES

December 3, 1980


Dr. K. Nagatani  
Texas A & M University  
Cyclotron Institute  
College Station, Texas 77843

Dear Ken:

Thanks for your letter and the enclosed preprint. It is an excellent job.

What about getting a new  ${}^7\text{Be}(p,\gamma){}^8\text{B}$  measurement done? Its badly needed after so many years.

Best regards,

  
John W. Bahcall.

JNB:lu

Fig. 29

- 274 -

Scenarios of Galaxy Formation\*

Humitaka SATO

Research Institute for Fundamental Physics

Kyoto University, Kyoto 606

Abstract

Two different approaches to galaxy formation is contrasted: One is the primordial scenario based on particle physics and the another is the pregalactic scenario from astronomical side.

---

\* A short summary of the review talk given at the Workshop "GUT and Astrophysics" at KEK in December 7-10, 1983.

1. Invasion of particle physics to astronomy

In the last five years, the early universe has been explored in a close contact with particle physics and it has strengthened the belief that the history of the universe could be traced back to the Planck era of  $10^{-44}$ sec after the big bang. The breakthrough of this rapid progress was mainly due to the advancement of particle physics since the middle 1970's, particularly, the establishment of quark model and unified gauge theory.<sup>1)</sup> For example, we had not a definite image about the state of matter before  $10^{-4}$ sec, in which hadrons overlap each other, and the further trace before the nucleosynthesis era had been blocked by this hadron barrier.<sup>2)</sup>

Overcoming this hadron barrier by QCD, we realized that the state of matter is so simple to treat even in such early universe; it is just a mixture of ideal gases. Therefore, the theorists have now become bold enough to trace back the history up to the GUT era of  $10^{-35}$ sec, only because nothing has been known to block it. They have discussed baryon number synthesis,<sup>3)</sup> astronomical roles of relic particles,<sup>4)</sup> clustering of massive neutrinos,<sup>5)</sup> magnetic monopoles and so forth. They are still concerned mainly with the origin of matter, which would be still the realm of the particle physics or its colony.

However, encouraged by the success of the inflation scenario,<sup>6)</sup> the theorists working these subjects have started to try to explain the formation of galaxy in terms of quantum fluctuation. That is an invasion to the realm of astronomy and the contrast between the two approaches has become evident, that will be discussed later.

Anyway, the contributions of the particle physics to the

cosmology is already enormous. This approach to the study of the universe is like to the cartoon (Fig.), in which he is studying galaxies through a microscope.

## 2. Large scale structure and dark matter — favorite menus of particle cosmology —

What is the major observational facts for our problem of galaxy formation? First of all, the galaxies and their system are decoupled dynamically from the universal cosmic expansion and their distribution is almost uniform in the scales larger than 100Mpc. In the transient scale between the clumpy and the uniform scales, the cluster-void structure, which is called "large scale structure" recently, has become more evident.<sup>7)</sup> Further, the density correlation function takes a power law form such as  $\gamma^{-1.7}$  not only for the galaxy-galaxy but also for supercluster-supercluster.<sup>8)</sup> The dark matter, firstly found in clusters of galaxies, has been detected more clearly in the halos of spiral galaxies and, possibly, also in the dwarf galaxies, the companions of our galaxy.<sup>9)</sup> Then, the average density of matter clustered with sizes less than 100Mpc is estimated as  $\Omega=0.1 \sim 0.3$ , which is larger than the baryon density from the nucleosynthesis but smaller than  $\Omega=1$  predicted by the new inflation scenario.

The above facts are favorite menus which the particle cosmology wants to cook. But, what is overlooked sometime by it is the actively shining sources in the early stage like quasars, the background radiation in X-ray and NIR, the extrem old stars, the chemical abundances on them and so forth. The astronomical

approach would pay more attention to these later facts.

## 3. Who is crazy?

Peoples will regard him a bit crazy if somebody answered to a question "who is your ancestor?" that "he was a coacervate in some muddy shore". But, what the particle cosmology is trying to do about the galaxy formation is just this sort of answer. Such an attitude comes from the theoretical prejudice that we have known all the components necessary for this subject and the problem is merely how to cook them. Sometime, this sort of innocent boldness is very effective and should not be laughed away. Historically, the big bang cosmology was this case.

The astronomical approach tries to answer in more moderate manner. One seeks to find what is a pregalactic objects which generates galaxy. You may criticize it saying that the pregalactic object needs further pre-pre-galactic object and such an answer is not final. But, in this approach, they are seeking new components which might be indispensable for galaxy formation. They look to be polite enough not to regard themselves as they have known all the components already, but, in other sense, they are bold enough to introduce any new objects which are not forbidden by the present observation.

We call this approach as "pregalactic scenario" and the previous one of the particle cosmology as "primordial scenario". Recent works in this direction is the isothermal fluctuation scenario, the population III stars<sup>10)</sup> and the explosion dominated universe.<sup>11)</sup> All the nonstandard models such as the "cold" or "tepid" scenario are included in this category.

4. Is the microwave fluctuation absolute restriction?

The angular fluctuation in the microwave background is considered to give the restriction on the primordial fluctuation. According to the primordial scenario, this amplitude of the fluctuation,  $\delta_0$ , is a very fundamental quantity and might be explained by the materialization of quantum fluctuation in the inflationary stage. The amplitude at the horizon size is scale independent and same for all the components of matter. The motto of this scenario is to decrease the number of parameters as strict as possible. Since the observed upper limit has decreased so far, it seems difficult already that the primordial fluctuation reaches to the dynamical decoupling, without introducing invisible light particles heavier than electron neutrino, "warm matter". They are very generous to increase the number of parameter within the particle physics but not in the universe model.

In the pregalactic scenario, the angular fluctuation of the microwave background is not directly related to the density fluctuation which had generated the pregalactic objects. We have much freedom to do in the smaller sizes less than the angular resolution. This objects, however, might generate the additional background radiation in other wave bands and would affect the thermal history of the extragalactic medium. In this respect, the clouds observed by multiple Ly- $\alpha$  absorption lines may be interesting.

Between "the extrem primordial scenario" and the pregalactic scenario, there are many intermediate scenarios. Although the smallness of the angular fluctuation restricts the variety of possible models to some extent, there might be some clever loop-hole such as "blue sky" effect where the inhomogeneous

radiation has been isotropized by random scattering.

4. What gives dimension?

The primordial spectrum of the density fluctuation is subject to the modification by the filter process such as free-streaming of collisionless particles and photon diffusion. In the baryonsynthesis, the energy density fluctuation is succeeded to that of baryon number and the isothermal fluctuation will not come out without introducing inhomogeneities in CP-violation parameter or in the cosmic expansion rate.

The filter process introduces the critical length scale determined by microscopic physics. The smaller scale fluctuations have been erased up to this length, which is much larger than the average separation among galaxies.

This length is

$$\lambda = (ct)_{T=mc^2} \frac{mc^2}{T_0} \text{ for collisionless particle with mass } m$$

and

$$\lambda = (ct)_{T=T_{Dec}}^{1/2} \frac{T_{Dec}}{T_0} \text{ for nucleonic matter,}$$

where  $\ell = (n_e \sigma_{Th})^{-1}$  is mean free path for Thomson scattering and  $T_{Dec} = 4000K$  is the decoupling temperature. It has been pointed out that these lengths may explain the characteristic size of the large scale structure. However, we should keep in mind also another possibility that it is derived from  $\delta_0$  and the cosmic age  $t_0$  as

$$\lambda = ct_0 \left( \frac{\delta_0}{Z_{PF}+1} \right)^{1/2},$$

where  $Z_{PF}$  is the  $Z$  factor at the epoch of "pancake" formation. If this explanation is true, theory would be completely scale-free. Anyhow, it might be worthwhile to consider why these lengths are coincide approximately.

If the smaller size fluctuations have been erased in their linear stage, the formation of those smaller objects is possible through the fragmentation from the larger structures which have reached the nonlinear stage. The primordial scenario concludes this "fragmentation scenario", which is called also "adiabatic scenario" (because the adiabatic fluctuation is subject to the filter process) or "pancake scenario".

The another scenario in this respect is the "hierarchical clustering scenario", in which the smaller objects such as globular cluster or dwarf galaxy bind earlier. The computer simulation reproduces the correlation function pretty well in this scenario. Since the possible seed of the pregalactic objects is the isothermal density fluctuation, this belongs also to "isothermal scenario".

#### 6. Pancake or void

The large scale structure is characterized by the nonlinear behaviour of density fluctuation. Zeldovich pointed out in 1971 that the collapse tends to be unidirectional and gave an approximate solution which describes the behavior both in the expanding and collapsing phases. The positively perturbed region

evolves to "pancakes" or to "filaments", forming a network structure as a whole, those are supposed to fragment into multiple galaxies.<sup>12)</sup>

On the other hand, the negatively perturbed region or the void expands rapidly and pushes the surrounding medium to form a compressed blast wave. This is an example of the "pancaking" and is a complementary view of the same process.<sup>13)</sup> The expansion of this dense shell tends to be isotropized in different from the positive perturbation.

Formation of the dense expanding shell due to a snowplow mechanism is also expected to occur by the exploding stars in the early universe as discussed in the explosion dominant scenario.<sup>11)</sup> In terms of the explosion picture, the void is a "clean bomb", which explodes at the nonlinear epoch and with energy of  $Mc^2(\lambda/ct)^2$ . Because of a rapid expansion in the void, the light traversed it suffers the larger redshift.<sup>14)</sup>

#### 7. The first star

Whatever the dark matter consists of, the nucleonic component should evolve to ordinary stars. When was the first star formed? What was its mass spectrum? This problem is related to the thermal history of the pregalactic medium and vice versa. The formation of hydrogen molecules which work as a coolant is crucial process to determine the mass ranges of stars: the previous estimation gave a large mass like  $M \sim 10^4 M_\odot$ ,<sup>15)</sup> but the recent estimation has derived a low mass like  $0.1 M_\odot$ , adding the molecular formation process of a triple collision.<sup>16)</sup>

If the first stars were a low mass star, they would have continued to shine until now. However, since the number of low mass stars are limited observationally by the background radiation, they could not be the main part of the dark matter. If the first stars were formed at  $z \sim 10^2$  with mass of  $10^2 M_{\odot}$ , they would generate the near infrared background which might have been observed by the rocket observation.<sup>17)</sup> The accretion to the black holes may generate the X-ray background as well as the NIR background.

The chemical abundance seems to be almost normal even in such extragalactic objects as quasars, diffuse X-ray source, Ly- $\alpha$  cloud and so forth. The extreme population II stars contain the heavy element by  $10^{-4}$  of the normal. These informations will give an important clue to consider the first star formation and the mixing of the synthesized elements.<sup>17)</sup> The first star problem has been studied to some extent in the explosion dominant scenario and these processes should be worked also in the primordial scenario.

#### References

- 1) See for example, A. Zee, "Unity of Forces in the Universe", World Scientific, 1983.
- 2) What we had thought before this is written in Chap.2 of our paper: H. Sato, T. Matsuda and H. Takeda, Prog. Theor. Phys. Supplement No.49, 11 (1971). See also, H. Sato, Prog. Theor. Phys. 63, 1971 (1980).
- 3) M. Yoshimura, Phys. Rev. Lett, 41, 281 (1978).
- 4) K. Sato and H. Sato, Prog. Theor. Phys. 54, 912; 1567 (1975).
- 5) H. Sato and F. Takahara, Prog. Theor. Phys. 64, 2029 (1980). H. Sato, Proc. 10th Texas Sympo., P.43, N. Y. Academy Press (1981).
- 6) See for example, M. S. Turner, in "Proc. of 4th Workshop on Grand Unification" (E. Fermi Institute Preprint No.83-37).
- 7) C. S. Frenk, S. D. M. White, M. Davis, Ap. J. 271, 417 (1983).
- 8) N. A. Bahcall and R. M. Soneira, Ap. J. 270, 20 (1983).
- 9) S. M. Faber and D. N. C. Lin, Ap. J. 266, L17 (1983).
- 10) B. J. Carr and M. Rees, "Can Pregalactic Objects Generate Galaxies?" SISSA-Trieste Preprint 19/83/A. See also B. J. Carr et al, Ap. J. 277, 445 (1984).
- 11) S. Ikeuchi, K. Tomisaka and P. J. Ostriker, AP. J. 265, 583 (1983).
- 12) S. F. Shandarin, A. G. Dorshkevich and Ya. B. Zeldovich, Soviet Phys. USPEKHI 26, 46 (1983).
- 13) H. Sato, Prog. Theor. Phys. 68, 236 (1982); K. Maeda and H. Sato, ibid, 70, 1276 (1983); Y. Suto, K. Sato and H. Sato, ibid, 71, No.5 (1984).
- 14) H. Sato, "Effect of the Voids on Redshift" RIPP-555 (1984).



- 15) T. Matsuda, H. Sato and H. Takeda, Prog. Theor. Phys. 42, 219 (1969).
- 16) F. Palla, E. E. Salpeter and S. W. Stahler, Ap. J. 271, 632 (1983).
- 17) B. J. Carr, J. McDowell and H. Sato, Nature, 306, 666 (1983).
- 18) B. J. Carr and S. Ikeuchi, RIFP-550 (1984).



Galaxy Formation Theory and Large Scale Structure  
in the Universe

S. Ikeuchi

Department of Physics, Hokkaido University,  
Sapporo 060, Japan

ABSTRACT

Two scenarios of galaxy formation are presented: one is the galaxy formation triggered by explosions of pregalactic objects, and the other is the growth of voids and surrounding ridges in a neutrino dominated universe. These scenarios are discussed in relation to the formation of large scale structure in the universe. Some related problems on intergalactic medium are discussed briefly.

---

\*) Talk presented at the Workshop on Grand Unified Theories and Cosmology,  
National Laboratory for High Energy Physics, Japan, December 7 - 10, 1983.

§1. Introduction

Till the end of '70s, two theories of galaxy formation are competitive: one is the fragmentation theory, which has been mainly developed by Zeldovich and his collaborators<sup>1)</sup>, and the other is the clustering theory, which has been studied by the western group<sup>2)</sup>. Both theories have fatal difficulties. In the former, the adiabatic perturbations of matter are assumed at the pre-decoupling era, so that the background radiation is also perturbed. In order to make galaxies at  $1+z \geq 5$ , the amplitude of density perturbation must be

$$\frac{\Delta\rho}{\rho} \geq 10^{-2}. \quad (1)$$

On the other hand, the fluctuation of micro-wave background radiation is limited to<sup>3)</sup>  $\Delta T/T \leq 10^{-4}$ , which gives the upper limit to the density perturbation of nucleonic matter

$$\frac{\Delta\rho_b}{\rho_b} = 3 \frac{\Delta T}{T} \leq 3 \times 10^{-4}. \quad (2)$$

Then, if  $\rho = \rho_b$ , the above two conditions (1) and (2) contradict with each other. The latter clustering theory failed to predict and reproduce the honeycomb structure of galaxy distribution in the universe<sup>4),5)</sup>. In this theory, the long wavelength perturbation corresponding to the honeycomb structure must be assumed in the initial condition.

In the present paper, we propose two scenarios of galaxy formation which, respectively, resolve the difficulties of above two theories, and the relations to the formation of large scale structure are discussed. At present, the galaxy formation theory must be the theory of the space devoid of galaxies (i.e., voids). As is seen in the later, galaxies are thought to be formed through fragmentation for both scenarios, even if the starting conditions are completely different.

§2. Growth of Density Contrast in a Neutrino Dominated Universe <sup>6)</sup>

To avoid the contradiction of conditions (1) and (2) in the adiabatic theory, it is natural to assume  $\rho \neq \rho_b$ . That is, the gravitating mass ( $\rho$ ) is not the nucleonic mass  $\rho_b$  (hereafter, we refer baryons). This assumption is supported from the following two facts. One is that the neutrinos may have a finite mass <sup>7)</sup> such as  $10 \sim 30$  eV. If it is true, the mass density of the universe is almost occupied by neutrinos, neutrino dominated universe. In order to form the astrophysical objects (galaxies) before  $1 + z > 5$ , the density perturbation of neutrinos must satisfy the condition (1), while that of baryons satisfies only the condition (2). The apparent contradiction is resolved.

The other is that the missing mass increases with the mass of astronomical objects. This fact indicates that the missing mass seems to be uniformly distributed and it dominates the gravity of baryons.

Then, we assume: (i) the mass densities of neutrinos and baryons are, respectively,  $\Omega_\nu = 0.9$  and  $\Omega_b = 0.1$ , (ii) at the decoupling epoch the perturbation of baryons is set as zero while that of neutrinos is taken finite, and (iii) the neutrinos are collisionless particles which can pass through freely, and baryons are treated by hydrodynamics.

In reality, at the decoupling epoch  $1 + z_d = 10^3$ , the radial Lagrangian coordinate of neutrinos is perturbed as

$$r_{\nu,i} = r_{\nu,i,0} (1 + \delta_{\nu,i}), \quad (3)$$

and its corresponding perturbed velocity becomes

$$v_{\nu,i} = H_i r_{\nu,i,0} (1 + 2\delta_{\nu,i}). \quad (4)$$

The functional form of perturbation is taken as <sup>8)</sup>

$$\delta_{\nu,i} = \begin{cases} \frac{\epsilon_\nu}{3} \cos^2 \left[ \frac{\pi}{2} \left( \frac{r_{\nu,i,0}}{R_{\nu,i}} \right) \right] & \text{at } r_{\nu,i,0} \leq R_{\nu,i} \\ 0 & \text{at } r_{\nu,i,0} \geq R_{\nu,i}. \end{cases} \quad (5)$$

From the mass conservation condition within the radius  $R_{\nu,i}$ ,

$$\int_0^{R_{\nu,i}} 4\pi r_{\nu,i}^2 \rho_{\nu,i} dr_{\nu,i} = 4\pi R_{\nu,i}^3 \bar{\rho}_{\nu,i} / 3, \quad (6)$$

we can obtain the initial density distribution  $\rho_{\nu,i}$ . In the above, the fact that the missing mass is due to neutrinos appears as the size of perturbation

$$R_{\nu,i} = 3.75 \times 10^{-4} m_\nu^{-1} (1 + z_d)^{-1} \text{ Mpc}, \quad (7)$$

where  $m_\nu$  is the neutrino mass in eV. It is to be notified that after the cosmic expansion the size of this neutrino perturbation becomes equal to that of a void at present,

$$R_\nu(0) = R_{\nu,i} (1 + z_d) = 12.5 (30 \text{ eV}/m_\nu) \text{ Mpc}. \quad (8)$$

Starting from the above initial condition, we investigate the evolution of density perturbations of baryons and neutrinos. In Figure 1, the density distributions of baryons (solid lines) and neutrinos (dotted lines) are illustrated in the comoving coordinate for several stages in the case of the initial amplitude  $\epsilon_\nu = 0.07$  of Eq. (5). During the epoch  $1 + z = 10^3 \sim 50$ , the baryons are dragged by the gravitational force of neutrinos. At the stage,  $1 + z = 20$ , both density distributions become almost equal to each other. At the stage  $1 + z = 11$ , the shell crossing of neutrinos occurs. In order to see this shell crossing in details <sup>9)</sup>, we show the radial distributions of expansion velocity  $V$ , its deviation from the Hubble law  $(V/Hr - 1)$  and the gravitational force  $|GM_\nu/r^2|$  in Figs. 2 and 3.

At the stage  $1 + z \approx 25$ , it is recognized that the reversal of velocity distribution near the dense shell occurs. Due to the strong gravitational force at the shell, the expansion is decelerated just outside the shell. On the other hand, inside the shell the matter overtakes the shell. As a result, the dense shell becomes more concentrated. Meanwhile, the neutrinos of inner region

can go ahead of those of outer region. This is the shell crossing.

After the first shell crossing, neutrinos do experience the shell crossing several times as is seen in Fig. 1 (c). On the other hand, baryons can not pass through each other. When neutrinos make shell crossings, the baryons are enforced to collide violently from the rear-end and their kinetic energy is dissipated to increase the baryon temperature quickly. In Fig. 4, we illustrate the time variations of temperature  $T(r_m)$  at the radius  $r_m$ , where the baryon density is maximum  $\rho_{b,max}$ , and the neutrino density at there, as well as  $T(r_m)$  for the case  $\epsilon_v = 0.03$ .

It is interesting that the baryons can be heated up naturally at the epoch  $1 + z \approx 10$  without any energy injection mechanisms. As the second point, we can claim that after the shell crossing the baryon density  $\rho_{b,max}$  always exceeds the neutrino one  $\rho_v(r_m)$ . This means that it is possible for baryons to exceed neutrinos locally.

As is seen in Figs. 1(b) and (c), a very thin layer of baryons and a fairly thick layer of neutrinos are formed after the shell crossing. This shell will ultimately fragment to pieces. Here, according to Ostriker and Cowie (1981) we estimate the mass of a fragment, which is gravitationally unstable. A pancake cut out of the shell, with radius  $a$ , has a kinetic, gravitational and internal energy per unit mass, which are, respectively, written as

$$E_K = \frac{1}{2} \left(\frac{a}{R_S}\right)^2 v_S^2, \quad E_G = -\kappa a G \Sigma_t \quad \text{and} \quad E_I = \frac{3}{2} \frac{\Sigma_I}{\Sigma_t}, \quad (9)$$

where  $\Sigma_t$  is the total surface mass density and  $\Sigma_I$  is the surface internal energy. If the total specific energy  $E = E_K + E_G + E_I$  is negative, this pancake is unstable and the shell can fragment into many pancakes. The condition  $E < 0$  is generally satisfied at the range  $a_1 < a < a_2$  and the energy becomes minimum at

$$a_m = \kappa \pi G \Sigma R_S^2 / v_S^2. \quad (10)$$

The characteristic masses of this pancake are calculated as

$$M_v = \pi a_m^2 \Sigma_v, \quad M_b = \pi a_m^2 \Sigma_b, \quad (11)$$

where  $\Sigma_v$  and  $\Sigma_b$  are the surface mass density of neutrinos and baryons. In Figure 5, we plotted the masses  $M_v$  and  $M_b$  as well as  $M_{t1} = \pi a_m^2 (\Sigma_v + \Sigma_b)$ . As is seen, if we take the size of the present void  $R_v(0) = 10$  Mpc, the mass of a fragment is in the range  $M_t = M_v + M_b = 7 \times 10^{11}$  to  $1.1 \times 10^{12} M_\odot$ . Also, it is to be noticed that this characteristic mass does not depend upon the epoch of shell fragmentation.

In Figure 6, we illustrate the shell crossing epoch  $1 + z_{cross}$  and the density ratio of the void to the average  $\rho_{b,c}(0) / \bar{\rho}_b(0)$  versus the initial amplitude of perturbation  $\epsilon_v$ . The case with  $\Omega_v(0) = 0.1$  and  $\Omega_b(0) = 0.05$  is also added for comparison. As is seen, the following simple relations are obtained

$$\begin{aligned} (\Omega_v = 0.9, \Omega_b = 0.1) \\ 1 + z_{cross} = 17 \left(\frac{\epsilon_v}{0.1}\right), \quad \frac{\rho_{b,c}(0)}{\bar{\rho}_b(0)} = 2 \times 10^{-3} \left(\frac{\epsilon_v}{0.1}\right)^{-1.4}, \end{aligned} \quad (12a)$$

$$\begin{aligned} (\Omega_v = 0.1, \Omega_b = 0.05) \\ 1 + z_{cross} = 10 \left(\frac{\epsilon_v}{0.1}\right)^{1.9}, \quad \frac{\rho_{b,c}(0)}{\bar{\rho}_b(0)} = 3 \times 10^{-2} \left(\frac{\epsilon_v}{0.1}\right)^{-1.12}. \end{aligned} \quad (12b)$$

If the epoch of galaxy formation is later than the shell crossing time, the condition for the galaxy formation  $z > 4$  gives the severe lower limit to the amplitude  $\epsilon_v \geq 3.2 \times 10^{-2}$ .

Summing up the above results, we may conclude that if we consider the neutrino mass is in the order of  $20 \sim 30$  eV the neutrino perturbation with the amplitude  $\epsilon_v \geq 0.03$  can well explain (i) the size of a void, (ii) the mass of a supercluster of galaxies, (iii) the mass of a galaxy and (iv) the missing mass.

### §3. Hierarchical Explosion Scheme

This unfamiliar name of a galaxy formation theory is recently named<sup>11)</sup> but its fundamental idea were presented several years ago<sup>10),12)</sup>. In the above neutrino dominated universe, the fluctuations of adiabatic form (= isentropic) are presupposed. In this case, the minimum non-dissipated mass is  $\sim 10^{13} M_{\odot}$ . On the other hand, another possible form of fluctuations is isothermal. Even before the decoupling era, the isothermal perturbation does not affect the microwave background radiation. This isothermal perturbation can rapidly grow after the decoupling at the Jeans mass region  $M_J \sim 10^6 M_{\odot}$ . The western clustering theory proposes that these  $10^6 M_{\odot}$  objects might form galaxies by their mutual gravitational clustering. This idea is completely denied by the discovery of the honeycomb structure.

However, it does not mean the first object in the universe was not the  $10^6 M_{\odot}$  object. We can suppose the following different scenario of galaxy formation. The  $10^6 M_{\odot}$  object would fragment to massive stars after collapse. The lifetime of these massive stars is as short as  $\sim 3 \times 10^6$  yr and they explode as supernovae at their final stage of evolution. As a whole, the energy of  $10^{55} \sim 10^{56}$  ergs is ejected and it induces the shock wave in the ambient medium.

With expanding the shock wave, the cooling processes (Compton cooling at  $z > 10$  and radiative cooling at  $z < 10$ ) become efficient at the shock front, so that the dense cooled shell is formed. Even if the cooling processes are not efficient, the matter is swept and concentrated to the shock front in the Einstein - de Sitter universe<sup>13)</sup>.

By examining the energy of a fragment of the shell, like Eq.(9), the most probable mass  $M_2$  of the objects which can be expected after the shell fragmentation is estimated. If this mass  $M_2$  is greater than the mass of the first object  $M_1 = 10^6 M_{\odot}$ , we may say the mass scale is amplified by explosions.

As same as the first objects, we may expect these second objects also

fragment to many stars, which explode as a whole to make shock waves. The total mass  $M_{s,1}$  swept to the shock front around the first object can make the second objects about  $n_2 = M_{s,1} / M_2$  ( $\approx 300 \sim 3000$ ). This brings more violent explosions by a factor of  $n_2 \times (M_2 / M_1) = (M_{s,1} / M_1)$ .

Similar to the above, the third objects will be formed by the fragmentation of the dense shell at the shock front. If the mass of the fragment  $M_3$  is higher than  $M_2$ , we can expect the above process to repeat. In this way, after several generations of objects, formed and exploded, galaxies would be formed as the n-th generation objects. Therefore, we name this model as Hierarchical Explosion Scheme (HES)<sup>11)</sup>.

HES becomes invalid, when either of following two conditions is not satisfied. One is the amplification condition,  $M_n > M_{n-1} > M_{n-2} \cdots > M_1$ . If this condition is not satisfied, we can not make galaxies from a smaller ( $10^6 M_{\odot}$ ) first object. The other is the fragmentation condition,  $E = E_K + E_G + E_I < 0$ . If the explosion energy is too high or the ambient gas density is too low, the energy of any fragment is always positive, and the shell can not be gravitationally unstable.

After the galaxies are formed at  $z \approx 5$ , the latter condition can not be satisfied, because of too rarefied ambient density or overlappings of shock waves. As a result, the HES stops.

Generally speaking, the void of galaxies corresponds to the cavity of a shock wave, and if it is  $\sim 20$  Mpc the explosion energy of  $\sim 10^{61}$  ergs must be supplied. The formation of void and surrounding ridge of galaxies by the energy injection of this order just corresponds to the formation of density fluctuation with amplitude 0.05 at the decoupling epoch. That is, from the viewpoint of formation of a void and cluster of galaxies the above two scenarios completely agree with each other.

#### 4. Discussion

We present two scenarios for the formation of galaxies starting from the adiabatic and isothermal fluctuations. In the final stage, two scenarios become similar, i.e., fragmentation of dense shells into galaxies. Here, we summarize the problems which are not resolved until now.

##### (a) Adiabatic Theory in a Neutrino Dominated Universe

The most important and unclarified problem is whether the density fluctuations of baryons can not be induced by the gravitational force of neutrinos at the pre-decoupling era. Generally, the fluctuations of baryons are to be erased by collisions of isotropic photons. However, the collision time is finite and several times of collisions are necessary. Therefore, a large scale and a relatively high amplitude fluctuation can not be erased. If so, the present scenario fails.

Similarly, at the decoupling era the smaller fluctuations of baryons than the mean free path of photons are thought to be dissipated away. From this condition, the minimum mass above which the clumps can survive is determined (Silk mass). It is necessary to treat accurately these photon-baryon interactions in the expanding and decoupling era.

In the next step, two- and three-dimensional calculations considering both the collisionless and collisional particles are necessary. At present, the usual N-body calculations can not succeed in reproducing the epoch of galaxy formation and the observed correlation function at the same time. The author believes that the hydrodynamical treatment for the formation process of galaxies is necessitated.

##### (b) Hierarchical Explosion Scheme

The most fatal problem of this model is "not fully examined in discussing the details". Especially, the following problems should be urgently made clarified. (i) How many generations are necessary to make galaxies successfully?

In that case, is the universe covered by shock waves? (ii) Whether or not is the universe polluted by heavy elements more than pop. II stars due to these hierarchical explosions? And (iii) at the same time does a huge energy input deform the blackbody spectrum of microwave background radiation?

In addition to the above problems, it is necessary to make simulations of galaxy formation and formation of large scale structures in an expanding universe. Different from the N-body simulations, we may have other constraints to our scenario.

The author would like to thank Mr. M. Umemura for his collaboration of the first half of this paper.

References

1. For a review, S. Shandarin, A. Doroshkevich and Ya.B. Zeldovich, Soviet Phys. - Uspekhi, 26 (1983) 46.
2. S. Aarseth, J. Gott and E. Turner, Ap. J. 228 (1979) 664.
3. R. Partridge, Physica Script. 21 (1980) 614.
4. R. Kirshner, A. Oemler, P. Schechter and S. Schectman, Ap. J. 248 (1981) L57.
5. Ya.B. Zeldovich, J. Einasto and S. Shandarin, Nature 300 (1982) 407.
6. S. Ikeuchi and M. Umemura, submitted to Prog. Theor. Phys. (1984, March).
7. V. Lyubimov, E. Novikov, V. Nozik, E. Tret'yakov and V. Kozik, Phys. Lett. B 94 (1980) 266.
8. P. Peebles, Ap. J. 257 (1982) 438.
9. E. Bertsinger, submitted to Ap. J. Suppl. (1984, Jan.).
10. J. Ostriker and L. Cowie, Ap. J. 243 (1981) L127.
11. B. Carr and S. Ikeuchi, in preparation.
12. S. Ikeuchi, Publ. Astron. Soc. Japan 33 (1981) 211.
13. S. Ikeuchi, K. Tomisaka and J. Ostriker, Ap. J. 265 (1983) 583.

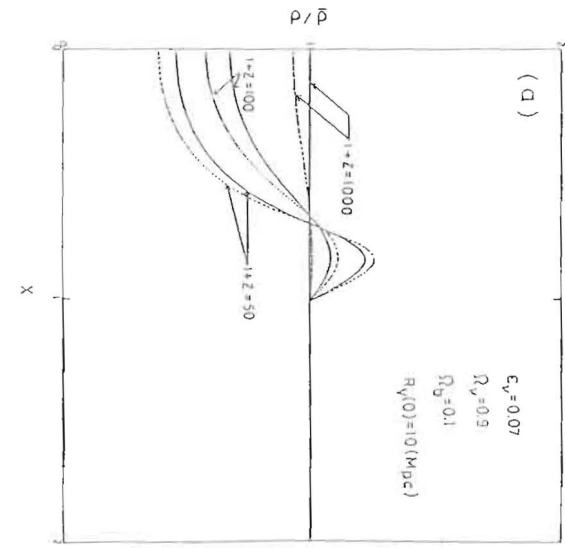


Fig. 1 (a)

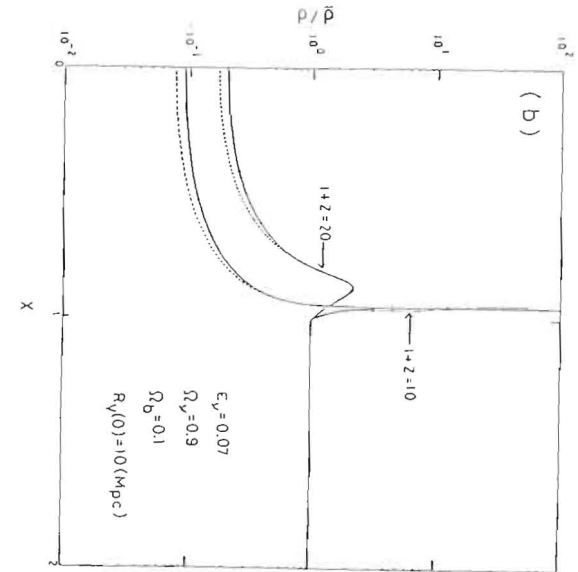


Fig. 1 (b)

Fig. 2

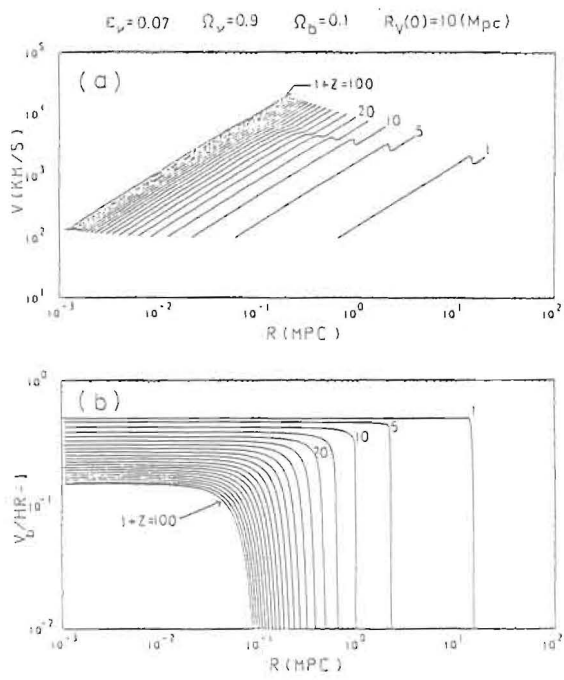


Fig. 3

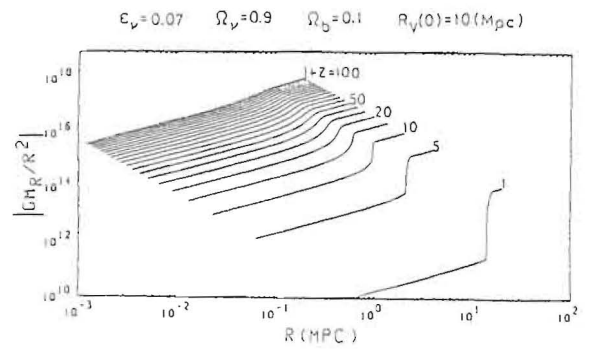


Fig. 1 (c)

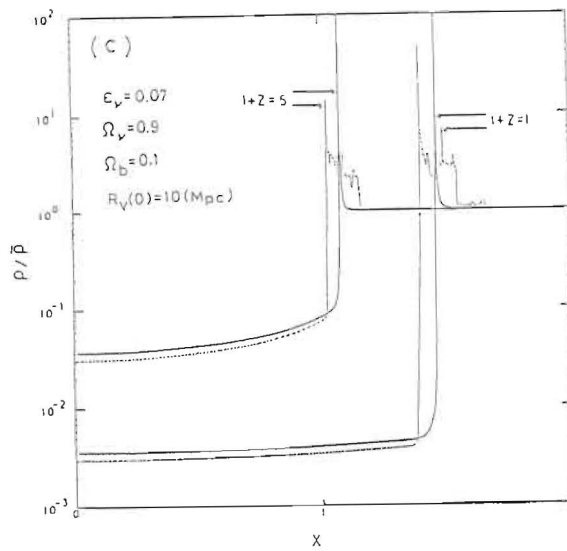




Fig. 6

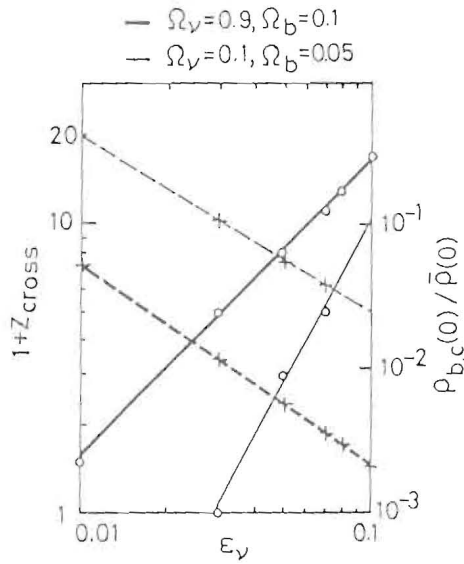


Fig. 4

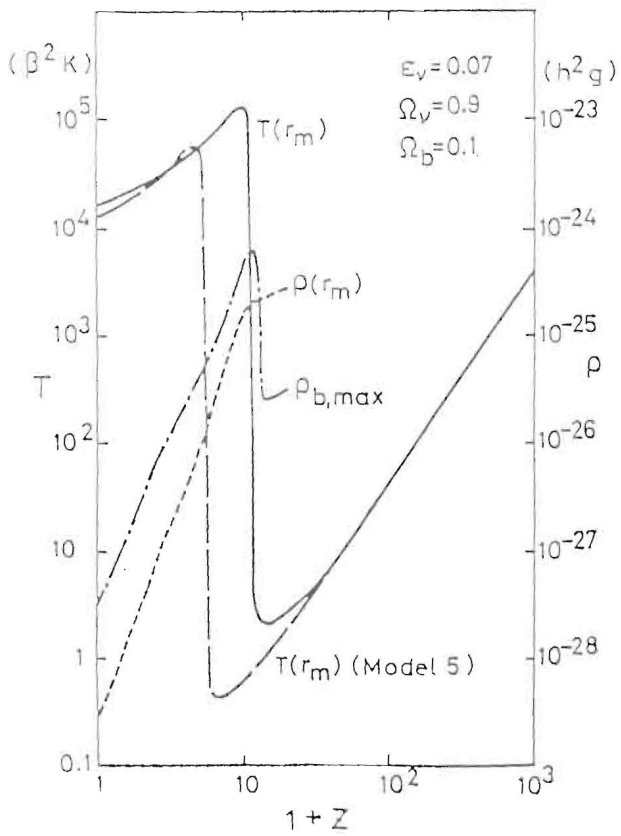
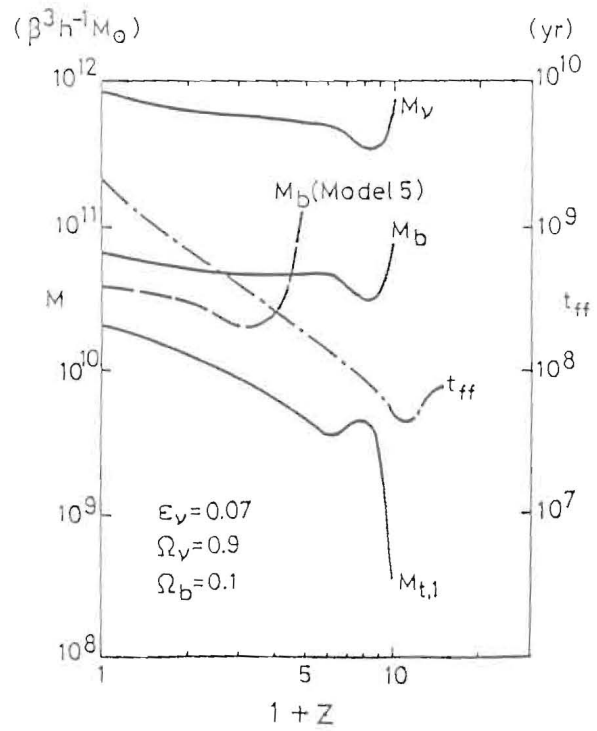


Fig. 5



# Inflationary Universe Model

Katsuhiko Sato

Department of Physics, Faculty of Science

University of Tokyo, Tokyo 113 Japan

## ABSTRACT

Recent studies of Inflationary universe model is reviewed. In § 1 essential idea and the merits are introduced. In § 2 details of the original inflationary model and its difficulties are shown. In § 3 the new inflationary model, which Linde and Albrecht et al proposed in order to avoid the difficulties of the original inflationary model, and the new versions are reviewed. the final section § 4 is devoted to remarks.

## § 1 First order phase transition and inflationary universe model

### 1) Exponential Expansion

About three years ago, an universe model was proposed, independently, by the present speaker<sup>1)-3)</sup> and Guth<sup>4),5)</sup>. In this model the universe expands exponentially in a period of phase transition if the GUT phase transition is of strongly first order (see, for example, reviews<sup>6),7)</sup>). This model is now called Inflationary universe model, which is named by Guth skillfully. As shown in Fig.1.1.1, the vacuum stays at the symmetric state  $\phi=0$  for a while even if the cosmic temperature decreases to lower than the critical temperature. This false vacuum  $\phi=0$  decays into the stable vacuum  $\phi_0$  by quantum fluctuations or by thermal fluctuations. Obviously if the potential barrier between two vacua  $\phi=0$  and  $\phi_0$  is high and the transition time-scale is much longer than that of cosmic expansion, the energy density of the vacuum becomes dominant and changes the expansion law of the universe, because the radiation energy density decreases with cosmic expansion, but the vacuum energy density  $\rho_v$  is a constant.

If we assume that our universe is homogeneous and isotropic, the metric takes the form

$$ds^2 = -dt^2 + R(t) \left( d\chi^2 + \begin{matrix} \sin^2 \chi \\ \chi^2 \\ \sinh^2 \chi \end{matrix} d\Omega^2 \right), \quad (1.1.1)$$

$$\text{for } k = \begin{cases} +1 & \text{;closed universe} \\ 0 & \text{;flat universe} \\ -1 & \text{;open universe} \end{cases} .$$

Then time evolution of a cosmic scale factor of the universe  $R(t)$  is

described by the following equation

$$\ddot{R}^2/R^2 + k/R^2 = 8\pi G(\rho_v + \rho_r(t))/3, \quad (1.1.2)$$

where a dot denotes time differentiation and  $\rho_v$  is the energy density of the false vacuum. The radiation energy density  $\rho_r(t)$  is given by

$$\rho_r(t) = N_b \pi^2 T^4(t)/30, \quad (1.1.3)$$

where  $N_b$  is the statistical weight of radiations before the phase transition in units of scalar bosons. We assume, for simplicity, the vacuum energy density is a constant and is independent of time before the phase transition. Now we introduce a characteristic scale factor of this model;

$$\ell = (8\pi G \rho_v/3)^{-1/2} \quad (1.1.4)$$

and a non dimension numerical factor which represents the "size" of the universe;

$$b = \rho_r(t) R^4(t)/\rho_v \ell^4, \quad (1.1.5)$$

which remains constant when the universe evolve adiabatically. Then Eq.(1.1.2) is rewritten as

$$\dot{x}^2/2 + V(x) = -k/2 (=E), \quad (1.1.6)$$

where  $x = R(t)/\ell$  and a dot denotes differentiation with respect to normalized

time  $\tau = t/\ell$ . The potential

$$V(x) = -(x^2 + b x^{-2})/2 \quad (1.1.7)$$

is displayed in Fig.(1.1.2). From the analogy of a particle motion in the potential, two types of the solution are obtained easily: The first type is

$$(R(t)/\ell)^2 = \begin{cases} b^{1/2} \sinh(2t/\ell) & \text{for } k=0, \\ k/2 + (b-1/4)^{1/2} \sinh(2t/\ell) & \text{for } k \neq 0, b > 1/4, \\ k/2 + \exp(2t/\ell) & \text{for } k \neq 0, b = 1/4, \\ k/2 + (1/4-b)^{1/2} \cosh(2t/\ell) & \text{for } k \neq 0, b < 1/4, \end{cases} \quad (1.1.8)$$

This is the solution that the universe can expand infinitely. The second type solution is

$$(R(t)/\ell)^2 = \begin{cases} 1/2 & \text{for } k=+1, b=1/4 \\ (1 - \exp(-2t/\ell))/2 & \text{for } k=+1, b=1/4 \\ 1/2 - (1/4-b)^{1/2} \cosh(2t/\ell) & \text{for } k=+1, b < 1/4 \end{cases} \quad (1.1.9)$$

This is the solution that the universe has a maximum scale factor which is less than  $\ell/2$ . Obviously, this type of solution can only exist for  $k=+1$  (see, Fig.1.1.2). It seems, therefore, hard to take the second type solution as a realistic present universe model, because the length  $\ell$  is an order of  $10^{-26}$  cm if we take a plausible value as for the vacuum energy density  $\rho_v = (10^{15} \text{ GeV})^4$ . Even if we take the first type solution, the situation is very different according whether  $b < 1/4$  or  $b > 1/4$ . In the limit

$b \gg 1/4$  or the case  $k=0$ , the Eq.(1.1.8) reduces to

$$R(t) = b^{1/4} \sinh^{1/2}(2t/\ell) = \begin{cases} b^{1/4} (2t/\ell)^{1/2} & ; t \ll \ell/2 & (1.1.10a) \\ b^{1/4} 2^{-1/2} \exp(t/\ell) & ; t \gg \ell/2 & (1.1.10b) \end{cases}$$

independently of the value of  $k$ . As seen from Eq.(1.1.10a), the universe expands as  $t^{1/2}$  in the standard big bang model when  $t \ll \ell/2$ , because the radiation energy density is dominant in this period. After the cosmic time  $\ell/2$ , however, the universe begins to expand exponentially because the vacuum energy density becomes to be dominant. This exponential expansion continues until the vacuum energy density disappears by the termination of the phase transition. Recall also that the temperature before the phase transition varies as  $T \sim 1/R$ , so the time-temperature relation can be written as

$$T = T_v N_b^{-1/4} \sinh^{-1/2}(2t/\ell) , \quad (1.1.11)$$

where  $T_v$  is defined by the relation

$$\pi^2 T_v^4/30 = \rho_v . \quad (1.1.12)$$

Evidently,  $T_v$  may be interpreted as the temperature equivalent of the vacuum energy density if this were transformed into massless, spin zero boson radiation.

When the phase transition finishes, the energy density between two vacua (the false and the true vacua)  $\rho_v$  is released and the universe is strongly heated up if the energy is thermalized. Then, the temperature immediately after the phase transition  $T_a$  is given as

$$T_a = (30 \rho_v / \pi^2 N_a)^{1/4} = N_a^{-1/4} T_v . \quad (1.1.13)$$

This temperature is higher than the Higgs boson mass but lower than the lepto-quark boson mass. This suggests that baryon number asymmetry is generated by Higgs boson decay as in the standard scenario. The subsequent evolution of the universe after the phase transition in this inflation model is essentially the same as that of the standard big bang model: The cosmic scale factor after the phase transition is given by

$$R(t) = R(t_f) (2(t-t_f)/\ell + 1)^{1/2} \\ = b^{1/4} \exp(t_f/\ell) (\ell t)^{1/2} , \text{ for } t > t_f > \ell , \quad (1.1.14)$$

where  $t_f$  is the termination time of the phase transition.

## 2) Horizon Problem

One of the most important differences between the standard model and the inflation model is the size of a particle horizon, which is defined as<sup>8)</sup>

$$r_H(t) = R(t) \int_0^t dt' / R(t') \quad (1.2.1)$$

This is the maximum radius of a region in which causal relations can exist at the cosmic time  $t$ . In the standard model ( $R \sim t^{1/2}$ ), the horizon is  $r_H(t) = 2t$  and very small compared with the scale factor of the universe in the early universe. This means most of the regions of the universe have never been in causal relation in spite of the homogeneity of the universe. On the other hand, the horizon in the inflation model, which is calculated as

$$r_H(t) = \begin{cases} F(\cos^{-1}(\exp(-t/\ell), 2^{-1/2}) \sinh^{1/2}(2t/\ell) & ; t \ll t_f \\ F(\pi/2, 2^{-1/2}) \exp(t_f/\ell) (t/\ell)^{1/2} + 2t & ; t \gg t_f \end{cases} \quad (1.2.2)$$

can become very large, where  $F(\cdot, \cdot)$  is the elliptic integral of the first kind and the value  $F(\pi/2, \sqrt{1/2})$  is about 1.85. As shown in Fig.1.2.1, the particle horizon also increases exponentially as the scale factor increases, and becomes very larger than that of the standard big bang model (SBBM). As is well known from the observation of the cosmic background radiation, the universe is surprisingly homogeneous and isotropic. This fact suggests that a homogenizing process had worked in the very early universe. In the SBBM, however, this process could never work in principle, because the particle horizon is very small compared with the size of the universe at and before the recombination time. In this review, we define the size of the universe  $r(t)$  at cosmic time  $t$  as the proper length of the region which we know observationally at present, i.e.,

$$r(t) = t_0 R(t)/R(t_0) = R(t)/(R(t_0)H_0), \quad (1.2.3)$$

where  $t_0$  is the present cosmic age and  $H_0$  is the Hubble constant. If we assume that the universe expands as  $R \propto t^n$  ( $n=1/2$  for radiation dominant universe, i.e.,  $t < t_{eq}$ ,  $n=2/3$  for matter dominant universe, i.e.,  $t > t_{eq}$ ,  $t_{eq}$  being the time when the universe becomes matter-dominant.) the ratio of the particle horizon to the size of the universe is calculated from Eq.(1.2.1) and Eq.(1.2.3) as

$$\begin{aligned} r_H(t)/r(t) &= 2t_0^{-1/3} t_{eq}^{-1/6} t^{1/2} = 2(t_{eq}/t_0)^{1/3} (t/t_{eq})^{1/2} \\ &= 10^{-2} (t/t_{eq})^{1/2}, \text{ for } t \ll t_{eq} \end{aligned} \quad (1.2.4)$$

This shows that a region which could be causally related by that time  $t$  becomes smaller and smaller as we go back to the early universe. The question why the universe was very homogeneous over such scales which were completely causally unrelated is called horizon problem. Obviously this difficulties is removed in the present inflation model, because the horizon can become very large. If we assume that the finishing time  $t_f$  is later than the cosmic time  $62\ell$ , the regions which were in the particle horizon and were homogenized by some causal processes before phase transition  $t = \ell/2$ , can expand as large as the size of the universe which we know now observationally ( $\sim 1/H_0$ ) (see, Fig.1.2.1 and Eq.(1.2.2)).

### 3) Generation of seeds of galaxies

Although the universe is very homogeneous in the very large scale ( $>100\text{Mpc}$ ), there exist large scale inhomogeneities, i.e., galaxies, cluster of galaxies and super cluster of galaxies. According to conventional theories of galaxy formation,<sup>8)</sup> these large scale structure were formed by the growth of the density fluctuation which existed already in the early universe. It is now generally accepted that the amplitude of the density fluctuations with the masses of galaxies or cluster of galaxies must be an order of  $10^{-3}$ , i.e.,  $\delta\rho/\rho \sim 10^{-3}$ , when the wave length of the fluctuations is equal to the particle horizon. This means that the fluctuations whose wave length is larger than the horizon must be formed in the early universe. How these density fluctuations were formed? Zeldovich<sup>9)</sup> investigated in detail whether density fluctuations whose wave length are larger than the particle horizon are generated by causal processes (the energy and momentum are conserved) or not. He showed that even if the fluctuations are formed, the spectrum of the fluctuations is  $M^{-7/6}$ , which is too steep to explain the origin of large scale structure in the standard big bang model, where  $M$  is the mass of the fluctuation. This is because the creation of fluctuations from homogeneous state is strongly limited by the existence of the particle horizon, i.e., in order to create a large scale in homogeneity, the energy or momentum must be transferred over the horizon. It is obvious that this difficulty is removed in the inflation model, because whole universe which we now observe was in

the particle horizon before or at the phase transition i.e.,  $t \leq \ell/2$ . We may, therefore, conclude that global homogeneity and the large scale inhomogeneities could be formed at this time from the view point of causality, in principle. Of course, in order to construct a scenario of the generation and growth of fluctuations, we must investigate the physical processes before or at the phase transition in detail. It seems, however, very possible that fluctuations were formed by GUT phase transition.<sup>2)</sup>

#### 4) Monopole problem

Magnetic monopoles are relic of the cosmological phase transition which is predicted by GUTs. As discussed by Kibble, a monopole is formed at a point which is surrounded by four domains if the winding number is nonvanishing (see, Fig.1.4.1). Kibble estimated the number of monopoles produced by the cosmological phase transition as

$$n_m = p_G / d_c^3, \quad (1.4.1)$$

where  $p_G (\sim 1/8)$  is the group theoretical probability that the winding number is non vanishing (unity) and  $d_c$  is the radius of the domains or the coherent length of the Higgs field. This coherent length could become infinite and no monopoles were formed if the universe cools infinitely slowly keeping the thermal equilibrium. Because the universe expands and cools very rapidly in the early universe, monopoles which are topological defects of vacuum are formed abundantly as lattice defects are formed in crystal when melted material is cooled very rapidly. Now we take the particle horizon  $d_H$  in the SBBM as for this domain radius,

$$d_H(t) = 2t = (45 / \pi^3 N)^{1/2} m_p / T^2. \quad (1.4.2)$$

Then the monopole /entropy ratio just after the phase transition is given by

$$r(t_c) = p_G (\pi^2/2) (\pi N/45)^{1/2} (T_c/m_p)^3, \\ \sim (T_c/m_p)^3 \sim 10^{-12} \quad (1.4.3)$$

where we assumed  $p_G=1/8$ ,  $N=100$  and  $T_c=10^{15}$  Gev. Because the particle horizon is the upper limit of the radius of domains, Eq(1.4.3) gives the lower limit of the monopole/entropy ratio. As pointed out by Preskill<sup>10)</sup> this value is extremely larger than the upper limit obtained from the energy density of the present universe, i.e.  $n n_M(t_0) \leq \rho_0$ , where  $m (\sim 10^{16}$  Gev) is the monopole mass and  $n_M(t_0)$  is the number density of monopoles at present universe and the upper limit of the energy density  $\rho_0$  is about  $4 \times 10^{-29}$  g/cm<sup>3</sup>. This overproduction of monopoles in the universe is called as monopole problem.

Obviously, inflation model, in which the particle horizon is very longer than that in the SBBM, give a possible solution to the monopole problem<sup>11)-12)</sup>: If we take Eq.(1.2.2) as for the particle horizon and the temperature after the phase transition  $T_a$  (Eq.(1.1.13)) instead of the critical temperature  $T_c$ , the monopole/entropy ratio is given by approximately

$$r(t_f) = (T_v/m_p)^3 \exp(-3 t_f / \ell), \quad (1.4.4)$$

in stead of Eq.(1.4.3). If we takes  $t_f > 62 \ell$ , which value is necessary to explain the horizon problem, and  $T_v=10^{15}$  GeV, the monopole/entropy ratio is  $10^{-90}$ . This results suggests that the monopole problem can be solved from the horizon discussion in principle (but see, § 2-3)).

#### 5) Flatness problem

In spite of many efforts to determine the curvature of the present universe from the observation of deceleration parameter  $q_0$  or the cosmic energy density parameter  $\Omega (= \rho_0 / \rho_c)$ , it is unclear whether our universe has negative, positive or zero curvature<sup>8)</sup>. This is because our universe is very flat and the curvature radius is greater than the Hubble length  $H_0^{-1}$ . The

flatness problem is to ask why present universe is so flat. For the people who believe Creator likes flat space and he made the universe flat, i.e.,  $k=0$  exactly, this problem is nonsense. But if we consider the flat case is the limiting case of the curvature radius infinity, this is an important problem to be solved in the SBBM: As is well known, the unique characteristic scale of the Einstein equation for a radiation-dominated universe is the Planck length or the Planck time  $m_p^{-1}$ . This means the natural scale of the universe should be the Planck length or the Planck time. Why the universe is greater than the Planck length  $m_p^{-1}$  by a factor  $m_p/H_0 \sim 10^{62}$ ? This extremely large number should be explained physically. Now we show more in detail. Let's rewrite the Einstein equation as follows,

$$R^2 = H_p^2 R_p^2 \left( (1 - \rho_p / \rho_{cp}) + (\rho_p / \rho_{cp})(R_p/R)^2 \right), \quad (1.5.1)$$

where the scale factor  $R(t)$  is made to represent the spatial curvature radius of the universe for  $k=0$ . The critical density at the Planck era  $\rho_{cp}$  is defined as  $\rho_{cp} = 3H_p^2 / 8\pi G$ , where  $H_p$  is the Hubble constant at  $R=R_p$ , i.e.,  $H_p = (R/R)_{R=R_p}$ . For making the problem clear, let's discuss the case of closed universe, i.e.,  $\rho_p > \rho_{cp}$ . As is well known, in the closed universe model the scale factor has a maximum, which is given by

$$R_{\max} = R_p \delta^{-1/2}, \quad (1.5.2)$$

where

$$\delta = (\rho_p - \rho_{cp}) / \rho_p. \quad (1.5.3)$$

The first point is that as for the scale factor at the Planck era  $R_p$ , we must take a extremely large value compared with the characteristic length of Einstein equation  $m_p^{-1}$ : In the SBBM,  $R_p$  is calculated as

$$R_p = (T_0/T_p) R_0 > T_0/T_p H_0 = (T_0/H_0) m_p^{-1} \sim 10^{30} m_p^{-1}, \quad (1.5.4)$$

from the ratio of the cosmic temperatures of present universe  $T_0 \sim 3K$  to the Planck era  $T_p \sim m_p$ , where the present scale factor  $R_0$  is greater than the Hubble distance  $H_0^{-1}$ . Generally speaking, the size of a physical system should be of the order of the characteristic length of the basic equation which describes the system. From this point of view, the ratio  $10^{30}$  in Eq.(1.5.4) is extremely large number and should be explained the reason. The second point is that the value of  $\delta^{-1/2}$  should be greater than  $10^{32}$  in order that the universe becomes larger than the Hubble length  $H_0^{-1}$ ,  $R_{\max} > H_0^{-1}$ , i.e.,

$$\begin{aligned} \delta^{-1/2} &= R_{\max}/R_p > 1/(H_0 R_p) > T_p/(H_0 T_0 R_0) \\ &= T_p/T_0 = m_p/T_0 = 10^{32}. \end{aligned} \quad (1.5.5)$$

This means that very fine tuning for the cosmic energy density  $\rho_p$  is necessary, i.e.,  $\delta = (\rho_p - \rho_{cp}) / \rho_p < 10^{-64}$ , in order to explain the present size of the universe, even if we admit the unnatural large initial scale factor  $R_p > 10^{30} m_p^{-1}$ . Otherwise the universe collapse within a time of the order of the Planck time. In order to solve the flatness problem, therefore, we must explain these two points. It should be noticed that these two points have not been clearly discussed separately,

The explanation of the flatness problem for the open universe ( $k=-1$ ) is also essentially the same as the above, but is more complicate.

We can now easily understand that the flatness problem is solved (but partially) in the inflation model<sup>4</sup>). Because the universe expands exponentially, the non-dimensional large number can be easily introduced by the exponential factor  $\exp(t_f/\ell)$  in this model (see Eq.(1.1.10) or Eq.(1.1.14)). It should be noticed, however, the flatness problem cannot be solved completely by inflation model with the GUT-scale energy,  $10^{15} \text{GeV}$  for the closed universe model. If the universe which starts at the Planck era with the size of Planck length  $R_p = m_p^{-1}$  collapses before GUT phase transition, the inflation mechanism cannot work. In order for the universe to survive

until GUT-temperature, fine-tuning on the density at Planck era is necessary. From Eq.(1.5.2), we obtain the condition

$$\rho = (R_p/R_{\max})^2 < (\ell_{\text{mp}})^{-2} = 10^{-16} . \quad (1.5.6)$$

## §2 Details of original inflation model and its difficulties

### 1) Nucleation of bubbles and progress of the phase transition

As discussed by Coleman<sup>13)</sup>, the first order phase transition is triggered by nucleation of bubbles of true vacuum. This first-order phase transition finishes when all the space is covered by bubbles (see Fig.( 2.1.1)).

Let  $r$  be the volume fraction of bubbles (true vacuum). Then the fraction at the cosmic time  $t$  is given by

$$r(t) = \int_0^t p(t') (R(t')/R(t))^3 V(t',t) dt' , \quad (2.1.1)$$

where  $p(t')$  is the nucleation rate of bubbles per unit time and unit volume in the false vacuum ( $\phi=0$ ) at the cosmic time  $t'$ . The volume of a bubble at the cosmic time  $t$ , which has been created at  $t'$  is

$$V(t',t) = 4 \pi R^3(t) \int_{t'}^t (1-r(t'')) \chi^2(t',t'') (d\chi(t',t'')/dt'') dt'' \quad (2.1.2)$$

where

$$\chi(t',t'') = \int_{t'}^{t''} dt / R(t) \quad (2.1.3)$$

Here we have assumed that a bubble expands spherically at the velocity of light and the increase in the volume of a bubble is depressed by the factor  $(1-r(t''))$  due to the overlapping of bubbles. The solution of this integral equation (2.1.1) is given by<sup>4),6)</sup>

$$u(t)=1-r(t)= \exp(-) \int_0^t p(t') R^3(t') (4 \pi/3) \left( \int_{t'}^t dt'' / R(t'') \right)^3 dt' \quad (2.1.4)$$

where  $u(t)$  is the volume fraction of false vacuum. Nucleations of bubbles of a stable vacuum are induced by two factors, one is the thermal fluctuations and the other is the quantum fluctuations. These factors give very different types of nucleation rates, and we therefore, in the following, discuss them separately in order to make clear the characteristic properties of the progress of the phase transition.

The nucleation rate  $p$  by thermal fluctuations is essentially the same as the probability of bubble formations in a boiling liquid, which is given as  $p \sim \exp[-S_3(T)/T]$ . The factor  $S_3(T)$  is the free energy of an  $O(3)$  symmetric bubble,<sup>14)</sup> which is given as  $S_3(T)=16 \pi a^3/3 \epsilon^2$  if we take a thin wall approximation, i.e, the radius of a bubble is much greater than the thickness of the wall, where  $a$  is the surface energy per unit area and  $\epsilon$  is the energy density difference between two vacua. The nucleation rate  $p(T)$  has a sharp peak at  $T_c'$  which is less than  $T_c$ , but close to  $T_c$ . The nucleation rate by quantum fluctuation is given as  $p \sim \exp(-S_4)$ , where  $S_4$  is the action of an  $O(4)$  symmetric bubble. If we take a thin wall approximation, the factor  $S_4$  is given as  $S_4=27 \pi^2 a^4/2 \epsilon^3$ . The quantum nucleation rate is almost independent of the temperature and has a constant value.

Now we assume the nucleation rate is given by the following simplified equation<sup>2)</sup>

$$p(t) = v_T \ell^{-3} \delta(t-t_N) + v_Q \ell^{-4} \theta(t-t_N) , \quad (2.1.5)$$



where the first term represents the nucleation by thermal fluctuations and the second term, the nucleation by quantum fluctuations. The coefficients  $\nu_T$  and  $\nu_Q$  are, respectively, nucleation rates of thermal and quantum fluctuations normalized by the characteristic scale  $\ell$ . This approximation holds when the nucleation time by thermal fluctuations and the rise time for growth of quantum nucleation rate are much shorter than the duration time of the phase transition.

Substituting Eqs.(1.1.10b) and (2.1.5), we obtain

$$u(t) = u_T(t) u_Q(t) \quad , \quad (2.1.6)$$

where

$$u_T(t) = \exp(-4\pi/3) \nu_T (1 - \exp(-(t-t_N)/\ell))^3 \quad (2.1.7)$$

and

$$u_Q(t) = \exp(-4\pi/3) \nu_Q (t-t_N)/\ell \quad . \quad (2.1.8)$$

We assumed that the ignition time of nucleation  $t_N$  is later than the characteristic cosmic time  $\ell/2$ , i.e.,  $t_N > \ell/2$ , and  $(t-t_N) \gg \ell$ .

## 2) Percolation of bubbles and fractal structure of vacuum

Now we discuss how phase transition finishes in the universe, using the results of preceding subsections.

Eq.(2.1.7) shows that  $u_T(t \rightarrow \infty) = \exp(-4\pi/3) \nu_T$ . This means that the volume fraction of bubbles  $r(t) (=1-u(t))$  stops increasing in spite that bubbles expand at light velocity. This is because that the event horizon exists in the de Sitter universe (exponentially expanding universe). From Eq.(2.1.3),

the radius of a bubble in comoving coordinate stops increasing in spite of the light speed expansion because the bubble expansion cannot catch up the cosmic expansion;

$$\chi(t, t_N) = \alpha^{-1} (\exp(-t_N/\ell) - \exp(-t/\ell)) + \alpha^{-1} \exp(-t_N/\ell) \quad (2.2.1)$$

According to theories of percolation<sup>15)</sup>, infinite networks of linked bubbles are formed when  $r(t)$  becomes greater than a critical value  $p_c (\sim 0.3)$  and when  $u(t)$  becomes less than  $p_c$  the infinite networks of the false vacuum disappear, i.e., the size of all the false vacuum regions becomes finite and are surrounded by bubbles. As shown in the papers<sup>16)-22)</sup>, the regions become black holes or wormholes and drop out of our universe if we take into account the effects of general relativity. We, therefore, may consider that the first order phase transition finishes effectively when  $u(t)$  becomes less than  $p_c$ , i.e.,  $u(t_f) = p_c \sim 0.3$ . The present result shows that the phase transition never terminates by the bubbles nucleated by thermal fluctuations, if the nucleation rate  $\nu_T$  is smaller than a critical value,

$$\nu_T^c = -(3/4\pi) \ln p_c = 0.29 \quad (2.2.2)$$

Although nucleation rate by quantum fluctuations is very small compared with that by thermal fluctuations, the volume fraction  $u(t)$  decreases steadily by bubbles created continuously by quantum fluctuations as shown by Eq.(2.1.8). The size of bubbles in comoving coordinate, however, becomes smaller and smaller when the bubbles nucleate later and later as shown by Eq.(2.2.1). no more conventional percolation theories, in which the size of bubbles is fixed, can be applied. As discussed by Guth and Weinberg<sup>23)</sup>, the critical value for percolation  $p_c$  increases with decreasing nucleation rate  $\nu_Q$  and no percolating network of linked bubbles is formed eventually if  $\nu_Q \leq 10^{-6}$ . It seems that the universe takes fractal structure<sup>24)</sup> displayed in Fig(2.2.1)(Kodama et al<sup>25)</sup>), i.e., in spite of the bubble fraction  $r$

$\approx 0.99999$ ....., no infinite network exists and the structure is self similar with respect to scale transformation. Because quantum nucleation rate is very small in conventional Higgs potential model for first order phase transition, the thermal nucleation rate must be almost equal or greater than the critical value  $\nu_T^c$  in order for the phase transition to terminate. On the other hand, nucleation rate must be small in order for sufficient inflation, otherwise the phase transition terminate quickly and it becomes very hard to get sufficient inflation time to explain the horizon and flatness problems. In order for these two condition to be satisfied, very fine tuning on nucleation rate,  $\nu_T^c \leq \nu_T \leq \nu_T^c (1 + \delta)$ ,  $\delta \ll 1$ , is necessary.

### 3) Difficulties of original model

As shown in § 1, the inflation model gave a plausible solution for many cosmological problems which could not be solved in the SBBM. But many difficulties were found in the first-order phase transition model (this is now called as the original inflation model) when we investigated it in detail. Now, we summarize them.

#### i) Homogeneity problem

As discussed in the preceding subsections, in a period of the first order phase transition, the universe is very bumpy and its spacetime structure becomes very complicate. Of course if all the black holes and wormholes evaporate away quickly, the universe can go back to homogeneous state. Sasaki et al.<sup>22)</sup> and Kodama et al.<sup>25)</sup> investigated in this scenario that seeds of galaxies (density fluctuations) can be induced by the number density fluctuation of the black holes and wormholes created by the first order phase transition. They showed, however, that very unnatural fine tunings on thermal and quantum nucleation rates,  $\nu_T$  and  $\nu_Q$  are necessary in order to explain the global homogeneity of the universe and the adequate amplitude of the density fluctuation for galaxy formation, otherwise, the universe remains at bumpy state or the phase transition never terminates and the vacuum takes fractal structure. These consequences conflict with the

present observations of the universe.

#### ii) Monopole problem

In the first order phase transition model, a monopoles are formed at a point surrounded by bubbles, because a bubble is a coherent domain in this model. In § 1, we calculated the monopole/entropy ratio assuming the domain radius is the particle horizon length. This gives, obviously, a very lower estimate to the monopole/entropy ratio. If very large number of bubbles are continuously nucleated by quantum nucleations (see Fig.(2.2.1)), monopoles are also created abundantly. It should be also noticed that monopoles are produced as magnetic black holes or magnetic wormholes in the strongly first order phase transition<sup>20),21)</sup>. As we have shown, the false vacuum region surrounded by bubbles becomes a black hole or a wormhole, which is the same position where a monopole is created. Therefore, usual monopoles remain after the evaporation of these holes. Sato<sup>20)</sup> and Izawa and Sato<sup>21)</sup> investigated the monopole production in the first order phase transition and showed that it is very hard to suppress the over production of monopoles without unnatural fine tunings on thermal and quantum nucleation rates.

#### iii) Problem of the thermalization of the energy released by the phase transition

In the preceding sections, it was assumed that the energy released by the phase transition is thermalized and the universe is heated up by this latent heat. As discussed by Coleman,<sup>13)</sup> the energy goes to the kinetic energy of bubble walls at first. It is expected that this kinetic energy is thermalized when bubbles collide with each other. Hawking, Moss and Stewart,<sup>26)</sup> and Wu<sup>27)</sup> investigated the collision of bubbles and the energy dissipation by computing the evolution of Higgs field classically. They showed that dissipation of the kinetic energy is hard, in particular, when the size of two colliding bubbles are very different. Sawyer<sup>28)</sup> and Kodama<sup>29)</sup> investigated the particle creation in a bubble as a thermalization process and showed that sufficient particles are not produced. At present no plausible thermalization process is not found.

§ 3 New Inflation model

1) Idea and Coleman-Weinberg potential

In order to avoid these difficulties, Linde<sup>30)</sup> and Albrecht and Steinhardt<sup>31)</sup> proposed a new version of inflationary universe scenario. This model is based on the Coleman-Weinberg potential, which is given by

$$V(\phi, T) = m^2(T)\phi^2 + B\phi^4(\ln(\phi/\sigma)^2 - 1/2) + B\sigma^4/2 \quad (3.1.1)$$

where  $m^2(T) = C T^2$  and  $B = 5625g^4/1024\pi^2$ ,  $C$  and  $g$  being a constant of an order of unity and the gauge coupling constant, respectively. In this potential the adjoint Higgs field,  $\phi$ , has been reexpressed as  $\phi(1, 1, 1, -3/2, -3/2)$ . The characteristic properties of this potential are, first, that the potential barrier between  $\phi = 0$  and  $\phi = \sigma$  is made by a thermal effect, i.e.,  $m^2(T) = C T^2$ , and, second, that the potential between  $\phi = 0$  and  $\phi = \sigma$  is very flat, as shown in Fig.(3.1.1). In this model, bubbles are nucleated at the temperature  $T \ll \sigma$ . The vacuum expectation value of Higgs field in the bubbles is, however, very small  $\phi_i \ll \sigma$  because of the small height of the potential barrier  $v_{\text{top}} \sim C T^4 \ll \sigma^4$ . Albrecht and Steinhardt<sup>31)</sup> claim that bubbles or domains with nonvanishing Higgs field  $\phi_i (\ll \sigma)$  are formed rather by spinodal decomposition not by nucleation of bubbles. In any way the essential difference to the original model is that bubbles have almost the same vacuum energy density as that of the symmetric state  $\phi = 0$ . This means that bubbles also expand exponentially. The exponential expansion continues until Higgs field rolls over until  $\phi = \sigma$ , where the vacuum energy density is vanishing. The roll over time is estimated from the equation of motion of Higgs field,

$$\ddot{\phi} + 3H\dot{\phi} + V' = 0 \quad (3.1.2)$$

where GUT-Hubble constant is defined as

$$H = \lambda^{-1} = (8\pi G V(0)/3)^{-1/2} = (4\pi G B\sigma^4/3)^{-1/2}, \quad (3.1.3)$$

If the motion of  $\phi$  is friction-dominated, i.e.,  $3H\dot{\phi} = -V'$ , the characteristic time for roll over is estimated as

$$\tau = (\phi/\dot{\phi})_i \sim 3H\phi_i/V'(\phi_i) \sim H/T_i^2 \quad (3.1.4)$$

where  $T_i$  is the cosmic temperature when bubbles created. As has been discussed in §1, the horizon- and the flatness problems are solved if the duration time of the exponential expansion,  $\tau$ , is longer than  $62\lambda$  because a bubble with the size of  $\lambda$  ( $\sim 10^{-26}$  cm) is stretched to  $\lambda \exp(\tau/\lambda) \sim 10$  cm when the phase transition finishes. The present size of this region is longer than the Hubble length  $H_0^{-1}$ . In this new scenario, therefore, we are now living in a bubble. Obviously no monopole problem exists, because monopoles are out of our universe which we know observationally, even if they have formed. Albrecht et al.<sup>32)</sup> demonstrated that the universe will be reheated almost to the GUT temperature  $T \sim \sigma$  also, by introducing a phenomenological viscosity term  $\Gamma\dot{\phi}$  in the equation of motion (3.1.2), provided  $\Gamma \gg H$ .

2) Gravitational effects on phase transition

In the new inflation model, a gravitational effect on the nucleation of bubble becomes very important because the radius of a critical bubble becomes comparable to the characteristic length of the de Sitter universe  $\lambda (= H^{-1})$ . In order to obtain a critical bubble solution, we must calculate the Euclidean Einstein equation<sup>33)</sup>,

$$R'^2 = 1 + 8\pi G R^2 (\phi'^2 - V)/3 \quad , \quad (3.2.1)$$

simultaneously with Euclidean equation of motion of Higgs field,

$$\phi'' + 3(R'/R) \phi' - dV/d\phi = 0 \quad , \quad (3.2.2)$$

where we have assumed the metric

$$ds^2 = d\xi^2 + R^2(\xi) d\Omega^2 \quad , \quad (3.2.3)$$

and a dash ' is the derivative with respect to  $\xi$ . In this metric,  $\xi$  is the radial coordinate and  $d\Omega$  is the element of length on a unit three sphere;  $R$  is the Euclidean analogue of the Robertson-Walker scale factor. In the new inflation model, however, the adequate approximate solution of Eq.(3.2.1) is obtained independent of Eq.(3.2.2) as<sup>34)</sup>

$$R(\xi) = H^{-1} \sin(H\xi), \quad (0 \leq \xi \leq \pi/H) \quad (3.2.4)$$

because the energy density everywhere remains roughly constant, i.e.,  $\phi'^2 \ll V(\phi) \sim V(0)$ . Then the equation of motion for Higgs fields is given by

$$\phi'' + 3H \cot(H\xi) \phi' - dV/d\phi = 0 \quad . \quad (3.2.5)$$

The Euclidean action  $S_4$  is given by

$$S_4 = 2\pi^2/H^3 \int_0^{\pi/H} d\xi (\sin^3 H\xi) (\phi'^2_b + V(\phi_b) - B\sigma^4/2) \quad (3.2.6)$$

where  $\phi_b(\xi)$  is a bounce solution of Eq.(3.2.5).

Note that no bounce solution exists if we take the boundary condition similar to that of Coleman's original paper;  $\phi_b'(0)=0$  and  $\phi_b(H/\xi)=0$ . As Jensen and Steinhardt<sup>34)</sup> and Albrecht et al<sup>35)</sup> claimed, the adequate boundary condition is  $\phi'(0)=\phi'(\pi/H)=0$ . Equation of motion for Higgs field Eq.(3.2.5) corresponds to the mechanical equation for a particle moving in a potential minus  $V$  and subject to a somewhat peculiar viscous damping force with Stokes's-low coefficient proportional to  $\cot(H\xi)$ , if we interpret  $\phi$  as a particle position and  $\xi$  as time. In Fig.(3.2.1), two typical bounce solutions for case i)  $H \ll m$  and ii)  $H \gg m$  are shown. Because the oscillation frequency at  $\phi = \phi_{top}$ ,  $\omega = (-d^2V/d\phi^2)^{1/2}$  is of the order of  $m$ , the Case i) corresponds to the limit that the de Sitter time  $H^{-1}$  is much longer the oscillation time. In this case, the bounce solution reduces to that given by Coleman originally, because the gravitational effect is negligible. This is easily understood that equation of motion (3.2.5) reduces to the original ones<sup>13)</sup> in the limit of  $H \rightarrow 0$ . On the other hand, Case ii) corresponds to the strong limit of the gravitational effect. Because the bounce time  $\pi/H$  is much smaller than the oscillation time, the movement of particle is strongly limited near the bottom of the minus potential ( $-V$ ),  $\phi \sim \phi_{top}$ . This means that the vacuum expectation value in a bubble  $\phi_c$  is almost equal to  $\phi_{top}$ , where the potential  $V$  has a maximum. Hawking and Moss<sup>36)</sup> claimed that the only Euclidean solution which can give a meaningful tunneling probability is the homogeneous solution  $\phi(\xi) = \phi_{top}$  for  $0 < \xi < \pi/H$ , and the action is

$$S_4 = (3/8\pi G)(1/V(0) - 1/V(\phi_{top})) \quad . \quad (3.2.7)$$

This is obviously the limiting case  $H \gg m$ . Evolution of bubbles or Higgs field is essentially the same as that of the preceding discussion, except the initial value of Higgs field is almost equal to  $\phi_{top}$ . If the time for rolling over is sufficiently long, the size of a bubble can become larger than the Hubble length  $H_0^{-1}$  at present.

3) Dynamical evolution of Higgs field in SU(5) model

In most of the investigation until now it has been assumed that the vacuum expectation value of the Higgs field evolves directly to the SU(3)xSU(2)xU(1) direction from the start of rolling down. Recently, however, some people have pointed out<sup>37)~41)</sup> that the Higgs field goes toward SU(4)xU(1) state, which may give rise to serious difficulties in the new inflationary scenario. Sato and Kodama<sup>41)</sup> investigated this point in more detail by examining the evolution of a Higgs field  $\Phi$  belonging to the adjoint representation of SU(5) in the full 24-dimensional space by numerical simulation and to elucidate whether the new inflationary scenario is consistent with cosmological observations or not.

In practical computations, however, it is not necessary to calculate the evolution of the full 24 components directly. It is quite natural to assume that the time derivative of the Higgs field vanishes,<sup>13)</sup> i.e.,  $\dot{\Phi}=0$ , just when the Higgs field acquires non-vanishing classical expectation values in the first stage of the phase transition. Then the Higgs field  $\Phi$  represented by an arbitrary 5x5 hermitian traceless matrix can be diagonalized in each coherent region by a global gauge transformation, keeping  $\dot{\Phi}=0$ . The equation of motion of the Higgs field guarantees that  $\Phi$  remains diagonal in the course of its evolution if  $\Phi$  is diagonal and  $\dot{\Phi}=0$  at the start. Thus in the calculation of evolution we can restrict the form of the Higgs field without loss of generality as

$$\Phi = \text{diag}[\phi_1, \phi_2, \phi_3, \phi_4, \phi_5] \quad (3.3.1)$$

with the constraint  $\text{Tr} \Phi = \sum_{i=1}^5 \phi_i = 0$ .

In this representation the Coleman-Weinberg potential with the one-loop correction by the gauge boson (the Higgs boson contribution to the one-loop correction is not included) is given by<sup>42)</sup>

$$V(\Phi) = \frac{3g^4}{256\pi^2} \left[ C \left( \sum_{i=1}^5 \phi_i^4 \right) - \frac{7}{30} \left( \sum_{i=1}^5 \phi_i^2 \right)^2 \right] + \sum_{i,j=1}^5 (\phi_i - \phi_j)^4 \left[ \ln \left( \frac{\phi_i - \phi_j}{\mu} \right) - \frac{1}{2} \right], \quad (3.3.2)$$

where  $C$  is an arbitrary parameter of this potential and  $\mu$  is a renormalization parameter related to the vacuum expectation value of the Higgs field at the SU(3)xSU(2)xU(1) minimum point  $\Phi = \sigma \times \text{diag}[1, 1, 1, -3/2, -3/2]$  ( $\sigma \sim 4.5 \times 10^{14} \text{GeV}$ ) as  $\mu = 5\sigma/2$ . The effect of temperature on the potential is neglected because the inflation begins after the cosmic temperature becomes less than the GUT temperature ( $\sim 10^{14} \text{GeV}$ ) and the essential fate of the Higgs field is determined before the universe is heated up again to the GUT temperature.

Because of the traceless condition  $\sum_{i=1}^5 \phi_i = 0$ , five

components of the Higgs field  $\phi_i$ ,  $i=1, 2, \dots, 5$ , are not independent. This makes the numerical computation complicated if we calculate the time evolution of these components directly. In order for the convenience of numerical computation, we introduce the following four components fields which are completely independent each other,

$$\psi_i = \phi_i + \sum_{j=1}^4 \phi_j / (1 + \sqrt{5}), \quad (i=1, 2, 3, 4) \quad (3.3.3)$$

In Fig.3.3.1, contours of the potential on a plane ( $x = 3\psi_1 = 3\psi_2 = 3\psi_3$ ,  $y = \psi_4$ ) are displayed for the case of the potential parameter  $C=1$ . On

this plane there are two SU(3)xSU(2)xU(1) minima and four SU(4)xU(1) minima. As is easily understood from the potential Eq.(3.3.2), local minima in the SU(4)xU(1) direction can exist for  $C < 15$ , and these minima become global minima for  $C < -15 \ln(1.5)$ .<sup>38)</sup>

The equation of motion for the fields  $\psi_i$  are given by

$$\ddot{\psi}_i + 3\dot{R}/R \dot{\psi}_i + V' \psi_i + C_{vis} |\psi_i| \dot{\psi}_i = 0, \quad (3.3.4)$$

where a viscosity term  $C_{vis} |\psi_i| \dot{\psi}_i$  is introduced in order to convert the energy of the Higgs field into thermal energy. The scale factor of the universe  $R$  is calculated by the expansion equation of the universe

$$(\dot{R}/R)^2 = 8\pi G(\rho_r + \rho_\phi)/3 \quad (3.3.5)$$

where we have assumed that the universe is spatially flat, which is adequate in the early universe even if it is not flat exactly. The change of radiation energy density  $\rho_r$  and the energy density of the Higgs field  $\rho_\phi$  are described, respectively, by

$$d(\rho_r R^4)/dt = -\sum_{i=1}^4 C_{vis} |\psi_i| \dot{\psi}_i^2 R^4, \quad (3.3.6)$$

and

$$\rho_\phi = \frac{1}{2} \sum_{i=1}^4 \dot{\psi}_i^2 + V. \quad (3.3.7)$$

The initial value of the Higgs field has four degrees of

freedom: one is the norm of Higgs field  $\|\psi\| = (\sum_{i=1}^4 \psi_i^2)^{1/2}$  and the others are the direction of the vector  $(\psi_1, \psi_2, \psi_3, \psi_4)$  in the four dimensional  $\psi$  space.

In the present investigation we take  $\epsilon = \|\psi\|_0 = 8 \times 10^{-6}$  as the initial value of  $\|\psi\|$ , which is about  $0.2H$ , where  $H = (8\pi G V(0)/3)^{1/2}$ . In order to parametrise the initial direction of  $\psi$  we utilize three angles  $\alpha$ ,  $\theta$  and  $\beta$ .  $\alpha$  represents the deviation angle from the SU(4)xU(1) direction  $\psi_1 = \psi_2 = \psi_3 = \psi_4 (> 0)$  on the plane  $\psi_1 = \psi_2 = \psi_3$  as shown in Fig.3.3.1 and  $\theta$  and  $\beta$  represent the deviation angles off this plane. Though we do not limit the range of  $\alpha$  essentially, we restrict  $\theta$  and  $\beta$  in the very narrow range  $|\theta| < 10^{-4}$  and  $|\beta| < 10^{-4}$  for the convenience of the analysis of the numerical computation as a first step.

In Fig.3.3.2a-3.3.2c, some results of numerical computation for the case of the potential parameter  $C=1$  are shown. As demonstrated in these figures, the Higgs field goes to the SU(4)xU(1) direction  $\psi_1 = \psi_2 = \psi_3 = \psi_4 > 0$  at first independent of the values of viscosity parameter  $C_{vis}$  and the initial angle  $\alpha$ , provided that  $-0.29\pi < \alpha < 0.21\pi$ . This is obviously a natural consequence of the fact that the potential has the steepest gradient along this direction where the norm  $\|\psi\|$  is small<sup>38)</sup> (see Fig.3.3.1). In Fig.3.3.3, the dependence of the degree of inflation on the initial angle  $\alpha$  is shown. Here we define the degree  $D$  as the ratio of the cosmic scale factors,  $D=R_2/R_1$ , where  $R_1$  is the value when the inflation begins, i.e., when the vacuum energy density becomes greater than that of the radiation, and  $R_2$  is the value when the Higgs field arrives near the SU(4)xU(1) minimum. Note that inflation begins again when the Higgs field settles down at an SU(4)xU(1) state because of the remaining vacuum energy density. Of course this inflation is not taken into account in this definition. As shown in Fig.3.3.3, the degree of inflation is very sensitive to the initial angle  $\alpha$ , but almost independent of the potential parameter  $C$  and the viscosity parameter  $C_{vis}$  provided that  $C_{vis} < 0.1$ . The evolution after the arrival to this minimum, of course, depends on the value of  $C_{vis}$ .

Generally speaking, the Higgs field settles down in a very direct way to the SU(4)xU(1) minimum after short time oscillation independent of the angle  $\alpha$  provided that  $C_{vis} > 1$  as is illustrated in Fig.3.3.2a. For the smaller values of the viscosity parameter  $C_{vis}$ , however, it can depart from this local minimum and further evolve to an SU(3)xSU(2)xU(1) minimum state. As shown in Fig.3.3.2b ( $C_{vis}=10^{-2}$  and  $\alpha=0.2\pi$ ), the Higgs field evolves to the nearest SU(3)xSU(2)xU(1) state via the SU(4)xU(1) state and settles down to this state after oscillation around it. Note that, however, details of the evolution of the Higgs field are different for the different initial angles even if the value of the viscosity parameter  $C_{vis}$  is the same. For example, if we take  $\alpha=-0.1\pi$ , the Higgs field settles down to the SU(4)xU(1) state after large amplitude oscillations. When we take the smaller values for  $C_{vis}$ , the Higgs field begins to circulate in this (x,y)-plane and wanders around a lot of SU(3)xSU(2)xU(1) states and SU(4)xU(1) states. After a few time circulations, the Higgs field goes out from this plane and begins to wander around more numbers of SU(3)xSU(2)xU(1) and SU(4)xU(1) states in the four dimensional space of the Higgs field. This result suggests that the eventual state of the Higgs field is changed greatly by the very small deviation of the initial direction of the Higgs field.

In Fig.3.3.4 final states of the Higgs field are summarized in the plane of the initial angle  $\alpha$  and the viscosity parameter  $C_{vis}$ . As has been discussed, final states of the Higgs field depend on the value of the viscosity parameter  $C_{vis}$  strongly. The final state is an SU(4)xU(1) minimum if  $C_{vis}$  is greater than a critical value. Although the critical value depends on the initial angle  $\alpha$  as shown in Fig.3.3.4, we may conclude that the final state is SU(4)xU(1) if  $C_{vis} > 10^{-2}$ . This state is, however, unstable if we take into account the quantum tunnelling effect. Although the Higgs field can reach a nearby SU(3)xSU(2)xU(1) state by this tunnelling, the same difficulties as appeared in the original inflation scenario<sup>1)~7)</sup> arise because this phase transition is of first order, i.e., large scale inhomogeneity is created by bubbles formed by this phase transition as pointed out by Breit, Gupta and Zacks.<sup>38)</sup>

When we take the value of the viscosity parameter in the range  $2 \times 10^{-3} < C_{vis} < 10^{-2}$ , the Higgs field can settle down to the nearest SU(3)xSU(2)xU(1) state steadily without traveling to other SU(3)xSU(2)xU(1) or SU(4)xU(1) states. In this case, the inflation scenario works well with no trouble as

has been investigated by many people.

On the other hand, if we take values  $C_{vis} < 2 \times 10^{-3}$ , the Higgs field travels around a lot of minimum states as illustrated in Fig.3.3.2c. Eventual states to which the Higgs field settles down are changed by fluctuations of initial values of the Higgs field. This result strongly suggests that a coherent region, which is formed by nucleation of bubbles or spinodal decomposition in the early stage  $\|\phi\| < H$ , will be fragmented into many SU(3)xSU(2)xU(1) and SU(4)xU(1) states by the fluctuations of the Higgs field associated with the initial state. Even if the classical fluctuations of the Higgs field are extremely small, the quantum ones might fragment the original coherent region into small pieces of different state.<sup>38)</sup> Thus this result suggests that large scale inhomogeneities also appear for the too small values of  $C_{vis}$ , which conflict with present observation.

Sato and Kodama<sup>41)</sup> have carried out numerical computations for the different values of the potential parameter C (Eq(3.3.2)) in the range  $0 < C < 14$ . The result displayed in Fig.3.3.1 (the case C=1), does not change qualitatively for the other values of C in this range except that the critical value of the viscosity parameter  $C_{vis}$ , which decides whether the final state is SU(4)xU(1) or not, depends more sensitively on the initial angle  $\alpha$ . We have also carried out the simulation by using the viscosity of the form  $C_{vis} \|\psi\| \dot{\psi}_i$  instead of  $C_{vis} |\psi_i| \dot{\psi}_i$ . The results were essentially the same both qualitatively and quantitatively except for the small change of the critical value of  $C_{vis}$ .

In order for an inflationary scenario of the universe consistent with observation to be constructed, the value of the viscosity must be in an adequate range, otherwise large scale inhomogeneities which conflict with observations arise. Recently Abbott et al.<sup>42)</sup> and Hosoya and Sakagami<sup>43)</sup> estimated the strength of viscosity. At present, it is hard to determine whether the value of viscosity lies in the range adequate for the new inflationary scenario or not, because the result of Abbott et al. is very qualitative and the viscosity obtained by Hosoya and Sakagami is a thermal viscosity. In order to make clear whether a consistent scenario can be constructed or not, more precise evaluation of viscosity is necessary.

#### 4) Primordial Inflation and Chaotic Inflation

New Inflationary universe model, which is proposed in order to avoid difficulties of the original models, however, has also some problems to be solved: i) As discussed in preceding sections initial value of Higgs field  $\phi_i$  must be lower than Hawking temperature  $T_H = H/2\pi$  in order for sufficient inflation. If the Hawking temperature<sup>44),45)</sup> is the amplitude of the quantum fluctuations in the de Sitter universe, the universe cannot cool lower than  $T_H$  and the initial value should be greater than or equal to  $T_H$ . Actually many people discussed the possibility that the phase transition begins at  $T=T_H$  from various points of view<sup>46)-50)</sup>. These two conditions conflict each other. ii) Hawking and many people observed that quantum fluctuations in the Higgs field during the roll over stage may give rise to classical inhomogeneities on large scales.<sup>51)-54)</sup> The amplitude of density perturbations when the wave length is equal to the particle horizon, was estimated as

$$(\delta\rho/\rho) \sim B^{1/2}(\ln(L/\ell))^{3/2} \quad (3.4.1)$$

where  $B=5625g^4/1024\pi \sim 0.1$  (see Eq.(3.1.1)) and  $L$  is the wave length at the finishing time of the inflation. If we take  $L \sim 10^{23}\ell$  for a scale of cluster of galaxies, the value of  $\delta\rho/\rho$  is about 50, which conflict with the homogeneity observed by the microwave background radiation. iii) Generally speaking, Higgs potential  $V(\text{Eq.}(3.1.1))$  may have nonvanishing mass term  $\mu^2/2$  at  $T=0$ . For Coleman-Weinberg model, the coefficient of the quadratic term  $\mu^2/2$  is set equal to zero and the Higgs mass is  $m_{CW}=2.7 \times 10^{14}\text{GeV}$ . Now, let the Higgs mass for the nonvanishing quadratic term be  $m_H$  and define  $\Delta_H = (m_H^2 - m_{CW}^2)/m_{CW}^2$ . As claimed by Albrecht and Steinhardt,<sup>35)</sup> the value of  $\Delta_H$  must be extremely small  $\Delta_H \ll 10^{-6}$  in order for sufficient inflation to occur. This means very fine tuning on the Higgs mass is necessary for the new inflation scenario to solve the flatness and the horizon problems.

Recently, the CERN group<sup>55)</sup> proposed an alternative version of the new inflationary model in order to avoid these problems on the basis of super symmetry GUTs. In this scenario, first, inflation (exponential expansion of the universe) is induced by a phase transition of a hypothetical scalar field, which they named inflaton. This phase transition occurs at the

temperature higher than GUT-scale  $10^{15}\text{GeV}$ , but lower than the Planck scale  $10^{19}\text{GeV}$ . Second,  $SU(5)$  symmetry is broken at the temperature  $T \sim \Lambda_5$ ,  $\Lambda_5 (\sim 10^9\text{GeV})$  being the strong coupling limit of  $SU(5)$  gauge interaction, which is much lower than the breaking scale  $10^{15}\text{GeV}$ . They claim all the difficulties are solved without fine tuning of parameters in the theory. Although this model is very interesting, it is to be noticed that the inflaton introduced newly in this model has nothing to do with the physical phenomena of particles at the present universe.

Inflation models based on the super GUTs or supergravity also have been investigated by many people<sup>56)-59)</sup>. Albrecht et al<sup>56)</sup> pointed out that the reheating of the universe is hard in a plausible super GUT.

Recently, Linde<sup>60)</sup> also proposed a new interesting model, which he named chaotic inflation model. In this model, no more exists the concept of phase transition, which was essential in the original inflation model, but a simple scalar field with the very flat potential

$$V(\phi) = \lambda/4 \phi^4, \quad \lambda \ll 1, \quad (3.4.2)$$

is introduced. At the Planckian era,  $T \sim 10^{19}\text{GeV}$ , it is believed that the universe was chaotic because of quantum and thermal fluctuations. One may expect that the field  $\phi$  may take any value between  $-m_p/\lambda^{1/4}$  and  $m_p/\lambda^{1/4}$  in different regions of space, so that  $V(\phi) = \lambda\phi^4/4 < m_p^4$ . Now we consider a domain with the size of  $m_p^{-1}$  or greater, in which the scalar field is almost constant. It is expected that such domains exist abundantly in the open (infinite) universe at the Planck time. Because the thermal energy and the energy connected with the inhomogeneity of the field rapidly decrease during the expansion of the universe, the potential energy density  $V(\phi)$  becomes dominant, and the domain becomes exponential expanding with the scale factor  $R \sim \exp(Ht)$ , where  $H = (8\pi V(\phi)/m_p^2)^{1/2}$ . As the result of this expansion, the universe becomes divided into many exponentially large domains containing an almost homogeneous field  $\phi$ . It is obvious, that<sup>4)</sup> the initial value is sufficiently larger than  $m_p^{-1}$  and the time for rolling over to  $\phi=0$  is sufficiently longer than  $m_p^{-1}$ , these domains can become greater than the Hubble length  $H_0^{-1}$  at present as discussed in the subsection 1). In this



scenario however, the space time between the universes (domains) would be extremely curved and the universes would be causally disconnected each other. Note that this scenario is quite similar to the idea of multiproduction of universes by first order phase transition. 19)

#### § 4 Summary and conclusion

In the present paper, we have reviewed the inflationary universe scenario which seems to give a plausible answer to the fundamental cosmological problems discussed in §1. The inflationary model, however, has been revised until now in order to avoid difficulties. It is not clear even now what is the most realistic and plausible model. Although the consistent scenario with no difficulties and no conflicts with observations can be constructed in the primordial inflation model<sup>55)</sup>, it seems that the necessity of the existence of hypothetical field inflaton is not definite from particle physics. Although the new inflation model<sup>30)-31)</sup> seems to be realistic, because this is based on GUTs which have been investigated in detail until now, and this is a natural extension of the Weinberg-Salam theory which is almost justified by the discovery of W and Z bosons in CERN, it seems not easy to remove difficulties discussed in §3-4. Chaotic inflation model<sup>60)</sup> is daring and interesting model in spite of its simplicity. In this theory, no more concept of spontaneous symmetry breaking or phase transition, by which present gauge theories could succeed, is necessary but a simple scalar field with shallow and flat potential is only necessary. But there is no definite reason of the existence of such field from particle physics.

There remain, however, important problems to be solved in all the inflation models. One is on the thermalization mechanism of the vacuum energy density which induced the exponential expansion of the universe. In the discussion §3, viscosity term  $C_{vis} \dot{\phi} \ddot{\phi}$  is introduced phenomenologically in order to convert the released energy into thermal energy. At present, no reliable theory exist to give this viscosity. If there were not sufficiently strong viscosity, the energy of scalar field is absorbed by the expansion of the universe and the universe is not heated up. In this case, it becomes hard to explain not only baryon number asymmetry, but also the 3K microwave background radiation.

The biggest and the most fundamental problem in the inflation model may be why the cosmological term or the vacuum energy density at the present universe is almost vanishing. As is well known, the vacuum energy density is strongly limited from the observations of the Hubble constant and the deceleration parameter as  $\rho_v^0 < 10^{-122} m_p^4$ . This implies that when the universe was created as a fire ball, the value of the cosmological constant had been adjust to cancel the present value  $V(\sigma)$ . Although it has been pointed out that super symmetry GUT may account for this problem, it has not yet become clear.

In spite of the existence of difficulties and unsolved problems, it is very clear that inflation stage (exponential expanding era) is inevitable to explain the two fundamental problems i.e., the horizon problem and the flatness problem. In order to make clear how this inflation occurs, more careful and precise investigations along the border line between particle physics and cosmology would be necessary.

#### Acknowledgement

The author would like to thank H. Sato, M. Yoshimura and A. Hosoya for stimulating discussion. Some subjects discussed here are collaborations with H. Kodama, M. Sasaki, K. Maeda and M. Izawa. He should like to thank them for important contributions and fruitful discussions. This research is supported in part by the Grant in Aid for Science Research Fund of the Ministry of Education, Science and Culuture No.56340021 and also by Asahi Scholastic Promotion Fund.

References

- 1) K. Sato, Mon. Not. Roy. Astron. Soc. **195** (1981), 467.
- 2) K.Sato, Phys. Lett. **99B** (1981), 66
- 3) K. Sato, Prog. in Part. & Nucl. Phys. **6** (1981),311
- 4) A.H. Guth, Phys. Rev. **D23** (1981), 347.
- 5) A.H. Guth and E.J. Weinberg, Phys. Rev. **D23** (1981), 876.
- 6) K. Sato, Grand Unified Theories and Related Topics, ed. by M. Konuma and T. Maskawa (World Science Pub. Co. Singapore, 1981)
- 7) M.Sasaki,Proceedings of Workshop on Grand Unified Theories and Early Universe. KEK83-13 ed.by M.Fukugita and M. Yoshimura (Japan), Jan., 1983.
- 8) S. Weinberg, Gravitation and Cosmology, (John Wiley & Sons, New York, 1972).
- 9) Ya B. Zel'dovich, Adv. Astron. Astrophys. **3** (1965), 241.
- 10) J.P. Preskill, Phys. Rev. Lett. **43** (1979), 1365.
- 11) A.H. Guth and S. -H. H. Tye, Phys. Rev. Lett. **44** (1980),631,963.
- 12) M.B. Einhorn and K. Sato, Nucl. Phys. **B180** (1981), 385.
- 13) S. Coleman, Phys. Rev. **D15** (1977), 2929.
- 14) A.D. Linde, Phys. Lett. **70B** (1977), 306.
- 15) D. Stauffer, Phys. Rep. **54** (1979), 1.
- 16) K. Sato, M. Sasaki, H. Kodama and K. Maeda, Prog. Theor. Phys. **65** (1981), 1443.  
K. Maeda, K. Sato, M. Sasaki and H. Kodama, Phys. Lett. **108B**(1982), 98.
- 17) H. Kodama, M. Sasaki, K. Sato and K. Maeda, Prog. Theor. Phys. **66** (1981), 2052.
- 18) V.A.Berezin, V.A.Kuzmin and I.I.Tkachev, Institute for Nuclear Research preprint 0258(1982)
- 19) K. Sato, H. Kodama, M. Sasaki and K. Maeda, Phys. Lett. **108B** (1981), 103.
- 20) K. Sato, Prog. Theor. Phys. **66** (1981),2287.
- 21) M. Izawa and K. Sato, Prog. Theor. Phys. **68** (1982), 1564.
- 22) M. Sasaki, H. Kodama and K. Sato, Prog. Theor. Phys. **68** (1982), 1561.
- 23) A.H. Guth and E. Weinberg, Nucl. Phys.**B212** (1982),321 .
- 24) B.B. Mandelbrot, Fractals, Form, Chance and Dimension (W.H. Freeman and Co. San Francisco, 1977)
- 25) H. Kodama, M. Sasaki and K. Sato, Prog. Theor. Phys. **68** (1982), 1979.
- 26) S.W. Hawking, I.G. Moss and J.M. Stewart, Phys. Rev. **D26** (1981), 2681.
- 27) Z.C. Wu, Cambridge preprint (1981)
- 28) R.F. Sawyer, Preprint UCSB TH-52 (1981)
- 29) H. Kodama, Prog. Theor. Phys. **68** (1982), 1157.
- 30) A.D. Linde, Phys. Lett. **108B** (1982), 389.
- 31) A. Albrecht and P.J. Steinhardt, Phys.Rev. Lett. **48** (1982), 1220.
- 32) A. Albrecht, P.J. Steinhardt, M.S. Turner and F. Wilczek, Phys. Rev. Lett. **48** (1982), 1437.
- 33) S. Coleman and F. De Luccia, Phys. Rev. **D21** (1980), 3305.
- 34) L.G. Jensen and P.J. Steinhardt, preprint UPR-0231T (1983).
- 35) A. Albrecht, L.G. Jensen and P.J. Steinhardt, Preprint UPR-0229T (1983).
- 36) S.W. Hawking and I.G. Moss, Phys. Lett. **110B** (1982), 35.
- 37) I.G. Moss, Phys. Lett. **128** (1983), 385.
- 38) J.P. Breit, S. Gupta and A. Zacks, Phys. Rev. Lett. **51**(1983), 1087.
- 39) J. Kodaira and J. Okada, Phys. Lett.**133B**(1983),291.
- 40) J.S.Kim and C.W.Kim, WUTP JHU-HET 8310(1983).
- 41) K. Sato and H. Kodama, in Proceeding of International Symposium G. Lemaitre (1983) to be published.
- 42) L.F. Abbott, E. Farhi and M.B. Wise, Phys. Lett. **117B** (1983), 29.
- 43) A. Hosoya and M. Sakagami, preprint OU-HET-57 (1983).
- 44) G.Gibbons and S. W. Hawking ,Phys. Rev. **D17** (1977),2738.
- 45) M.Horibe and A.Hosoya ,Prog. Theor. Phys. **67** (1982),816.
- 46) P. Hut and F.R. Klinkhamer, Phys. Lett. **104B** (1981),439.
- 47) M.D.Pollock and M.Calvani, Phys. Lett.**117B** (1983),392.
- 48) L.F. Abott, Nucl.Phys. **B185** (1981),233.
- 49) Y. Fujii, Phys. Lett. **103B** (1981),29 ,**107B** (1981),51.
- 50) M. Sher, Phys. Rev. **D24** (1981),1699.
- 51) S.W. Hawking, Phys. Lett. **115B** (1982), 295.
- 52) A. Guth and S.-Y. Pi, Phys. Rev. Lett.**49** (1982), 1110.
- 53) A.A. Starobinskii, Phys. Lett. **117B** (1982), 175.
- 54) J.M. Bardeen, P.J. Steinhardt, and M.S. Turner, Phys. Rev. D, in press (1983)
- 55) J. Ellis, D.V. Nanopoulos, K.A. Olive and K. Tamvakis, Phys. Lett. **118B** (1982), 335, **120B** (1983), 331, and Nucl. Phys.**B221** (1983), 524.  
D.V. Nanopoulos, K.A. Olive, M. Srednicki and K. Tamvakis, Phys. Lett. **123B** (1983), 41, **124B** (1983), 243.

- D.V. Nanopoulos, K.A. Olive and M. Srednicki, Phys. Lett.127B (1983), 30.  
 G.B. Gelmini, D.V. Nanopoulos and K. Olive, Phys. Lett.131B (1983) 53.  
 D.V. Nanopoulos and M. Srednicki, Phys. Lett. 133B (1983), 287.
- 56) A. Albrecht, S. Dimopoulos, W. Fisher, E.W. Kolb, S. Raby and P.J. Steinhardt, Nucl. Phys. B, to be published (1983)
- 57) A. Albrecht and P. Steinhardt, Phys. Lett. 131B (1983), 45.
- 58) B.A. Ovrut, and P.J. Steinhardt, RU-83/B/65.
- 59) A.D. Linde, Phys. Lett. 131B (1983), 330 and 132B (1983), 317.
- 60) A.D. Linde, Phys. Lett. 129B (1983), 177.

Figure Captions

- Fig.1.1.1: The schematic diagram of the effective potential of Higgs field in first order phase transition model.
- Fig.1.1.2: The Einstein equation for the expansion of the universe corresponds to the mechanical equation for a particle moving in the potential  $V(x)$  if we interpret  $x$  as a particle position .
- Fig.1.2.1: Evolutions of the particle horizon and the cosmic scale factor in the inflationary universe model. For comparison, those of the standard big bang model (SBBM) also are displayed. If the finishing time of the phase transition is about sixty,  $t_f > 60 t_p$ , the particle horizon at present becomes greater than the Hubble length  $H_0^{-1} \sim 10^{28}$  cm and the homogeneity of the universe may be explained.
- Fig.1.4.1: A monopole is formed, a point surrounded by domains if the winding number is nonvanishing.
- Fig.2.1.1: Progress of the first order phase transition. In the early stage, bubbles are isolated each other (a). With increasing volume fraction of bubbles, they begin to overlap (b), and eventually all the false vacuum regions are surrounded by bubbles (c).

- Fig.2.2.1: Fractal structure of the universe when quantum nucleation rate is very small.
- Fig.3.1.1: The Coleman-Weinberg potential at finite temperature. A small potential barrier is formed by thermal effect. The gradient at  $\phi \ll \phi_0$  is very small.
- Fig.3.2.1: Two typical bounce solutions: Case i) The oscillation time at  $\phi \sim \phi_{top}$  is very shorter than the de Sitter time  $\pi/H$ , i.e.,  $H \ll m$ . Case ii) The opposite case  $H \gg m$ . In the second case, gravitational effects become essential.
- Fig.3.3.1: Contour map of the Coleman-Weinberg potential for the case  $C=1$  is displayed on the plane ( $x = 3\psi_1 = 3\psi_2 = 3\psi_3, y = \psi_4$ ). On this plane, there exist four  $SU(4) \times U(1)$  minima ( $x$ ) and two  $SU(3) \times SU(2) \times U(1)$  minima ( $*$ ). In this plane,  $\alpha = 0$  means  $SU(4) \times U(1)$  direction,  $\alpha = -0.29\pi$   $SU(3) \times SU(2) \times U(1)$  direction and  $\alpha = 0.21\pi$  the direction vertical to the  $SU(3) \times SU(2) \times U(1)$  direction. Numerical computation is carried out in the range  $-0.29\pi < \alpha < 0.21\pi$ .
- Fig.3.3.2a: Time evolution of the Higgs field for the case  $C=1, C_{vis}=0.1$  and  $\alpha=0.1\pi$  is projected on the same plane as shown in Fig.3.2.1.
- Fig.3.3.2b: Same as Fig.3.3.2a, but for the case  $C=1, C_{vis}=0.01$  and  $\alpha=-0.2\pi$ .
- Fig.3.3.2c: Same as Fig.3.3.2a, but for the case  $C=1, C_{vis}=0.001$  and  $\alpha=0.2\pi$ .
- Fig.3.3.3: The dependence of the degree of inflation on the initial angle  $\alpha$ . Initial absolute value of the Higgs field is assumed to be  $(\psi_1^2 + \psi_2^2 + \psi_3^2 + \psi_4^2)^{1/2} = 8 \times 10^{-6} \sigma$ .
- Fig.3.3.4: Summary of final states of the Higgs field in the time evolution calculations.
- Fig.4.1: Inflation universe models

Fig.1.1.1

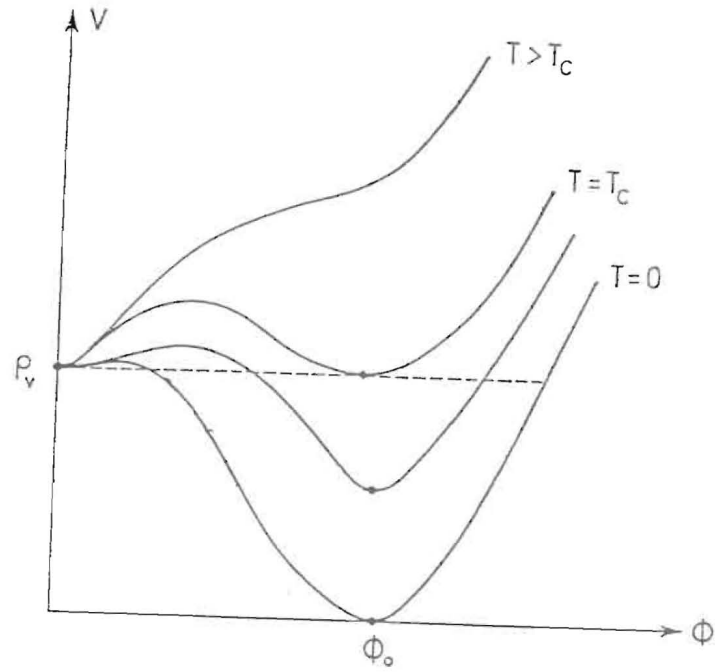


Fig.1.1.2

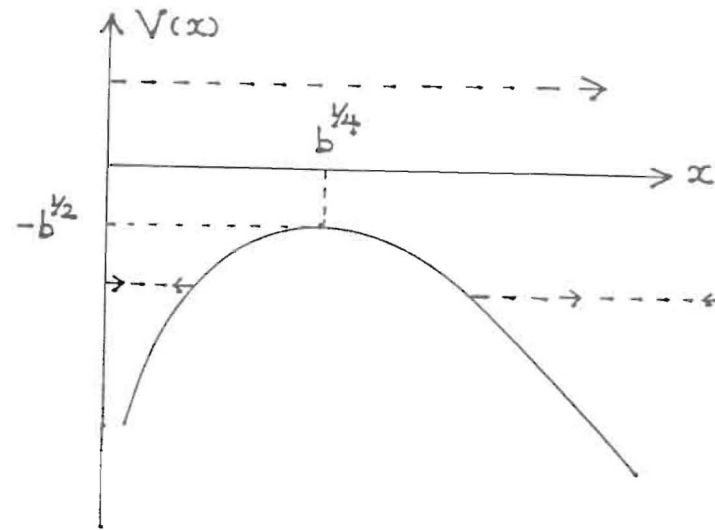


Fig.1.2.1

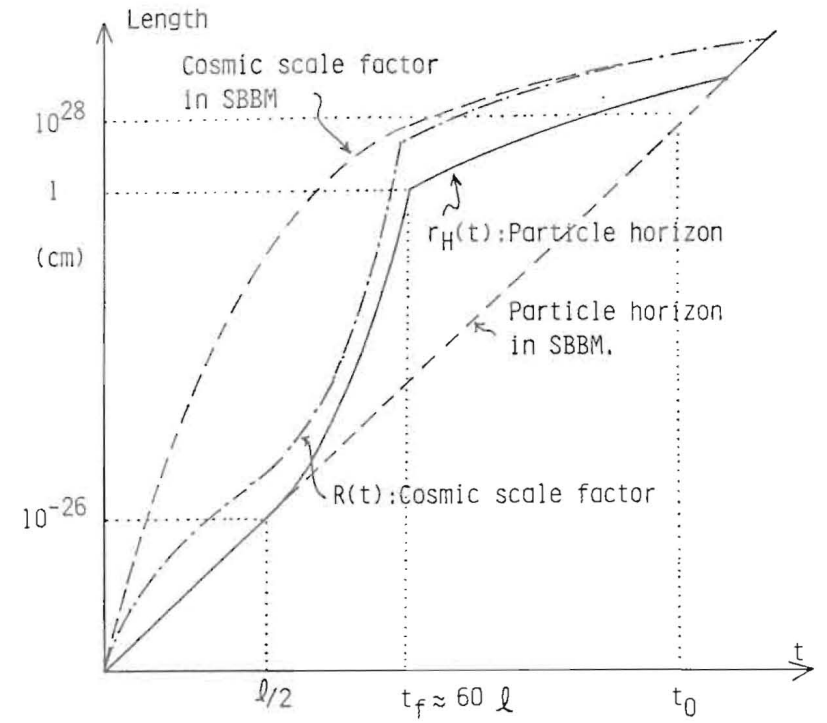


Fig. 2.2.1

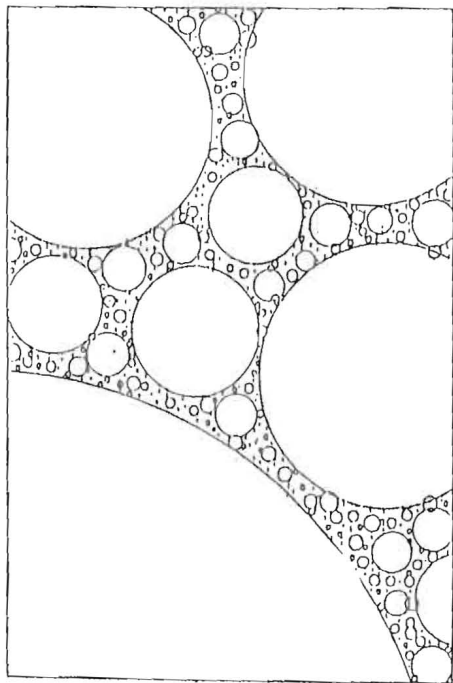


Fig. 1.4.1

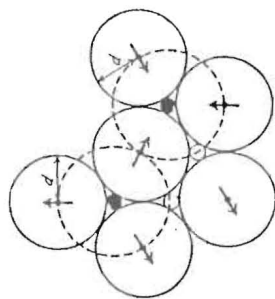


Fig. 2.1.1

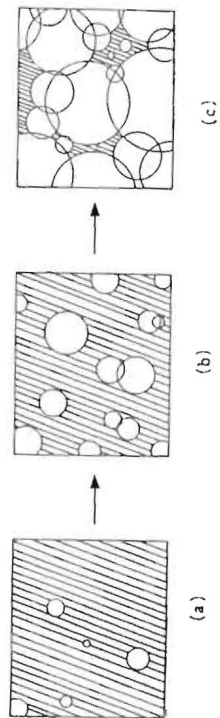


Fig. 3.1.1

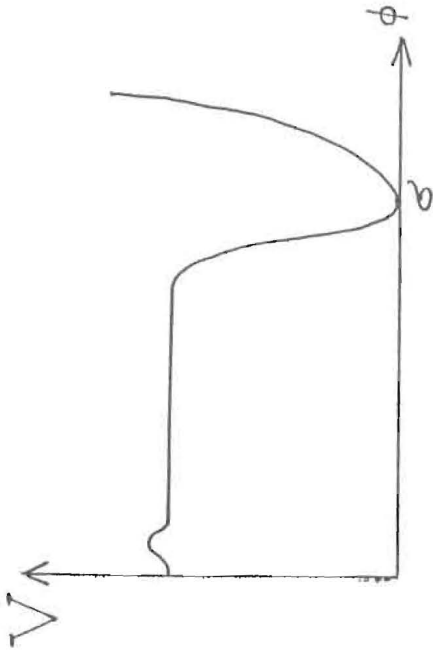


Fig. 3.2.1

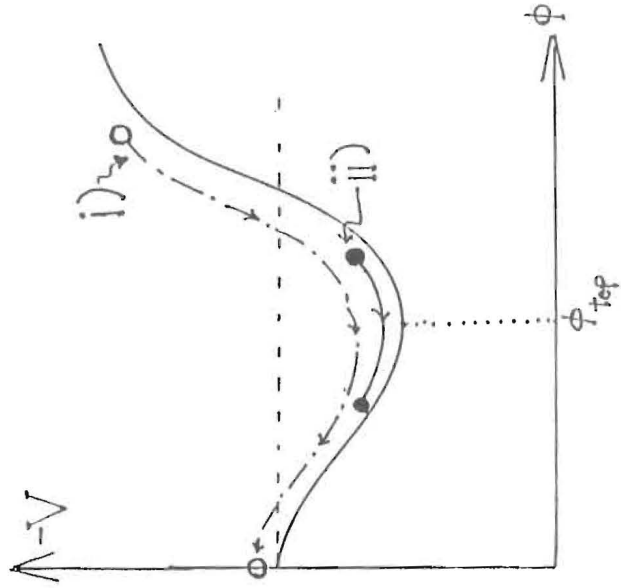


Fig. 3.3.1

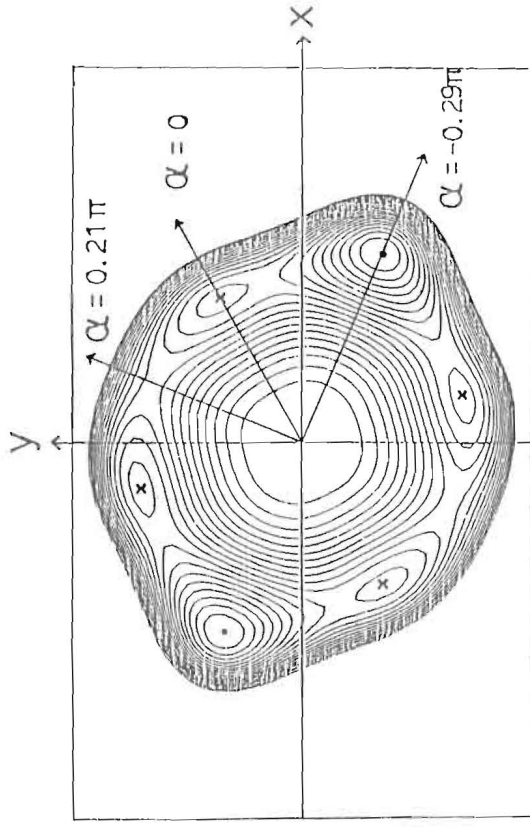


Fig.3.3.2a

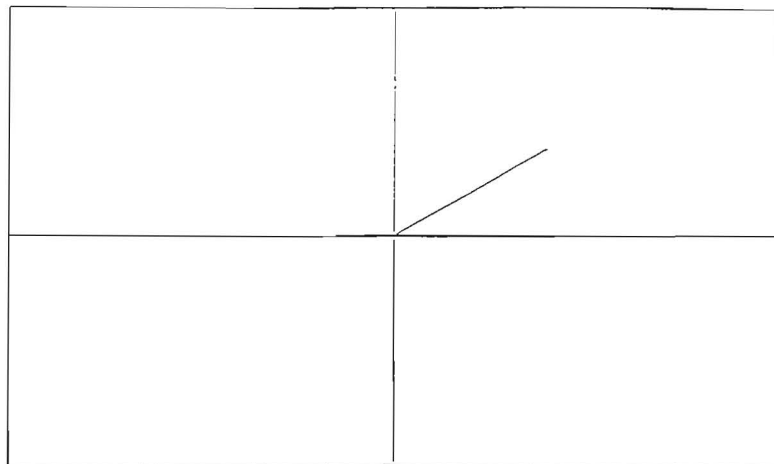


Fig.3.3.2b

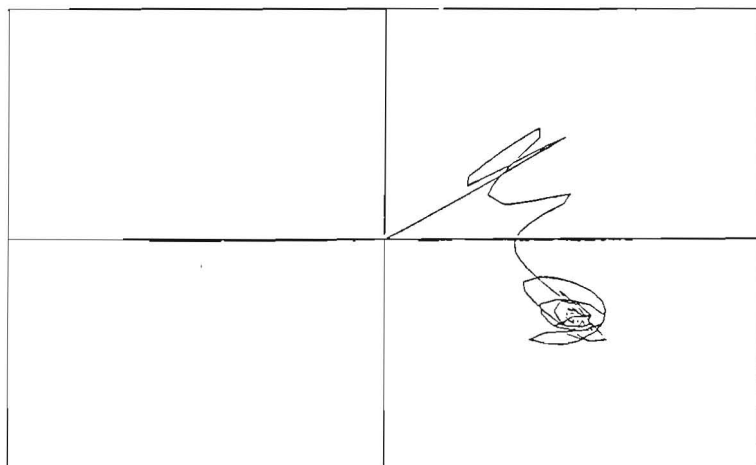


Fig.3.3.2c

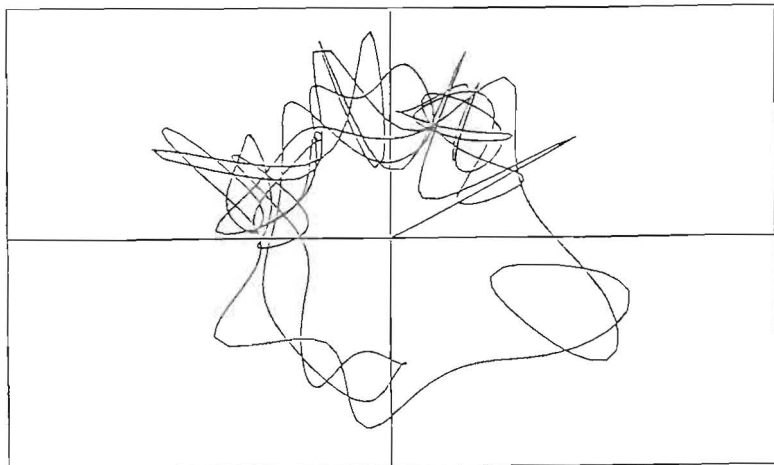
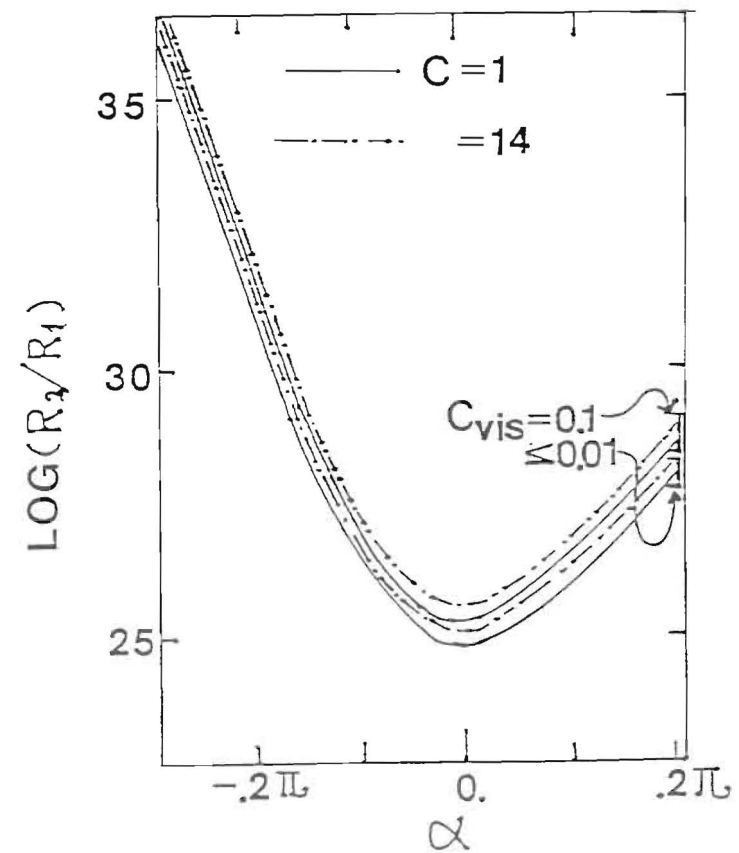
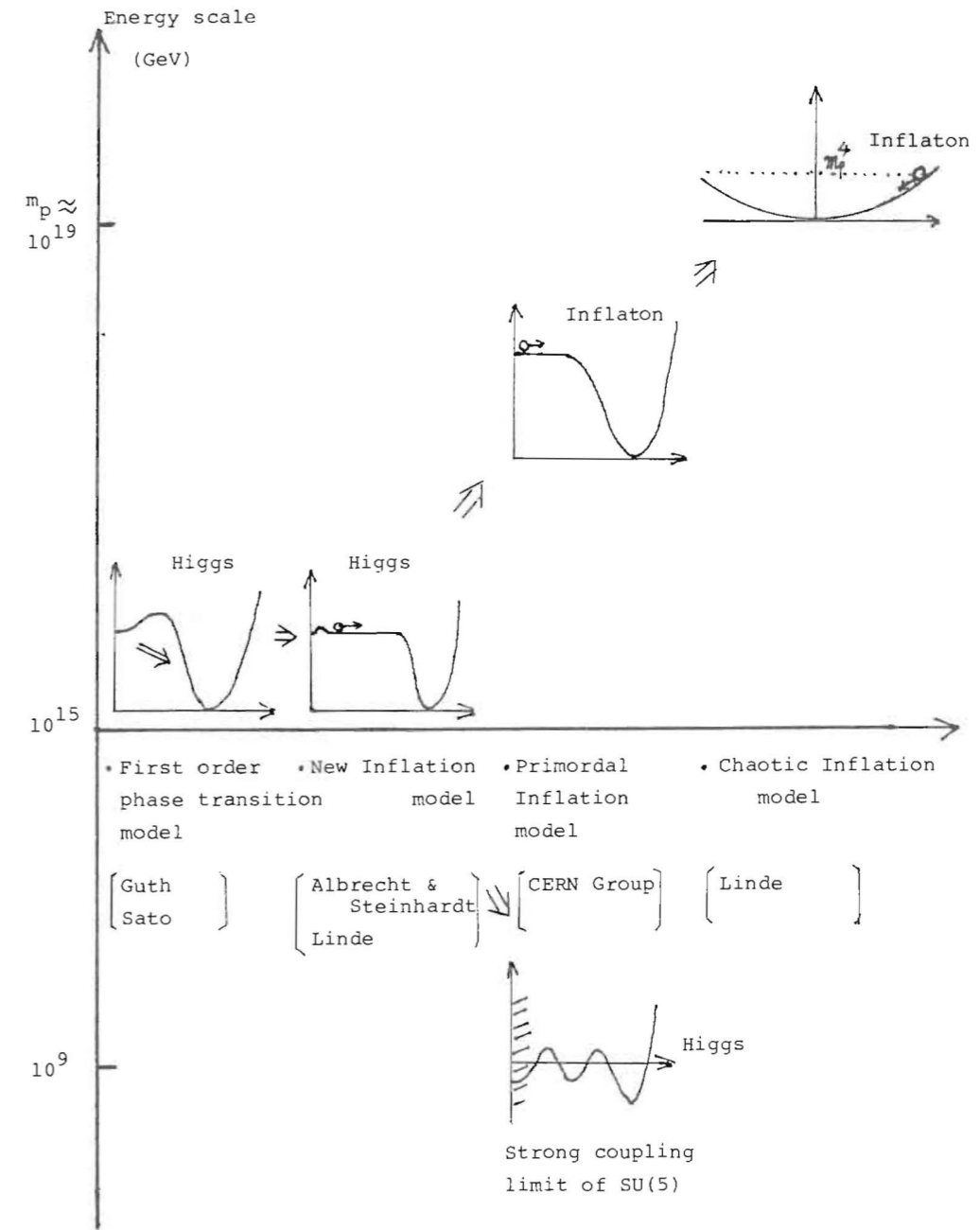
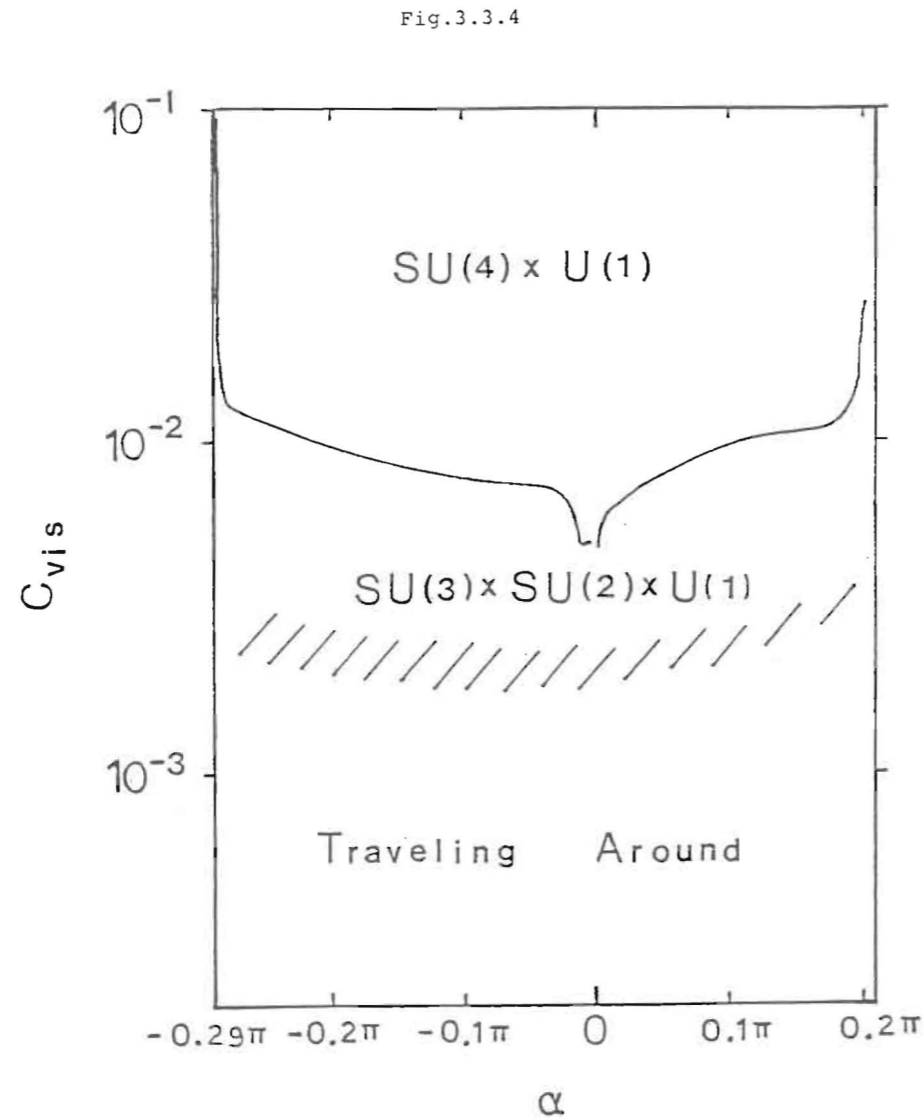


Fig.3.3.3







Problems in the Perturbation Analysis of a Universe  
 Dominated by a Coherent Scalar Field

Misao SASAKI

Department of Physics, Kyoto University, Kyoto 606

ABSTRACT

The temporal behavior of density fluctuations in a universe dominated by a coherent scalar field is investigated. Our analysis strongly indicates that the coherent nature of the scalar field restrains the growth of density inhomogeneities. This raises a serious problem against a cosmological model which is dominated by a coherent scalar field such as the invisible axion.

1. Introduction

It has been suggested that the axion which is introduced to solve the strong CP problem might be the dominant matter in the present universe.<sup>1)</sup> It has been further suggested recently that the axion fluctuations in an axion dominated universe might be the seeds of structures in the universe.<sup>2)</sup>

The scenario of an axion dominated universe is roughly as follows:<sup>1)</sup> The spontaneous breakdown of global  $U(1)_{PQ}$  symmetry at a temperature  $T \approx f_A$  produces a massless Nambu-Goldstone boson called the axion which is defined by  $\varphi = f_A \theta$  where  $\theta$  is the  $U(1)_{PQ}$  angle. Especially when  $f_A$  is much greater than the Weinberg-Salam scale  $\Lambda_{EW}$ , it is called the invisible axion and in the present scenario  $f_A \lesssim 10^{12} \text{ GeV}$  is assumed. When the temperature drops to  $\Lambda_{EW}$ , the axion begins to acquire its mass due to QCD instanton effects and at  $T \approx \Lambda_{QCD}$  it settles down to the zero temperature value  $m_A \approx f_\pi m_\pi / f_A \approx 10^{-5} (10^{12} \text{ GeV} / f_A) \text{ eV}$ . Then the coherent energy of the axion behaves as

$$\rho_A \approx \left( \frac{T}{\Lambda_{QCD}} \right)^3 m_A^2 f_A^2 \theta^2 \propto a^{-3}, \quad (1.1)$$

where  $a$  is the cosmic scale factor. Therefore, it behaves as the energy density of a non-relativistic dust matter and eventually dominates the energy density of the universe.

Naively, since the energy density of the axion field behaves as that of a dust matter, one may think that its perturbation is gravitationally unstable for every wavelength. However, it should be noted that this naive extrapolation neglects the coherent nature of the axion field completely. Hence let us consider whether the coherent nature is important or not.

When the  $U(1)_{PQ}$  breaks down spontaneously, the angle  $\theta$  (modulo  $2\pi$ ) would attain some finite expectation value which would be coherent over at least the horizon size of the universe,  $t_0$ , which is just the cosmic time of that epoch. This means the axion field,  $\varphi = f_A \theta$ , is realized classically and its energy momentum tensor contributes to the right hand side of the Einstein equations directly. Now since  $\varphi$  should be regarded as a classical field at the horizon scale, its spatial fluctuations on scales  $\lambda \gg t_0$  should be also regarded as real classical fluctuations. Therefore when we perform the perturbation analysis by Fourier transforming the fluctuations, the Fourier amplitudes  $\varphi_{\vec{k}}$  (where  $\vec{k}$  is the comoving wavenumber) for  $k/a(t_0) < t_0^{-1}$  are classically meaningful (observable) quantities by themselves. This is in contrast with the case when the field is highly incoherent and its Fourier amplitudes themselves cannot be well-defined classically. Thus in the

present case. Fourier amplitudes  $\varphi_{\vec{k}}$  should be added "coherently" to describe the spatial dependence of  $\varphi$  even though their phases may be quite random. This happens often in cosmological perturbation theories in which fluctuation scales of interests are so large that each perturbation directly reacts to the evolution of the universe in the region the perturbation is present.

## 2. Coherent scalar field and its fluctuations

In this section, we briefly review several properties of a coherent scalar field and give some basic ideas which will be used later when discussing the growth of perturbations of the field. For simplicity, we consider a free scalar field.

A quantum scalar field operator is represented by

$$\hat{\varphi} = \int \frac{d^3k}{(2\pi)^{3/2}} (\hat{a}_{\vec{k}} f_{\vec{k}}(t) + \hat{a}_{-\vec{k}}^+ \overline{f_{\vec{k}}(t)}) e^{-i\vec{k}\cdot\vec{x}}, \quad (2.1)$$

where  $\hat{a}_{\vec{k}}$  and  $\hat{a}_{\vec{k}}^+$  are the annihilation and creation operators, respectively, and  $f_{\vec{k}}(t)$  is the normalized positive frequency function. A coherent state is defined by the condition that it is an eigen state of  $\hat{a}_{\vec{k}}$ ; <sup>3)</sup>

$$\hat{a}_{\vec{k}} |\alpha_{\vec{k}}\rangle = \alpha_{\vec{k}} |\alpha_{\vec{k}}\rangle, \quad (2.2)$$

where  $\alpha_{\vec{k}}$  is a complex number. A general state which has certain coherence is now represented by a density matrix

$$\rho = \int P(\{\alpha_{\vec{k}}\}) \prod_{\vec{k}} |\alpha_{\vec{k}}\rangle \langle \alpha_{\vec{k}}| d^2\alpha_{\vec{k}}, \quad (2.3)$$

where  $\{\alpha_{\vec{k}}\}$  denotes the set of all amplitudes  $\alpha_{\vec{k}}$ . The weight function  $P(\{\alpha_{\vec{k}}\})$  is not positive definite due to the over-completeness of the states  $|\alpha_{\vec{k}}\rangle$ . <sup>3)</sup> However, it can be interpreted as a probability distribution function when amplitudes of the field are large or frequencies of the field are low enough. Then the expectation value  $\varphi \equiv \text{Tr}(\rho \hat{\varphi})$  behaves as a classical field and the correlation function  $G(x, y) \equiv \text{Tr}(\rho \hat{\varphi}(x) \hat{\varphi}(y))$  reduces to that of the classical field  $\varphi$ . In the case of the axion, since the expectation value of the axion field is probably homogeneous enough over the horizon scale with large amplitudes  $\varphi \sim f_A$  at the time of the spontaneous  $U(1)_{PQ}$  breaking, we may regard it to be a classical field with a strong coherence when considering the fluctuations on scales larger than the horizon.

Now let us consider a classical field whose low frequency amplitudes are appreciably large. Its Fourier decomposition has the same form as Eq. (2.1) except that  $\hat{a}_{\vec{k}}$  is replaced by a complex number  $A_{\vec{k}} \equiv \text{Tr}(\rho \hat{a}_{\vec{k}})$  and  $\hat{a}_{\vec{k}}^+$  by  $\bar{A}_{\vec{k}}$ . Therefore the Fourier amplitude  $\varphi_{\vec{k}}$  is given by

$$\varphi_{\vec{k}} = A_{\vec{k}} f_{\vec{k}}(t) + \bar{A}_{-\vec{k}} \overline{f_{\vec{k}}(t)} \quad (2.4)$$

Although the spatial average of  $\varphi$  is determined by the  $\vec{k}=0$  amplitude alone in the strict sense, in cosmological problems it is more meaningful to define the spatial average of  $\varphi$  over a scale  $L$  which is comparable to or greater than the present horizon size of the universe. Denoting such an average by  $\bar{\varphi}$ , it is given by

$$\bar{\varphi} \equiv \frac{1}{V} \int_V d^3x \varphi(x) \approx \int_{kL < 1} \frac{d^3k}{(2\pi)^{3/2}} \varphi_{\vec{k}}, \quad (2.5)$$

where  $V=L^3$  ( $L$  is a comoving coordinate length for the cosmological case). Thus the spatial average is determined by the spectrum of the low frequency amplitudes  $\varphi_{\vec{k}}$  ( $kL < 1$ ). Note that if  $\bar{\varphi}$  varies little when  $L$  is varied (which we expect in the cosmological case),  $\varphi_{\vec{k}} \propto k^{-3}$  for small  $k$ . This gives the so-called Zel'dovich spectrum on large scales  $kL < 1$ . <sup>4)</sup> Now with a fixed  $L$ , the background average and the perturbations on scales  $kL > 1$  (but not for so large  $k$  that the classical picture breaks down) of the energy momentum tensor are given by

$$T_{BQ}^{\mu\nu} = T^{\mu\nu}[\bar{\varphi}, \bar{\varphi}], \quad (2.6a)$$

$$(\delta T^{\mu\nu})_{\vec{k}} = 2 T^{\mu\nu}[\bar{\varphi}, \varphi_{\vec{k}}], \quad (2.6b)$$

where  $T^{\mu\nu}[\varphi, \varphi]$  is the value of the bilinear energy momentum tensor operated on a free scalar field  $\varphi$ . As noted above Eq.(2.6a) would give the Zel'dovich spectrum on large scales by itself if  $L$  is to be varied and if the universe is dominated by some other matter such as radiation. <sup>4)</sup> However when the energy density of the universe is dominated by the scalar field, one must define the background energy density somehow. Then it is more relevant to regard Eq.(2.6a) as the background by fixing  $L$  and consider fluctuations only on scales smaller than  $L$  which is given by Eq.(2.6b). Thus, in particular, the density perturbations are directly (linearly) related to the fluctuating field amplitudes  $\varphi_{\vec{k}}$ .

### 3. Temporal behavior of density perturbations

Now we investigate the temporal behavior of density perturbations in a universe dominated by a coherent scalar field. For definiteness we consider the case of the invisible axion with its symmetry breaking scale  $f_A \simeq 10^{12}$  GeV.

The background universe is assumed to be spatially homogeneous and flat;

$$ds^2 = -dt^2 + a^2(t) d\vec{x}^2 . \quad (3.1)$$

The background field  $\bar{\Phi}$  obeys the field equation given by

$$\ddot{\bar{\Phi}} + 3H\dot{\bar{\Phi}} + m^2\bar{\Phi} = 0 , \quad (3.2)$$

where  $H = \dot{a}/a$  and  $m$  is the mass of the axion. Since we are interested in the axion dominated stage of the universe, we assume  $m$  to be constant with time ( $m \simeq 10^{-5}$  eV). Now assuming  $m \gg H$  Eq.(3.2) can be solved approximately to yield

$$\bar{\Phi} \simeq \left(\frac{a_0}{a}\right)^{3/2} \bar{\Phi}_0 \sin mt , \quad (3.3)$$

where  $\bar{\Phi}_0$  is the amplitude of  $\bar{\Phi}$  at  $a=a_0$ . Note that for  $T < \Lambda_{\text{QCD}}$ , the inequality  $m \gg H$  is well satisfied. Then the background energy momentum tensor takes a perfect fluid form with the energy density  $\rho$  and the pressure  $P$  given by

$$\rho = \frac{1}{2} \dot{\bar{\Phi}}^2 + \frac{1}{2} m^2 \bar{\Phi}^2 , \quad (3.4a)$$

$$P = \frac{1}{2} \dot{\bar{\Phi}}^2 - \frac{1}{2} m^2 \bar{\Phi}^2 . \quad (3.4b)$$

Thus if one averages  $\rho$  and  $P$  over a time scale  $\gg m^{-1}$  one finds  $\langle \rho \rangle \propto a^{-3}$  and  $\langle P \rangle = 0$ , which is the same as the case of a dust-like fluid. This led some people to conclude that the Jeans length for the axion density perturbations is zero, i.e., the axion is gravitationally unstable for perturbations at any wavelength.<sup>2)</sup>

However we should remember that the behavior of perturbations depend essentially on the "sound velocity" of the background which is defined by the ratio of the time derivative of  $P$  and that of  $\rho$ ,  $\dot{P}/\dot{\rho}$ , but not  $P/\rho$ . From Eqs.(3.2) and (3.4) we have

$$\dot{\rho} = -3H\dot{\bar{\Phi}}^2 , \quad (3.5a)$$

$$\dot{P} = -3H\dot{\bar{\Phi}}^2 - 2m^2\bar{\Phi}\dot{\bar{\Phi}} . \quad (3.5b)$$

Therefore the square of the "sound velocity" is given by

$$C_s^2 \equiv \frac{\dot{P}}{\dot{\rho}} = 1 + \frac{2m\bar{\Phi}}{3H\dot{\bar{\Phi}}} \simeq 1 + \frac{2m}{3H} \tan mt , \quad (3.6)$$

which is clearly different from the value of  $P/\rho$ . Although there appears the singularities due to the function  $\tan mt$  in Eq.(3.6), they should not be the real ones. It disappears if one takes account of the presence of some other fluid like components such as radiation in the universe, however small their contribution to the energy density of the universe is. So, we assume  $C_s^2$  to be regular in reality.

We have shown in the previous section that density fluctuations of the axion field in the axion dominated stage are represented by Eq.(2.6b). Then combining it with the Einstein equation, one obtains the time evolution equation for  $\varphi_k$  or equivalently for  $\delta\rho_k$ . This has been done in a gauge-invariant way in Ref. 5). With a slight change of notations and variables it takes the form

$$\ddot{f} + (2+3C_s^2)H\dot{f} + \left(\frac{k^2}{a^2} - \frac{\chi}{2}(\rho+P)\right)f = 0 , \quad (3.7)$$

where  $\chi = 8\pi G$  and  $f = \rho a^3 \epsilon_m$  with  $\epsilon_m$  being the density contrast  $\delta\rho/\rho$  measured on the so-called comoving hypersurfaces.<sup>6)</sup> In the present case, since the amplitude of  $\rho a^3$  is approximately constant with time,  $f$  represents the amplitude of density contrast directly except for the small but rapidly oscillating part. Note that if the coefficient in front of  $(k/a)^2$  were  $C_s^2$  instead of unity, Eq.(3.7) would be exactly the same as the equation governing perturbations of a usual fluid.<sup>6)</sup> If this were so and if  $C_s^2=0$ , we would obtain the result which is exactly equivalent to the case of a dust-like fluid. However, the fact that we actually have the equation that differs completely from the dust-like fluid case indicates that we have a different answer.

Now let us investigate the temporal behavior of  $f$ . Since we are interested in cosmological density fluctuations we consider perturbations with wavenumber  $(k/a)^2 \ll m^2$  only. Note that for the invisible axion of mass of order  $10^{-5}$  eV, this corresponds to wavelengths  $\gg 1$  cm. For wavelengths of order 1 cm or less, the classical picture breaks down. We approach the problem in two different ways whose results are seemingly contradictory to each other. However there exists a possible resolution to this contradiction which will be discussed at the end of this section.

The first approach is to assume that the rapidly oscillating parts of  $C_s^2$ ,  $\rho$  and  $P$  at time scales of  $m^{-1}$  are inessential to the behavior of  $f$  over

an expansion time  $H^{-1}$ . Then we may take the temporal average of the coefficients in Eq.(3.7) for  $f$ . Since  $\langle C_s^2 \rangle = 1$ ,  $\langle \rho \rangle \propto a^{-3}$  and  $\langle P \rangle = 0$ , we obtain

$$\ddot{f} + 5 \langle H \rangle \dot{f} + \left( \frac{k^2}{a^2} - \frac{\chi}{2} \langle \rho \rangle \right) f = 0. \quad (3.8)$$

Thus the critically stable wavenumber is given by

$$\frac{k_c}{a} = \sqrt{\frac{\chi}{2} \langle \rho \rangle} = \sqrt{\frac{3}{2}} \langle H \rangle, \quad (3.9)$$

where we have used the Friedmann equation  $H^2 = \kappa \rho / 3$  in the latter equality. This implies the Jeans length of order  $H^{-1}$  (i.e. the horizon scale), which is analogous to the case of a relativistic fluid. The difference is that in the present case the effective sound velocity is equal to the light velocity, while for a relativistic fluid it is  $1/\sqrt{3}$  times smaller. The growth rate of  $f$  for  $(k/a)^2 \ll H^2$  can be easily estimated. Noting  $\langle H \rangle = 2/3t$  we obtain

$$f \propto a^m \propto t^{2m/3}; \quad m = \frac{\sqrt{73} - 7}{4} \approx 0.386. \quad (3.10)$$

However when the wavelength comes inside the horizon,  $f$  commences a damped oscillation given by

$$f \propto a^{-2} \cos\left(\left(\frac{k}{a}\right) H^{-1}\right). \quad (3.11)$$

This implies that no density fluctuations in the axion field can survive.

The second approach is to take account of the oscillatory parts of  $C_s^2$ ,  $\rho$  and  $P$  but consider the behavior of  $f$  only over a time interval of order  $m^{-1}$ . Then we may neglect the effect of expansion and Eq.(3.7) becomes

$$\ddot{f} + 2m \tan mt \dot{f} + \left( \left(\frac{k}{a}\right)^2 - \chi \langle \rho \rangle \cos^2 mt \right) f = 0, \quad (3.12)$$

where  $a$  and  $\langle \rho \rangle$  are assumed to be constant with time. Except for the presence of  $\tan mt$  in front of  $\dot{f}$ , this equation takes the same form as the well-known Mathieu equation<sup>7)</sup> which is a Schrödinger type equation with a periodic potential. Therefore we expect there are certain unstable modes similar to the case of the Mathieu equation. As noted before, the singularities of  $\tan mt$  in front of  $\dot{f}$  must be artificial. Hence we should regard  $f$  to be a sufficiently smooth and regular function. Then Eq.(3.12) can be analyzed by a method similar to the one employed in analyzing the Mathieu equation<sup>7)</sup> (i.e. expanding  $f$  in terms of relevant periodic functions). In the present case, since we are considering the behavior of  $f$  over a time interval of order  $m^{-1}$ , in addition to the assumptions  $(k/a)^2 \ll m^2$  and  $H^2 \ll m^2$ ,

we should assume  $(k/a)^2 \gg H^2$ ; otherwise the expansion effect cannot be neglected. Under these assumptions, Kodama and I have found that the first critically stable mode is given by<sup>8)</sup>

$$\left( \frac{k_{c1}}{a} \right)^2 = 6 m H, \quad (3.13)$$

and for  $k < k_{c1}$   $f$  becomes unstable. Although there are many other critically stable modes at higher wavenumbers, they appear at wavenumbers of order  $m$ ; hence quantum effects become essential for them and our classical picture is not applicable.

Compared with the result of the first approach, the value of  $k_{c1}$  is  $2\sqrt{m/H}$  ( $\gg 1$ ) times greater than  $k_c$  of Eq.(3.9). Thus the results seem to be contradictory. However, we must remember that Eq.(3.12) is derived by neglecting the expansion effect and therefore correct only for a time interval of  $\Delta t \gtrsim m^{-1}$ . Although we have not succeeded to obtain the growth rate of unstable modes for  $k < k_{c1}$  due to the complexity of Eq.(3.12), a preliminary analysis indicates that it has the form  $\exp(\gamma mt)$  when  $\gamma = p(H/m)$  with some positive constant  $p$ . By extrapolating this result to a time interval of order  $H^{-1}$ , we would have

$$f \propto \exp\left(\int^t p H dt'\right) \propto a^p. \quad (3.14)$$

However at this time scale, the expansion effect becomes important and in particular the coefficient  $5H$  in front of  $\dot{f}$  in Eq.(3.7) (where  $(2+3C_s^2)H \approx 5H + 2m \tan mt$ ) cannot be neglected. If this is taken into account, it will induce an adiabatic damping of  $f$  of the form

$$f \propto a^{-q} \quad (q > 0), \quad (3.15)$$

Therefore the total growth rate of  $f$  will be given by  $f \propto a^{(p-q)}$  and if we can prove the inequality  $p < q$ , the contradiction disappears. Although I have not yet proved this conjecture, it is plausible that the first approach is qualitatively correct. This conjecture is supported by another example which will be discussed in the next section.

#### 4. Gravitationally bound system of a coherent scalar field

In the previous section we have investigated the behavior of linear perturbations which indicates that density perturbations of a coherent scalar field cannot grow. In this section, we consider whether the non-linearity of gravity can make a coherent scalar field to form a gravitationally bound

system.

The possibility of a star composed of a scalar boson was discussed by Ruffini and Bonazzola.<sup>9)</sup> They considered a quantum scalar field  $\hat{\varphi}$  coupled to the classical gravity through the Einstein equation of the form

$$G^{\mu\nu} = \kappa \langle \psi | : \hat{T}^{\mu\nu} : | \psi \rangle, \quad (4.1)$$

where  $G^{\mu\nu}$  is the Einstein tensor,  $:$  denotes the normal ordering operator and  $|\psi\rangle$  is the lowest energy state of a static spherical star they looked for. Because of the staticity and spherical symmetry,  $\hat{\varphi}$  is expanded as

$$\hat{\varphi} = \sum_{n,\ell,m} \{ \hat{a}_{n\ell m} \varphi_{n\ell}(r) Y_{\ell m}(\theta, \varphi) e^{-iE_{n\ell} t} + h.c. \}, \quad (4.2)$$

where  $\hat{a}_{n\ell m}$  is the annihilation operator of a  $(n, \ell, m)$ -quantum. Then they assumed that  $|\psi\rangle$  is a perfectly incoherent state composed of  $(n=1, \ell=0)$ -quanta, i.e.,

$$|\psi\rangle = \frac{1}{\sqrt{N!}} (\hat{a}_{100}^{\dagger})^N |0\rangle. \quad (4.3)$$

Perfect incoherence is apparent since this is an eigen state of the number operator so that the phase is completely undetermined. Then with a given  $N$ , Eq.(4.1) becomes an eigen value equation for  $E_{10}$  with  $\varphi_{10}(r)$  being an eigen function. Ruffini and Bonazzola found a non-vanishing eigen value for  $E_{10}$  as one naturally expects.

Now let us consider if a perfectly coherent field can form a static spherical star. Then by replacing  $|\psi\rangle$  of Eq.(4.3) by the coherent state  $|\alpha\rangle$ ;<sup>3)</sup>

$$|\alpha\rangle = \exp(\alpha \hat{a}_{100}^{\dagger} - \bar{\alpha} \hat{a}_{100}) |0\rangle, \quad (4.4)$$

it is easy to write down the Einstein equation corresponding to Eq.(4.1). Then the resulting equation has the same form as the one considered by Ruffini and Bonazzola if one sets  $E_{10}=0$  and regard  $\varphi_{10}(r)$  as the coherent field  $\varphi(r)$ . This implies that a coherent scalar field cannot bound by itself unless the coherence is effectively destroyed by some mechanism. However for a super weakly interacting field such as the invisible axion, there seems to be no such mechanism.

## 5. Summary

We have investigated the behavior of density perturbations in a universe dominated by a coherent scalar field. The linear analysis indicates that

density fluctuations cannot grow inside the horizon, while the consideration on a gravitationally bound system of a coherent scalar field leads us to conclude that a coherent scalar field cannot bound gravitationally by itself. The growth of density perturbations on scales greater than  $H^{-1}$  is not really dynamical as discussed by Bardeen.<sup>6)</sup> Therefore we conclude that in a universe dominated by a coherent scalar field, no density inhomogeneities can be induced by fluctuations in the coherent scalar field.

## Acknowledgement

I would like to thank Prof. M. Fukugita, Prof. H. Sato and Dr. H. Kodama for valuable discussions.

## References

- 1) J. Preskill, M.B. Wise and F. Wilczek, Phys. Lett. 120B (1983) 127; L.F. Abbott and P. Sikivie, Phys. Lett. 120B (1983) 133; M. Dine and W. Fischler, Phys. Lett. 120B (1983) 137.
- 2) M.S. Turner, F. Wilczek and A. Zee, Phys. Lett. 125B (1983) 35; M. Fukugita and M. Yoshimura, Phys. Lett. 127B (1983) 181; J. Ipser and P. Sikivie, Phys. Rev. Lett. 50 (1983) 925.
- 3) R.J. Glauber, Phys. Rev. 131 (1963) 2766.
- 4) A. Vilenkin, Phys. Rev. Lett. 48 (1981) 59.
- 5) M. Sasaki, Prog. Theor. Phys. 70 (1983) 394.
- 6) J.M. Bardeen, Phys. Rev. D22 (1980) 1882.
- 7) N.W. McLachlan, Theory and application of Mathieu functions (Oxford University Press, Oxford, 1947).
- 8) H. Kodama and M. Sasaki, unpublished.
- 9) R. Ruffini and S. Bonazzola, Phys. Rev. 187 (1969) 1767.

## Comments on the Chaotic Inflation

Hideo Kodama

Department of Physics, Faculty of Science, University of Tokyo

113 Tokyo, Japan

### ABSTRACT

Some aspects of the chaotic inflation scenario proposed by Linde are discussed. Especially the two problems, whether sufficient reheating of the universe occurs and whether the large scale inhomogeneities produced by quantum fluctuation of the inflation field are small enough, are studied in detail. It is shown that rather strong constraints on the model parameters should be satisfied in order that the chaotic inflation scenario is a viable one.

### § 1. Introduction

The inflationary universe scenario was born as a natural consequence of the application of the grand unified gauge theories (GUTs) to the very early stage of the universe [1]. Since the original version of the scenario based on the strongly first-order phase transition was shown not to work well [2], a revised version based on the Coleman-Weinberg symmetry breaking mechanism was soon proposed [3]. Though it looked like a very promising one, the detailed investigation has revealed that this new inflationary universe scenario also has some serious difficulties as well as various fascinating features. The main difficulties are summarized as follows:(see [4] for a detailed account)

- 1) Extreme fine tuning of the parameters of the theory is necessary so as for a sufficient inflation to be achieved [5],[6].
- 2) Even with an appropriate fine tuning it is not guaranteed that the desired symmetry breaking pattern of the gauge group is realized since the effective potential of the symmetry breaking field ( the Higgs field) has very complex structure. If the desired pattern is not realized, too strong large scale inhomogeneities or too large number of monopoles may be produced by the GUT phase transition [7]-[9].
- 3) Quantum fluctuations of Higgs fields enlarged and amplified by the rapid expansion of the universe in the inflation stage may produce too strong large scale density irregularities to be consistent with the observation of the anisotropy in the cosmic microwave radiation and the large scale structure of the present universe [10]-[12].

Though these difficulties are quite serious, the inflationary universe scenario is too fascinating to be abandoned. The difficulties essentially come from the use of GUTs, whereas the fascinating features are brought about by the mere existence of the rapidly expanding stage in the early universe. With these points in mind some people proposed to separate the inflation from GUTs and constructed so-called primordial inflation scenarios [13]-[15]. The liberation from GUTs automatically resolves the difficulty 2) and also relaxes the difficulty 3) because the self-coupling constant of the scalar field which brings about the inflation is no longer constrained by the gauge coupling constant. The term "primordial" comes from that the inflationary stage appears around the Planck time in these scenarios. This primordialness of the inflation resolves the difficulty 1). Linde's chaotic inflation

scenario is one which exhibits the essential features of the primordial inflation scenario most clearly [16]. In this paper we examine whether Linde's scenario is a viable one or not as a model of our universe by studying the time evolution of the universe taking into account the reheating process and estimating the amplitude of the density fluctuations produced during the inflation stage.

## § 2. Chaotic Inflation Scenario

Let  $\phi$  be a singlet scalar field with a self interaction given by a potential  $V(\phi)$  which produces the inflation stage. Then  $\phi$  follows the time evolution equation

$$\ddot{\phi} + 3H \dot{\phi} + V'(\phi) = 0, \quad (1)$$

where a dot denotes the differentiation by time and a prime the one by  $\phi$ .  $H$  is the expansion rate of the universe given by

$$(\dot{a}/a)^2 = H^2 = (8\pi G/3)(\rho_r + \rho_\phi), \quad (2)$$

where  $\rho_r$  is the energy density of matter other than the  $\phi$  field and  $\rho_\phi$  is the energy density of  $\phi$  field given by

$$\rho_\phi = (1/2)\dot{\phi}^2 + V(\phi). \quad (3)$$

Let  $m=m(\phi)$  be the curvature of the potential  $V(\phi)$ :

$$m^2 = |V''(\phi)|. \quad (4)$$

If  $H \gg m$ , then a simple argument shows that the characteristic time scale of change of  $\phi$  is given by  $\Delta t \sim H/m^2$ . Within this time scale  $\rho_\phi$  remains nearly constant, hence if  $H\Delta t$  is sufficiently larger than unity,  $\phi$  field gives rise to the inflation. This condition is written as

$$H\Delta t \sim (H/m)^2 \gg 1. \quad (5)$$

Thus if the potential is sufficiently flat, the inflation occurs.

Noticing this point, Linde proposed the following simple primordial inflation scenario [16]. Let  $V(\phi)$  be given by

$$V = (1/4)\lambda\phi^4. \quad (\lambda \ll 1) \quad (6)$$

At the Planck time  $t_{p1}$  when the quantum effect on the gravitational field becomes negligible it is natural to assume that the  $\phi$  field is in a chaotic

state and the average energy density of the  $\phi$  field is about the Planck energy density  $\rho_{p1}$ . Then the  $\phi$  field has fluctuations of order  $\phi_0 = m_{p1}/\lambda^{1/4} \gg m_{p1}$ . Since the characteristic time scale of the change of  $\phi$ ,  $H/m(\phi)^2 \sim m_{p1}/\lambda\phi^2 > 1/\lambda^{1/2}m_{p1}$ , is much larger than the Planck time, the  $\phi$  field can be treated as a classical field. Hence the universe is divided randomly into domains in each of which the  $\phi$  field is coherent and takes a uniform expectation value between  $-\phi_0$  and  $\phi_0$  and it is most probable that  $\phi = 0(\phi_0)$  (see Fig.1). The typical size  $L$  of this coherent domain is estimated by assuming that the average kinetic energy and the average potential energy of the  $\phi$  field is nearly equal, namely  $(\partial_\mu\phi)^2 \sim m_{p1}^4$ :

$$L \sim \phi/m_{p1}^2 \sim (1/m_{p1})\lambda^{-1/4}. \quad (7)$$

Note that  $L$  is much larger than the horizon scale at the Planck time  $1/m_{p1}$ . Since the potential is very flat, the typical roll down time scale of  $\phi$  is very long and each domain begins inflation as time goes on (see Fig.2). The typical inflation rate is now given by

$$\exp(H\Delta t) \sim \exp(4\pi/9\lambda^{1/2}). \quad (8)$$

Hence if  $\lambda < 0.0001$ , most of the domains suffer sufficient inflation such that each domain expands to a scale large enough for the horizon problem and the flatness problem to be resolved. The most charming point of Linde's scenario consists in that no metastable state of the inflation field is required.

Though Linde's scenario is very simple and attractive, two problems should be solved for it to be accepted as a viable universe scenario. One is to construct a realistic elementary particle model in which a singlet scalar field with a weak self-coupling is a natural ingredient. The other is to show that the universe is reheated up to a temperature high enough for the sufficient baryosynthesis to occur. In the remainder of this paper I investigate the latter problem in some detail.

## § 3. Reheating in Linde's Scenario

In this paper it is assumed that the reheating process is phenomenologically described by adding a viscosity term of form  $f_{vis} = -\gamma|\dot{\phi}|^2$  with a dimensionless constant  $\gamma$  to the right hand side of Eq.(1) and a corresponding source term to the energy conservation equation. Hence the fundamental equations are given by

$$\ddot{\phi} + \lambda \phi^3 = -(3H + \gamma|\phi|) \dot{\phi}, \quad (9)$$

$$(\dot{a}/a)^2 = H^2 = \rho_r + (1/2)\dot{\phi}^2 + (1/4)\lambda\phi^4, \quad (10)$$

$$\dot{\rho}_r = -4H\rho_r + \gamma|\phi|\dot{\phi}^2, \quad (11)$$

Throughout this paper the units  $c = \hbar = 8\pi G/3 = 1$  are used. Further it is assumed that  $\rho_\phi \approx \rho_r \approx (1/2)\rho_0 (\sim (8\pi/3)^2/2)$  initially at  $t=t_0 (\sim t_{p1})$ , hence  $\phi_0 = (2\rho_0/\lambda)^{1/4}$ .

In the stage  $\rho_\phi > \rho_r$ , Eqs.(9)-(11) have the following four characteristic time scales:

$$\text{cosmic expansion: } \tau_{\text{exp}}^{-1} = H \approx (1/2)\lambda^{1/2}\phi^2, \quad (12)$$

$$\text{viscosity: } \tau_{\text{vis}}^{-1} = \gamma|\phi|, \quad (13)$$

$$\text{dynamical: } \tau_{\text{dyn}}^{-1} = \frac{\lambda\phi^2}{3H + \gamma|\phi|} \quad \text{and} \quad \lambda^{1/2}|\phi|, \quad (14)$$

$$\text{reheating: } \tau_{\text{reh}}^{-1} = \frac{\gamma|\phi|\dot{\phi}^2}{\lambda|\phi|^4} = \frac{\gamma}{\lambda} \frac{1}{|\phi|} \tau_{\text{dyn}}^2. \quad (15)$$

Especially the dynamics of the  $\phi$  field is governed by the former three time scales. According to which is the shortest, three stages are generally possible: the one in which  $\tau_{\text{exp}}$  is the shortest and  $\ddot{\phi}$  is negligible, the one in which  $\tau_{\text{vis}}$  is the shortest and again  $\ddot{\phi}$  is negligible, and finally the one in which  $\tau_{\text{dyn}}$  is the shortest and  $\ddot{\phi}$  is no longer negligible. Depending on the values of  $\lambda$ ,  $\gamma$  and  $\phi_0$ , three combinations of these stages occur actually. We consider them separately.

$$(i) \quad \gamma/\lambda^{1/2} > \phi_0 \gg 1$$

In this case the viscosity dominates the field dynamics and  $\ddot{\phi}$  term is negligible throughout until the reheating time  $t_*$ . Hence during  $t_0 < t < t_*$  the solutions of Eqs.(9)-(10) for  $\phi$  and the cosmic scale factor  $a$  are expressed as

$$\phi \approx [\lambda t/\gamma + \phi_0^{-1}]^{-1}, \quad (16)$$

$$a/a_0 \approx \exp[\gamma(\phi_0 - \phi)/2\lambda^{1/2}], \quad (17)$$

Putting these expressions into Eq.(11), we obtain the following estimate for  $\rho_r$ :

$$\begin{aligned} \rho_r &= \rho_{r0}(a/a_0)^{-4} \\ &\quad + \lambda \exp(2\gamma\phi/\lambda^{1/2}) \int_{\phi_0}^{\phi} \phi'^3 \exp(-2\gamma\phi'/\lambda^{1/2}) d\phi' \\ &\approx \frac{\lambda^{3/2}}{2\gamma} \phi^3. \end{aligned} \quad (18)$$

The reheating time is determined by the condition  $\rho_r = \rho_\phi$ . From Eqs.(16)-(18) the values of various quantities at the reheating time (represented by the suffix \*) are estimated as

$$t_* \approx \gamma^{2/2} \lambda^{3/2}, \quad (19)$$

$$\phi_* \approx 2^{-1/2} \gamma, \quad (20)$$

$$\rho_{r*} \approx \rho_{\phi*} \approx 2\lambda^{3/2} \gamma^4, \quad (21)$$

$$a_*/a_0 \approx \exp(\gamma\phi_0/2\lambda^{1/2}) \quad (22)$$

After the reheating time  $\phi$  is still described by Eq.(16), and its modulus changes approximately in inverse proportion to time. As a consequence  $\rho_\phi$  decreases in proportion to  $t^{-4}$ , while  $\rho_r$  decreases in proportion to  $t^{-2}$ . The temporal behavior of the cosmic scale factor and the energy densities are schematically shown in Figs.3-a and 3-b.

$$(ii) \quad \phi_0 > \gamma/\lambda^{1/2} > 1$$

In this case two stages appear before reheating: first the cosmic expansion dominated stage and second the viscosity dominated stage. In the first stage the adiabatic damping term balances with the potential force term in Eq.(9) and  $\phi$  is given by

$$\phi \approx \phi_0 \exp[-(2\lambda^{1/2}/3)t]. \quad (23)$$

Substituting this expression into Eqs.(10) and (11), we obtain the following approximate expressions for the cosmic scale factor and the radiation energy density:

$$a/a_0 \approx \exp[(3/8)(\phi_0^2 - \phi^2)], \quad (24)$$

$$\rho_r \approx (2/9)\lambda^{1/2}\gamma\phi. \quad (25)$$

This stage terminates at the time  $t_1$  when  $\tau_{\text{vis}} = \tau_{\text{exp}}$ . The values of various quantities at  $t_1$  is give as

$$t_1 \approx (3/2\lambda^{1/2}) \ln(\lambda^{1/2}\phi_0/\gamma), \quad (26)$$



$$\phi_1 = \gamma/\lambda^{1/2}, \quad (27)$$

$$\rho_{r1} \approx (2/9)\gamma^2, \quad (28)$$

$$\rho_{\phi 1} \approx \gamma^4/2\lambda, \quad (29)$$

$$a_1/a_0 \approx \exp[(3/8)\phi_0^2(1-(\gamma/\phi_0\lambda^{1/2})^2)]. \quad (30)$$

After  $t_1$  the viscosity term balances with the potential force. Hence the behavior of various quantities are the same as in the case (i) and the expressions for them are obtained by simply replacing the suffix 0 by 1 and the time  $t$  by  $t-t_1$ . The energy densities and the cosmic scale factor at the reheating time  $t_*$  are given as follows:

$$\rho_{r*} \approx \rho_{\phi*} \approx 4\lambda^3/\gamma^4 \quad (31)$$

$$a_*/a_0 \approx \exp[(3/8)\phi_0^2 + \gamma^2/8\lambda]. \quad (32)$$

The temporal behavior of the cosmic scale factor and the energy densities in this case are shown in Figs.4-a and 4-b.

(iii)  $\phi_0 \gg 1 > \gamma/\lambda^{1/2}$

In this case the viscosity term does not play an important role until the reheating time. Hence the viscosity dominated stage does not appear, but instead the oscillatory stage during which  $\tau_{dyn}$  is the shortest appears. The first stage is the cosmic expansion dominated stage and the behavior of the various quantities in this stage are given by the same expressions as in the case (ii). This stage terminates at the time  $t_2$  when  $\tau_{exp} = \tau_{dyn}$ . The values at this time are given as

$$t_2 \approx (3/2\lambda^{1/2})\ln\phi_0, \quad (33)$$

$$\phi_2 \approx 1, \quad (34)$$

$$\rho_{r2} \approx 2\gamma\lambda^{1/2}/9, \quad (35)$$

$$\rho_{\phi 2} \approx \lambda/4, \quad (36)$$

$$a_2/a_1 \approx \exp[(3/8)\phi_0^2]. \quad (37)$$

After  $t_2$  neither the viscosity term nor the adiabatic damping term can balance with the potential force term, hence the second time derivative term can not be neglected any longer in Eq.(9). Therefore the time evolution of

$\phi$  is determined by

$$\ddot{\phi} + \lambda\phi^3 = 0. \quad (38)$$

The solution of this equation is well approximated by

$$\phi \approx f(t)\text{cn}(f(t)\lambda^{1/2}(t-t_2)|2^{1/2}), \quad (39)$$

where

$$f(t) = [3\alpha\lambda^{1/2}(t-t_2) + 1]^{-1/2}, \quad (40)$$

$\alpha$  is the average of the function  $[\text{sn}(x|2^{-1/2})\text{cn}(x|2^{-1/2})\text{dn}(x|2^{-1/2})]^2$  and  $\text{sn}(x|k)$ ,  $\text{cn}(x|k)$  and  $\text{dn}(x|k)$  are the Jacobi's elliptical functions of modulus  $k$ . Substituting this expression into Eqs.(10) and (11), we obtain the following approximate expressions for the cosmic scale factor and the radiation energy density:

$$a/a_2 \approx f^{-1/3\alpha}, \quad (41)$$

$$\rho_r \approx \frac{2\beta}{4-9\alpha} \lambda^{1/2} \gamma f^3, \quad (42)$$

where  $\beta$  is the average of  $\text{sn}^2(x)|\text{cn}(x)|\text{dn}^2(x)$ . As time goes on  $\rho_r$  decreases slower than  $\rho_{\phi}$  and at some time  $t_*$  they become equal. The values of various quantities at this reheating time is given by

$$t_* - t_2 \approx \lambda/3\alpha\gamma^3, \quad (43)$$

$$f_* \approx \gamma/\lambda^{1/2}, \quad (44)$$

$$a_*/a_2 \approx [\lambda^{1/2}/\gamma]^{1/\alpha}, \quad (45)$$

$$\rho_{r*} \approx \rho_{\phi*} \approx \gamma^4/4\lambda. \quad (46)$$

The total inflation rate is give by

$$a_*/a_0 \approx \exp[(3/8)\phi_0^2 + (1/3\alpha)\ln(\lambda^{1/2}/\gamma)]. \quad (47)$$

The temporal behaviors of the cosmic scale factor and the energy densities are shown in Figs.5-a and 5-b.

Now we summarize the results. The first interesting feature of the reheating process in Linde's inflation scenario is that the radiation energy density decreases monotonically through almost all of the stage. This behavior is to be contrasted with that in the ordinary GUT inflation scenario in which the radiation energy density decreases first in the adiabatic stage but begins to increase in the late stage when the reheating process becomes

important. The second point to be noted is the dependence of the radiation energy density at the reheating time,  $\rho_{r*}$ , on the viscosity coefficient  $\gamma$ . From Eqs.(21),(31) and (46)  $\rho_{r*}$  is given by

$$\rho_{r*} \approx \begin{cases} \lambda^3/\gamma^4 & \text{for } \gamma/\lambda^{1/2} > 1, \\ \gamma^4/\lambda & \text{for } \gamma/\lambda^{1/2} < 1. \end{cases} \quad (48)$$

It is easy to see that  $\rho_{r*}$  becomes maximum when  $\gamma = \lambda^{1/2}$ . Hence the large viscosity does not always imply the large reheating temperature.

#### § 4. Amplitude of Density Fluctuations

We normalize the cosmic scale factor as  $a_*=1$  in this section. Let  $t_k$  be the time when the proper size of the region with a comoving scale  $1/k$  coincides with the Hubble horizon size  $1/H$ . Then as shown by some people the amplitude of the density fluctuation with comoving scale  $1/k$  generated from the quantum fluctuation of the  $\phi$  field is given by

$$\delta_k \sim 4H^2/\dot{\phi} \text{ at } t=t_k, \quad (49)$$

when the fluctuation reenters the horizon [10]. From this formula and the temporal behavior of  $H$  and  $\dot{\phi}$  found in § 3, the amplitude and the spectrum of the density fluctuation are estimated as follows:

(i)  $\gamma/\lambda^{1/2} > \phi_0 > 1$

$$k \approx (e \lambda^{1/2}/2\gamma) \delta_k \exp[-(\gamma \delta_k/4\lambda)^{1/2}]. \quad (50)$$

(ii)  $\phi_0 > \gamma/\lambda^{1/2} > 1$

$$\text{For } (3/2)\lambda^{1/2} > \delta_k > (4/9)\gamma^3/\lambda$$

$$k \approx (1/2)\lambda^{1/2}(2\delta_k/3\lambda^{1/2})^{2/3} \exp[-(3/8)(2\delta_k/3\lambda^{1/2})^{2/3}], \quad (51)$$

$$\text{and for } 4\lambda/\gamma < \delta_k < 4\gamma^3/9\lambda$$

$$k \approx (\lambda^{3/2}/2\gamma^2)(\gamma \delta_k/\lambda) \exp[-(1/2)(\gamma \delta_k/\lambda)^{1/2}]. \quad (52)$$

(iii)  $\phi_0 > 1 > \gamma/\lambda^{1/2}$

$$\text{For } (3/2)\lambda^{1/2}\phi_0^{1/2} > \delta_k > \lambda^{1/2}$$

$$k \approx (1/2)(\lambda/\gamma^2)^{-1/6\alpha} \lambda^{1/2}(2\delta_k/3\lambda^{1/2})^{2/3} \times \exp[(3/8)\{1-(2\delta_k/3\lambda^{1/2})^{2/3}\}] \quad (53)$$

and for  $3\gamma/2\alpha < \delta_k < \lambda^{1/2}$

$$k \approx (1/2)\lambda^{-1/2}\gamma^{1/3\alpha} (3\alpha/2)^{-(1/3\alpha-2)} \delta_k^{-(1/3\alpha-2)}. \quad (54)$$

From these equations it is seen that the spectrum of the density fluctuations when they reenter the horizon is approximately given by the near Zeldovich spectrum

$$\delta_k \propto (\log k^{-1})^n, \quad (55)$$

where  $n=2$  for the fluctuations produced during the viscosity dominated stage and  $n=3/2$  for those produced during the cosmic expansion dominated stage.

#### § 5. Constraints

In order for an inflationary universe scenario to be a viable one, the following three conditions should be satisfied at least [15]: sufficient inflation, appropriate reheating and not too large density fluctuations.

Since the radiation energy density decreases in proportion to  $a^{-4}$ , the proper scale of a region with a fixed comoving scale is written as

$$L_* = (\rho_r/\rho_{r*})^{1/4} L. \quad (56)$$

Hence the first condition is expressed as

$$L_{h0}(a_*/a_0) \gg (\rho_{r,pr}/\rho_{r*})L_{h*}, \quad (57)$$

where  $L_h$  denotes the Hubble horizon size and the suffix  $pr$  represents the values at the present universe. This condition can be written in a more convenient form as

$$\ln(a_*/a_0) \gg 60 - (1/4)\ln(\rho_*/4\rho_0) + \ln[(4\rho_0)^{-1/4}L_0^{-1}(L_{h,pr}/3000\text{Mpc})]. \quad (58)$$

The second condition comes from the requirement that an appropriate number of baryons consistent with the observation are produced after the reheating of the universe. Hence in order to write down a precise condition one must specify the baryon production processes. Since the conventional GUTs can not be used now, some modified or new mechanism should be invented [17]. Unfortunately, however, there exists no satisfactory such mechanism. What can be said from the arguments made by various people so far[12],[13] is that the reheating temperature  $T_*$  should be probably larger than  $10^9$  GeV at least. This condition is expressed by  $\rho_*$  as

$$\rho_{*/4} \rho_0 < 10^{-40}. \quad (59)$$

We can not say any definite thing about the upper bound on  $T_*$  in general. However, if the conventional GUTs is correct, the reheating temperature should be low enough for the symmetry breaking of the grand unified group not to be restored after the reheating. This condition is expressed as  $T_* < 10^{15} \text{ GeV}$  or

$$\rho_{*/4} \rho_0 > 10^{-16}. \quad (60)$$

Finally from the observation of the microwave cosmic radiation and the result of the galaxy formation theories, the third condition is written as

$$\delta \sim 0(10^{-3}) \quad \text{for } L_{\text{pr}} \sim 100 \text{ Mpc}. \quad (61)$$

The constraints on the values of the self-coupling constant  $\lambda$  and the viscosity coefficient  $\gamma$  are summarized in Fig.6.

### § 6. Discussion and Conclusion

From Fig.6 one sees that there is a parameter region in which all the three conditions mentioned in §5, namely sufficient inflation, sufficient reheating and appropriate density fluctuations, are satisfied even if a natural constraint  $\gamma \leq \lambda$  is further imposed. Therefore Linde's chaotic inflation scenario can be a viable one, although the allowed parameter region is rather narrow and the corresponding coupling constant is very small:  $\lambda \sim 10^{-12}$ .

Finally I should make some comments. The first point is the possibility of constructing a realistic elementary particle model which leads naturally to Linde's chaotic inflation scenario. Probably the most promising model is the one based on the supergravity theory since the allowed self-coupling of the inflation field is extremely weak [13]. However, there occurs a very serious problem in a model based on supergravity. In the supergravity theory the interaction of a scalar field  $\phi$  is determined by a superpotential  $W(\phi)$

$$W(\phi) = \mu^3 [ a_0 + \sum (a_n/n) \phi^n ], \quad (62)$$

where  $a_n$  are dimensionless constants. The interaction potential  $V(\phi)$  is expressed by  $W(\phi)$  as

$$V(\phi) = \exp(|\phi|^2/2) [ |D_\phi W|^2 - (3/2) |W|^2 ], \quad (63)$$

where

$$D_\phi W \equiv \partial_\phi W + \phi^* W. \quad (64)$$

Due to the appearance of the exponential factor  $V(\phi)$  begins to increase very rapidly as  $|\phi|$  becomes greater than unity if  $W(\phi)$  is a polynomial of finite degree. Hence even if there exists an appropriate superpotential yielding a potential which is flat for sufficiently large  $|\phi|$ , one must take into account the contributions from very high degree terms in calculation.

The second point is concerned with the behavior of the temperature in the inflation stage. Recently some people argued that even if the conventional GUT phase transition occurs, in the primordial inflation scenario, it does not give rise to such difficulties as plagued the inflation scenario based on GUTs because the GUT phase transition occurs in the very early phase of the inflation stage due to the rapid decrease of temperature there. This argument does not hold in the Linde type scenario since the temperature decreases monotonically due to the viscosity effect throughout the inflationary stage from  $t_{\text{pl}}$  to  $t_*$ . This result seems to hold also in other primordial inflation scenarios as long as the reheating process is expressed by the viscosity term as is considered in this paper. Thus unless a model is constructed in which the reheating process is switched on only after a sufficient inflation has occurred, the primordial inflation scenario can not be liberated from the difficulties that plagued the GUT inflation scenario.

## References

- [1] K.Sato, Phys. Lett. **99B** (1981) 66; Month. Not. R. Astron. Soc. **195**(1981) 467.  
A.H.Guth, Phys. Rev. **D23** (1981) 347.  
A.H.Guth and E.J.Weinberg, Phys. Rev. Lett. **45** (1980) 1131; Phys. Rev. **D23** (1981) 876.
- [2] K.Sato, M.Sasaki, H.Kodama and M.Maeda, Phys. Lett. **108B** (1982) 103.  
H.Kodama, M.Sasaki and K.Sato, Prog. Theor. Phys. **68** (1982) 1979.  
S.W.Hawking, I.G.Moss and J.M.Stewart, Phys. Rev. **D26** (1982) 2681.
- [3] A.D.Linde, Phys. Lett. **108B** (1982) 389.  
A.Albrecht and P.J.Steinhardt, Phys. Rev. Lett. **48** (1982) 1220.
- [4] K.Sato, Inflationary Universe Model, in this Proceedings.
- [5] A.D.Linde, Phys. Lett. **116B** (1982) 335, 340.
- [6] A.Vilenkin and L.Ford, Phys. Rev. **D26** (1982) 1231.  
A.Vilenkin, Nucl. Phys. **B226** (1983) 504.
- [7] J.D.Breit, S.Gupta and A.Zachs, Phys. Rev. Lett. **51** (1983) 1007.
- [8] I.G.Moss, Phys. Lett. **128B** (1983) 385.
- [9] K.Sato and H.Kodama, preprint UTAP-3 (1984)
- [10] S.W.Hawking, Phys. Lett. **115B** (1982) 295.  
A.H.Guth and S-Y.Pi, Phys. Rev. Lett. **49** (1982) 1110.
- [11] J.Bardeen, P.J.Steinhardt and M.S.Turner, Phys. Rev. **D28** (1982) 679.
- [12] A.Vilenkin, Nucl. Phys. **B226** (1983) 527.  
S.W.Hawking and I.G.Moss, Nucl. Phys. **B224** (1983) 180.
- [13] J.Ellis, D.V.Nanopoulos, K.A.Olive and K.Tambakis, Phys. Lett. **118B** (1982) 335; **120B** (1983) 331; Nucl. Phys. **B221** (1983) 524.  
D.V.Nanopoulos, K.A.Olive, M.Srednicki and K.Tambakis, Phys. Lett. **123B** (1983) 41; **124B** (1983) 171; **127B** (1983)  
D.V.Nanopoulos and M.Srednicki, Phys. Lett. **133B** (1983) 287.  
C.B.Gelmini, D.V.Nanopoulos and K.A.Olive, Phys. Lett. **131B** (1983) 53.
- [14] A.D.Linde, Phys. Lett. **131B** (1983) 330; **132B** (1983) 317.
- [15] B.A.Ovrut and P.J.Steinhardt, Phys. Lett. **133B** (1983) 161.
- [16] A.D.Linde, Phys. Lett. **129B** (1983) 177.
- [17] D.V.Nanopoulos and K.Tambakis, Phys. Lett. **110B** (1982) 449.  
D.V.Nanopoulos, K.A.Olive and K.Tambakis, Phys. Lett. **115B** (1982) 15.  
R.Barbieri, D.V.Nanopoulos and A.Masiero, Phys. Lett. **B58** (1981) 191.

## Figure Captions

- Fig.1 Schematic diagram to show that it is most probable for  $\phi$  to take a value of  $O(\phi_0)$ . If the axis of the potential energy is divided into intervals of equal length, the length of the corresponding interval in the  $\phi$  axis decreases as the modulus of  $\phi$  increases. Hence if the values of  $V$  in each domain distribute randomly, it is most probable for the modulus of  $\phi$  to take the largest value.
- Fig.2 Schematic diagram illustrating the inflationary expansion of domains in the chaotic inflation scenario. A little fraction of the domains (represented by shading) may recollapse within a time scale of  $O(t_{p1})$ , whereas the most of them suffer inflation and become extremely large and flat.
- Fig.3 Fig.3a shows typical temporal behaviors of the energy densities of the radiation and the  $\phi$  field for the case (i) ( $\lambda=10^{-4}$  and  $\gamma=1$ ), and Fig.3b shows those of the cosmic scale factor and the Hubble horizon size for the same case. In Fig.3b the values on the vertical axis is for the cosmic scale factor  $a$ , and the scale for  $H^{-1}$  is different.
- Fig.4 The temporal behaviors of the same quantities as in Fig.3 for the case (ii) ( $\lambda=10^{-4}$  and  $\gamma=5 \times 10^{-2}$ ). The power in a parenthesis on the vertical axis of Fig.4b is for  $H^{-1}$ .
- Fig.5 The temporal behaviors of the same quantities as in Fig.4 for the case (iii) ( $\lambda=10^{-4}$  and  $\gamma=10^{-5}$ ).
- Fig.6 The allowed region in the parameter plain of  $\lambda$  and  $\gamma$ . The two dashed lines represent the boundaries of the three cases and the dash and dot line represents the line on which  $\gamma = \lambda / (4\rho_0)^{1/2}$ . The shaded regions are forbidden by the requirements of sufficient inflation and appropriate reheating. If the requirement on the amplitudes of density fluctuations is imposed, the allowed region is restricted to the region between the two curves corresponding to  $\delta=10^{-4}$  and  $\delta=10^{-2}$ . Probably the allowed region is further restricted to below the dash and dot line.

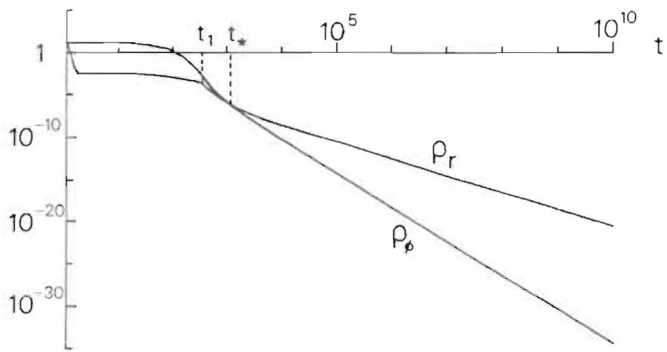


Fig. 4a

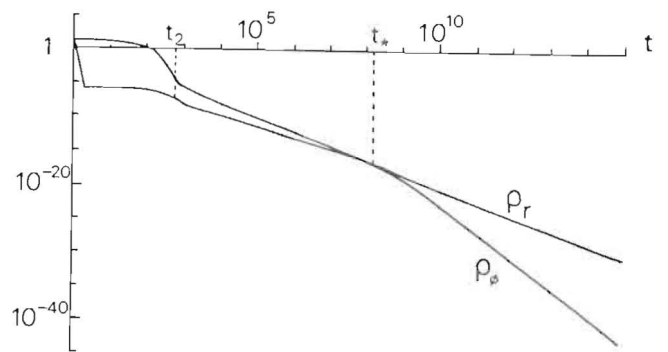


Fig. 5a

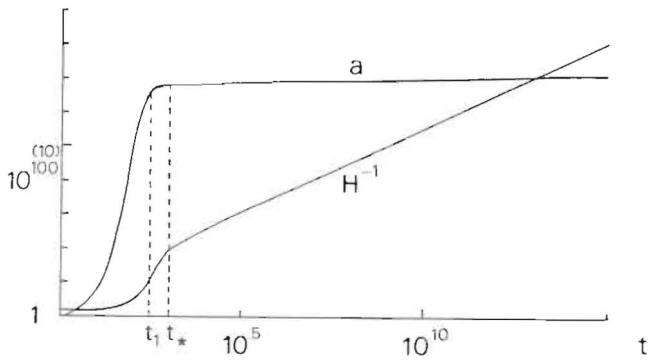


Fig. 4b

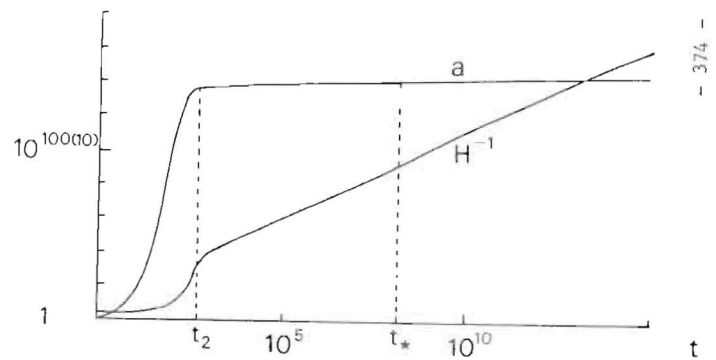


Fig. 5b

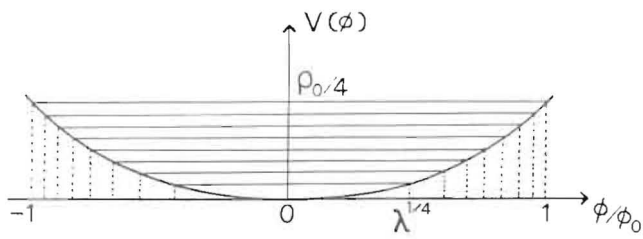


Fig. 1

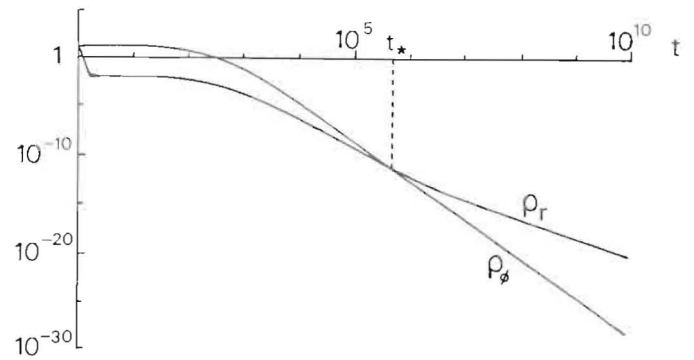


Fig. 3a

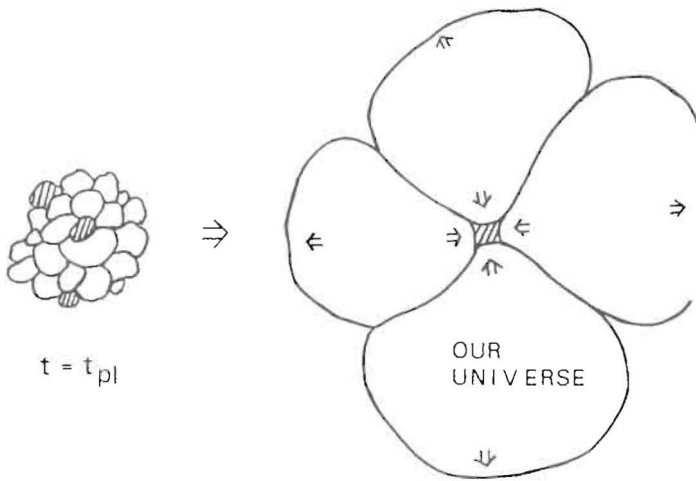


Fig. 2

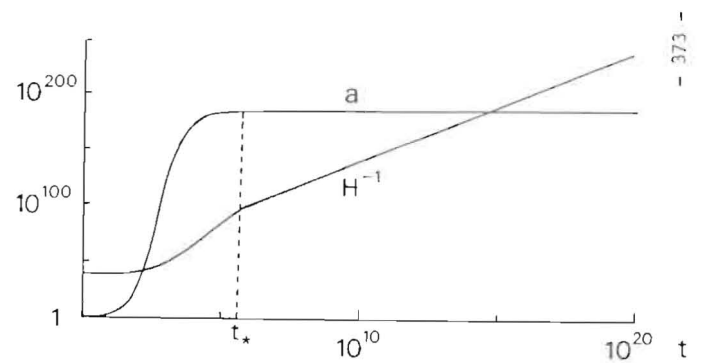


Fig. 3b

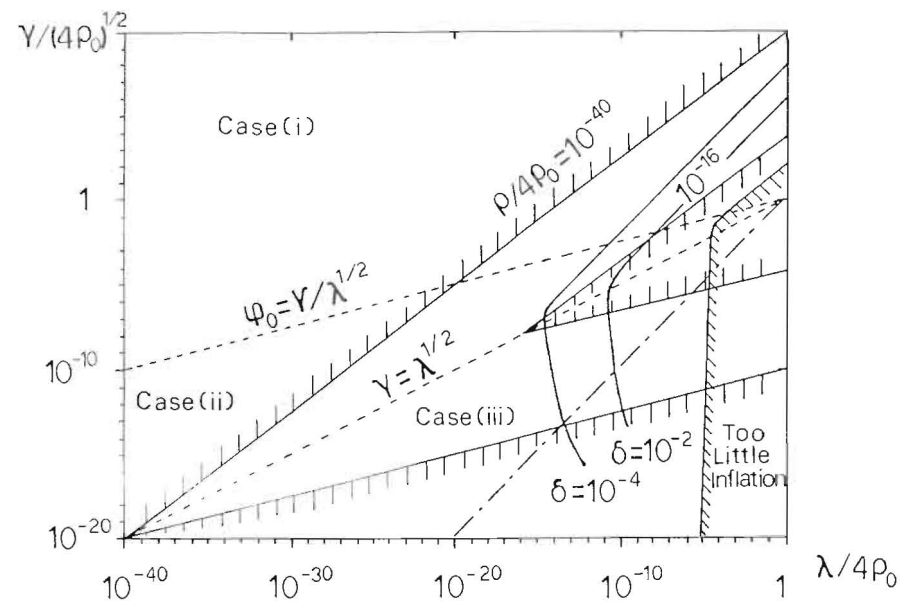


Fig. 6

Proton Decay in Supersymmetric Models

N. Sakai

Department of Physics, Tokyo Institute of Technology  
Oh-okayama, Meguro, Tokyo 152, Japan

ABSTRACT

Supersymmetric unified models and their implications to proton decay are reviewed. Particular emphasis is given to grand unified models embedded in the  $N=1$  supergravity. Dimension-five operators generally give dominant contributions in supersymmetric models. In the supergravity model with "the hidden sector" as the source of supersymmetry breaking, each dimension-five operator is found to accompany  $\Delta B \neq 0$  four-scalar interactions. The Higgs fermion exchange for loop diagrams at low energies can be as important as the gauge fermion exchange, if the associated Yukawa coupling is significant as suggested by the radiatively induced  $SU(2) \times U(1)$  breaking mechanism. Experimental bound for  $p \rightarrow K^0 \mu^+$  gives the lower bound of the order of  $10^{16}$  GeV for the mass of the baryon-number violating Higgs particle.

Electroweak interactions at present accelerator energies are beautifully described by the  $SU(2) \times U(1)$  gauge model[1]. The most dramatic confirmation has been obtained by the recent discovery of W and Z bosons at the CERN  $\bar{p}p$  collider[2]. It is now an experimentally established fact that nature has a mass scale  $M_W$  of order 100 GeV. In order to truly unify the electroweak interaction with QCD of strong interactions, grand unified theories necessarily predict an enormous mass scale  $M_{GUT}$  where the three gauge coupling constants merge into a single one[3]. Even if one does not accept the premises of the grand unification, one certainly believes in the existence of the mass scale  $M_{pl}$  for gravitational interactions. Apart from the possibility of relying upon the uncalculable non-perturbative effects, the only possibility to explain the enormous ratio between  $M_W$  and  $M_{GUT}$  (or  $M_{pl}$ ) from symmetry reasons seems to be supersymmetric models[4,5]. Namely the supersymmetry together with chiral symmetry can protect the Higgs scalars from acquiring the huge mass of order  $M_{GUT}$  (or  $M_{pl}$ ).

An extensive search to construct a realistic supersymmetric unified theory has been performed in last two or three years[6]. One of the most attractive class of models are the grand unified models embedded in the  $N=1$  supergravity[7-17]. In these models, supersymmetry is broken by vacuum expectation values of scalar fields in a "hidden sector", and the effect of the breaking is transmitted to ordinary matter ("the observable sector") only through gravity. Low energy effective theories of these models contain terms which break supersymmetry explicitly and softly. These terms, although numerous, are systematically given by only a few parameters derivable from supergravity and the hidden sector. Besides the incorporation of gravity, this control of supersymmetry breaking is the most attractive feature of the models. Extensive phenomenological analyses suggest that relatively large mass (about 100 GeV or so) for top quark or higher generation fermion may be needed to induce  $SU(2) \times U(1)$  breaking radiatively[14,17,18].

Proton decay should offer the most spectacular and important information on the grand unification. In supersymmetric models,  $\Delta B \neq 0$  operators with dimension-five were found[19,20]. They contribute to nucleon decay through one-loop diagram, which were explicitly evaluated for the  $SU(5)$  model[21-23]. On the other hand, a model-independent operator analysis showed that there are  $\Delta B \neq 0$  four-scalar interactions if the scale of supersymmetry breaking is intermediate between  $M_W$  and  $M_{GUT}$ [24]. They contribute to nucleon decay through two-loop diagrams and can be as important as the dimension-five operators. Since the supergravity model with the hidden sector also requires an intermediate

scale for supersymmetry breaking, one expects the existence of the  $\Delta B \neq 0$  four-scalar interactions.

We have examined the proton decay in the supergravity model in detail in ref. 25. There we have found that each dimension-five operator accompanies the  $\Delta B \neq 0$  four-scalar interaction with a definite ratio. The ratio is determined by the parameters which characterize the explicit soft breaking and are derivable from supergravity and the hidden sector. In evaluating loop diagrams for nucleon decay, we need to take account of Higgs fermion exchange, since the top quark or higher generation fermions may have relatively large Yukawa coupling in order to induce  $SU(2) \times U(1)$  breaking radiatively. The supergravity model also predicts a left-right transition term in scalar quark (and lepton) propagators as a soft breaking of supersymmetry. This term affects the chirality structure of the loop diagrams for nucleon decay, which is crucial for supersymmetric models.

In ref. 25 nucleon decay operators and loop diagrams are explicitly evaluated using  $SU(5)$  as the grand unification group in the supergravity model, in order to obtain a maximal predictive power. We find that most of the new features in the supergravity model are suppressed either by the specific Higgs coupling structure of  $SU(5)$  and/or symmetry associated with identical particles.

A possible exception is the Higgs fermion exchange diagram. Depending on the magnitude of the Kobayashi-Maskawa angles[26], the diagram can dominate over the usual gauge fermion exchange diagrams, if the mass of the top quark or higher generation fermions are 100 GeV or larger. It has been noted that in supersymmetric models,  $p \rightarrow K^+ \bar{\nu}_\tau$  and  $n \rightarrow K^0 \bar{\nu}_\tau$  are the dominant decay modes of nucleon and  $p \rightarrow K^0 \mu^+$  is most promising for observation, whereas the pion modes  $p \rightarrow \pi^0 e^+$  or  $p \rightarrow \pi^0 \mu^+$  are too small to be observed[21-23]. The Irvine-Michigan-Brookhaven collaboration recently gave a stringent experimental bound on the partial proton decay rate  $\Gamma(p \rightarrow K^0 \mu^+)$  into the final state  $K^0 \mu^+$ [27]

$$1/\Gamma(p \rightarrow K^0 \mu^+) > 2.0 \times 10^{31} \text{ years.} \quad (1)$$

With reasonable estimates for the Kobayashi-Maskawa angles and for other theoretical uncertainties, we obtain, from the experimental bound, a lower bound of the mass of the superheavy Higgs particle which mediates proton decay

$$M \gtrsim 10^{16} \text{ GeV.} \quad (2)$$

An improvement of factor 2-3 can be obtained if one uses more recent data from the experiment at Kamioka[28].

A more complete account of proton decay in the supergravity model can be found in ref. 25.

#### REFERENCES

1. S. Weinberg, Phys. Rev. Lett. 19 (1967) 1264; A. Salam, in "Elementary Particle Theory", ed. N. Svartholm (Almqvist and Wiksells, Stockholm, 1969) p. 367.
2. UA1 Collaboration, G. Arnison et al., Phys. Lett. 122B (1983) 103; *ibid* 126B (1983) 398.
3. H. Georgi and S. L. Glashow, Phys. Rev. Lett. 32 (1974) 438; H. Georgi, H. R. Quinn and S. Weinberg, *ibid* 33 (1974) 451.
4. N. Sakai, Z. f. Phys. C11 (1981) 153; S. Dimopoulos and H. Georgi, Nucl. Phys. B193 (1981) 150.
5. E. Witten, Nucl. Phys. B185 (1981) 513; M. Dine, W. Fischler and M. Srednicki, Nucl. Phys. B189 (1981) 575; S. Dimopoulos and S. Raby, Nucl. Phys. B192 (1981) 353.
6. A brief review of supersymmetric models before the summer of 1982 may be found in N. Sakai, in Proceedings of the International Symposium on Gauge Theory and Gravitation, Nara, Japan 1982 (Springer Lecture Notes in Physics 176).
7. A. H. Chamseddine, R. Arnowitt, and P. Nath, Phys. Rev. Lett. 49 (1982) 970; L. Ibanez, Phys. Lett. 118B (1982) 73.
8. R. Barbieri, S. Ferrara, and C. A. Savoy, Phys. Lett. 119B (1982) 343; J. Ellis, D. V. Nanopoulos, and K. Tamvakis, Phys. Lett. 121B (1983) 123.
9. H. P. Nilles, M. Srednicki, and D. Wyler, Phys. Lett. 120B (1983) 346.
10. S. Ferrara, D. V. Nanopoulos, and C. A. Savoy, Phys. Lett. 123B (1983) 214.
11. J. Leon, M. Quiros, and M. Ramon Medrano, Madrid preprint.
12. B. Ovrut and S. Raby, Phys. Lett. 121B (1983) 381.
13. L. Hall, J. Lykken, and S. Weinberg, Phys. Rev.
14. L. Alvarez-Gaume, J. Polchinski, and M. Wise, Harvard preprint HUTP-82/A063.
15. A. H. Chamseddine, P. Nath and R. Arnowitt, Harvard preprint HUTP-82/A057.
16. E. Cremmer, P. Fayet, and L. Girardello, Ecole Normale Supérieure preprint LPTENS 82/30.
17. L. Ibanez and C. Lopez, Madrid preprint, FTUAM/83-2.
18. K. Inoue, A. Kakuto and S. Takeshita, Kyushu Univ. preprint 83-HE-6, to appear in Prog. Theor. Phys.
19. N. Sakai and T. Yanagida, Nucl. Phys. B197 (1982) 533.
20. S. Weinberg, Phys. Rev. 26D (1982) 287.
21. J. Ellis, D. V. Nanopoulos, and S. Rudaz, Nucl. Phys. B202 (1982) 43.



22. P. Salati and J. C. Wallet, Nucl. Phys. B209 (1982) 389.
23. W. Lucha, Wien Univ. preprint, UWThPh-1982-29.
24. N. Sakai, Phys. Lett. 121B (1983) 130; one type of  $\Delta B \neq 0$  four-scalar interaction was found earlier in a model by J. P. Derendinger and C. A. Savoy, Phys. Lett. 118B (1982) 347.
25. N. Sakai, Tokyo Institute of Technology preprint, TIT/HEP-78 to appear in Nucl. Phys. B.
26. M. Kobayashi and T. Maskawa, Prog. Theor. Phys. 49 (1973) 652.
27. Irvine-Michigan-Brookhaven collaboration, result quoted in E. Fiorini's talk at the Lepton-Photon Conf. at Cornell (1983).
28. Private communications from M. Koshihara.

H. Komatsu

Institute for Nuclear Study, University of Tokyo,  
Tanashi, Tokyo 188

Abstract

The soft-breaking mass parameters in the supersymmetric grand unified theories based on N=1 supergravity are discussed to obtain mass relations of superpartners.

In this talk, I discuss the soft-breaking mass parameters to obtain mass relations of superpartners in the framework of the supersymmetric grand unified models based on spontaneously broken N=1 supergravity coupled with matter.<sup>1)</sup> The renormalization effects are taken into account by means of the renormalization group equations.<sup>2)</sup>

The low energy parameters including the soft-breaking masses are calculated by the renormalization group equations in the effective  $SU(3)_C \times SU(2) \times U(1)$  supersymmetric gauge theory with boundary conditions at  $\mu = M_X$  given by N=1 supergravity. The boundary conditions are parametrized as follows<sup>1)</sup>;

$$(\text{mass of squarks and sleptons}) = m_{3/2} ,$$

$$(\text{mass of gauginos}) = M_0 ,$$

$$(\text{mass of Higgsino}) = m_H ,$$

$$(\text{cubic scalar coupling}) = Am_{3/2} \quad (\text{Yukawa coupling}) ,$$

$$m_1^2 = m_2^2 = m_{3/2}^2 + m_H^2 ,$$

$$m_3^2 = Bm_{3/2}m_H ,$$

where  $m_1^2$ ,  $m_2^2$  and  $m_3^2$  are defined as soft-breaking mass parameters for Higgs scalars by

$$L' = -m_1^2 \phi_1^\dagger \phi_1 - m_2^2 \phi_2^\dagger \phi_2 - m_3^2 (\phi_1^\dagger \phi_2 + \phi_2^\dagger \phi_1) ,$$

In order for Higgs scalars to have desired vacuum expectation values,  $m_1^2$ ,  $m_2^2$  and  $m_3^2$  should satisfy the following conditions at  $\mu = M_W$ <sup>3)</sup>;

$$m_1^2 + m_2^2 > 2|m_3^2| , \quad m_1^2 m_2^2 < m_3^4 ,$$

$$M_Z^2 = \frac{-m_1^2 + m_2^2}{\cos 2\theta} - m_1^2 - m_2^2 ,$$

$$\sin 2\theta = \frac{2m_3^2}{m_1^2 + m_2^2}$$

$$\cos 2\theta \begin{cases} > 0 & (m_1^2 < m_2^2) \\ < 0 & (m_1^2 > m_2^2) . \end{cases}$$

In case where the top-quark Yukawa coupling is much larger than others,  $\cos 2\theta$  is negative.

Neglecting all Yukawa couplings, we can represent the masses of sleptons and squarks as follows;

$$m^2(\tilde{\nu}_L) = m_{3/2}^2 + 0.73M_0^2 + 0.50M_Z^2 \cos 2\theta ,$$

$$m^2(\tilde{e}_L) = m_{3/2}^2 + 0.73M_0^2 + 0.27M_Z^2 \cos 2\theta ,$$

$$m^2(\tilde{p}_L) = m_{3/2}^2 + 2.70M_0^2 - 0.35M_Z^2 \cos 2\theta ,$$

$$m^2(\tilde{p}_R) = m_{3/2}^2 + 2.62M_0^2 - 0.15M_Z^2 \cos 2\theta ,$$

$$m^2(\tilde{n}_L) = m_{3/2}^2 + 2.70M_0^2 + 0.42M_Z^2 \cos 2\theta ,$$

$$m^2(\tilde{n}_R) = m_{3/2}^2 + 2.61M_0^2 + 0.08M_Z^2 \cos 2\theta .$$

In these representations, the second terms are caused by the renormalization effects and the third terms by the  $SU(2) \times U(1)$  breaking effects.

In Figs. 1 and 2, the allowed region for masses of the lightest charged gauge-Higgs fermion and scalar electron ( $\tilde{e}_R$ ), respectively, are shown as functions of  $m_{3/2}$ . These calculations are made within the following region of parameter space;

$$m_t = 30 \sim 40 \text{ GeV} , \quad |A| < 3 ,$$

$$m_H/m_{3/2} < 1 , \quad B = (A^2+1)/(A-1) ,$$

$$m_H^2 + m_{3/2}^2 > |B m_H m_{3/2}| .$$

The last inequality is the vacuum stability condition at  $\mu = M_X$ . From these figures, we can conclude that one of charged gauge-Higgs fermions is lighter than scalar electrons if  $m(\tilde{e}_R)$  is larger than about 80 GeV.

More detailed discussion on the low energy parameters are given in ref. 4).

References

- 1) E. Cremmer, B. Julia, J. Scherk, S. Ferrara, L. Girardello and P. van Nieuwenhuizen, Phys. Lett. 79B(1978) 231; Nucl. Phys. B147(1979) 105;  
E. Cremmer, S. Ferrara, L. Girardello and A. van Proyen, Nucl. Phys. B212(1983) 413;  
T. Kugo and S. Uehara, Nucl. Phys. B222(1983) 125;  
P. Nath, R. Arnowitt and A. H. Chamsedine, Phys. Rev. Lett. 49(1982) 970;  
R. Barbieri, S. Ferrara and C. A. Savoy, Phys. Lett. 119B(1982) 343.
- 2) K. Inoue, A. Kakuto, H. Komatsu and S. Takeshita, Prog. Theor. Phys. 68(1982) 927;  
Prog. Theor. Phys. 71(1984) 413.
- 3) K. Inoue, A. Kakuto, H. Komatsu and S. Takeshita, Prog. Theor. Phys. 67(1982) 1889.
- 4) H. Komatsu, in "Proceedings of the 5th INS Winter Seminar" ed. A. Iwazaki (INS, 1984).

Fig. 1 Allowed region of the lightest charged gauge-Higgs fermion mass.

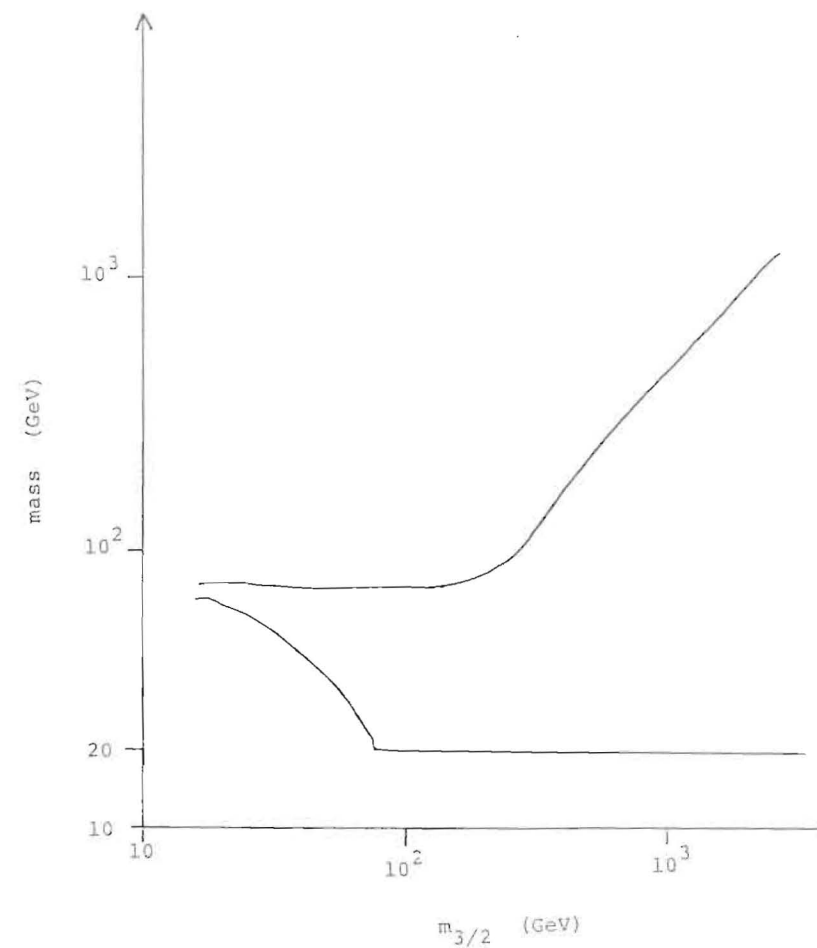
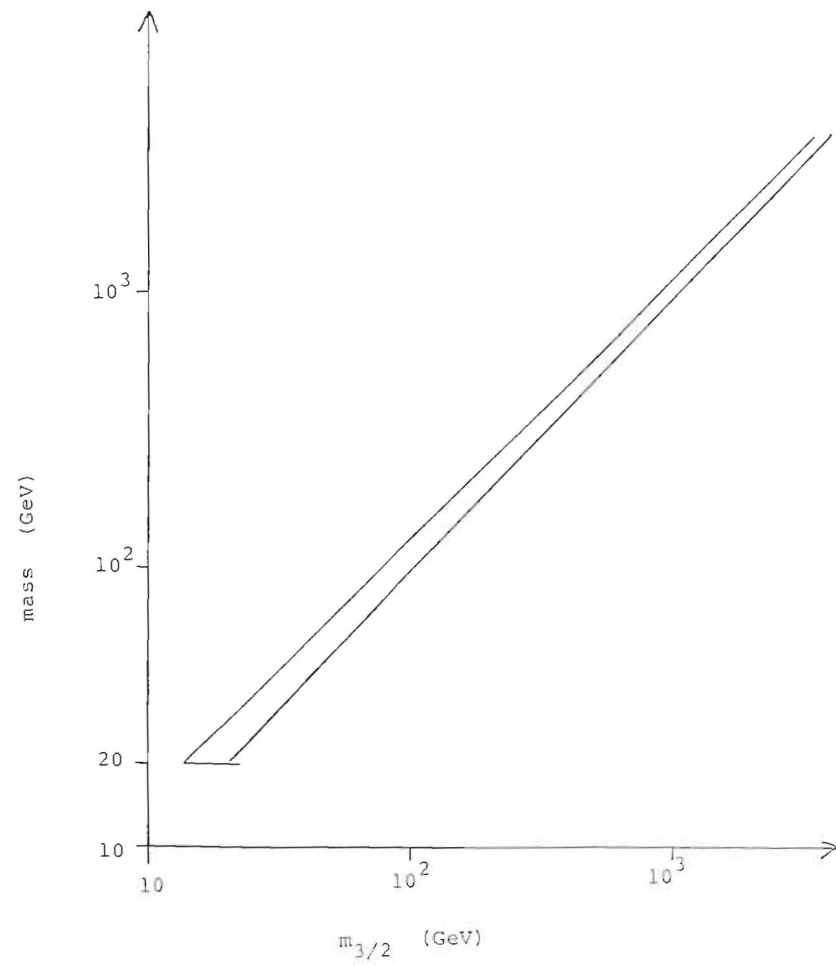


Fig. 2 Allowed region of scalar electron ( $e_R$ ) mass.



On Kaluza-Klein Theories

Yasunori Fujii

Institute of Physics  
University of Tokyo, Komaba  
Meguro-ku, Tokyo 153

Abstract

An introductory review is given on Kaluza-Klein theories with an emphasis on the notion of spontaneous compactification. Some future problems are also discussed.

§1. Introduction

The Kaluza-Klein theory [1], a theory of gravitation in higher dimensions part of which are compactified, is almost as old as its mother theory, Einstein's general relativity theory. After more than half a century of hibernation, it is now being rediscovered. The Kaluza-Klein theory in 1920's was one of many ambitious attempts to unify gravitation and electromagnetism into a geometrical theory of spacetime.

In spite of its truly ingenious idea it, like the other unified field theories to which Einstein himself devoted vigorous efforts, has failed to receive full recognition as a successful physical theory. Most probably this was due to the fact that the electromagnetic field with which gravity was intended to unify seemed to be an unlikely representative of the real matter system which appeared much more complicated. It was only a decade or so ago that the gauge principle, enriched with the nonabelian extension finally surfaced as the underlying theory, elucidating the basic interactions probably over the whole range of physical phenomena. The fundamental physical laws now appear much simpler than they used to, suggesting that the time is ripe to try again a geometrical unification.

The move also picked up momentum because the advent of Grand Unified Theories extended the energy scale of physical interest to as high as  $10^{15}$ - $10^{17}$  GeV, which is close enough to the Planck mass of  $\sim 10^{19}$  GeV to make the contact with gravity something beyond a mere speculation.

The strongest motivation of higher-dimensional theories has come, however, from the remarkable work of Cremmer and Julia [2] who showed that the attractive theory of N=8 extended supergravity can be formulated by dimensionally reducing the theory of simple supergravity in 11 dimensions. In what follows I try to give an introductory review of the modern versions of the Kaluza-Klein theories.

The basic idea of Kaluza-Klein theories [3-6] is that our 4-dimensional spacetime is part of a larger spacetime in  $D=4+n$  dimensions; the  $n$ -dimensional part is a compact manifold which is too small to be observed directly but manifests itself as internal space on which internal symmetry transformations of elementary particles act. ( $n=1$  in Kaluza's original theory).

In what is now called a pure Kaluza-Klein theory one starts with the Einstein theory in  $D$ -dimensional spacetime and derives a unified theory of Einstein gravity and other gauge theories, both in 4 dimensions. The process

in which D-dimensional spacetime is split into 4-dimensional spacetime and n-dimensional internal space is called dimensional reduction. In the most naive approach one may impose this splitting by fiat. In a more elaborate and plausible approach, however, one expects that this splitting occurs as a physical process; remember that spacetime geometry is a dynamical entity in the general theory of relativity.

To be more specific, let us ask what the "ground state," or the vacuum of D-dimensional spacetime  $M_D$  is. One of the obvious candidates is  $M_D^b$ , D-dimensional flat Minkowski spacetime which is maximally symmetric. It might happen, however, that a direct product  $M_4 \times M_n$  is the true ground state where  $M_4$  is 4-dimensional spacetime, most hopefully flat  $M_4^b$ , while  $M_n$  is some compact space in n dimensions. This product space is less symmetric than  $M_D^b$ . As is well known in other areas of physics, a less symmetric state could be a ground state if some symmetries are broken spontaneously. For this reason, the occurrence of  $M_4 \times M_n$  is often called a spontaneous compactification [7-16].

Unfortunately this notion of spontaneous compactification has never been formulated at the same level of rigor as other examples of spontaneous symmetry breaking; questions like how Nambu-Goldstone bosons result in apparent breaking of symmetries or how Higgs bosons supply nonzero masses to some gauge fields, are yet to be answered. As of now the only necessary condition for spontaneous compactification to take place is that  $M_4 \times M_n$  is a classical solution of the Einstein equation in D dimensions. (Of course,  $M_D^b$  is a trivial classical solution if the D-dimensional cosmological constant is absent.)

In §2 we explain how the Einstein theory in D dimensions is dimensionally reduced to the Einstein-Yang-Mills theory plus some scalar fields in 4 dimensions. We choose a specific example in which  $M_n$  is an n-dimensional sphere  $S^n$ . Important relations are derived among the Planck mass, the gauge coupling constant and the radius of  $S^n$ . We also show that the classical solution of the scalar field which represents the radius of  $S^n$  is unstable in this pure Kaluza-Klein theory.

The following §3 discusses the above mentioned stability in the theories which include an antisymmetric tensor field. We thereby deviate from the spirit of the original Kaluza theory. As was shown by Cremmer and Julia [2], however, the rank-3 antisymmetric tensor field is an essential ingredient in the supergravity theory in 11 dimensions, which gives N=8 supergravity in 4 dimensions.

The final Section 4 contains discussions on future problems. Three appendices will serve as an easy access to some of the technicalities. In Appendix A we give part of the explicit calculations illustrating how dimensional reduction is formulated. Appendix B shows how the number of physical degrees of freedom is counted for a massless antisymmetric tensor field. Appendix C contains a brief account of the relevance of the Kaluza-Klein approach to the extended supergravity theory of N=8.

## §2. Pure Kaluza-Klein Theory in D=4+n dimensions

Let us consider the Einstein theory in D=4+n dimensions. The coordinates are denoted by  $x^{\hat{\mu}}$ , with  $\hat{\mu} = 0, 1, \dots, D-1$ . The tangent space metric is  $\text{diag}(-, +, \dots, +)$ ; we assume that there is only one time coordinate [f1]. The Einstein-Hilbert action is given by

$$I = \int d^D x \mathcal{L}_D = \int d^D x \sqrt{-\hat{g}} \frac{1}{2\hat{\kappa}^2} \hat{R}_D, \quad (2.1)$$

where  $\hat{R}_D$  is the D-dimensional curvature scalar constructed from the metric  $\hat{g}_{\hat{\mu}\hat{\nu}}$  with its determinant  $\hat{g}$ , and  $\hat{\kappa}^2$  the gravitational constant in D dimensions.

We now assume, without asking why for the moment, that the coordinates  $x^{\hat{\mu}}$  with  $\hat{\mu} = 4, 5, \dots, D-1$  are compactified. It is convenient to introduce the notation,

$$x^{\hat{\mu}} = \begin{cases} x^\mu, & (\mu = 0, 1, 2, 3) \\ y^\alpha, & (\alpha = 1, 2, \dots, n) \end{cases}$$

for dimensional reduction  $M_D \rightarrow M_4 \times M_n$ . Examples of compact manifolds include  $T^n$  (n-torus),  $S^n$  (n-sphere),  $CP_{n/2}$  (complex projective space), or their products [f2]. A coordinate  $y^\alpha$  may be chosen to be a dimensionless angular variable. Correspondingly, the components  $\hat{g}_{\mu\alpha}$  and  $\hat{g}_{\alpha\beta}$  are given dimension of length and length squared, respectively.

For the sake of illustration we choose the example of  $S^n$ , an n-dimensional sphere for  $M_n$ . We are going to show that the following metric (and the inverse metric) gives a classical solution of the D-dimensional Einstein equation: (zero-mass ansatz)

$$\hat{g}_{\rho\sigma} = \begin{pmatrix} \tilde{\phi}^n g_{\mu\nu} + e^2 \phi^{-2} \Gamma_{AB} \Lambda_\mu^A \Lambda_\nu^B & e \phi^{-2} \xi_{AB} \Lambda_\mu^A \\ e \phi^{-2} \xi_{A\alpha} \Lambda_\nu^A & \phi^{-2} \gamma_{\alpha\beta} \end{pmatrix},$$

$$\hat{g}^{\rho\sigma} = \begin{pmatrix} \tilde{\phi}^{-n} g^{\mu\nu} & -e \tilde{\phi}^{-n} \xi_A^\beta \Lambda^{\mu A} \\ -e \tilde{\phi}^{-n} \xi_A^\alpha \Lambda^{\nu A} & \phi^2 \gamma^{\alpha\beta} + e^2 \tilde{\phi}^{-n} \xi_A^\alpha \xi_B^\beta \Lambda^{\lambda A} \Lambda^\lambda{}_\beta \end{pmatrix}, \quad (2.2)$$

where

- $g_{\mu\nu}(x)$ : 4-dimensional metric which describes Einstein gravity.  
 $\gamma_{\alpha\beta}(y)$ : Standard metric on  $S^n$  of unit radius.  
 $\phi(x)$ : Scalar field in 4 dimensions giving the inverse radius of  $S^n$ . Also  $\tilde{\phi}(x) = \ell \phi(x)$ , where  $\ell$  is given by the stationary value  $\phi_0$  of  $\phi(x)$ ;  $\ell = \phi_0^{-1}$ , if it exists.  
 $\xi_A^\alpha(y)$ : Killing vector.  $A = 1, \dots, K$ , with  $K$  the number of parameters of a group  $G$  which acts on  $M_n$ . In the case of  $S^n$ , we have  $S^n = SO(n+1)/SO(n)$ , so that  $K = n(n+1)/2$ . Also  $\Gamma_{AB}^\gamma(y) = \gamma_{\alpha\beta} \xi_A^\alpha \xi_B^\beta$ .  
 $\Lambda_\mu^A(x)$ : Massless vector field which turns out to be a Yang-Mills field in 4 dimensions. The index  $\mu$  is lowered or raised in terms of  $g_{\mu\nu}$  or its inverse  $g^{\mu\nu}$ .  
 $e$ : Gauge coupling constant.

Notice that  $\hat{g}_{\mu\nu} = \tilde{\phi}^n g_{\mu\nu} \neq g_{\mu\nu}$ ; the reason we have applied a Weyl rescaling will become clear shortly. The Killing vectors satisfy the conditions

$$\nabla_\alpha \xi_{A\beta} + \nabla_\beta \xi_{A\alpha} = 0, \quad (2.3a)$$

$$\xi_A^\beta (\partial_\beta \xi_B^\alpha) - \xi_B^\beta (\partial_\beta \xi_A^\alpha) = \xi_C^\alpha f^C{}_{AB}, \quad (2.3b)$$

where  $f^C{}_{AB}$  is the structure constant of  $G$ , and

$$\xi_{A\alpha} = \gamma_{\alpha\beta} \xi_A^\beta. \quad (2.3c)$$

Instead of showing directly that the above metric (2.2) solves the Einstein equation, we substitute (2.2) into the action (2.1) obtaining the 4-dimensional action which describes the system of Einstein gravity and the Yang-Mills field together with the scalar field. After a series of lengthy and complicated calculations (see Appendix A) we finally obtain

$$\mathcal{L}_D = \sqrt{\gamma} \sqrt{-g} \frac{\ell^n}{2\kappa^2} \left[ R_4 + \tilde{\phi}^{n+2} \ell^{-2} R_n - \frac{1}{2} n(n+2) g^{\mu\nu} \Phi_{,\mu} \Phi_{,\nu} - \frac{1}{4} e^2 \ell^2 \tilde{\phi}^{-n-2} \Gamma_{AB} F_{\mu\nu}^A F^{B\mu\nu} \right], \quad (2.4a)$$

where  $R_4$  and  $R_n$  are the curvature scalars in 4 and  $n$  dimensions, respectively. For  $S^n$  of unit radius we find

$$R_n = n(n-1). \quad (2.4b)$$

Also

$$F_{\mu\nu}^A = \partial_\rho A_\nu^A - \partial_\nu A_\rho^A - e f^A{}_{BC} A_\rho^B A_\nu^C, \quad (2.5a)$$

$$\Phi_{,\mu} = \partial_\mu \ln \tilde{\phi}, \quad (2.5b)$$

$$g = \det(g_{\mu\nu}), \quad \gamma = \det(\gamma_{\alpha\beta}). \quad (2.5c)$$

The 4-dimensional Lagrangian may be obtained by integrating (2.4a) over  $M_n = S^n$ . We use

$$\int d^n y \sqrt{\gamma} = \Omega_n = \ell^{-n} \mathcal{V}_n, \quad (2.6a)$$

where  $\mathcal{V}_n$  is the volume of  $M_n$ . For  $S^n$  we have

$$\Omega_n = 2\pi^{n/2+1/2} / \Gamma(\frac{n}{2} + \frac{1}{2}). \quad (2.6b)$$

We can also verify

$$\int d^n y \sqrt{\gamma} \Gamma_{AB} = \frac{n}{K} \delta_{AB} \Omega_n. \quad (2.6c)$$

By using these results we obtain

$$\begin{aligned} \mathcal{L}_4 &= \int d^4 x \mathcal{L}_D \\ &= \sqrt{-g} \left[ \frac{1}{2\kappa^2} R_4 - \frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - \frac{1}{4} \tilde{\phi}^{-n-2} F_{\mu\nu}^A F^{A\mu\nu} - V(\phi) \right], \end{aligned} \quad (2.7a)$$

where

$$V(\phi) = \frac{1}{\kappa} \sqrt{\frac{n(n+2)}{2}} \ln \tilde{\phi}, \quad (2.7b)$$

$$V(\phi) = -\frac{\ell^n}{2\kappa^2} n(n-1) \phi^{n+2}. \quad (2.7c)$$



In deriving (2.7) we have chosen

$$\hat{\mu}^2 = \kappa^2 \mathcal{U}_m, \quad (2.8)$$

in order to arrive at the correct Einstein-Hilbert action  $(1/2\kappa^2)R_4$  in 4 dimensions.

The advantage of applying Weyl rescaling in (2.2) is now obvious. Had we not applied this conformal transformation, the  $R_4$  term would have been multiplied by  $\phi^n$  coming from  $\sqrt{-g}$ . This implies the presence of mixing between  $\phi$  and the scalar component of the gravitational field. Only after removing this mixing by means of Weyl rescaling, can one discuss stability of the  $\phi$  field based solely on  $V(\phi)$ , disregarding the  $R_4$  term.

Some important remarks are in order. The coefficient  $e$  of  $\hat{g}_{\mu\alpha}$  and  $\hat{g}^{\mu\alpha}$  in (2.2) reappears in the second term of the Yang-Mills field strength (2.5a), thus allowing the interpretation of  $e$  as the gauge coupling constant. In other words the coefficient in (2.2) is uniquely determined in this way. Due to this determination, the Yang-Mills term (the last term) of (2.4a) carries the coefficient  $-e^2 \mathcal{L}^2/4$ . Combining (2.6) and (2.8), we obtain the coefficient  $-(n/K)(1/2\kappa^2)(e^2 \mathcal{L}^2/4)$  in front of the third (FF) term in (2.7a). We should adjust this coefficient to be equal to  $-1/4$  so that the field  $A_\mu^A$  has the correct normalization for  $\phi(x) \rightarrow \ell^{-1}$ , i.e.  $\tilde{\phi}(x) \rightarrow 1$ . It thus follows that  $e$  must obey the condition

$$e = \frac{\kappa}{\ell} \sqrt{\frac{2K}{n}}, \quad (2.9a)$$

which can also be put into the form

$$\ell = 2\ell_* \alpha^{-1/2} \sqrt{K/n}, \quad (2.9b)$$

where  $\alpha = e^2/4\pi$  and  $\ell_* = \kappa/\sqrt{8\pi} = 1.61 \times 10^{-33}$  cm = Planck length. Since  $K/n \gtrsim 1$  and  $\alpha \lesssim 10^{-4}$ , we find

$$\ell \gtrsim \ell_*.$$

We thus see that the "size" of compactified space should be of the order of, or slightly larger than the Planck length, which is extremely small compared with any other physical length scale.

The field  $\phi(x)$  defined by (2.7b) is a canonical scalar field;  $V(\phi)$  coming from the curvature scalar of  $S^n$  is its "potential." As one finds from (2.7c), however, the potential  $V(\phi)$  is negative without a minimum, allowing no stationary classical solution for  $\phi(x)$ , except for the meaningless value

$\phi = 0$ , i.e. the infinitely large radius. For this reason, the pure Kaluza-Klein theory fails to admit a solution of spontaneous compactification, except for  $D = 5$ ,  $n = 1$  for which  $V(\phi)$  vanishes. We come back later to this question of stability of the scalar field solution.

As another unique feature of the Kaluza-Klein theories, we show that the Yang-Mills gauge transformation is part of general coordinate transformations in  $D$  dimensions.

Consider a coordinate transformation

$$x^{\hat{\mu}} \rightarrow x'^{\hat{\mu}} = x^{\hat{\mu}} + \zeta^{\hat{\mu}},$$

with

$$\begin{cases} \zeta^{\hat{\mu}} = 0, \\ \zeta^\alpha = -\epsilon^A(x) \xi_A^\alpha(v), \end{cases} \quad (2.10)$$

This is an  $x$ -dependent isometry transformation in internal space. The Lie derivative of the metric is given generally by

$$\delta_* \hat{g}_{\rho\sigma} = -(\partial_\rho \zeta^{\hat{\lambda}}) \hat{g}_{\hat{\lambda}\sigma} - (\partial_\sigma \zeta^{\hat{\lambda}}) \hat{g}_{\rho\hat{\lambda}} - \zeta^{\hat{\lambda}} (\partial_{\hat{\lambda}} \hat{g}_{\rho\sigma}). \quad (2.11)$$

We apply this to

$$\hat{g}_{\mu\beta} = e \phi^{-2} \xi_{AB} A_\mu^A,$$

with  $\zeta^{\hat{\lambda}}$  given above. On the left-hand side we have

$$\delta_* \hat{g}_{\mu\beta} = e \phi^{-2} \xi_{AB} \delta_* A_\mu^A.$$

(Notice that the Killing vector itself remains unchanged under isometry transformations.) On the right-hand side of (2.11) we use (2.3) obtaining finally

$$\delta_* A_\mu^A = e^{-1} \partial_\mu \epsilon^A + f^A_{BC} \epsilon^B A_\mu^C. \quad (2.12)$$

This is precisely the Yang-Mills local gauge transformation which is now endowed with a geometrical meaning in higher-dimensional spacetime.

### §3. Classical stability of the size of internal space

As we saw in the preceding section, the potential of the scalar field

( $\phi$  or  $\sigma$ ) expressing the radius of  $S^n$  has no minimum. This represents a failure of the idea of spontaneous compactification, which requires the existence of a classical solution giving a finite size to internal space. In the exceptional case of 5 dimensions or D dimensions with  $T^n$  as internal space, the classical potential is zero everywhere. Then possible quantum effects might be important. The one-loop calculations indicate that internal space would collapse to the size of the Planck length [18][f3].

In more general cases, however, something must be done perhaps at the classical level. It has been proposed to introduce additional gauge fields in D dimensions to start with [7,9]. This is certainly a deviation from the spirit of Kaluza whose main purpose was to derive the 4-dimensional gauge field (the Maxwell field) from the 5-dimensional Einstein theory with nothing else added. To some extent, however, it might be justifiable to sacrifice simplicity in order to give a high priority to the realization of spontaneous compactification.

One could also argue that confining oneself to the radius of internal space is far from being sufficient: the configuration of internal space must be stable against every possible fluctuation. But we simply emphasize that stability of the radius mode is a necessary condition.

It is also true that we should be prepared for the possibility that the true ground state is provided by a configuration other than a sphere. But perhaps the most troublesome is the fact that no one knows exactly what one should show to establish spontaneous compactification. With all these reservations in mind, we pursue in the following the question of classical stability of the radius mode of the sphere  $S^n$  assumed for internal space [14,16]. The essential ingredients of the foregoing proposals of additional gauge fields are shown in a most typical way in the models having a totally antisymmetric tensor field with a vacuum expectation value which carries some topological charge.

Consider a totally antisymmetric tensor field  $A_{\hat{\mu}_1 \dots \hat{\mu}_r}$  of rank r in D dimensions. The field strength is a generalized curl, also totally antisymmetric and of rank r+1: [f4]

$$\begin{aligned} F_{\hat{\mu}_1 \dots \hat{\mu}_{r+1}} &= (r+1) \partial_{[\hat{\mu}_1} A_{\hat{\mu}_2 \dots \hat{\mu}_{r+1}]} \\ &= \partial_{\hat{\mu}_1} A_{\hat{\mu}_2 \dots \hat{\mu}_{r+1}} - \partial_{\hat{\mu}_2} A_{\hat{\mu}_1 \hat{\mu}_3 \dots \hat{\mu}_{r+1}} + \dots \end{aligned} \quad (3.1)$$

Examples:

$$F_{\hat{\mu}\hat{\nu}} = \partial_{\hat{\mu}} A_{\hat{\nu}} - \partial_{\hat{\nu}} A_{\hat{\mu}},$$

$$F_{\hat{\mu}\hat{\nu}\hat{\lambda}} = \partial_{\hat{\mu}} A_{\hat{\nu}\hat{\lambda}} - \partial_{\hat{\nu}} A_{\hat{\mu}\hat{\lambda}} + \partial_{\hat{\lambda}} A_{\hat{\mu}\hat{\nu}}. \quad (3.2)$$

The field strength is invariant under an abelian gauge transformation,

$$A_{\hat{\mu}_1 \dots \hat{\mu}_r} \rightarrow A_{\hat{\mu}_1 \dots \hat{\mu}_r} + \partial_{[\hat{\mu}_1} \Lambda_{\hat{\mu}_2 \dots \hat{\mu}_r]}, \quad (3.3)$$

where the gauge function  $\Lambda_{\hat{\mu}_1 \dots \hat{\mu}_r}$  is again totally antisymmetric. Obviously this gauge field is a generalization of the Maxwell field.

We then assume that the fundamental action in D dimensions is, apart from the possible matter term, a generalized Einstein-Maxwell action:

$$I = \int d^D x \sqrt{-g} \left[ \frac{1}{2\kappa^2} \hat{R}_D - \frac{1}{2} \frac{1}{(r+1)!} F_{\hat{\mu}_1 \dots \hat{\mu}_{r+1}} F^{\hat{\mu}_1 \dots \hat{\mu}_{r+1}} \right], \quad (3.4)$$

where the tensor indices of the gauge field are raised or lowered by  $\hat{g}_{\hat{\mu}\hat{\nu}}$  or  $\hat{g}^{\hat{\mu}\hat{\nu}}$ . From this action one derives the Einstein equation and the generalized Maxwell equation:

$$\begin{aligned} \kappa^2 \hat{G}_{\hat{\mu}\hat{\nu}} &= \frac{1}{r!} F_{\hat{\mu}\hat{\mu}_1 \dots \hat{\mu}_r} F_{\hat{\nu}}^{\hat{\mu}_1 \dots \hat{\mu}_r} \\ &\quad - \frac{1}{2} \frac{1}{(r+1)!} \hat{g}_{\hat{\mu}\hat{\nu}} F_{\hat{\lambda}_1 \dots \hat{\lambda}_{r+1}} F^{\hat{\lambda}_1 \dots \hat{\lambda}_{r+1}}, \end{aligned} \quad (3.5a)$$

$$\frac{1}{\sqrt{-g}} \partial_{\hat{\mu}} (\sqrt{-g} F^{\hat{\mu}\hat{\mu}_1 \dots \hat{\mu}_r}) = 0. \quad (3.5b)$$

Since we are here mainly interested in the vacuum, we drop  $\Lambda_{\hat{\mu}}^{\hat{\alpha}}(x)$ , which has no vacuum expectation value in 4 dimensions, in (2.2), hence choosing the ansatz

$$\hat{g}_{\hat{\mu}\hat{\nu}} = \begin{pmatrix} \tilde{\Phi}^m g_{\mu\nu} & 0 \\ 0 & \Phi^{-2} \gamma_{\alpha\beta} \end{pmatrix}. \quad (3.6)$$

A special situation occurs if  $r = n-1$  [9,14]. Then the field strength  $F_{\hat{\mu}_1 \dots \hat{\mu}_n}$  can have the following nonzero vacuum expectation value:

$$\begin{cases} F_{\alpha_1 \dots \alpha_n} = f \sqrt{\gamma} \epsilon_{\alpha_1 \dots \alpha_n}, \\ \text{other components} = 0, \end{cases} \quad (3.7)$$

where f is a constant while  $\epsilon_{\alpha_1 \dots \alpha_n} = 0, \pm 1$  is the Levi-Civita symbol in n dimensions, so that  $F_{\alpha_1 \dots \alpha_n}$  is a tensor on  $S^n$ . It is crucial to observe that

the ansatz (3.7) solves the generalized Maxwell equation (3.5b). (Notice  $F^{\alpha_1 \dots \alpha_n} = \phi^{2n} (f/\sqrt{g}) \epsilon^{\alpha_1 \dots \alpha_n}$ ,  $\sqrt{-g} = l^{2n} \phi^n \sqrt{g}$  so that  $\sqrt{-g} F^{\alpha_1 \dots \alpha_n} = f l^{2n} \phi^{3n} \times \sqrt{g}$  which is independent of  $y^\alpha$ .)

The vacuum expectation value (3.7) contributes

$$-\frac{1}{2} f^2 \phi^{2n},$$

to the second term in the square brackets in (3.4). Combining this with  $\sqrt{-g}$  given above, and rearranging the result in the form of (2.7a), we obtain the potential

$$V(\phi) = -\frac{l^n}{2\kappa^2} n(n-1) \phi^{n+2} + \frac{1}{2} l^{2n} \Omega_n f^2 \phi^{3n}. \quad (3.8)$$

The additional second term is positive and has a higher power  $3n$  than the first term arising from  $S^n$ . Consequently  $V(\phi)$  has a (global) minimum at  $\phi = \phi_0$  given by

$$\phi_0^{2n-2} = \frac{1}{3} f^{-2} l^{-2} \frac{1}{\kappa^2 \Omega_n} (n-1)(n+2).$$

By identifying  $\phi_0$  with  $l^{-1}$  introduced in (2.2), we obtain a relation expressing  $l$  in terms of  $f$ :

$$l = \left[ \frac{(n-1)(n+2)}{3f^2 \kappa^2 \Omega_n} \right]^{\frac{1}{2-n}}. \quad (3.9)$$

In this way we arrive at a necessary condition for spontaneous compactification to be realized.

Most crucial in the above model is the nonzero value of the constant  $f$  in (3.7). One may naturally ask the question what the origin of this constant is. It may come from the matter interaction which we have neglected. Or it may have a cosmological origin, as is suggested from the observation that the minimum of the potential which depends on  $f^2$  gives an effective cosmological constant in 4 dimensions. For the moment, we simply point out that the nonzero  $f^2$  allows an interpretation in terms of a "monopole charge." [9,14]

One imagines  $(n+1)$ -dimensional Euclidean space in which  $S^n$  is embedded. The radial component of the dual of  $F_{\alpha_1 \dots \alpha_n}$  may be defined by

$$\tilde{F}^r = \frac{1}{n!} r^{-n} \frac{1}{\sqrt{g}} \epsilon^{\alpha_1 \dots \alpha_n} F_{\alpha_1 \dots \alpha_n},$$

which is  $fr^{-n}$  according to (3.7). Hence its surface integral gives

$$\int d^n y \sqrt{g} r^n \tilde{F}^r = f,$$

which implies that the monopole charge  $f$  is inside the sphere. It is

interesting to point out that all the previously proposed models [7,9] resulting in stability of the size of internal space share the common property of having charges of topological nature.

Another interesting possibility was discovered by Freund and Rubin [8] in connection with supergravity in 11 dimensions. As will be explained briefly in Appendix C, the supergravity theory in 11 dimensions contains the graviton, the gravitino and an antisymmetric tensor field of rank 3. The corresponding field strength is then of rank 4. Then, in contrast to (3.7), the field strength may have a vacuum expectation value in 4-dimensional spacetime rather than in internal space:

$$\begin{cases} F_{\mu\nu\lambda\sigma} = f \sqrt{-g} \epsilon_{\mu\nu\lambda\sigma}, \\ \text{other components} = 0, \end{cases} \quad (3.10)$$

where  $\epsilon_{\mu\nu\lambda\sigma} = 0, \pm 1$  is a tensor density,  $f$  a constant. This ansatz is possible precisely because the theory contains a rank-3 tensor field and hence a field strength of rank 4; the number 4 has no immediate connection with dimensionality of spacetime at the beginning. In this sense it can be said that the theory of supergravity in 11 dimensions explains why our spacetime is 4-dimensional.

Establishing classical stability of the solution based on the ansatz (3.10) requires some care, as will be discussed below [14,16]. If one allows the  $x$ -dependence in  $\hat{g}_{\alpha\beta}$  as in (3.6), the first expression of (3.10) must be modified to give

$$F_{\mu\nu\lambda\sigma} = Q \tilde{\phi}^{3n} \sqrt{-g} \epsilon_{\mu\nu\lambda\sigma}, \quad (3.11a)$$

where  $Q$  is a true constant. In fact one derives

$$\begin{aligned} F^{\mu\nu\lambda\sigma} &= \hat{g}^{\mu\mu'} \dots \hat{g}^{\sigma\sigma'} F_{\mu'\nu'\lambda'\sigma'} = \tilde{\phi}^{-4n} g^{\mu\mu'} \dots g^{\sigma\sigma'} F_{\mu'\nu'\lambda'\sigma'} \\ &= Q \tilde{\phi}^{-n} \sqrt{-g} g^{\mu\mu'} \dots g^{\sigma\sigma'} \epsilon_{\mu'\nu'\lambda'\sigma'} = Q \tilde{\phi}^{-n} \frac{1}{\sqrt{-g}} \epsilon^{\mu\nu\lambda\sigma}, \end{aligned} \quad (3.11b)$$

and then

$$\begin{aligned} \sqrt{-g} F^{\mu\nu\lambda\sigma} &= l^n \tilde{\phi}^n \sqrt{g} \sqrt{-g} F^{\mu\nu\lambda\sigma} \\ &= l^n Q \sqrt{g} \epsilon^{\mu\nu\lambda\sigma}, \end{aligned} \quad (3.11c)$$

which is  $x$ -independent. One sees why  $\tilde{\phi}^{3n}$  is necessary in (3.11a) in order to

solve (3.5b). One then computes the contribution from the second term of (3.4)(r=3):

$$\begin{aligned} -\sqrt{-g} \frac{1}{2} \frac{1}{4!} F_{\mu\nu\lambda\sigma} F^{\mu\nu\lambda\sigma} &= -\frac{1}{48} l^n \tilde{\phi}^{3n} \sqrt{g} \sqrt{-g} Q^2 \epsilon_{\mu\nu\lambda\sigma} \epsilon^{\mu\nu\lambda\sigma} \\ &= \frac{1}{2} l^n \sqrt{g} \sqrt{-g} Q^2 \tilde{\phi}^{3n}. \end{aligned} \quad (3.12a)$$

The overall plus sign arises because  $\epsilon_{\mu\nu\lambda\sigma} \epsilon^{\mu\nu\lambda\sigma} = -4! < 0$  due to the presence of one time component. If one naively interprets (3.12a) as giving a potential

$$-\Omega_n^{-1} \sqrt{g} \sqrt{-g} V',$$

then one would obtain

$$V'(\phi) = -\frac{1}{2} l^n \Omega_n Q^2 \tilde{\phi}^{3n}. \quad (3.12b)$$

In contrast to the second term of (3.8), this potential is negative. The total potential, the sum of the first term of (3.8) and (3.12b), is hence negative, having no minimum. A more careful analysis shows, however, that the sign in (3.12b) should be reversed in the correct treatment.

One in fact finds that the Einstein equation in D dimensions contains the  $\phi$  field equation which with (3.11) inserted results in what one would expect from the potential of the opposite sign of (3.12b). This complication arises obviously through the process of eliminating the auxiliary field  $A_{\mu\nu\lambda}$ , which has no physical degree of freedom in 4 dimensions (see (B.15)).

In general situations in which one has an auxiliary field, one can proceed via two routes. In one of the routes one first derives the field equations by varying the Lagrangian considering all the fields as independent. Then the equation for the would-be auxiliary field can be used to express this field in terms of other independent fields. This result is then substituted into the equations of the independent fields. There is no ambiguity in these processes. In the alternative route, the relation which expresses the auxiliary field in terms of the independent fields is substituted back to the Lagrangian from which the field equations finally follow. There are many known examples in which the above two routes lead to the same end result for the field equations of the independent fields. In our present calculation, however, we encounter an exception. A detailed investigation has been given in ref. 16. See also ref. 15 for a somewhat different approach. Here we simply illustrate how subtle the calculations

are.

Suppose we vary (3.12a) with respect to  $A_{\mu\nu\lambda}$ . We obtain

$$\delta \mathcal{L}_{FF} = -\frac{1}{24} \sqrt{-g} F^{\mu\nu\lambda\sigma} \delta F_{\mu\nu\lambda\sigma}, \quad (3.13a)$$

which, on using (3.11c), is put into

$$\delta \mathcal{L}_{FF} = -\frac{1}{24} l^n Q^2 \sqrt{g} \epsilon^{\mu\nu\lambda\sigma} \delta F_{\mu\nu\lambda\sigma}. \quad (3.13b)$$

This is a 4-divergence if one regards it as a variation with respect to  $A_{\mu\nu\lambda}$ :  $\delta F_{\mu\nu\lambda\sigma} = \partial_\mu \delta A_{\nu\lambda\sigma} - \dots$ . If one further uses (3.11a) for  $\delta F_{\mu\nu\lambda\sigma}$ , however, one obtains

$$\delta \mathcal{L}_{FF} = l^n Q^2 \sqrt{g} \delta(\sqrt{-g} \tilde{\phi}^{3n}), \quad (3.13c)$$

which is no longer a 4-divergence, hence giving a nonzero value when it is integrated over spacetime.

This observation shows that the variation  $\delta A_{\mu\nu\lambda}$  should not vanish on the spacetime boundary. This is indeed true because the solution (3.11a) gives the derivative of  $A_{\mu\nu\lambda}$  in terms of  $\tilde{\phi}$ . The field  $A_{\mu\nu\lambda}$  itself is expressed as an integral of  $\tilde{\phi}$ . The same applies also to their variations. It thus follows that  $\delta A_{\mu\nu\lambda}$  fails to vanish on some part of the boundary, at  $t = +\infty$ , say, even if it is chosen to vanish at  $t = -\infty$ , with  $\delta \tilde{\phi}$  vanishing at  $t = \pm\infty$ .

In order to be able to carry out the usual method of variation, one must add an extra term whose variation on the boundary cancels the contribution from (3.13b). As one finds, the required term is given by

$$\mathcal{I}_\Delta = \frac{1}{6} \partial_\mu (\sqrt{-g} F^{\mu\nu\lambda\sigma} A_{\nu\lambda\sigma}). \quad (3.14)$$

This is a 4-divergence so that no contribution occurs if it is varied with respect to  $A_{\mu\nu\lambda}$  regarded as independent. If one performs the differentiation  $\partial_\mu$  and uses (3.5b), one obtains

$$\mathcal{I}_\Delta = \frac{1}{24} \sqrt{-g} F_{\mu\nu\lambda\sigma} F^{\mu\nu\lambda\sigma}, \quad (3.15a)$$

which is further put into

$$\mathcal{I}_\Delta = -l^n Q^2 \sqrt{g} \sqrt{-g} \tilde{\phi}^{3n}, \quad (3.15b)$$

on substituting from (3.11)(cf.(3.12a)). As one sees this would give a potential which is -2 times (3.12b). By adding this contribution to (3.12b) one finally obtains the correct potential

$$V'' = \frac{1}{2} l^n Q^2 \Omega_m \tilde{\phi}^{2n}. \quad (3.16)$$

The total potential is now given by

$$V(\phi) = -\frac{l^n}{2\kappa^2} m(m-1)\phi^{m+2} + \frac{1}{2} l^{4n} \Omega_m Q^2 \phi^{2n}, \quad (3.17a)$$

which has an absolute minimum at  $\phi_0 = l^{-1}$ , where

$$l = \left[ \frac{(n-1)(n+2)}{3Q^2 \kappa^2 \Omega_m} \right]^{\frac{1}{2+n}}. \quad (3.17b)$$

It is interesting to notice that the additional term  $\mathcal{L}_\Delta$  in the form of

$$\mathcal{L}_\Delta = \frac{1}{24} l^n Q \sqrt{\gamma} \epsilon^{\mu\nu\lambda\sigma} F_{\mu\nu\lambda\sigma}, \quad (3.18)$$

as obtained by using (3.11c) in (3.15a), is such that the ansatz (3.11b) can be obtained as a field equation [20,21]:

$$\begin{aligned} 0 &= \frac{\delta}{\delta F_{\mu\nu\lambda\sigma}} \left[ \text{second term of (3.4) (with } r=3) + (3.18) \right] \\ &= -\frac{1}{24} \left( \sqrt{-\hat{g}} F^{\mu\nu\lambda\sigma} - l^n Q \sqrt{\gamma} \epsilon^{\mu\nu\lambda\sigma} \right). \end{aligned}$$

The value of the potential (3.17a) at the minimum is obviously negative:

$$V(l^{-1}) = -\frac{(n-1)^2}{\kappa^2 l^2},$$

with  $l$  given by (3.17b). This implies that, if there is no intrinsic cosmological term in D dimensions as in 11-dimensional supergravity, then 4-dimensional spacetime is anti-de Sitter spacetime with a negative effective cosmological constant. Furthermore, the value of this cosmological constant is huge --- of the order of  $l_*^{-4}$ . This is about 120 orders too large compared with the observed upper limit in the present universe. There might be an intrinsic D-dimensional cosmological constant [f5] or a contribution from quantum effects [f6]. It is still true that one needs an extremely fine tuning of the order of 120 to have a realistic theory. This is perhaps the most formidable difficulty of this type of unified theory [f7]. We come back again to this problem in the next section.

We have examined two types of solution for which the antisymmetric tensor field has a vacuum expectation value either on  $S^n$  or in 4-dimensional spacetime. Another solution is also known to exist for which the rank-3 antisymmetric tensor field has vacuum expectation values both in 4-dimensional spacetime and on  $S^7$  [12]. The vacuum expectation value on  $S^7$  is also related to parallelizable torsion, the existence of which is a unique feature of  $S^7$ .

The magnitude of this torsion field occurs as another scalar field in 4 dimensions.

To these modes, the radius of  $S^7$  and the magnitude of parallelizable torsion (as well as the magnitude of squashing of  $S^7$ ), one can apply basically the same technique in terms of the potential [14-16][f8]. One of the advantages of computing the potential is that one can discuss stability against any disturbance. Unfortunately, however, this method has been applied only to a limited class of modes. More general classes of modes can be analyzed only by linearizing equations and using a harmonic expansion on  $M_4$ , and hence restricting oneself to the discussion of small perturbations. For the details of this type of analysis on  $S^7$ , interested readers should consult refs. 11,13, in which mass spectra of  $0^\pm$ ,  $1^\pm$  fields are obtained. It was shown that no tachyonic modes appear in the background of anti-de Sitter spacetime.

At this point we add a remark on a massless field in anti-de Sitter spacetime. Anti-de Sitter (or de Sitter) spacetime is a constant-curvature spacetime with  $R = 4\Lambda$ , where the cosmological constant  $\Lambda$  is negative (or positive). It is known that these spacetimes can be transformed to flat Minkowski spacetime  $M_4^b$  by a series of conformal transformations [25,26]. On the other hand,  $M_4^b$  is the only spacetime which provides a sound basis for defining a massless field.

Combining these two observations, one concludes that the theory of a massless field in anti-de Sitter (or de Sitter) spacetime can be obtained by a conformally invariant theory which reduces to a massless theory in  $M_4^b$ . For a spinless field in 4 dimensions one thus arrives at

$$\mathcal{L} = \sqrt{-g} \left( -\frac{1}{12} \sigma^2 R - \frac{1}{2} g^{\mu\nu} \partial_\mu \sigma \partial_\nu \sigma \right), \quad (3.19)$$

which is invariant under conformal transformations

$$g_{\mu\nu} \rightarrow \Omega^{-2} g_{\mu\nu}, \quad \sigma \rightarrow \Omega \sigma,$$

with an arbitrary spacetime function  $\Omega(x)$ .

The Lagrangian (3.19) certainly describes a massless field for  $R = 0$ . For  $R = \text{const.} \neq 0$ , however, the first term of (3.19) acts as a mass term. For this reason, a massless field in anti-de Sitter (or de Sitter) spacetime is said to have an effective mass squared

$$m^2 = R/6 = (2/3)\Lambda.$$

This is already negative in anti-de Sitter spacetime. But the critical value below which the particle would behave as a tachyon is still smaller,  $(9/8)(R/6)$ , as was shown by Breitenlohner and Freedmann [27] who took a peculiar property of anti-de Sitter spacetime fully into account. It is remarkable that this critical value coincides with the lowest eigenvalue of the Casimir operator of the anti-de Sitter group  $SO(3,2)$  [25]. It is also noticed that some of the scalar modes on  $S^7$  hit precisely this critical value [13]. (Absence of tachyonic fields is shown in ref. 13 by verifying that no modes have mass squared below this critical value.)

The modes corresponding to the  $S^7$  radius and the amount of parallelizable torsion obtained above are massive of the order of the Planck mass. Such super-heavy modes are not very interesting at the moment unless they are tachyonic. In fact they are not. On  $S^7$  there are a number of massless scalar modes in the above sense [11,13]. It is yet to be seen if they can be good candidates for realistic Higgs scalars of the masses of the order of GeV, which could be viewed as massless compared with the Planck mass.

#### §4. Discussion

I have surveyed some of the key elements of the Kaluza-Klein theories, showing how promising they appear as attempts toward a geometrical unification of particle physics and gravitation. Although I have not covered all the related topics, I now turn to discussing problems we face.

First of all, we still do not know what the true ground state of D-dimensional spacetime is. We even do not know precisely how we can determine that. Putting aside this question of principle for the moment, we must also determine what internal space should be chosen to derive the internal symmetries which have been established for particles in the energy region below 100 GeV.

In this connection we reiterate that the basic energy scale of Kaluza-Klein theories is supplied by the Planck mass which is more than 17 orders higher than energy scales we are familiar with in ordinary particle physics. Particles of ordinary mass scales should be regarded as massless in the first approximation in this approach. Masses of order less than 100 GeV could be due to small violation of certain invariance principles of Kaluza-Klein theories.

The zero-mass ansatz (2.2) results in massless Yang-Mills fields for any types of internal space assumed. Mass spectra of scalar fields are, however, more complicated. Only for some specific examples of internal space, e.g. round  $S^7$ , massless scalar fields have been shown to emerge [11,13]. In a sense we are still far away from what Cho and Freund dreamed [4]: Kaluza-Klein theories would yield Nambu-Goldstone-Higgs bosons on the same footing as gauge fields.

The situation is even worse with fermions. Suppose one introduces massless fermions in D dimensions. A Dirac (or Majorana) field has  $2^{\lfloor D/2 \rfloor}$  components. This object then describes  $2^{\lfloor D/2 \rfloor - 2}$  of 4-component spinor fields in 4 dimensions. The Dirac operator  $\not{D}_D$  would split into  $\not{D}_4 + \not{D}_n$ ; the second part made of gamma matrices, derivatives and spin connections in n dimensions would give a mass term in 4 dimensions. It is known (the Lichnerowicz theorem [28]) that the mass squared arising in this way is proportional to the curvature scalar of internal space. It is then natural that the fermion mass would be of the order of the Planck mass unless a special zero mode exists. It was shown that massless fermions arise from 11-dimensional supergravity with the round  $S^7$  for internal space [10,11], due to the N=8 supersymmetry which remains in 4 dimensions. It is not clear, however, if these fermions stay massless after the supersymmetry is broken.

The question of the presence of fermion zero modes is also related to the question if chiral fermions emerge in 4 dimensions [29]. Chirality can be defined only in even dimensions because the equivalent of  $\gamma_5$  does not exist in odd dimensions. Basically for this reason one cannot expect chiral fermions in 4 dimensions starting from Kaluza-Klein theories of odd D, if the only interaction of the fermions is gravitational. In Cremmer and Julia's theory of 11-dimensional supergravity, chiral fermions nevertheless occur because the interaction which includes the terms of the antisymmetric tensor field singles out one of the chiralities. Unfortunately the gauge field of the "hidden symmetry"  $SU(8)$  is no longer an elementary field, and the complete dynamical implementation of the idea of composite gauge fields has never been worked out [19].

It was also shown [29] that no chiral fermions appear if  $D = 4m$  because fermions of both chiralities occur in internal space of  $4(m-1)$  dimensions. Only by starting from  $D = 4m+2$ , can one evade this type of "no-go theorem." It is still necessary to introduce an additional gauge field, as was demonstrated in the 10-dimensional theories including an abelian (vector)

gauge field [31][f10].

Another area in which the Kaluza-Klein theories may have interesting consequences is cosmology. Since the size of internal space is given by the spacetime dependent scalar field, one can expect that the size of internal space may have changed as the universe has evolved. More specifically, the radius  $b(t) = \phi^{-1}$  of  $S^n$  would behave as a function of the cosmic time  $t$ , just like the cosmological scale factor  $a(t)$  in the Robertson-Walker metric. One can even try a conjecture that our universe and internal space might have started (probably at the Planck time) with more or less the same size, but thereafter the former has expanded while the latter has contracted or stayed nearly the same, hence explaining why internal space is so small compared with the size of the universe at the present time.

Kasner-type solutions have been analyzed in this context [33,34]. It has turned out that, taking the effect of particle creation into account, the growth of the desired asymmetry with time is rather unlikely; its decay is more likely [34]. This tendency, however, may be counter-balanced by the "Casimir effect" due to quantum loop effects which can favor smaller internal space [18]. Other effects due to the temperature, for example, should also be studied [35].

The possibility of time variation of the gravitational constant has also been discussed [36]. This can be anticipated rather naturally because the ansatz (2.2) without Weyl rescaling easily leads to a scalar-tensor theory; the  $R_4$  term in (2.4a) would have been multiplied by  $\phi^n$ . The effective gravitational constant  $\tilde{\kappa}^2$  would be proportional to  $\phi^{-n} \sim b^m(t)$ .

One must exercise care, however, when one investigates whether the gravitational constant changes with time. As was emphasized by Dirac [37], possible time variability of the gravitational constant can be detected by measuring gravitational phenomena using atomic clocks. Notice that the masses of elementary particles, specifically of the electron, fix the fundamental time scale of atomic clocks. The effect could be due to possible different time dependences of the gravitational constant and the particle masses.

From this point of view, one must know how particle masses depend on time, or spacetime coordinates, before one can discuss the problem. In the conformal frame in which the gravitational constant is a true constant after Weyl rescaling as in (2.4a), time dependence of particle masses would imply a time variability of the gravitational constant [f11]. As was pointed out, we are still not in a position to calculate the masses of ordinary light

elementary particles. In this sense it seems premature to try to understand Dirac's Large Numbers hypothesis from the point of view of Kaluza-Klein theories.

Discussing possible time variation of the gauge coupling constant is also interesting. As was emphasized before, the gauge coupling constant  $e$  occurring in (2.5a) is a true constant; one cannot allow  $e$  to depend on spacetime coordinates through  $\phi$  without violating gauge invariance. In spite of this the observed gauge coupling constant may vary for a different reason.

In our analysis in §2, we normalized the gauge potential assuming that the  $\phi$  field takes its stationary value  $\ell^{-1}$ . For varying  $\phi$  before it reaches  $\ell^{-1}$  we cannot normalize  $A_\mu^A$  in a strict sense. We reasonably assume, however, that  $\phi$  changes so slowly that it can be considered as a constant for any microscopic processes. In this approximate sense we rescale  $A_\mu^A$  by multiplying it by  $\tilde{\phi}^{n/2+t}$  taken to be a constant:  $A_\mu^A \rightarrow \tilde{\phi}^{n/2+t} A_\mu^A$ , so that the FF term of (2.7a) recovers the correct normalization factor  $-1/4$  for  $\tilde{\phi} \neq 1$ . Applying this rescaling to (2.5a) implies a rescaling  $e \rightarrow \tilde{\phi}^{-n/2-t} e$ . Consequently, the fine structure constant  $\alpha$  changes in effect as the radius raised to the power  $-(n+2)$ . The reported high accuracy of the constancy of [39] hence tells us that the size of internal space has stayed unchanged to the same accuracy for the considerable part of the cosmological past.

The final question is on the size of the cosmological constant. Many of the classical solutions discovered so far result in 4-dimensional spacetime having an extremely large cosmological constant. Setting aside a more serious question, if any of these solutions is a true ground state, it is quite clear that we are desparately far from reality. One can conceive at least two ways out: either re-adjusting the cosmological constant or interpreting the theory as applicable only in some early epoch of the universe.

In the first way, one needs a fine tuning of parameters, classical or quantum, to an incredible accuracy. One notices too that the 11-dimensional supergravity theory admits no such freedom. In any case one must confess ignorance.

The second way, on the other hand, may sound like seeking a sanctuary. Nevertheless, like in the inflationary scenario [40], one hopes to obtain a deep insight into fascinating interplays between cosmology and particle physics. Then comes the crucial question: can one discover in the framework of Kaluza-Klein theories a mechanism which causes a cosmological transition to  $M_4^b$  from a spacetime of (perhaps negatively) large curvature?

Acknowledgements

I thank Masaaki Kato for many useful conversations, and Janis Niedra for carefully reading the manuscript.

Appendix A. Some details of deriving Eq. (2.4a) [5,41]

Computations are simpler if one uses the vielbein instead of the metric tensor directly.

The vielbein  $\hat{b}_{\hat{\rho}}^{\hat{m}}$  gives a coordinate basis in the local Lorentz frame such that

$$\hat{g}_{\hat{\rho}\hat{\sigma}} = \hat{\eta}_{\hat{m}\hat{n}} \hat{b}_{\hat{\rho}}^{\hat{m}} \hat{b}_{\hat{\sigma}}^{\hat{n}} = \hat{e}_{\hat{\rho}}^{\hat{m}} \hat{e}_{\hat{\sigma}}^{\hat{n}}, \quad (\text{A.1})$$

where the local Lorentz frame is labelled by the latin indices ( $\hat{m}, \hat{n}, \dots$  for D-dimensional;  $m, n, \dots$  for 4-dimensional;  $a, b, \dots$  for n-dimensional coordinates, respectively) which are raised or lowered by the Minkowski metric  $\hat{\eta}_{\hat{m}\hat{n}} = \hat{\eta}^{\hat{m}\hat{n}} = \text{diag}(-+++ \dots +)$ . The inverse vielbein  $\hat{b}_{\hat{m}}^{\hat{\rho}}$  is introduced so that

$$\hat{b}_{\hat{m}}^{\hat{\rho}} \hat{b}_{\hat{\rho}}^{\hat{n}} = \delta_{\hat{m}}^{\hat{n}}, \quad \hat{b}_{\hat{\rho}}^{\hat{m}} \hat{b}_{\hat{m}}^{\hat{\nu}} = \delta_{\hat{\rho}}^{\hat{\nu}}. \quad (\text{A.2})$$

It gives the inverse metric by

$$\hat{g}^{\hat{\rho}\hat{\sigma}} = \hat{\eta}^{\hat{m}\hat{n}} \hat{b}_{\hat{m}}^{\hat{\rho}} \hat{b}_{\hat{n}}^{\hat{\sigma}} = \hat{e}_{\hat{m}}^{\hat{\rho}} \hat{e}_{\hat{n}}^{\hat{\sigma}}. \quad (\text{A.3})$$

It is easy to derive

$$\sqrt{-\hat{g}} = \hat{e} = \det(\hat{b}_{\hat{\rho}}^{\hat{m}}). \quad (\text{A.4})$$

The zero-mass ansatz (2.2) can be reproduced by

$$\hat{b}_{\hat{\rho}}^{\hat{m}} = \begin{pmatrix} \tilde{\phi}^{-n/2} b_{\mu}^m & e \phi^{-1} \xi_A^{\alpha} A_{\rho}^A \\ 0 & \phi^{-1} e_{\alpha}^a \end{pmatrix},$$

$$\hat{b}_{\hat{m}}^{\hat{\rho}} = \begin{pmatrix} \tilde{\phi}^{-n/2} b_{\mu}^{\mu} & -e \tilde{\phi}^{-n/2} \xi_A^{\alpha} A_m^A \\ 0 & \phi e_{\alpha}^a \end{pmatrix}, \quad (\text{A.5})$$

where  $b_{\mu}^m(x)$  is the purely 4-dimensional vierbein giving  $g_{\mu\nu} = b_{\mu}^m b_{\nu m}$ , whereas

$e_{\alpha}^a$  is the purely n-dimensional vielbein giving  $\gamma_{\alpha\beta} = e_{\alpha}^a e_{\beta a}$ . We also use  $A_m^A(x) = b_{\mu}^m(x) A_{\mu}^A(x)$  and  $\xi_A^{\alpha}(y) = e_{\alpha}^a(y) \xi_A^{\alpha}(y)$ . Given the metric, the vielbein is unique only up to a local Lorentz transformation. We have exploited this arbitrariness to put  $\hat{b}_{\hat{\rho}}^{\hat{m}} = \hat{b}_{\hat{\rho}}^{\hat{a}} = 0$ , which simplifies computations considerably.

The Ricci rotation coefficient is defined by

$$\hat{c}_{\hat{\rho}, \hat{m}\hat{n}} = (\hat{e}_{\hat{m}}^{\hat{\rho}} \hat{e}_{\hat{n}}^{\hat{\nu}} - \hat{e}_{\hat{n}}^{\hat{\rho}} \hat{e}_{\hat{m}}^{\hat{\nu}}) \partial_{\hat{\sigma}} \hat{e}_{\hat{\rho}}^{\hat{\sigma}} = -\hat{c}_{\hat{\rho}, \hat{n}\hat{m}}, \quad (\text{A.6})$$

and the spin connection with Latin indices is given by

$$\hat{A}_{\hat{m}\hat{n}, \hat{\rho}} = \frac{1}{2} (\hat{c}_{\hat{\rho}, \hat{m}\hat{n}} - \hat{c}_{\hat{m}, \hat{\rho}\hat{n}} + \hat{c}_{\hat{n}, \hat{\rho}\hat{m}}) = -\hat{A}_{\hat{n}\hat{m}, \hat{\rho}}, \quad (\text{A.7})$$

which gives the Einstein-Hilbert action (2.1) in the form

$$\mathbb{I} = \int d^D x \hat{e} \frac{1}{2\kappa^2} (\hat{A}_{\hat{m}\hat{n}, \hat{\rho}} \hat{A}^{\hat{m}\hat{n}, \hat{\rho}} + \hat{A}_{\hat{m}\hat{n}, \hat{\rho}} \hat{A}^{\hat{m}\hat{\rho}, \hat{n}}). \quad (\text{A.8})$$

By substituting (A.5) into (A.6) one obtains

$$\hat{c}_{\hat{\rho}, mn} = \tilde{\phi}^{-n/2} [c_{\hat{\rho}, mn} + \frac{n}{2} (\eta_{lm} \Phi_m - \eta_{ln} \Phi_m)],$$

$$\hat{c}_{a, mn} = -e \tilde{\phi}^{-n-1} F^A{}_{mn} \xi_A^a,$$

$$\hat{c}_{\hat{\rho}, am} = 0,$$

$$\hat{c}_{a, am} = -\tilde{\phi}^{-n/2} \Phi_m \eta_{\beta a} - e \tilde{\phi}^{-n/2} A_m^A e_{\alpha}^a (\partial_{\alpha} \xi_A^{\beta} e_{\beta a} + \xi_A^{\beta} \partial_{\beta} e_{\alpha a}),$$

$$\hat{c}_{m, ab} = 0,$$

$$\hat{c}_{c, ab} = \phi c_{c, ab}, \quad (\text{A.9})$$

where

$$\Phi_m = b_m^{\mu} \Phi_{\mu} = b_m^{\mu} \partial_{\mu} \ln \tilde{\phi},$$

$$F^A{}_{mn} = b_m^{\mu} b_n^{\nu} F^A{}_{\mu\nu}. \quad (\text{A.10})$$

The Ricci rotation coefficient  $c_{\hat{\rho}, mn}$  is purely 4-dimensional and derived from  $b_{\mu}^m$ , while  $c_{c, ab}$  is purely n-dimensional and derived from  $e_{\alpha}^a$ .

Substitution of (A.9) further into (A.7) yields

$$\hat{A}_{mn, \hat{\rho}} = \tilde{\phi}^{-n/2} [A_{mn, \hat{\rho}} + \frac{n}{2} (\eta_{lm} \Phi_m - \eta_{ln} \Phi_m)],$$



$$\hat{A}_{mn,a} = -\frac{1}{2} \ell e^{\tilde{\phi}^{-n-1}} F_{mn}^A \xi_{Aa},$$

$$\hat{A}_{am,n} = -\hat{A}_{mn,a},$$

$$\hat{A}_{ab,\ell} = -\tilde{\phi}^{-n/2} \Xi_m \eta_{ab},$$

$$\hat{A}_{ab,m} = -\frac{1}{2} e^{\tilde{\phi}^{-n/2}} A_m^A \left[ \xi_A^\beta (e_\alpha^\alpha \partial_\beta e_{\alpha\alpha} - e_\alpha^\alpha \partial_\beta b_{\alpha\alpha}) + (e_\alpha^\alpha e_{\beta\alpha} - e_\alpha^\alpha e_{\beta\alpha}) \partial_\alpha \xi_A^\beta \right],$$

$$\hat{A}_{ab,c} = \phi A_{ab,c}. \quad (\text{A.11})$$

These are finally substituted into (A.8) to give (2.4a), where one uses

$$-A_{mn,\ell} A^{m\ell,n} + A_{m,n}^m A^{m\ell,\ell} = R_4,$$

$$-A_{ab,c} A^{ac,b} + A_{a,b}^a A^{ac,c} = R_m,$$

up to surface terms which are dropped.

#### Appendix B. A simplified counting of the physical degrees of freedom of a massless totally antisymmetric tensor field

For a massless totally antisymmetric tensor field  $A_{\mu_1 \dots \mu_r}$  in D-dimensional flat spacetime, the classical Lagrangian is given by

$$L_0 = -\frac{1}{2} \frac{1}{(r+1)!} F_{\mu_1 \dots \mu_{r+1}} F^{\mu_1 \dots \mu_{r+1}}, \quad (\text{B.1})$$

where

$$F_{\mu_1 \dots \mu_{r+1}} = (r+1) \partial_{[\mu_1} A_{\mu_2 \dots \mu_{r+1}]}. \quad (\text{B.2})$$

In this Appendix we suppress hats on the indices which are raised or lowered with the aid of  $\eta_{\mu\nu} = \eta^{\mu\nu} = \text{diag}(-++ \dots +)$ . The field equation derived from (B.1) has no inverse, thus requiring to add a gauge fixing term. The total Lagrangian is then given by

$$L = -\frac{1}{2} \frac{1}{(r+1)!} F_{\mu_1 \dots \mu_{r+1}} F^{\mu_1 \dots \mu_{r+1}} - \frac{1}{2} \alpha^{-1} \frac{1}{r!} G_{\mu_2 \dots \mu_r} G^{\mu_2 \dots \mu_r}, \quad (\text{B.3})$$

where  $\alpha$  is a gauge parameter, and

$$G_{\mu_2 \dots \mu_r} = \partial^K A_{\mu_2 \dots \mu_r}. \quad (\text{B.4})$$

The Euler-Lagrange equation yields

$$\partial_\rho F^{\rho\mu_1 \dots \mu_r} + \alpha^{-1} \partial^{[\mu_1} G^{\mu_2 \dots \mu_r]} = 0. \quad (\text{B.5})$$

Applying  $\partial_{\mu_1}$  to this equation one obtains

$$\alpha^{-1} \partial_{\mu_1} \partial^{[\mu_1} G^{\mu_2 \dots \mu_r]} = 0,$$

which can be put into

$$\alpha^{-1} \square G^{\mu_2 \dots \mu_r} = 0. \quad (\text{B.6})$$

This, with a finite  $\alpha$ , allows one to impose a subsidiary condition

$$G^{\mu_2 \dots \mu_r (+)} | \rangle = 0, \quad (\text{B.7})$$

for the physical state  $| \rangle$ , where (+) means a positive frequency part.

Eq.(B.5) can also be put into the form

$$\square A^{\mu_1 \dots \mu_r} - (1 - \alpha^{-1}) \partial^{[\mu_1} G^{\mu_2 \dots \mu_r]} = 0.$$

The choice  $\alpha = 1$  is particularly convenient because (B.5) then simplifies to

$$\square A^{\mu_1 \dots \mu_r} = 0. \quad (\text{B.8})$$

The corresponding Lagrangian is

$$L = -\frac{1}{2} \frac{1}{r!} \partial_\nu A_{\mu_1 \dots \mu_r} \partial^\nu A^{\mu_1 \dots \mu_r}. \quad (\text{B.9})$$

In this Fermi gauge, the Hamiltonian in the momentum representation is simply given by

$$H(k) = \frac{1}{r!} |k| \eta^{\mu_1 \nu_1} \dots \eta^{\mu_r \nu_r} a_{\mu_1 \dots \mu_r}^\dagger(k) a_{\nu_1 \dots \nu_r}(k), \quad (\text{B.10a})$$

which can be written as

$$H(k) = \sum_{\{\mu_1 \dots \mu_r\}} |k| \epsilon_{\{\mu_1 \dots \mu_r\}} a_{\mu_1 \dots \mu_r}^\dagger(k) a_{\mu_1 \dots \mu_r}(k), \quad (\text{B.10b})$$

where the sum is taken over independent combinations of the antisymmetrized indices. The symbol  $\epsilon_{\{\mu_1 \dots \mu_r\}}$  is -1 if 0 is included among  $\mu_1 \dots \mu_r$ , while it is +1 otherwise, as dictated from  $\eta^{\mu_1 \nu_1} \dots \eta^{\mu_r \nu_r}$ .

The indices may be split into  $\mu = 0$ ,  $\mu = 1$  for the direction of  $\vec{k}$ , and  $\mu = 2, \dots, D-1$  for the perpendicular directions. For these transverse components we use letters  $p, q, \dots$ . One then obtains

$$H(k) = |k| \sum_{\{p_2 \dots\}} \left( -a_{0_1 p \dots}^{\dagger} a_{0_1 p \dots} - a_{0 p_2 \dots}^{\dagger} a_{0 p_2 \dots} \dots \right. \\ \left. + a_{1 p_2 \dots}^{\dagger} a_{1 p_2 \dots} + a_{p_2 \dots}^{\dagger} a_{p_2 \dots} \right). \quad (B.11)$$

The same kind of splitting is also applied to the subsidiary condition (B.7). One first finds that  $G^{0_1 p \dots} | \rangle = 0$  yields nothing, because  $G^{0_1 p \dots}(k)$  vanishes trivially due to  $k^{\dot{p}} = 0$ :

$$G^{0_1 p \dots}(k) = i k_{\dot{p}} a^{0_1 p \dots}(k) = 0.$$

Consider next  $G^{0 p_2 \dots} | \rangle = 0$ . One obtains

$$G^{0 p_2 \dots}(k) = i k_{\dot{p}_2} a^{0 p_2 \dots}(k) = i |k| a^{0 p_2 \dots}(k).$$

The above condition therefore gives

$$a^{0 p_2 \dots}(k) | \rangle = 0,$$

or

$$a_{1_0 p \dots}(k) | \rangle = 0. \quad (B.12)$$

The same result follows also from  $G^{1 p_2 \dots} | \rangle = 0$ . Finally one observes

$$G^{p_2 \dots}(k) = i(k_0 a^{0 p \dots} + k_1 a^{1 p \dots}) \\ = i |k| (-a^{0 p \dots} + a^{1 p \dots}).$$

The condition  $G^{p_2 \dots} | \rangle = 0$  hence gives

$$a^{0 p \dots}(k) | \rangle = a^{1 p \dots}(k) | \rangle,$$

or

$$a_{0 p \dots}(k) | \rangle = -a_{1 p \dots}(k) | \rangle. \quad (B.13)$$

One can now apply these conditions to the physical matrix element  $\langle H(k) | \rangle$ . The first term in (B.11) drops out because of (B.12), while the second and third terms cancel due to (B.13) and its hermitian conjugate. One is thus left only with the transverse components:

$$\langle H(k) | \rangle = |k| \sum_{\{p_2 \dots\}} \langle a_{p_2 \dots}^{\dagger}(k) a_{p_2 \dots}(k) | \rangle. \quad (B.14)$$

The number of the independent combinations  $\{p_2 \dots\}$  is clearly  $\binom{D-2}{r}$ , which is the number of physical degrees of freedom:

$$\gamma = \binom{D-2}{r}. \quad (B.15)$$

### Appendix C. Relation between N=8 supergravity in 4 dimensions and simple supergravity in 11 dimensions

Generators of simple supersymmetry are provided by a Majorana spinor S whose 4 components are Grassmann numbers. For massless fields in 4 dimensions, one constructs a helicity-raising operator  $S_+$  and a helicity-lowering operator  $S_-$  which have nonzero norms. Given a helicity eigenstate  $|\lambda\rangle$ , these operators act like

$$S_{\pm} |\lambda\rangle = |\lambda \pm \frac{1}{2}\rangle.$$

Since  $S_+^2 = S_-^2 = 0$ , one obtains a supersymmetric doublet  $(|\lambda\rangle, |\lambda + \frac{1}{2}\rangle)$  or  $(|\lambda\rangle, |\lambda - \frac{1}{2}\rangle)$ .

Extended supersymmetries are characterized by N such spinorial generators  $S_{\pm}^a$  ( $a = 1, \dots, N$ ). Given a state  $|\lambda\rangle$ , the result of applying k of  $S_-$ 's is

$$S_-^{a_1} \dots S_-^{a_k} |\lambda\rangle = |\lambda - \frac{1}{2}k\rangle_{\{a_1 \dots a_k\}}.$$

Since  $S_-^{a_i}$  with different  $a_i$  anticommute, one obtains  $\binom{N}{k}$  states with the helicity  $\lambda - k/2$ . The minimum helicity arrived at in this way is  $\lambda - N/2$ .

If  $N \geq 9$ , the helicity values in a supersymmetric multiplet would span a range  $N/2 \geq 9/2$ , which implies the occurrence of fields with spin greater than 5/2. Accepting the common view that no consistent interacting field theory can be constructed for spins beyond 2, the highest possible value of N is 8.

The simplest supersymmetric multiplet of N=8 supersymmetry is the one which contains the spin 2 graviton. This is the multiplet of N=8 supergravity. According to the binomial coefficients for the lower spin states mentioned above, one finds the following spin content: 1 graviton ( $2^+$ ), 8 gravitinos ( $3/2$ ), 28 vectors ( $1^-$ ), 56 spinors ( $1/2$ ) and 70 scalars ( $0^{\pm}$ ).

Cremmer and Julia discovered that the above spin content is precisely the same as what is obtained by dimensionally reducing simple (N=1) supergravity in 11 dimensions. They first demonstrated that 11-dimensional supergravity contains one graviton (described by the metric  $\hat{g}_{\mu\nu}$  or the elfbein  $\hat{b}_{\mu}^{\hat{a}}$ ), one gravitino (a Rarita-Schwinger field  $\psi_{\hat{\mu}}$ ) and an antisymmetric tensor field of rank 3. The necessity of introducing the last type of field is understood on

the basis of matching the physical degrees of freedom of bosons and fermions.

Consider  $\hat{g}_{\hat{\rho}\hat{\sigma}}$  in D dimensions. This symmetric tensor has  $D(D+1)/2$  components to start with. One subtracts D, the number of gauge transformations, corresponding to general coordinate transformations. One further subtracts another D, the number of gauge-fixing conditions, e.g.  $\partial_{\hat{\rho}}(\sqrt{-\hat{g}} \hat{g}^{\hat{\rho}\hat{\sigma}}) = 0$ . For the physical degrees of freedom, one is thus left with  $D(D-3)/2$ , which is 2 for  $D = 4$ , and 44 for  $D = 11$ .

We next turn to  $\psi_{\hat{\rho}}$ . Each spinor field has  $2^{\lfloor D/2 \rfloor}$  components. Hence a Majorana vectorial spinor has  $D \times 2^{\lfloor D/2 \rfloor}$  components. A local supersymmetry transformation is a gauge transformation with a spinorial parameter  $\epsilon$  of  $2^{\lfloor D/2 \rfloor}$  components. This number should be subtracted. Two gauge-fixing conditions must be imposed:  $\partial_{\hat{\rho}}\psi^{\hat{\rho}} = \gamma_{\hat{\rho}}\psi^{\hat{\rho}} = 0$ . These represent  $2 \times 2^{\lfloor D/2 \rfloor}$  conditions. Subtracting these also, one is left with  $(D-3) \times 2^{\lfloor D/2 \rfloor}$ . Now recall that  $\psi_{\hat{\rho}}$  obeys a first-order differential equation, but each dynamical degree of freedom corresponds to a harmonic oscillator which obeys a second-order equation. For this reason the above number is divided by 2, giving  $[(D-3)/2] \times 2^{\lfloor D/2 \rfloor}$  for the physical degrees of freedom of the gravitino. This equals 2 for  $D = 4$ , and 128 for  $D = 11$ .

One has thus  $128 - 44 = 84$  for the excess of fermionic degrees of freedom for  $D = 11$ . The number 84 turns out to be  ${}_{11-2}C_3 = {}_9C_3$  for the physical degrees of freedom of a massless antisymmetric tensor field of rank 3, according to (B.15). With this matching, Cremmer and Julia explicitly constructed the action of simple supergravity in 11 dimensions.

We finally outline how this theory gives the same field content as what one expects from N=8 supergravity in 4 dimensions. Among the components of  $\hat{g}_{\hat{\rho}\hat{\sigma}}$ , the purely 4-dimensional part  $g_{\mu\nu}$  describes 1 graviton as usual. The off-diagonal part  $\hat{g}_{\mu\alpha}$  represents 7 vector fields, while the remaining  $\hat{g}_{\alpha\beta}$  gives  $7 \times 8/2 = 28$  scalar fields.

The components of the totally antisymmetric tensor field  $A_{\hat{\rho}\hat{\sigma}\hat{\lambda}}$  are also split into four classes:  $A_{\mu\nu\lambda}$ ,  $A_{\mu\nu\alpha}$ ,  $A_{\mu\alpha\beta}$ , and  $A_{\alpha\beta\gamma}$ . The first one has no physical degrees of freedom, as is suggested by the meaningless result  ${}_2C_3$  from (B.15). The second class gives 7 antisymmetric tensor fields of rank 2. Each has the degree of freedom  ${}_2C_2 = 1$ , corresponding to a scalar field. Altogether one has 7 scalar fields. The third class provides  ${}_7C_2 = 21$  vector fields, while the last class represents  ${}_7C_3 = 35$  spinless fields in 4 dimensions. Summing up, one obtains 1 graviton,  $7+21 = 28$  vector fields and  $28+7+35 = 70$  spinless fields, in agreement with the number obtained from N=8

supergravity in 4 dimensions.

The same type of analysis leads to agreement also for the fermion fields.

#### Footnotes

- f1 One may wonder if spacetime of n+1 time coordinates, i.e.  $SO(3, n+1)$  for tangent space, will also do. One can construct n-dimensional compact internal space exclusively out of the timelike coordinates. A negative curvature scalar  $R_n$  is not unacceptable. One cannot accept, however, that the Yang-Mills field is a ghost; the sign of the last term of (2.4a) comes out wrong.
- f2 With  $D=11$ , Cremmer and Julia chose  $T^7$ , resulting only in abelian gauge symmetries. As a candidate giving nonabelian symmetries,  $S^7$  has been a target of intensive studies [10-13]. Witten pointed out [17], on the other hand, that  $n=7$  is a minimum dimensionality to realize  $SU(3) \times SU(2) \times U(1)$  parametrized on  $CP_2 \times S^2 \times S^4$ , where 4-dimensional  $CP_2$  is homogeneous space  $SU(3)/SU(2) \times U(1)$ , while  $S^2 = SO(3)/SO(2) = SU(2)/U(1)$ .
- f3 Possible semiclassical instability was also discussed [19].
- f4 The coefficient (r+1) has been so chosen that each independent term occurs only once on the far-right-hand side.
- f5 This is the point of view in ref. 9, for example.
- f6 See, for example, ref. 22.
- f7 An entirely different approach was suggested by Hawking [23].
- f8 For applications to more general cases, see ref. 24.
- f9 See, however, refs. 30.
- f10 For other attempts, see refs. 32.
- f11 For more details, see refs. 38 and papers cited therein.

#### References

- 1. Th. Kaluza, Sitzungber. Preuss. Akad. Wiss. Berlin, Math. Phys. (1921), 966; O. Klein, Z. Phys. 37(1926) 895.
- 2. E. Cremmer and B. Julia, Nucl. Phys. B159(1979) 141.
- 3. B. DeWitt, in Relativity, Groups and Topology (Gordon and Breach, 1964).
- 4. Y.M. Cho and P.G.O. Freund, Phys. Rev. D12(1975) 1711.

5. J. Scherk and J. Schwarz, Nucl. Phys. 153(1979) 61.
6. S. Tanaka, Prog. Theor. Phys. Suppl. No 67(1979) 282; A. Salam and J. Strathdee, Ann. Phys. (N.Y.) 141(1982) 316.
7. E. Cremmer and J. Scherk, Nucl. Phys. B108(1976) 409; B118(1977) 61; J.F. Luciani, Nucl. Phys. B135(1978) 111.
8. P.G.O. Freund and M.A. Rubin, Phys. Lett. 97B(1980) 233.
9. S. Randjbar-Daemi, A. Salam and J. Strathdee, Nucl. Phys. B214(1983) 491.
10. M.A. Awada, M.J. Duff and C.N. Pope, Phys. Rev. Lett. 50(1983), 294; M.J. Duff, B.E.W. Nilsson and C.N. Pope, Phys. Rev. Lett. 50(1983) 2043; L. Castellani, R. D'Auria and P. Fre, Univ. of Turin preprint, IFTT 427 (1983); M.J. Duff, B.E.W. Nilsson and C.N. Pope, Imperial College preprint TP/83-84/18.
11. M.J. Duff and C.N. Pope, in Supersymmetry and Supergravity 82 (World Science Publishing Co., 1983).
12. F. Englert, Phys. Lett. 119B(1982) 339.
13. B. Biran, A. Casher, F. Englert, M. Rooman and P. Spindel, Universite Libre de Bruxelles preprint (1983).
14. Y. Fujii, T. Inami, M. Kato and N. Ohta, Phys. Lett. 131B(1983) 319.
15. D.N. Page, Phys. Rev. D28(1983) 2976.
16. Y. Fujii, T. Inami and M. Kato, Research Institute for Fundamental Physics preprint, RIFP-544(1984).
17. E. Witten, Nucl. Phys. B186(1981) 412.
18. T. Appelquist and A. Chodos, Phys. Rev. Lett. 50(1983) 141; T. Inami and O. Yasuda, Phys. Lett. 133B(1983) 180.
19. E. Witten, Nucl. Phys. B195(1982) 481.
20. A. Aurilia, H. Nicolai and P.K. Townsend, Nucl. Phys. B176(1980) 509.
21. N. Ohta, Prog. Theor. Phys. 66(1981) 1408.
22. P. Candelas and D.J. Raine, Center for Theoretical Physics, University of Texas preprint (1983).
23. S.W. Hawking, Nucl. Phys. B144(1978) 349.
24. N. Ohta, Phys. Lett. 138B(1984) 63; Osaka Univ. preprint OS-GE 84-06 (1984).
25. C. Fronsdal, Phys. Rev. D12(1975) 3819.
26. P. Candelas and J.S. Dowker, Phys. Rev. D19(1979) 2902; N.D. Birrel and P.C.W. Davies, Quantum Fields in Curved Space (Cambridge University Press, 1982).
27. P. Breitenlohner and D.Z. Freedmann, Phys. Lett. 115B(1982) 197; Ann. Phys. (N.Y.) 144(1982) 249.
28. A. Lichnerowicz, in Proceedings of the 1963 Les Houches Summer School.
29. E. Witten, Princeton University preprint (1983).
30. A.C. Davis, A.J. Macfarlane and J.W. van Holten, Phys. Lett. 125B(1983) 151; N. Ohta, Phys. Lett. 134B(1984) 75.
31. S. Watamura, Phys. Lett. 129B(1983) 188; 136B(1984) 245.
32. S. Tanaka, Prog. Theor. Phys. 70(1983) 563; Yong-Shi Wu and A. Zee, University of Washington preprint (1983).
33. A. Chodos and S. Detweiler, Phys. Rev. D21(1980) 2167; P.G.O. Freund, Nucl. Phys. B209(1982) 146; Y. Yamaguchi, University of Tokyo preprint, (1982).
34. K. Maeda, Research Institute for Fundamental Physics preprint, RIFP-532 (1983).
35. S. Randjbar-Daemi, A. Salam and J. Strathdee, Phys. Lett. 135B(1984) 388; M. Yoshimura, National Laboratory for High Energy Physics preprint KEK-TH 75 (1984).
36. W.J. Marciano, Phys. Rev. Lett. 52(1984) 489.
37. P.A.M. Dirac, Proc. Roy. Soc. London A165 (1938) 199; A333(1973) 403; A338(1974) 439.
38. Y. Fujii, Phys. Rev. D26(1982) 2580; Y. Fujii and J.M. Niedra, Prog. Theor. Phys. 70(1983) 412.
39. P.C.W. Davies, J. Phys. A5(1972) 1296; A.I. Shlyakhter, Nature 26 (1976) 340.
40. K. Sato, Mon. Not. R. Astron. Soc. 195(1981) 467; A. Guth, Phys. Rev. D23(1981) 347.
41. N. Ohta, Osaka Univ. preprint, OS-GE 84-04(1984).

Quantum Effects in Generalized Kaluza-Klein Theories<sup>\*)</sup>

Osamu Yasuda

Department of Physics, University of Tokyo

Bunkyo-ku, Tokyo 113, Japan

ABSTRACT

One-loop effective potential is evaluated in Kaluza-Klein theories of topology  $M^4 \times T^n$ , treating the divergences carefully. It is shown that some of the physical circumferences of  $T^n$  shrink to sizes on the order of the Planck length.

---

<sup>\*)</sup> This work was done in collaboration with T. Inami (RIFP, Kyoto Univ.).

We shall consider here the gravitational Casimir effect in Kaluza-Klein theories. In Kaluza-Klein theories the sizes of the extra space are expected to be of the order of the Planck length  $\kappa$ . This is because gauge coupling constants are given by the ratio of the Planck length  $\kappa$  to the sizes of the extra space and this factor should be of order unity if Kaluza-Klein theories contain grand unified theories at all. Hence the question arises: Why are the sizes of the extra space so small?

Recently Appelquist and Chodos<sup>1)</sup> have argued in five-dimensional Kaluza-Klein theory that the circumference of the extra space contracts to a size on the order of the Planck length  $\kappa$  due to the gravitational Casimir effect. Here we investigate the quantum effects in Kaluza-Klein theories compactified on  $M^4 \times T^n$ , where  $M^4$  and  $T^n$  are a four-dimensional Minkowski space-time and an  $n$ -dimensional torus, respectively. The quantum effects in these models turn out to be natural extension of that of Appelquist and Chodos.

We consider the Einstein action in  $D(=4+n)$ -dimensions

$$I = \int d^4x \int d^n y \Omega_n^{-1} \frac{-1}{\kappa^2} (-g)^{1/2} R$$

with

$$0 \leq y^\alpha \leq 2\pi R_\alpha \quad (\alpha=1, \dots, n), \quad \Omega_n = \prod_{\alpha=1}^n (2\pi R_\alpha)$$

which leads to the Einstein equation

$$R_{MN} - \frac{R}{2} g_{MN} = 0 \quad (M, N = 1, \dots, D).$$

The metric  $\overset{\circ}{g}_{MN}$  for  $M^4 \times T^n$

$$\overset{\circ}{g}_{MN} = \Phi \text{diag}(1, 1, 1, -1, \phi_1^2, \dots, \phi_n^2)$$

with

$$\Phi = \left( \prod_{\alpha=1}^n \rho_{\alpha} \right)^{-\frac{2}{D-2}}$$

is a solution of the Einstein eq., where  $\phi_{\alpha}$  ( $\alpha=1, \dots, n$ ) are variables independent of coordinates,  $\Phi$  makes the coefficient of the reduced Einstein action in four dimensions the same as the usual one and  $R_{\alpha}$  is a radius of each  $S^1$ . The physical circumference of each  $S^1$  is given by  $L_{\alpha} = 2\pi\phi_{\alpha}R_{\alpha}$  and we take  $L_{\alpha}$  as a field variable hereafter.

We decompose the metric  $g_{MN}$  into the background metric and the fluctuation field

$$g_{MN} = \overset{\circ}{g}_{MN} + \kappa R_{MN}.$$

The effective potential at the tree level vanishes, since  $M^4 \times T^n$  has vanishing curvature. Hence it is necessary to consider the one-loop effective potential in order to lift the vacuum degeneracy. When we examine divergences of the one-loop effective potential carefully, the path integral measure of gravity becomes important. The measure for gravity is given by Fujikawa<sup>2)</sup> from the point of view of BRS symmetry associated with general covariance :

$$\begin{aligned} Z &= \int D[(-g)^{\frac{D-4}{4D}} g_{MN}] DB_M DC_N D[(-g)^{\frac{D+2}{4D}} \bar{C}^M] \\ &\times \exp i \int d^4x \int d^n y \Omega_n^{-1} \left[ -\frac{(-g)^{1/2}}{\kappa^2} R + \frac{\alpha}{2} B_M B_N \overset{\circ}{g}^{MN} \right. \\ &\quad \left. + i(-g)^{1/2} B_M (\partial^N R^M_N - \frac{\beta}{2} \partial^M R^N_N) - (-g)^{1/4} \bar{C}^M \Delta^N_M C_N \right] \end{aligned}$$

with

$$\Delta^N_M = \delta^N_M \partial^L \partial_L + (1-\beta) \partial^N \partial_M + O(R_{MN}),$$

where  $\alpha$  and  $\beta$  are gauge parameters,  $\bar{C}^M$  and  $C_N$  are the Faddeev-Popov ghost associated with the gauge fixing of general covariance.

Since  $Z$  is a sourceless partition function,  $Z$  is independent of gauges. Therefore we can put  $\alpha=\beta=1$  which diagonalizes mass matrix of gravity. Performing integration over  $B_M$ ,  $g^{(D-4)/4D} g_{MN}$ ,  $C_N$  and  $g^{(D+2)/4D} \bar{C}^M$  (up to two and higher loops)

$$\int DB_M \exp(-\int d^Dx \frac{1}{2} B_M B_N \overset{\circ}{g}^{MN}) = \det[\overset{\circ}{g}^{1/2}],$$

$$\begin{aligned} &\int D[\overset{\circ}{g}^{\frac{D-4}{4D}} R_{MN}] \exp -\int d^Dx R_{MK} \overset{\circ}{g}^{MN} \overset{\circ}{g}^{KL} (-\square) R_{NL} \\ &= \det[\overset{\circ}{g}^{\frac{(D-4)(D+1)}{8}}] \cdot \det^{-1/2}[\overset{\circ}{g}^{\frac{(D-4)(D+1)}{4}} (-\square)^{\frac{D(D+1)}{2}}] \\ &= \det[(-\square)^{-\frac{D(D+1)}{4}}], \end{aligned}$$

$$\begin{aligned} &\int DC_N D[\overset{\circ}{g}^{\frac{D+2}{4D}} \bar{C}^M] \exp i \int d^Dx \bar{C}^M \overset{\circ}{g}^{1/4} (-\square) C_M \\ &= \det[\overset{\circ}{g}^{-\frac{D+2}{4}}] \cdot \det^D[\overset{\circ}{g}^{1/4} (-\square)] = \det[\overset{\circ}{g}^{-1/2}] \cdot \det[(-\square)^D], \end{aligned}$$

we finally obtain

$$Z = \det[(-\square)^{-\frac{D(D-3)}{4}}]$$

with no factor of  $\det \overset{\circ}{g}$  <sup>3)</sup>.

Now we evaluate the one-loop effective potential.

$$\begin{aligned} \int d^4x V_{\text{eff}} &= \frac{D(D-3)}{4} \text{tr} \ln(-\square) \\ &= \frac{D(D-3)}{4} \int d^4x \int \frac{d^4k}{(2\pi)^4} \sum_{\{n_\alpha\}} \ln \left[ \Phi^{-1} \left( k_\mu^2 + \sum_{\alpha=1}^n \frac{4\pi n_\alpha^2}{L_\alpha^2} \right) \right] \end{aligned}$$

Using Schwinger's proper time formulation with a cutoff  $\Lambda$ , we have

$$\begin{aligned} \frac{1}{\text{vol}_4} \text{tr} \ln(-\square) &= - \int_{1/\Lambda^2}^{\infty} \frac{dt}{t} \cdot \frac{1}{\text{vol}_4} \text{tr} e^{\square t} \\ &= - \frac{2}{D} \left( \frac{\Phi}{4\pi} \right)^{D/2} \left( \prod_{\alpha=1}^n L_\alpha \right) \Lambda^D \\ &\quad - \frac{\Gamma(D/2)}{\pi^{D/2}} \sum'_{\{n_\alpha\}} \left( \prod_{\alpha=1}^n L_\alpha \right) \left( n_1^2 L_1^2 + \dots + n_n^2 L_n^2 \right)^{-D/2}, \end{aligned}$$

where  $\text{vol}_4 = \int d^4x$  and  $\Sigma'$  denotes all summation over  $n_\alpha$  except  $n_1 = \dots = n_n = 0$ . The divergence of the first term can be renormalized by the cosmological counter term  $I_\lambda$ :

$$\frac{1}{\text{vol}_4} I_\lambda = \left( -\frac{\overset{\circ}{g}}{2} \right)^{1/2} \left( \prod_{\alpha=1}^n 2\pi R_\alpha \right) \delta\lambda = \Phi^{D/2} \left( \prod_{\alpha=1}^n L_\alpha \right) \delta\lambda.$$

Note that this renormalization of cosmological constant can be carried out with  $\delta\lambda$  which is independent of field variables.

This feature is important, since we shall compare the free energy at different field variables. In general  $\text{tr} e^{\square t}$  has the asymptotic expansion as  $t \rightarrow +0$  <sup>4)</sup>

$$\begin{aligned} \frac{1}{\text{vol}_D} \text{tr} e^{\square t} &\underset{t \rightarrow +0}{=} c_0 t^{-D/2} + c_1 t^{1-D/2} R + c_2 t^{2-D/2} RR + \dots \\ &\quad + c_{D/2} \underbrace{R \dots R}_{D/2} + o(t^{-1}). \end{aligned}$$

Substituting this series into the formula of proper time formulation, we get

$$\begin{aligned} \frac{1}{\text{vol}_D} \text{tr} \ln(-\square) &= c_0 \Lambda^D + c_1 \Lambda^{D-2} R + c_2 \Lambda^{D-4} RR + \dots \\ &\quad + c_{D/2} \ln \Lambda \underbrace{R \dots R}_{D/2} + (\text{finite}) \end{aligned}$$

In the present case, only the first and the last term survive, for  $M^4 \times T^n$  has vanishing curvature. Field-independent renormalization mentioned above is, therefore, possible only for manifolds with vanishing curvature. We note in passing that the topology of the Casimir effect <sup>5)</sup> in QED is  $M^3 \times S^1$  and that this manifold also has vanishing curvature.

Now suppose that the physical circumference of each  $S^1$  shrinks (or expands) due to the Casimir force in the same way as that of QED. Then we obtain the conclusion as follows.

In five dimensions we get the same effective potential as that of Appelquist and Chodos <sup>1)</sup> and the circumference of the extra space  $S^1$  shrinks to a size on the order of the Planck length where the parameter of perturbation becomes of order unity and perturbation ceases to be reliable.

In six dimensions the effective potential is given by

$$V_{\text{eff}} = - \frac{2}{\pi^3} \sum'_{n,m} \frac{L_1 L_2}{(n^2 L_1^2 + m^2 L_2^2)^3},$$

whose form is depicted in FIG.1 and FIG.2. The behavior of  $L_1$  and  $L_2$  can be traced by the analogy of the trajectory of a small ball moving in the potential  $V_{\text{eff}}$  (FIG.3). By this analogy we arrive at the following conclusion. The circumference which has the smaller

initial value shrinks to a size on the order of the Planck length, while another circumference reaches a certain size much larger than the Planck length unless the two initial values are almost the same.

In general dimensions the conclusion is derived immediately. The circumferences which have the smallest initial value shrink to a size on the order of the Planck length, while the other circumferences remain of sizes much larger than the Planck length. This fact follows from the observation in mechanics that the force  $-\partial V_{\text{eff}}/\partial L_{\alpha}$  is predominantly in the direction of axes which have the smallest initial value of  $L_{\alpha}$ . The conclusion can be summarized schematically in the following way.

$$\begin{array}{ccc}
 \left( \begin{array}{l} L_{i_1} = \dots = L_{i_k} \\ < L_{i_{k+1}} = \dots = L_{i_{k+m}} \\ < \dots \\ < L_{i_l} = \dots = L_{i_n} \end{array} \right. & \Rightarrow & \left( \begin{array}{l} L_{i_1} = \dots = L_{i_k} \rightarrow \kappa \\ \text{others} \simeq \text{others} \Big|_{\text{initial}} \end{array} \right. \\
 \text{(initial)} & & \text{(final)}
 \end{array}$$

For further details, see ref.6).

#### References

- 1) T. Appelquist and A. Chodos, Phys. Rev. Lett. 50('83)141.
- 2) K. Fujikawa, Nucl. Phys. B226('83)437.
- 3) O. Yasuda, UT-414('83)(to appear in Phys. Lett. B).
- 4) P. B. Gilkey, J. Diff. Geom. 10('75)601.
- 5) H. B. G. Casimir, Proc. Kon. Ned. Akad. Wetenschap. 51('48)793.
- 6) T. Inami and O. Yasuda, Phys. Lett. 133B('83)180.

The same problem is considered also in

T. Appelquist, A. Chodos and E. Myers, Phys. Lett. 127B('83)51, but they have reached a different conclusion.



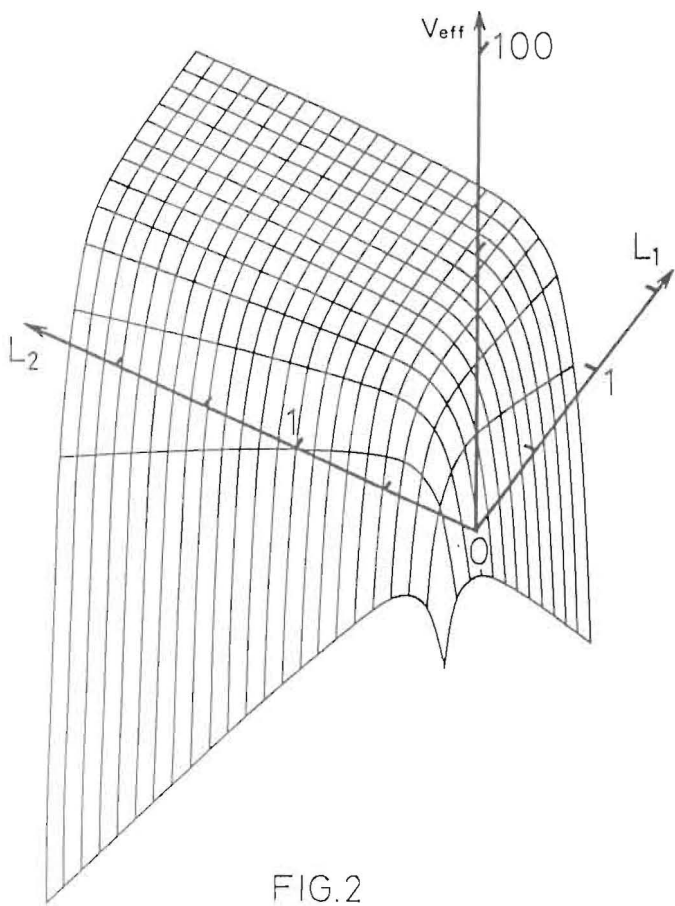


FIG.2

A bird's-eye view of the effective potential  $V_{\text{eff}}$  as a function of two circumferences  $L_1$  and  $L_2$ .

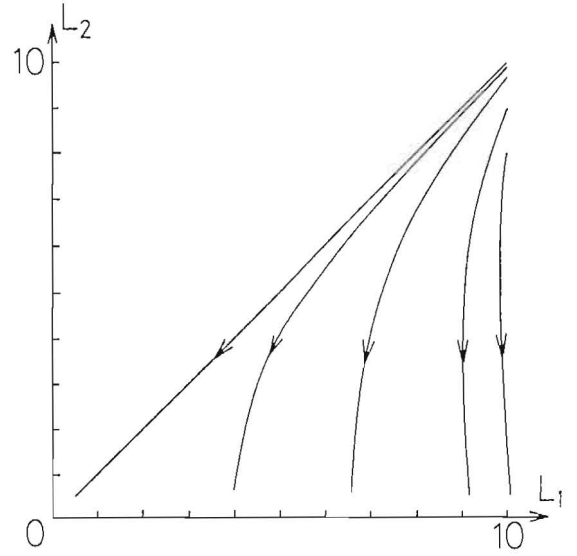


FIG.3

Trajectories of a ball moving due to the force  $F_\alpha = -\partial V_{\text{eff}}/\partial L_\alpha$  ( $\alpha = 1,2$ ) for various initial value of  $L_\alpha$ .

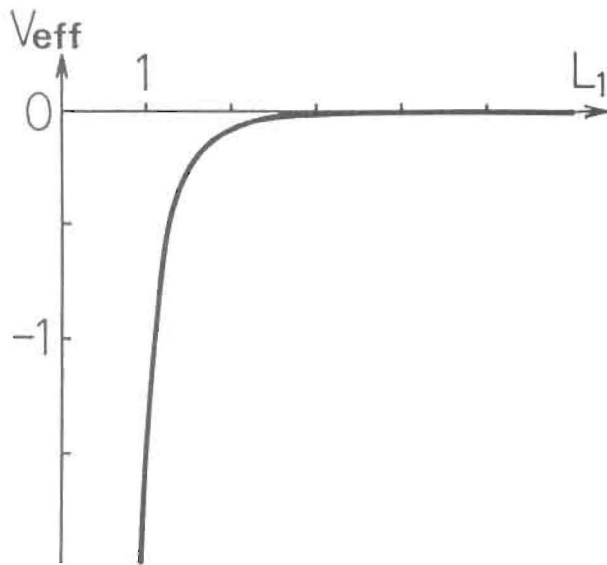


FIG.1a

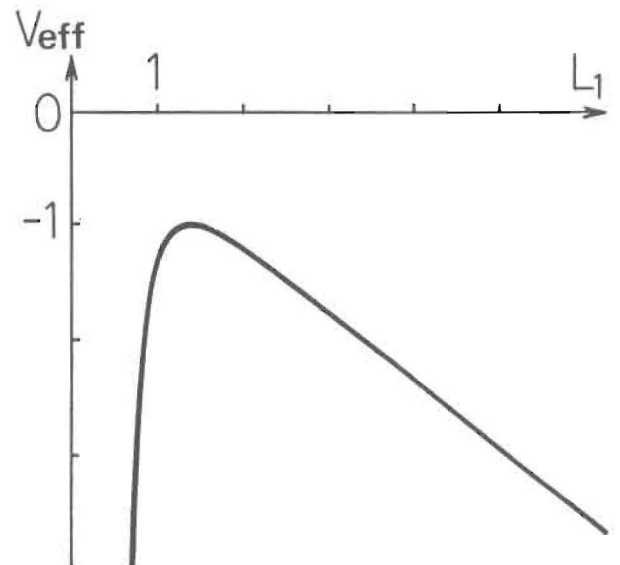


FIG.1b

The effective potential  $V_{\text{eff}}$  as a function of  $L_1$ :  
 a) along the direction  $L_2 = L_1$ , b) along the line  $L_2 = 1$ .  
 (Throughout all figures we use the unit in which  $\kappa = 1$ .)

Masami Yamasaki\*

Department of Physics  
Osaka University  
Toyonaka 560, Japan

Abstract

In a generalized Kaluza-Klein theory, we have investigated solutions to the Einstein equation including matter fields in the background geometry  $M_4 \times S_N \times S_M$ . There are solutions compactified to  $M_4 \times S_1 \times S_M$  due to quantum fluctuation of matter fields in one-loop level. We also calculated the gauge coupling constant of  $SO(N)$ , which is the isometry group of  $S_N$ .

---

\*) This talk is based on the works carried out in collaboration with K. Kikkawa, T. Kubota and S. Sawada, Osaka University.

1. Introduction

The Kaluza-Klein theory<sup>1)</sup> can provide a possible origin of gauge symmetry. It gives a geometrical interpretation of the gauge symmetry under the assumption that the space-time is a  $4+N$  dimensional one and the extra  $N$ -dimensional space is compactified. The geometrical interpretation provides a possibility of the unification of gravity and gauge theories.

Some models provide compact extraspaces instead of noncompact 4-dimensions as a result of the equation of motion. It is said to be spontaneous compactification.

One example of the compactification is the Freund-Rubin<sup>2)</sup> solution in an 11-dimensional super gravity, where the compactification occurs in classical level. On the other hand Appelquist and Chodos<sup>3)</sup> pointed out a possible quantum compactification in 4+1 dimensional Kaluza-Klein model. Along this line Candelas and Weinberg<sup>4)</sup> showed the spontaneous compactification by treating matter fields treated in the quantum one-loop level but the gravitational fields in classical level.

In order to make the theory realistic, the gauge symmetries induced from the extra dimensions should contain  $SU(3) \times SU(2) \times U(1)$  as a subgroup in the isometry groups of extra manifolds.<sup>5)</sup> The manifolds can be generally complicated, but the simplest one is  $CP_2 \times S_2 \times S_1$ . For this reason, it is very important to consider whether the spontaneous compactification is possible in the case that the extra space is direct products of some number of symmetry spaces. We consider in this paper an extension of Candelas-Weinberg model, i.e. the internal space is a direct product of two symmetry spaces instead of a single space as in Candelas-Weinberg.

2. Geometrical Interpretation of Gauge Symmetry

We consider  $4+N$  dimensional space with coordinates  $x^\mu$  and  $y^n$  representing a four dimensional space  $M_4$  and an  $N$ -dimensional compact manifold  $G_N$ , respectively.  $[(x^A) = (x^\mu, y^n)]$ . Suppose that the metric tensor of the  $4+N$  dimensional space has the following form, (Kaluza-Klein ansatz)

$$g^{AB}(x, y) = \begin{bmatrix} g^{\mu\nu}(x) & g^{\mu p}(x) A_p^\alpha(x) \tilde{\xi}_\alpha^n(y) \\ g^{\mu\sigma}(x) A_\sigma^\alpha(x) \tilde{\xi}_\alpha^m(y) & g^{mn}(y) \end{bmatrix}, \quad (2.1)$$

where  $\tilde{\xi}_\alpha^n(y)$  is a Killing vector on  $G_N$ .

If we take the Einstein action in  $4+N$  dimensions, then in four dimensions we obtain the Einstein-Yang-Mills action,

$$S = -\frac{1}{16\pi G} \int d^4x d^N y \sqrt{-g_{4+N}} (R_{4+N} + \Lambda) \\ = \int d^4x \sqrt{-g_4} \mathcal{L}_4$$

where  $R_{4+N}$  is the scalar curvature in  $4+N$  dimensions and  $\mathcal{L}_4$  is the effective Lagrangian in four dimensions. The ansatz (2.1) gives

$$\mathcal{L}_4 = -\frac{1}{16\pi G} \int d^N y \sqrt{-g_N} [ R_4(x) + R_N(y) + \Lambda \\ + \frac{1}{4} g_{nm}(y) \xi_\alpha^n(y) \xi_\beta^m(y) F_{\mu\nu}^\alpha F_{\lambda\rho}^\beta g^{\mu\lambda}(x) g^{\nu\rho}(x) ],$$

where

$$F_{\mu\nu}^\alpha = \partial_\mu A_\nu^\alpha - \partial_\nu A_\mu^\alpha - A_\mu^\beta A_\nu^\gamma C_{\beta\gamma}^\alpha.$$

The structure constant  $C_{\alpha\beta}^\gamma$  is defined by the commutator of the operators of Killing vectors,

$$[ \xi_\alpha^m \partial_m, \xi_\beta^n \partial_n ] = -C_{\alpha\beta}^\gamma \xi_\gamma^p \partial_p.$$

This shows that the symmetry of the gauge fields is of the isometry group of  $G_N$ . The correct normalization of the gravitational constant is

$$\frac{1}{16\pi G} = \frac{1}{16\pi G} \int d^N y \sqrt{\tilde{g}(y)} \quad (2.2)$$

and the canonical normalization condition of the gauge fields are given by

$$\frac{\int d^N y \sqrt{\tilde{g}(y)} \tilde{g}_{nm}(y) \xi_\alpha^n(y) \xi_\beta^m(y)}{\int d^N y \sqrt{\tilde{g}(y)}} = 16\pi G \delta_{\alpha\beta}. \quad (2.3)$$

These conditions determine the gauge coupling constant with matter fields. In this type theory, the gauge covariant derivative is induced from the geometrical one.

If we define the charge by the coefficient of  $A_\mu^\alpha$  in the covariant derivative, then its magnitude is fixed through (2.3).<sup>6)</sup> To see this, we rewrite the norm of the Killing vector as the following,

$$\tilde{g}_{nm}(y) \xi_\alpha^n(y) \xi_\beta^m(y) e^\alpha e^\beta = \frac{g_e^2 N_e s^2(e, y)}{(2\pi)^2} \quad (2.4)$$

where  $g_e$  is the unit charge and  $N_e$  is the degree of degeneracy of the representation with which the coordinate system  $[y^N]$  of  $G_N$  is constructed.

The  $s(e, y_0)$  is circumference of the envelope which is produced by the Killing vector field  $e^\alpha \xi_\alpha^n(y)$  corresponding to the isometry transformation  $\exp[i\sigma e^\alpha \tau_\alpha^n(1)]$ , where  $e$  is a fixed arbitrary vector normalized to 1. Substituting (2.4) into (2.3), one obtains the relation between the charge and the scale of the extraspace manifold,

$$1 = \frac{g_e^2 N_e^2}{(2\pi)^2 16\pi G} \int d^N y \sqrt{\tilde{g}(y)} s^2(e, y) / \int d^N y \sqrt{\tilde{g}(y)}. \quad (2.5)$$

The Kaluza-Klein theory gives Einstein-Yang-Mills theory, under the assumption that the extra space  $G_N$  is compact. The scale of  $G_N$  must have the order of Planck length, in order to make the theory consistent with the observation, for instance, to make the magnitude of gauge coupling constant of order 1.

### 3. Effective Action on a Curved Space

Consider now induced components of gravity and Yang-Mills action by quantum fluctuation of matter fields. The one-loop effective action of scalar and fermion fields is given by

$$S_{\text{eff}} = \frac{i}{2} \ln \det (\square + m^2) / \mu^2 - i \ln \det (i\not{D} - m) / \mu.$$

where we have inserted the mass-dimensional constant  $\mu$  for the purpose of treating dimensionless logarithm.

The  $\zeta$ -function regularization is convenient to treat the effective action.<sup>7)</sup> Let  $\zeta_b$  and  $\zeta_f$  be generalized  $\zeta$ -function respectively of the operators  $\square + m^2$  and  $\not{D}^2 - m^2$ , and the effective action is given by

$$S_{\text{eff}} = -\frac{i}{2} n_b (\zeta_b'(0) + \zeta_b(0) \ln \mu^2) + \frac{i}{2} n_f (\zeta_f'(0) + \zeta_f(0) \ln \mu^2),$$

where  $n_b$  and  $n_f$  are the number of bosons and fermions respectively. The generalized  $\zeta$ -function of operator  $K$  is defined as the follows;

$$\zeta(s) = \sum_n b_n^{-s} = \text{tr } K^{-s} = \int d^4x \sqrt{-g} \langle x | K^{-s} | x \rangle,$$

where  $b_n$ 's are eigenvalues of the operator  $K$ . Using the heat kernel representation, one obtains the  $\zeta$ -function,

$$\zeta(s) = \frac{i}{(4\pi)^2} \frac{1}{\Gamma(s)} \int d^4x \sqrt{-g(x)} \int_0^\infty dt (it)^{s-3} e^{-im^2 t} F(x, x; it) ,$$

where F is assumed to be given by the expansion series,

$$F(x, y; it) = \sum_{k=0}^{\infty} a_k(x, y) (it)^k .$$

The coefficients  $a_k$  should consist of covariant quantities. (e.g.  $R_{\mu\nu\rho\sigma}$ ,  $F_{\mu\nu}^\alpha$  and these covariant derivatives.) They are expressed as the follows, for the operator  $\square + m^2$ ,

$$\begin{aligned} a_0(x, x) &= 1 , \\ a_1(x, x) &= \frac{1}{6} R , \\ a_2(x, x) &= -\frac{1}{12} q^2 FF + \dots , \end{aligned}$$

for the operator  $\not{D}^2 - m^2$ ,

$$\begin{aligned} a_0(x, x) &= 1 , \\ a_1(x, x) &= -\frac{1}{12} R \mathbf{1} , \\ a_2(x, x) &= \frac{1}{6} q^2 FF \mathbf{1} + \dots , \end{aligned}$$

where  $\mathbf{1}$  denotes 4x4 identity matrix.

Finally we obtain the effective action of matter fields,

$$S_{\text{eff}} = \int d^4x \sqrt{-g} L + \frac{F}{\delta_1} \int d^4x \sqrt{-g} R - \frac{1}{4} \sum_e g_e^2 D_e \int d^4x \sqrt{-g} F_{\mu\nu}^e F_{\mu\nu}^e ,$$

where

$$\begin{aligned} L &= -n_b \frac{m_b^4}{64\pi^2} \left( \ln \frac{m_b^2}{\mu^2} - \frac{3}{2} \right) + n_f \frac{m_f^4}{16\pi^2} \left( \ln \frac{m_f^2}{\mu^2} - \frac{3}{2} \right) , \\ E &= -n_b \frac{m_b^2}{192\pi^2} \left( \ln \frac{m_b^2}{\mu^2} - 1 \right) + n_f \frac{m_f^2}{96\pi^2} \left( \ln \frac{m_f^2}{\mu^2} - 1 \right) , (3.1) \\ g_e^2 D_e &= -n_b \frac{q_e^2}{96\pi^2} \ln \frac{m_b^2}{\mu^2} - n_f \frac{q_e^2}{12\pi^2} \ln \frac{m_f^2}{\mu^2} . \end{aligned}$$

This effective action is induced by massive matter fields in four dimensions. In the Kaluza-Klein theory, dimensional reduction produces an infinite set of massive multiplets in four dimensions. Hence sum over the masses and charges of multiplets makes the induced effective action from the

matter parts in higher dimensional Kaluza-Klein theory.<sup>8)</sup>

This induced Einstein-Yang-Mills effective action from matter loop correction changes the normalization condition (2.2) and (2.5), respectively, to

$$\begin{aligned} \frac{1}{16\pi G} &= \frac{1}{16\pi G} \int d^N y \sqrt{\tilde{g}} + \frac{1}{p^2} E , \\ 1 &= \frac{g_e^2 N_e^2}{(2\pi)^2 16\pi G} \int d^N y \sqrt{\tilde{g}(y)} s^2(e, y) / \int d^N y \sqrt{\tilde{g}(y)} + D_e g_e^2 , \end{aligned}$$

where G is the gravitational constant in 4 dimensions.

#### 4. Field Equation and Stability

We consider the theory of gravity having a  $M_4 \times G_N \times G_M$  background geometry with a number of massless matter fields. The Einstein field equation is

$$R^{AB} - \frac{1}{2} (g^{AB} + \Lambda) = -8\pi G \langle T^{AB} \rangle ,$$

where  $g_{AB}$  is metric tensor given by

$$g^{AB} = \begin{bmatrix} \eta^{\mu\nu} & & 0 \\ & g^{ab} & \\ 0 & & g^{nm} \end{bmatrix} .$$

$R^{AB}$  is a Ricci tensor, and  $\langle T^{AB} \rangle$  is an energy-momentum tensor of matter fields; " $\langle \rangle$ " denotes the vacuum expectation value. By Belinfante method,  $\langle T^{NM} \rangle$  is given by

$$\langle T^{AB} \rangle = \frac{\delta I_{\text{eff}}}{\delta g^{AB}} \frac{2}{\sqrt{-g}} , \quad (4.1)$$

$$g = \det(g_{AB}) ,$$

where  $I_{\text{eff}}$  is effective action defined by

$$\begin{aligned} \exp(iI_{\text{eff}}) &= \int \mathcal{D}\varphi e^{iS[\varphi]} , \\ I_{\text{eff}} &= \int d^4x d^N y d^M z \sqrt{-g} L_{\text{eff}} . \end{aligned} \quad (4.2)$$

The field equation can be separated into the parts of subspaces  $M_4, S_N$  and  $S_M$

$$\begin{aligned}
2R_4 + 4R_N + 4R_M + 4\Lambda &= 16\pi GT_4, \\
NR_4 + (N-2)R_N + NR_M + N\Lambda &= 16\pi GT_N, \\
MR_4 + MR_N + (M-2)R_M + M\Lambda &= 16\pi GT_M,
\end{aligned} \tag{4.3}$$

where  $R_4, R_N, R_M$  are scalar curvatures, respectively, of subspaces  $M_4, G_N$  and  $G_M$ , and  $T_4, T_N$  and  $T_M$  are traces of energy-momentum tensors in each subspace. Using (4.1) and (4.2), one obtains

$$\begin{aligned}
T_N &= \frac{2}{\sqrt{-g}} \bar{g}_{nm} \frac{\delta}{\delta \bar{g}_{nm}} I_{\text{eff}} \\
&= \frac{2}{\sqrt{-g}} g_{nm}(y) \frac{\delta}{\delta g_{nm}(y)} \int d^N y \sqrt{g_N(y)} L_{\text{eff}},
\end{aligned}$$

where  $g_N(y)$  is the determinant of the metric of the subspace  $G_N$ . Integration of both sides by  $y$  with  $g_N(y)$  gives

$$T_N = \frac{1}{\int d^N y d^M z \sqrt{g_N} \sqrt{g_M}} \int d^N y d^M z \frac{\partial}{\partial J_N} L_{\text{eff}}.$$

In obtaining above equations, we used the following properties: a tensor with two indices  $T_{nm}$  in maximal symmetric spaces (eg.  $S^N, CP^N$ ) can be written as  $T_{nm}(x) = A g_{nm}(x)$  ( $A$  is independent of  $x$ ).

For  $T_4$ , in a noncompact flat space  $M_4$ , we get

$$T_4 = 4 L_{\text{eff}}.$$

The scalar curvatures  $R_N$  and  $R_M$  are expressed as

$$\begin{aligned}
R_N &= -\frac{a_N}{J_N^2}, \\
R_M &= -\frac{a_M}{J_M^2}.
\end{aligned}$$

where the constants  $a_N$  and  $a_M$  are determined by the topology of  $G_N$  and  $G_M$ . Let  $M_4$  be a flat Minkowski space, then

$$R_4 = 0.$$

The field equations (4.3) become the algebraic ones with respect to the variables  $J_N$  and  $J_M$ . If we choose the dynamical variables  $J_N$  and  $\kappa = \frac{J_M}{J_N}$  as

independent parameters then the equations provide

$$f = \frac{a_N \kappa^2 - a_M(N+4)}{a_N \kappa^2 + a_M}, \tag{4.4a}$$

$$\begin{aligned}
J_N &= \left\{ \frac{-8}{a_N} (4+N+f) \bar{G} J(\kappa) / \int d^N y d^M z \sqrt{g_N} \sqrt{g_M} \right\}^{1/2}, \\
\Lambda &= \frac{a_N}{J_N^2} \frac{2+N+M}{4+N+f},
\end{aligned} \tag{4.4}$$

where

$$\begin{aligned}
\frac{1}{J_N^4} J(\kappa) &= \int d^N y d^M z L_{\text{eff}}, \\
f &= \kappa \frac{\partial}{\partial \kappa} \ln J(\kappa).
\end{aligned} \tag{4.4b}$$

The third equation implies that the cosmological constant must be fine tuned in order that the solutions exist.

One of the conditions that meaningful solutions exist is

$$\bar{G} J(\bar{\kappa}) < 0, \tag{4.5}$$

where  $\bar{\kappa}$  is the solution of (4.4).

In order to study the stability of solutions against small perturbations of  $J_N$  and  $J_M$ , it is enough to consider the total effective potential

$$\begin{aligned}
I_{\text{total}} &= \int d^4 x (-V_{\text{total}}), \\
V_{\text{total}} &= \frac{1}{16\pi G} \int d^{N+M} y \sqrt{\tilde{g}} [\tilde{R} + \Lambda] + V_m[\tilde{g}],
\end{aligned}$$

where  $\tilde{R}$  and  $V_m[\tilde{g}]$  are given by

$$\begin{aligned}
\tilde{R} &= -\frac{a_N}{J_N^2} - \frac{a_M}{J_M^2}, \\
V_m[\tilde{g}] &= -J(\kappa) \frac{1}{J_N^4}, \\
\kappa &= J_M / J_N.
\end{aligned} \tag{4.6}$$

On substituting (4.6) into  $V_{\text{total}}$ , one obtains

$$V_{\text{total}} = \frac{1}{16\pi G} \int d^{N+M}y \sqrt{g} \left[ -\frac{a_N}{\mathcal{P}_N^2} - \frac{a_M}{\mathcal{P}_M^2} + \Lambda \right] - \frac{1}{\mathcal{P}_N^4} J(\kappa) .$$

In order to evaluate the deviation of  $V_{\text{total}}$  from the value at  $\mathcal{P}_N$  and  $\mathcal{P}_M$ , let us introduce parameters  $x, y$

$$\mathcal{P}_N = \bar{\mathcal{P}}_N(1+x) ,$$

$$\mathcal{P}_M = \bar{\mathcal{P}}_M(1+y) ,$$

where " - " denotes the solutions of (4.4). Use of the equation of motion gives

$$V_{\text{total}} = -\frac{J(\bar{\kappa})}{2} \frac{1}{\bar{\mathcal{P}}_N^4} \left[ (N+M+2)(N+M+4)x^2 + 2(M-\bar{f})(N+M+2)xy + (M-\bar{f})(2M-3)-M(M-1)+\bar{\kappa}^2 \frac{J'(\bar{\kappa})}{J(\bar{\kappa})} y^2 \right] + \mathcal{O}(x^2, y^3, x^2y, xy^2) ,$$

where  $\bar{\mathcal{P}}_N$ ,  $\bar{\kappa}$  and  $\bar{f}$  denotes the solutions of (4.4).

Hence the local stability of solutions requires

$$\bar{\kappa}^2 \frac{J'(\bar{\kappa})}{J(\bar{\kappa})} > \frac{N+M+2}{N+M+4} \left[ \bar{f}^2 + \bar{f} \frac{N-3M-12}{N+M+2} + \frac{2M(N+4)}{N+M+12} \right] . \quad (4.7)$$

The asymptotic behaviour of  $V_{\text{total}}$  tells us the condition of global stability.

i) If  $\mathcal{P}_N$  is fixed and  $\mathcal{P}_M \rightarrow 0$ , the dominant term of  $V_{\text{total}}$  is

$$-\frac{1}{\mathcal{P}_N^2} J(\kappa) .$$

Then the stability condition is

$$J(\kappa) < 0 \quad \text{for } \kappa \rightarrow 0 . \quad (4.8)$$

ii) If  $\mathcal{P}_M$  is fixed and  $\mathcal{P}_N \rightarrow 0$ , the dominant term of  $V_{\text{total}}$  is

$$-\frac{1}{\mathcal{P}_N^2} J(\kappa) .$$

Then the stability condition is

$$J(\kappa) < 0 \quad \text{for } \kappa \rightarrow \infty . \quad (4.9)$$

iii) If  $\mathcal{P}_N$  is fixed and  $\mathcal{P}_M \rightarrow \infty$

$$V_{\text{total}} \propto \left[ N+M+2 - \left( \frac{\bar{\mathcal{P}}_N}{\mathcal{P}_N} \right)^2 (f+N+4) + \frac{2\bar{\mathcal{P}}_N^{N+4}\bar{\mathcal{P}}_M^M}{\mathcal{P}_N^{4+N+M}} \left( \frac{J_\infty}{J(\bar{\kappa})} \right) \right] ,$$

where

$$J_\infty = \lim_{\kappa \rightarrow \infty} J(\kappa) / \kappa^M .$$

The positivity of the right-hand side for  $\mathcal{P}_N > 0$  requires

$$\left( \frac{J_\infty}{J(\bar{\kappa})} \right) > \bar{\kappa}^{-M} \left( \frac{f+N+4}{N+M+4} \right)^{2+\frac{N+M}{2}} . \quad (4.10)$$

iv) If  $\mathcal{P}_N \rightarrow 0$  and  $\mathcal{P}_M$  is fixed

$$V_{\text{total}} \propto \left[ N+M+2 - \left( \frac{\bar{\mathcal{P}}_M}{\mathcal{P}_M} \right)^2 (M-f) + \frac{2\bar{\mathcal{P}}_N^{N+4}\bar{\mathcal{P}}_M^M}{\mathcal{P}_M^{4+N+M}} \left( \frac{J_0}{J(\bar{\kappa})} \right) \right] ,$$

where

$$J_0 = \lim_{\kappa \rightarrow 0} J(\kappa) / \kappa^{-N-4} .$$

In order that the right-hand side is positive for  $\mathcal{P}_M > 0$

$$\left( \frac{J_0}{J(\bar{\kappa})} \right) > \bar{\kappa}^{N+4} \left( \frac{M-f}{N+M+4} \right)^{2+\frac{N+M}{2}} . \quad (4.11)$$

If the solutions of (4.4) satisfy these all conditions, the model has a stable compactified solution, with both  $\mathcal{P}_N$  and  $\mathcal{P}_M$  being finite.

## 5. One-loop Effective Action of Matter Fields

We now calculate the effective action induced by matters on the background geometry  $M_4 \times S_{2N} \times S_{2M+1}$ . It can be performed by sum over the mass multiplets of the expression (3.1). The mass spectra for scalar and fermion are respectively,

$$M_{k\ell}^{(0)2} = \frac{1}{\mathcal{P}_1^2} k(k+2N-1) + \frac{1}{\mathcal{P}_2^2} \ell(\ell+2M) ,$$

$$M_{k\ell}^{(1/2)2} = \frac{1}{\mathcal{P}_1^2} (k+N)^2 + \frac{1}{\mathcal{P}_2^2} (\ell + \frac{2M+1}{2})^2 ,$$

and the degeneracies are

$$D_{k,l}^{(0)} = \frac{\Gamma(k+2N-1)\Gamma(l+2M)}{\Gamma(2N)\Gamma(2M+1)} \frac{1}{l!k!} (2k+2N-1)(2l+2M) ,$$

$$D_{k,l}^{(1/2)} = 2^{N+M+2} \frac{\Gamma(k+2N)\Gamma(l+2M+1)}{\Gamma(2N)\Gamma(2M+1)} \frac{1}{l!k!} ,$$

where  $\rho_1$  and  $\rho_2$  are the radii of  $S_{2N}$  and  $S_{2M+1}$  respectively.

The renormalization term to gauge coupling constants can be also obtained by summing over the mass and charge multiplets. All we need are the mean squared charge, that is

$$\bar{q}(\ell)^2 = \sum_{m,e} q_{e,m}^{(\ell)2} / (\text{multiplicity}) = \sum_e \text{tr } t_e^{(\ell)2} / (\text{multiplicity}) .$$

Scalar and fermion multiplets on  $S_N$  respectively have the mean squared charge;

$$\bar{q}(\ell)_{(0)}^2 = g^2 \ell(\ell+N-1) / [N(N+1)/2] ,$$

$$\bar{q}(\ell)_{(1/2)}^2 = g^2 [\ell^2 + \ell N + N(N+1)/8] / [N(N+1)/2] .$$

The infinite sum

$$\mathfrak{S}_M^{(\text{spin})}(s) = \sum_{k,l} D_{k,l}^{(\text{spin})} [M_{k,l}^{(\text{spin})}]^{-2s} ,$$

$$\mathfrak{S}_q^{(\text{spin})}(s) = \sum_{k,l} D_{k,l}^{(\text{spin})} q_{(\text{spin})}^2 [M_{k,l}^{(\text{spin})}]^{-2s} ,$$

are regularized by the analytic continuation in  $s$ .

These summations are converted to complex integrals by the Sommerfeld-Watson transformation. In the case of scalar, singularities of the integrand in the region of  $\text{Re } x > 0$  requires carefull treatment of the contour deformation and the phase of the integrand. The analytic continuation in  $s$  is the same of the Beta fuction. These techniques make the sum finite. The finite expressions of  $\mathfrak{S}_M^{(0)}$  can be written as

$$\begin{aligned} \mathfrak{S}_M^{(0)}(s) = & \frac{4\sin(\pi s)}{\Gamma(2N)\Gamma(2M+1)} \rho_1^{2s} [ \\ & \frac{(-1)^{M+s+1}}{2} \sum_{n=1}^M (-1)^n c_n B(s-2n+1, 1-s) \kappa^{2n+1} I_1 \\ & + 2(-1)^M \kappa^{2s} \sum_{k=k_0+1}^{\infty} k(k^2-(\frac{1}{2})^2) \dots (k^2-(N-\frac{3}{2})^2) I_2 \\ & + 2(-1)^M \kappa^{2s} \sum_{k=N-\frac{1}{2}}^{k_0} k(k^2-(\frac{1}{2})^2) \dots (k^2-(N-\frac{3}{2})^2) I_3 \\ & - \kappa^{2s} \sum_{k=N-\frac{1}{2}}^{k_0} k(k^2-(\frac{1}{2})^2) \dots (k^2-(N-\frac{3}{2})^2) I_4 \\ & + (\text{vanishing terms as } s \rightarrow \text{integers}) ] , \end{aligned}$$

where

$$I_1 = P \int_0^{\rho_1 \rho_2} dx x(x^2-(\frac{1}{2})^2) \dots (x^2-(N-\frac{3}{2})^2) [1^2 - x^2]^{2M+1-s} \tan(\pi x) ,$$

$$I_2 = \int_{y_k^+}^{\infty} dy y^2(y^2+1) \dots (y^2+(M-1)^2) [y^2-(y_k^+)^2]^{-s} \frac{1}{e^{2\pi} y-1} ,$$

$$I_3 = \int_0^{\infty} dy y^2(y^2+1) \dots (y^2+(M-1)^2) [y^2+(y_k^-)^2]^{-s} \frac{1}{e^{2\pi} y-1} ,$$

$$I_4 = \int_0^{y_k^-} dy y^2(y^2-1) \dots (y^2-(M-1)^2) [(y_k^-)^2 - y^2]^{-s} \cot(\pi y) ,$$

$$\rho_m = [ (N-\frac{1}{2})^2 / \rho_1^2 + M^2 / \rho_2^2 ]^{1/2} ,$$

$$y_k^+ = \kappa [k^2 - (\rho_1 \rho_m)^2]^{1/2} ,$$

$$y_k^- = \kappa [(\rho_1 \rho_m)^2 - k^2]^{1/2} ,$$

$$c_n : \sum_{n=1}^M y^{2n} c_n = y^2(y^2+1) \dots (y^2+(M-1)^2) ,$$

and  $k_0$  is the maximum half integer less than  $\rho_1 \rho_m$ . The other  $\mathfrak{S}$ -functions  $\mathfrak{S}_M^{(1/2)}(s)$ ,  $\mathfrak{S}_q^{(0)}(s)$ , and  $\mathfrak{S}_q^{(1/2)}(s)$  can be written as similar expressions.

With these finite sum, the effective action and the renormalization terms are expressed as the follows,

$$\begin{aligned} \frac{J(\kappa)}{\rho_1^4} = & \frac{n_b}{64\pi^2} [ \mathfrak{S}_M^{(0)}(-2) + (3/2 + \ln \mu^2) \mathfrak{S}_M^{(0)}(-2) ] \\ & - \frac{n_f}{16\pi^2} [ \mathfrak{S}_M^{(1/2)}(-2) + (3/2 + \ln \mu^2) \mathfrak{S}_M^{(1/2)}(-2) ] , \end{aligned}$$

$$\frac{E}{\beta_1^4} = -\frac{n_b}{192\pi^2} [ \zeta_M^{(0)}(-1) + (1 + \ln \mu^2) \zeta_M^{(0)}(-1) ]$$

$$- \frac{n_f}{96\pi^2} [ \zeta_M^{(1/2)}(-1) + (1 + \ln \mu^2) \zeta_M^{(1/2)}(-1) ] ,$$

$$g^2 D = \frac{n_b}{96\pi^2} [ \zeta_q^{(0)}(0) + \zeta_q^{(0)}(0) \ln \mu^2 ]$$

$$+ \frac{n_f}{12\pi^2} [ \zeta_q^{(1/2)}(0) + \zeta_q^{(1/2)}(0) \ln \mu^2 ] .$$

## 6. Results and Discussion

We calculated the  $\zeta$ -functions  $\zeta_M^{(0)}(s)$ ,  $\zeta_M^{(1/2)}(s)$ ,  $\zeta_q^{(0)}(s)$  and  $\zeta_q^{(1/2)}(s)$  numerically, and obtained  $J(\kappa)$ ,  $E$  and  $D_e$ . The results of calculation of  $-J(\kappa)$  are schematically presented in Fig. 1. for bosons and fermions.

The solutions are not stable, if the matter field is either boson or fermion. However, in some types of the extra spaces, appropriate fermion-boson ratio  $n_f/n_b$  makes them stable. The stable solutions are presented in the Table (1). These satisfy the all conditions to be stable. The scales of the extraspace are as small as in the case that the space is  $M_4 \times S_N$  type manifold.<sup>4)</sup> So there must be much many matters to get reasonably small values for the gauge couplings.

We have found also some solutions that give the same order scale of extraspace as the Planck length by a little number of matters in the cases of  $M_4 \times S_8 \times S_5$  and  $M_4 \times S_{10} \times S_5$ . Although these are locally stable, they are not globally stable.

It looks like that this type model which makes balance between the curvatures of extraspace and the quantum fluctuations of matters, always need enormous number of matter fields.

After this calculation of the effective action was made, we found that a contribution from the Pauli interaction term of fermion was missing, which exists in the fermion action in 4-dimensions when reduced from the higher dimensions in Kaluza-Klein theory. Hence our conclusions may be modified if the term makes large contributions to the renormalizations of gauge coupling constants. These contributions are now being investigated.

## References

- 1) T.Kaluza, Sitz. Preuss. Akad. Wiss. Berlin, Math. Phys. K1(1921),966.  
O. Klein, A.Phys. 37 (1926), 895.  
A. Salam and J. Strathdee, Ann. Phys. 141 (1982), 316
- 2) P.G.O. Freund and M.A. Rubbin, Phys. Lett. 97B (1980), 233  
See also, F. Englert, Phys. Lett. 119B (1982), 339
- 3) T. Appelquist and A. Chodos, Phys. Rev. Lett. (1983), 141; Phys. Rev. D28 (1983), 772;  
For the case including finite temperature; M.A. Rubin and B.D. Roth, Nucl. Phys. B226 (1983), 444.  
For the case of  $M_4 \times T_n$ , T. Appelquist, A. Chodos and E. Myers, Phys. Lett. 127B (1983), 51 ;  
T. Inami and O. Yasuda, Phys. Lett. 133b (1983), 180.
- 4) P. Candelas and S. Weinberg, Preprint UTG-6-83 (1982) (Revised)
- 5) E. Witten, Nucl. Phys. B186 (1981), 412
- 6) S. Weinberg, Phys. Lett. 125B (1983), 55
- 7) See, e.g., N.D. Birrel and P.C.W. Davies, "Quantum Fields in Curved Space" (Cambridge University press, 1982)
- 8) D.J. Toms, Phys. Lett. 129B (1983), 31



**Table 1.** The values of parameters of the stable solutions a boson field when the fermion-boson ratio is fixed. The items with an asterisk are propotional to the number of boson fields.

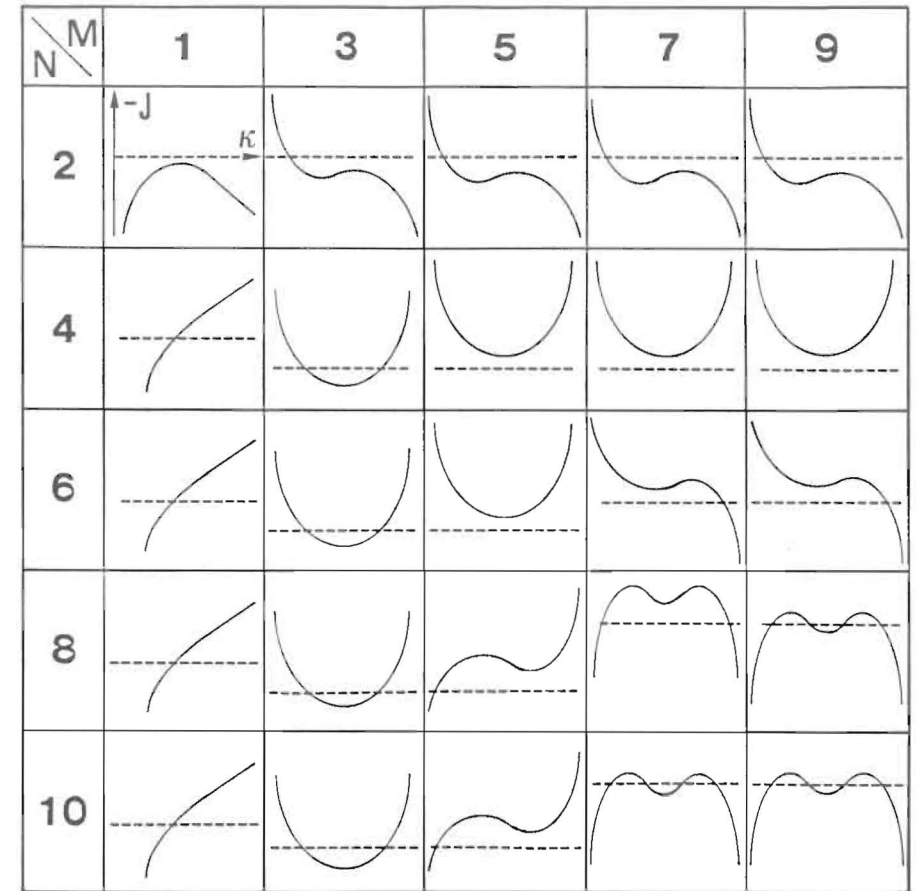
**Fig. 1.** The contribution to  $-J(\kappa)$  due to each boson (upper graph) and fermion (bottom graph) in the case of  $M_4 \times S_{2N} \times S_{2M+1}$ . The dashed line corresponds to  $J = 0$ . The behaviour of fermion contribution is the same for the case of  $N = 2, 4, 6, 8$  and  $10$ . The matter part of the effective potential  $-J(\kappa)/(\mathcal{F}_1)^4$  is given by superposing the upper and bottom graphs weighted by  $n_b$  and  $n_f$ , respectively.

External Space solutions	$S^1 \times S^4$	$S^1 \times S^6$	$S^1 \times S^8$	$S^1 \times S^{10}$
$n_f/n_b$	2.47	1.41	1.06	1.41
$\kappa$	0.581	0.459	0.394	0.367
$f$	1.00	1.00	1.00	1.00
$J$ *	-3.97	-11.3	-18.6	-32.3
$E$ *	2.67	4.55	6.77	8.68
$\mathcal{F}_1^2/16 G$ *	4.16	6.61	9.04	11.4
$D_{SO(2N+1)}$ *	2.03	4.28	7.82	11.2
$1/g_{SO(2N+1)}^2$ *	2.33	7.23	8.06	11.4
$D_{SO(2M+2)}$ *	36.9	39.7	43.7	48.7
$1/g_{SO(2M+2)}^2$ *	37.6	40.7	44.8	50.0

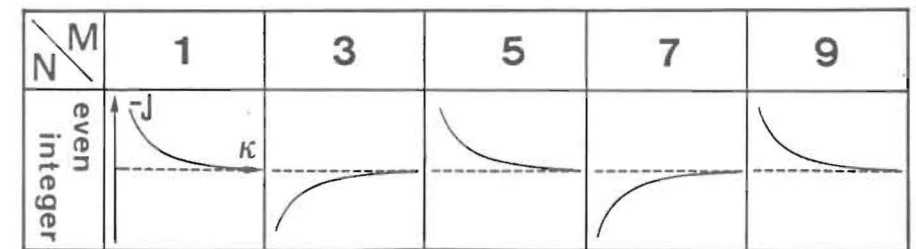
\*)  $\times 10^{-4}$

**Table 1.**

**BOSON**



**FERMION**



**Fig. 1.**

## Fermions in a Modified Kaluza-Klein Model

T. Maehara, T. Muta, J. Saito and K. Shimizu

Department of Physics

Hiroshima University

Hiroshima 730

The difficulty of introducing light fermions in the Kaluza-Klein scheme has been occasionally emphasized by many authors;<sup>1,2,3)</sup> the mass operator  $M$  which arises from the kinetic term in extra dimensions satisfies the relation  $M^2 \propto \tilde{R}$  where  $\tilde{R}$  is the curvature scalar of the extra space.<sup>4)</sup> Thus all the fermions in 4 dimensions acquire masses of order of the Planck mass  $M_{Pl}$  if the extra space is compactified with simple topology for which  $\tilde{R} > 0$ .

Several ways out of the difficulty have been known: 1) In 11-dimensional supergravity, spin-1/2 fields appear as 4-dimensional subcomponents of the Rarita-Schwinger field by the dimensional reduction and do not suffer from the above difficulty. 2) It is possible to introduce an extra matter gauge field so that the relation  $M^2 \propto \tilde{R}$  is modified and some zero-mass modes may be obtained.<sup>5)</sup> 3) By introducing torsion on the extra space one may obtain massless modes.<sup>6)</sup>

In the present paper we look for yet another possibility out of the fermion mass problem in the Kaluza-Klein scheme. For this purpose we consider the problem of introducing fermions in

a modified Kaluza-Klein model recently proposed by Rubakov and Shaposhnikov.<sup>7)</sup> In this model the metric tensor  $g_{MN}$  in  $D$  dimensions<sup>\*</sup>) is, after compactification, identified with the 4-dimensional metric  $g_{\mu\nu}$  with the factor  $\sigma(x^\ell)$  which may depend on extra variables  $x^\ell$  ( $\ell=4, 5, \dots, D-1$ ),

$$\hat{g}_{\mu\nu}(x^L) = \sigma(x^\ell) g_{\mu\nu}(x^\lambda), \quad (1)$$

where gauge fields are assumed to be absent for simplicity.

The Dirac equation for fermions in this model takes the form,

$$\hat{\gamma}^M \hat{\nabla}_M \psi = 0. \quad (2)$$

According to Eq.(1) the vielbein  $\hat{e}_M^A$  satisfies the following relations,<sup>\*\*)</sup>

$$\hat{e}_\mu^\alpha = \sqrt{\sigma} e_\mu^\alpha, \quad \hat{e}_\mu^a = 0, \quad \hat{e}_m^\alpha = 0, \quad e_m^a = \tilde{e}_m^a(x^\ell). \quad (3)$$

Hence operator  $\hat{\gamma}^M \hat{\nabla}_M$  in Eq.(2) reduces to

---

\*) Capital Latin letters  $L, M, N, \dots$  run from 0 to  $D-1$ , Greek letters  $\lambda, \mu, \nu, \dots$  from 0 to 3 and small Latin letters  $\ell, m, n$ , from 4 to  $D-1$ .

\*\*) Early Latin and Greek letters refer to the local Lorentz frame.

$$\hat{\gamma}^M \hat{\nabla}_M = \frac{1}{\sqrt{\sigma}} \gamma^\mu \nabla_\mu + \tilde{\gamma}^m \tilde{\nabla}_m . \quad (4)$$

Substituting Eq.(4) in Eq.(2) and rewriting Eq.(2) in the 4-dimensional form we find that the mass operator M for the 4-dimensional fermion fields is given by

$$M = -i\sqrt{\sigma} \tilde{\gamma}^m \tilde{\nabla}_m \quad (5)$$

We square the mass operator (5) to have the following form,

$$M^2 = \sigma \left( \frac{\tilde{R}}{4} - \tilde{\nabla}^m \tilde{\nabla}_m \right) - \frac{1}{2} \tilde{\gamma}^m \tilde{\gamma}^n (\tilde{\nabla}_m \sigma) \tilde{\nabla}_n . \quad (6)$$

In the model under consideration the factor  $\sigma$  is not an arbitrary function of  $x^\ell$  but is constrained by equations arising from the D-dimensional Einstein equation,

$$\hat{R}_{MN} - \frac{1}{2} \hat{g}_{MN} \hat{R} = \Lambda \hat{g}_{MN} + c \hat{T}_{MN} , \quad (7)$$

where  $\Lambda$  is the cosmological constant in D dimensions and  $\hat{T}_{MN}$  the D-dimensional energy-momentum tensor of the matter field  $\psi$  with coupling strength  $c$ . The cosmological constant  $\Lambda$  was introduced in the original paper by Rubakov and Shaposhnikov<sup>7)</sup> for the purpose of obtaining a small effective 4-dimensional cosmological constant  $\Lambda_P$ . In Eq.(7) the constant  $c$  is supposed to be small and the term  $cT_{MN}$  may be neglected as the first approximation.

Then Eq.(7) is essentially the same as the one considered in Ref.7. After the spontaneous compactification the 4-dimensional part of Eq.(7) reduces to the ordinary Einstein equation with the effective cosmological constant  $\Lambda_P$ . The extra-dimensional part of Eq.(7) reduces to the equation

$$\tilde{R}_{mn} = -\frac{2}{D-2} \Lambda \tilde{g}_{mn} + \frac{1}{\sigma} \{ 2(\tilde{\nabla}_m \tilde{\nabla}_n \sigma) - \frac{1}{\sigma} (\tilde{\nabla}_m \sigma) (\tilde{\nabla}_n \sigma) \} , \quad (8)$$

and the definition of  $\Lambda_P$  places a constraint on  $\sigma$ ,

$$(\tilde{\nabla}^m \tilde{\nabla}_m \sigma) + \frac{1}{\sigma} (\tilde{\nabla}^m \sigma) (\tilde{\nabla}_m \sigma) = \frac{4\Lambda\sigma}{D-2} - 2\Lambda_P . \quad (9)$$

From Eqs.(8) and (9) we obtain

$$\tilde{R} = 2 \frac{8-D}{D-2} \Lambda - \frac{4}{\sigma} \Lambda_P - \frac{3}{\sigma} (\tilde{\nabla}^m \sigma) (\tilde{\nabla}_m \sigma) . \quad (10)$$

If the extra space is compactified to the size of the Planck length and  $R$  is of order  $M_{Pl}^{-2}$ , then  $\Lambda$  is of order  $M_{Pl}^{-2}$  as long as  $\sigma$  is of the normal size.

If  $\sigma$  and  $\tilde{\nabla}_m \sigma$  are extremely small, the eigenvalues of  $M^2$  may also be very small, and we have effectively light fermions. Let us examine whether there is such a solution to Eqs.(8) and (9) which has the above property. For simplicity we restrict ourselves to the case  $D = 6$  which was fully studied in Ref.7. We look for a solution of the form,

$$\tilde{g}_{44} = -1, \quad \tilde{g}_{55} = -f(\theta), \quad \sigma = \sigma(\theta), \quad (11)$$

where we set  $x^4 = r\theta$  and  $x^5 = r\phi$  with  $r$  the length scale and  $0 \leq \theta \leq \pi$ ;  $0 \leq \phi < 2\pi$ . In this case Eqs.(8), and (9) reduce to

$$f(\theta) = \frac{1}{\sigma} \left( \frac{d\sigma}{d\theta} \right)^2, \quad (12)$$

$$\frac{d^2\sigma}{d\theta^2} + \frac{1}{4\sigma} \left( \frac{d\sigma}{d\theta} \right)^2 + \frac{r^2}{2} \Lambda \sigma - r^2 \Lambda_P = 0. \quad (13)$$

As noted in Ref.7, Eq.(13) may be rewritten in the form familiar in the classical mechanics:

$$\frac{d^2z}{d\theta^2} = - \frac{\partial U}{\partial z}, \quad (14)$$

with  $z = \sigma^{5/4}$  and

$$U(z) = \frac{15}{16} r^2 \Lambda z^{6/5} - \frac{10}{3} \frac{\Lambda_P}{\Lambda} z. \quad (15)$$

The shape of the "potential"  $U(z)$  depends on the sign of  $\Lambda$  and  $\Lambda_P$ . If  $\Lambda < 0$ , we do not have "stable" solutions except for the one restricted to the region near  $z=0$  for  $\Lambda_P < 0$ . If  $\Lambda > 0$ , there are "stable" solutions both for  $\Lambda_P \leq 0$  and  $\Lambda_P > 0$ .

We examine whether there is a solution to Eq.(14) which has  $S^2$  symmetry, i.e.,

$$f(\theta) = \sin^2 \theta. \quad (16)$$

Substituting Eq.(16) into Eq.(12) we have

$$\sigma(\theta) = \sigma(0) \cos^2 \theta, \quad (17)$$

where the constant  $\sigma(0)$  is to be determined later. We insert  $\sigma(\theta)$  of Eq.(17) in Eq.(14) and see whether this form of solutions is allowed. We then find that the solution (17) is allowed if

$$\sigma(0) = \frac{10\Lambda_P}{3\Lambda}; \quad r^2 \Lambda = 10. \quad (18)$$

Let us see what consequence may be obtained on the mass operator squared (6) with the above solution. For the solution (17) with Eq.(18), the curvature scalar  $\tilde{R}$  for the extra $\checkmark$ dimensional space reads

$$\tilde{R} = \frac{\Lambda}{5}. \quad (19)$$

Substituting Eqs.(17) and (19) into Eq.(6) we obtain

$$M^2 = 2\Lambda_P \left( -\frac{1}{3} \cos^2 \theta - \frac{5}{3\Lambda} \cos^2 \theta \tilde{\nabla}_m^2 - \frac{1}{3\sqrt{5}\Lambda} \sin \theta \cos \theta \tilde{\nabla}^4 \tilde{\nabla}_n^2 \right). \quad (20)$$

Obviously the mass operator squared given by Eq.(20) is of order  $\Lambda_P$  which is extremely small experimentally, and so the eigenvalues of  $M^2$  should be negligibly small. Thus we have light fermions in this solution.

Unfortunately, however, the  $S^2$  symmetry which our solution enjoys does not manifest itself in the 4-dimensional space because the factor  $\sigma$  relating  $\hat{g}_{MN}$  to  $g_{\mu\nu}$  does depend on the extra variable  $\theta$ . Hence we can not utilize spherical harmonics on  $S^2$  to expand fermion fields  $\psi(x^M)$ . At any rate all the eigenmodes of Eq.(20) have their mass of order  $\Lambda_P^{1/2}$  and thus we have proliferation of light fermions. This conclusion seems to reflect the fact that the size of the extra space in our solution, even though it is of order of the Planck length, is not practically small when viewed from the 4-dimensional world.

The difficulty that the symmetry of the extra space is not reflected in the 4-dimensional space is a rather common feature of the Rubakov-Shaposhnikov model and may become more serious when we take into account gauge fields which we have been neglecting. A possible way out of the difficulty may be to go over to the higher extra dimensions where one may find a solution whose symmetry property is partially realized in the 4 dimensions.

#### Acknowledgements

The authors would like to thank Prof. S. Tanaka and Prof. K. Kikkawa for useful conversations and the members of our laboratory for encouraging discussions.

#### References

- 1) A. Zee, Proc. 4th Kyoto Summer Institute on Grand Unified Theories and Related Topics (ed. M. Konuma and T. Maskawa, World Scientific, Singapore, 1981), p.143.
- 2) M.J. Duff and C.N. Pope, "Supersymmetry and Supergravity '82" p.193.  
S. Tanaka, Prog. Theor. Phys. 70(1982) 563.
- 3) W. Mecklenberg, Phys. Rev. D26(1982) 1327.
- 4) E. Schrödinger, Sitzungsber. Phys. Math. Kl 25(1932) 102.  
A. Lichnerowicz, Proceedings of the 1963 Les Houches Summer School "Relativity, Groups and Topology" p.849.
- 5) S. Randjbar-Daemi, A. Salam and J. Strathdee, Nucl. Phys. B214(1983) 491; Phys. Lett. 132B(1983) 56.
- 6) C. Destri, C.A. Orzalesi and P. Rossi, Ann. Phys. 147(1983) 321.  
Y.S.Wu and A. Zee, Univ. of Washington Preprint 40048-25 P3.
- 7) A. Rubakov and N.E. Shaposhnikov, Phys. Lett. 125B(1983) 139.

Composite Quasi Nambu-Goldstone Fermions

T. Yanagida

Physics Department, College of General Education  
Tohoku University, Sendai 980

Recently there has been a growing interest in composite models of quarks and leptons based on supersymmetric confining forces.<sup>1)-6)</sup> The reason for this interest is that the supersymmetry provides us with new kinds of dynamical possibilities to generate massless composite fermions; namely Nambu-Goldstone fermions arising from the spontaneous breakdown of the supersymmetry<sup>1)</sup> or quasi Nambu-Goldstone fermions which are fermion superpartners of Nambu-Goldstone bosons.<sup>2)</sup> In this talk we will concentrate on the latter case.

If a global symmetry  $G$  in the theory is broken down to some subgroup  $H$  -for instance by preon condensation-, then there will necessarily appear massless fermion boundstates as superpartners of Nambu-Goldstone bosons (called quasi Nambu-Goldstone fermions) as long as the supersymmetry is kept unbroken. It seems quite natural that such massless particles are not only Nambu-Goldstone ones but also composite objects of underlying preons.

However, it seems likely that the quasi Nambu-Goldstone (N-G) fermions become massive when the supersymmetry breaking is turned on and will disappear in low-energy spectrum. Fortunately, it is not always the case. The number and quantum numbers of quasi N-G fermions depend crucially on particular dynamical constraints.<sup>3,7)</sup>

Especially if the coset space  $G/H$  itself is a Kähler manifold, all quasi N-G fermions can be complex (chiral)<sup>8)</sup> under the unbroken subgroup  $H$  and remain massless, independent of the supersymmetry breaking scale, since they are protected from obtaining the masses by also the unbroken  $H$  symmetry.

Those massless fermions may be identified with the usual quarks and leptons if the  $H$  contains the standard gauge group  $SU(3)_C \times SU(2)_L \times U(1)$ . In this talk, we will show some examples where the quasi N-G fermions have precisely required  $SU(3)_C \times SU(2)_L \times U(1)$  quantum numbers for the observed quarks and leptons. This talk is basically a brief review of the recent works in collaboration with Buchmüller and Peccei<sup>3,4)</sup>

Our first example is based on a  $U(6)$  global symmetry since the Kähler manifold  $U(6)/U(4) \times U(2)$  is known to be the smallest coset space<sup>3,6)</sup> accommodating the 8 left-handed quarks and leptons of one generation as the quasi N-G fermions. The simplest way to realize the global  $U(6)$  symmetry in a preon model is to assume 6 preon chiral multiplets  $\chi_\alpha^i$  ( $i=1 \sim 6$ ) that transform as the fundamental representation of a hypercolor gauge group  $G_{HC}$ . Here  $\alpha$  is the hypercolor index. The minimal candidate for anomaly free gauge groups with this type of matter fields is  $SU(2)$  and thus we choose  $G_{HC} = SU(2)$ .

Let us now discuss some details of this preon model.<sup>4)</sup> The total symmetry group is

$$G = SU(2)_{HC} \times U(6) \times U(1)_R ,$$

where the last  $U(1)_R$  represents R-invariance. The standard strong and electroweak gauge group,  $SU(3)_C \times SU(2)_L \times U(1)$ , will be embedded into the  $U(6)$ . Therefore, the  $U(6)$  is the approximate global symmetry which is realized only when all gauge coupling constants of the flavour groups are taken to be zero.

What is the symmetry breaking pattern? The simplest hypercolor-invariant composite operator constructed from the preon multiplets  $\chi_\alpha^i$  is given by

$$\phi^{ij} = \epsilon_{\alpha\beta} (\chi_\alpha^i \chi_\beta^j - \chi_\alpha^j \chi_\beta^i)$$

which transforms as the antisymmetric  $\underline{15}$  of  $U(6)$ . If the  $\phi^{ij}$  has the vacuum expectation values, the following two breaking patterns are possible:

$$(a) \quad U(6) \rightarrow U(4) \times SU(2)$$

$$(b) \quad U(6) \rightarrow SP(6)$$

The case (b) is not interesting since the  $SU(3)_C \times SU(2)_L \times U(1)$  gauge group can not be embedded in  $SP(6)$ . We will, therefore, concentrate our discussion on the case (a), although it is not clear to us which option is realized. Perhaps the breaking (a) may be more economically favored than the breaking (b) because the  $SU(3)_C$ -charge condensation is required additionally in the case of (b).

Before proceeding further, we would like to remark that the breaking  $U(6) \rightarrow U(4) \times SU(2)$  is the only one allowed in the complementary picture of this model. Our preon model contains scalar fields (the first components of  $\chi_\alpha^i$ ) in the fundamental representation of the hypercolor group. It is believed in such theories that there is no phase transition between the confining and Higgs phases.<sup>1</sup> If it is the case, certain features of a confining theory may be analyzed by studying the complementary Higgs picture. We have found<sup>4)</sup> that there exist only one Higgs phase which corresponds to the symmetry breaking  $U(6) \rightarrow U(4) \times SU(2)$ , in the complementary picture of our model.

Corresponding to the spontaneous breakdown of  $U(6) \rightarrow U(4) \times SU(2)$ , we have 9 Nambu-Goldstone chiral multiplets  $\phi^{ai}$  ( $a=1 \sim 4$ ,  $i=5,6$ ) and  $\phi = \phi^{56}$  which transform as  $(4, 2)$  and  $(\underline{1}, \underline{1})$  under the unbroken  $U(4) \times SU(2)$ . It is easy to check that the fermion components of  $\phi^{ai}$  have precisely the required quantum numbers with respect to the  $SU(3)_C \times SU(2)_L \times U(1)$  for the usual 8 left-handed quarks and leptons of one generation and  $\phi$  is called as "novino". Here we have identified the unbroken  $SU(4)$  with the Pati-Salam like one and obviously the  $SU(2)$  with the weak gauge group.

It is gratifying that our breaking pattern is perfectly consistent with the 'tHooft's anomaly matching conditions. All chiral anomalies of the unbroken groups  $U(4) \times SU(2) \times U(1)_R$  at the preon level are completely saturated with only the quasi N-G fermions  $\psi^{ai}$  and  $\psi$ .<sup>4)</sup> Therefore, we do not need any additional massless chiral multiplets other than the Nambu-Goldstone modes. The masslessness of the quasi N-G fermions is guaranteed here by the Nambu-Goldstone nature of their scalar superpartners and also by the unbroken chiral symmetry. This double protection is important to keep the quasi N-G fermions massless even after the supersymmetry breaking.<sup>5)</sup> (See Ref.(4) for the detailed argument on the anomaly-matching.)

The residual interactions of the quasi Nambu-Goldstone fermions are identifiable with the weak interactions in this models.<sup>4)</sup> Thus the weak bosons observed at CERN may not be fundamental gauge particles but rather boundstates of preons like rho-mesons in QCD.

The straightforward way of including composite right-handed quarks and leptons is to replace the hypercolor  $SU(2)_{HC}$  by  $SU(2) \times SU(2)$  and introduce new matter multiplets  $\chi_\alpha^{ij}$  ( $j=1 \sim 6$ ) transforming as doublets of the second  $SU(2)$  in addition to the  $\chi_\alpha^i$ . Similarly to the previous case, the global symmetry  $U(6) \times U(6)$  is presumably broken down to  $[U(4) \times SU(2)]^2$ , which produces 16 quarks and leptons as the quasi NG fermions. However, it is very difficult to cause the breakdown of the electroweak symmetry in this model. If we add another preon field  $\chi^{\alpha\beta} = (2, 2)$ , 3-body bound-states  $H^{ij} = \chi_\alpha^i \chi^\alpha \chi_\beta^j$  ( $i=5,6$ ) may play a role of the Higgs chiral multiplets responsible for the  $SU(2)_L \times U(1)$  breaking. We need, however, direct Yukawa couplings of the preon  $\chi^{\alpha\beta}$  to generate effective ones of  $H^{ij}$  at the bound-state level.

In this point there is a more natural model, though the situation of dynamics is more puzzling to us. The model is based on the  $SU(4)$  hypercolor group, where the matter fields  $\chi_\alpha^i$  and  $\chi_i^\alpha$  ( $i=1 \sim 6$ ) are  $\underline{4}$  and  $\underline{4}^*$  of the  $SU(4)_{HC}$ . The complementary picture of this model suggests the breaking of the global group to be  $U(6) \times U(6) \rightarrow U(4) \times U(2)_L \times U(2)_R$ , which corresponds to the following 2-body condensation in the confining phase;

$$\chi_\alpha^i \chi_j^\alpha = \begin{pmatrix} v & & & & & \\ & v & & & & \\ & & v & & & \\ & & & v & & \\ & & & & v & \\ & & & & & 0 \\ & & & & & & 0 \end{pmatrix}$$

A small supersymmetry breaking will trigger the further small breakdown of the unbroken symmetry,  $U(2)_L \times U(2)_R$ , which is induced by, for instance, 2-body fermion-fermion condensation  $\psi_\alpha^i \psi_j^\alpha = v \delta_{ij}$ . The breaking pattern

$$U(6) \times U(6) \rightarrow U(4) \times U(2)_L \times U(2)_R \\ \rightarrow U(4) \times U(2)$$

may be the desirable one for constructing a realistic model, although whether such a breaking does occur or not is not obvious at all.

The effective Yukawa couplings (quark-lepton masses) are simply given by the 'tHooft planer diagram. However, we are still far away from being able to explain the observed mass spectrum of quarks and leptons even within one generation, because, for example, no other symmetry survives at all than the gauge  $SU(3)_C \times U(1)_{em}$  in the quark-lepton mass matrices. Who believes such a complicate breaking starting with the  $U(6) \times U(6)$ ? Do  $SU(3)_C$  interactions help us? In any way a more elaborate study on this problem is required toward the challenging calculation of the fermion masses.

#### References

- 1) H. Terazawa, Prog. Teor. Phys. 64(1980)1763.  
W.A. Bardeen and V. Visnjic, Nucl. Phys. B194(1982)422.
- 2) W. Buchmüller, R.D. Peccei and T. Yanagida, Physics Lett. 124B (1983) 67; see also W. Buchmüller, S.T. Love, R.D. Peccei and T. Yanagida, Physics Lett. 115B (1982) 233.
- 3) W. Buchmüller, R.D. Peccei and T. Yanagida, MPI-PAE/PTh 28/83 (to be published in Nucl. Phys.).
- 4) W. Buchmüller, R.D. Peccei and T. Yanagida, MPI-PAE/PTh 41/83 (to be published in Nucl. Phys.).
- 5) R. Barbieri, A. Masiero and G. Veneziano, CERN-TH-3561 (1983).
- 6) O.W. Greenberg, R.N. Mohapatra and M. Yasuè, University of Maryland preprint (1983).  
F. Bordi, R. Casalbuoni, D. Dominici and R. Gatto, UGVA-DTP 1983/04-387 (1983).  
P.H. Frampton and G. Mandelbaum, HUTP-83/A042 (1983).
- 7) G. 'tHooft, in Recent Developments in Gauge Theories, Cargèse 1979, eds. G. 'tHooft et al. (Plenum, New York and London).
- 8) K. Osterwalder and E. Seiler, Ann. Phys. 110 (1978) 440.  
E. Fradkin and S. Shenker, Phys. Rev. D19 (1979) 3682.  
T. Banks and E. Rabinovici, Nucl. Phys. B160 (1979) 349.
- 9) B. Schrepp and F. Schrepp, DESY 83-024 (1983).
- 10) W. Buchmüller, R.D. Peccei and T. Yanagida, unpublished (1983).